UNIVERSITÉ DE STRASBOURG

ÉCOLE DOCTORALE DES SCIENCES DE LA VIE ET DE LA SANTE

CNRS – IBMC UPR 9002

Université de Leipzig

THÈSE présentée par :

Kévin ROLLET

soutenue le : **17 mars 2022**

pour obtenir le grade de : **Docteur de l'université de Strasbourg**

Discipline/ Spécialité : Biologie structurale

From functional study of tRNA CCA-adding enzymes toward their characterization by serial X-ray crystallography

THÈSE co-dirigée par :

M SAUTER Claude	Directeur de recherches, Université de Strasbourg
M MÖRL Mario	Professeur des universités, Université de Leipzig

RAPPORTEURS :

Mme LE DU Marie-Hélène	Directrice de recherches, Université Paris-Saclay
M SCHNEIDER André	Professeur des universités, Université de Bern

---

AUTRES MEMBRES DU JURY :

M SCHULTZ Patrick	Directeur de recherches, Université de Strasbourg
Mme COIN Irene	Professeur des universités, Université de Leipzig

From functional study of tRNA CCA-adding enzymes toward their characterization by serial X-ray crystallography

Von der Fakultät für Lebenswissenschaften

der Universität Leipzig

genehmigte

D I S S E R T A T I O N

zur Erlangung des akademischen Grades

Doctor rerum naturalium

Dr. rer. nat

vorgelegt

von Herr Master of Science Kévin Rollet

geboren am 30.12.1994 in Mulhouse

Strasbourg, den 16.02.2022

Acknowledgments / Remerciements

La thèse, quelle aventure !! Dix ans après le BAC, je quitte enfin les rangs de l'école. Un tel parcours n'est possible qu'avec des rencontres, des amitiés, des échanges, des collaborations… Tellement de personnes à remercier, d'anecdotes à raconter ou de discussions à évoquer pour clore ce chapitre des études et plus particulièrement celui de la thèse. La pandémie de COVID-19 aura bien entendu impacté ce projet. Par chance, mon entourage, autant professionnel que personnel, aura permis de traverser cette période particulière en gardant la tête haute et en regardant toujours vers l'avenir avec espoir.

Pour commencer, Claude. Tu m'as accueilli à bras ouverts pour mon stage de master où tu m'as initié au monde de la cristallographie biologique avec Raphaël. Tu m'as ensuite proposé ce sujet novateur pour la thèse et tu m'as fait confiance pour le mener à bien. Tu m'as transmis cette passion qui doit animer chaque chercheur, celle qui entretient en permanence la curiosité et dont on a plaisir à partager avec la communauté scientifique ou le public. Au-delà d'un directeur de thèse, c'est une personne, ou plutôt une personnalité, unique avec laquelle j'ai eu la chance d'évoluer. Du fond du cœur, je te remercie pour ta bienveillance, ton enthousiasme inébranlable et ton enseignement.

Mario, I would like to thank you for welcoming me in your team and co-supervising this ambitious project. From the very first day we met, you were friendly, helpful and you always ensured everything was working fine. In Leipzig, I could benefit not only from your exceptional scientific expertise but also I could enjoy conversations regarding many different topics. Unfortunately, the pandemic did not permit the second period in you lab but be sure that I would love to come back if I could!

Great working atmosphere is essential to mature a project and learn good practice. I would like to thank members of my both teams from Strasbourg and Leipzig. Magali, you accepted me in the team and helped me a lot, especially before my defense when COVID kept members of my jury at home. Philippe, Joëlle, Anne, Caroline, many thanks for your patience, explanations, and good atmosphere at the 4th floor. Bernard, even you suddenly left the lab, thank you for (long) stories about science or various topics. Raphaël (my Jedi master of serial crystallograhy), Marta and José, your were like my "tutors" when I arrived to do my internship and more than that you became my friends. It was you who convinced me to sign for a PhD contract ("it's a trap!") and that become one of the best experience in my life! Thank you.

From the other side of the Rhine river, I would like to thank all the “Mörlies”. Heike, you have been such a caring person, always present for me (and for all your students), from my first experiments until my defense. Susanne, Marie, Christina, Karolin, Sonja, Richard, Liselotte, and Janett, it has been a pleasure to share the lab and its good atmosphere with you. Alexandra, we had started this work together in Strasbourg (many jokes, my funny mistakes…) and then we met again in Leipzig. Thank you for all the nice moments and coming back to Strasbourg especially for my defense. In Hamburg, special thoughts for Arwen (and her team) who brought new perspective on to my PhD project and my future job.

Je tiens également à remercier Joern pour les opportunités et le soutien que tu m’as apporté, de mon départ à Montréal en 2015 aux formalités administratives franco-allemandes liées à ma thèse. Un grand merci également à Danièle, Anaïs et Léandra pour avoir géré mes nombreux déplacements (congrès, formations, synchrotron etc.).

Léna, je pourrais te consacrer un chapitre entier. D’une simple rencontre en master, tu es devenue mon amie et ma collègue inséparable de labo. Ton humour, tes conseils, nos interminables conversations ont été un moteur essentiel de ces quatre dernières années. Mention spéciale pour la préparation de la soutenance (répétitions, cadeaux, coordination, etc...), tu as assuré !! J’ai entièrement confiance en toi pour la suite de tes projets =D.

À ma famille, qui a toujours cru en moi, vous m’avez appuyé dans mes décisions, vous avez toujours été présents pour moi, aussi bien dans les plus dures épreuves que dans les moments de joie. Maman, Papa, Morgane, je n’y serais jamais parvenu sans vous, merci pour tout. Mes chères tantes et mon cousin que je n’oublie pas et qui ont fait le déplacement pour ma soutenance, merci ! Enfin, à mes deux grand-pères qui auraient été fiers. Papi Claude, j’ai bien retenu ce que tu me répétais : « tu n’es pas plus bête qu’un autre ! ».

Pour garder le moral et prendre du recul, rien ne vaut des amis en or. Wylliam, mon ami de longue date, malgré la distance, nous avons gardé cette amitié indéfectible. À toute l’équipe de «l’étang», en particulier Alexis, je vous remercie pour la bulle d’oxygène et les fous rires qui m’ont fait repartir du bon pied. Une autre équipe que je me dois forcément de remercier, mon clan viking Vestrgarðr, ma seconde famille. Valentin, qui est comme mon grand-frère, avec Eir et Bjorn, vous m’avez fait une place dans ce groupe qui m’a poussé à me battre (au sens propre comme figuré) pour atteindre mes objectifs. Et puis Sarah, 12 ans plus tôt, toi qui voyais en moi « un futur chercheur », nous y voilà !

Je tiens également à remercier toutes les personnes qui ont su réveiller cette curiosité qui m’a porté jusqu’au grade de docteur. Mes enseignants, particulièrement Caroline Pernin, professeur de SVT au lycée qui a capté mon attention sur le monde du vivant, ainsi que

Hubert Becker, qui m’a accordé sa confiance pour intégrer le master BGM dont la promo 2016-2018 est devenue un groupe d’amis soudés. Pour finir, j‘adresse toute ma gratitude envers les membres de mon jury de thèse pour avoir accepté d‘évaluer mes travaux.

À tout le monde, To all of you, An alle,

Merci, Thank you, Danke !!!!

Table of Contents

Acknowledgments / Remerciements.....

I. Introduction.....1

1. Crystallography in biology.....2

1.1. History.....2

1.2. Structure and function of biological macromolecules.....4

1.3. Crystallization process.....7

1.3.1. The phase diagram.....7

1.3.2. One-step nucleation model.....8

1.3.3. Two-step nucleation.....10

1.3.4. Crystal growth.....12

1.3.5. Ripening phenomenon.....13

2. From standard to advanced crystallization methods.....14

2.1. Conventional methods for macromolecules crystallization.....14

2.1.1. Batch method.....14

2.1.2. Vapor diffusion.....14

2.2. Convection-less environments.....16

2.2.1. Free-interface and counter diffusion.....17

2.2.2. From seeding to Microseed Matrix Screening.....17

2.2.3. Nucleating agent – case of Tb-Xo4 crystallophore.....19

2.2.4. Trace-fluorescent labeling.....20

2.3. Phase diagram and Xtal Controller XC900.....21

3. Sample quality control.....23

3.1. Analytic gel and Size Exclusion Chromatography.....23

3.2. Dynamic and static light scattering.....24

3.3. Detecting crystallization hits.....26

4. Strategies for crystal analysis by X-ray diffraction.....29

4.1. Standard crystal analysis.....30

4.2. Serial synchrotron crystallography.....32

4.3. Time-resolved X-ray crystallography.....32

5. Structural study of tRNA maturation enzymes.....35

5.1. Transfer RNA.....35

5.2. CCA-adding enzymes.....38

5.2.1. Classification.....38
5.2.2. Structure and mode of action.....39
5.2.3. *Planococcus halocryophilus* – adaptation to the cold.....41
5.2.4. *Romanomermis culicivorax* – adaptation to minimal tRNAs.....41
5.2.5. Ancestral CCA-adding enzyme – Evolutive divergence of class II enzymes…43
6. Goal of the thesis.....46
6.1. Biochemical characterization of targeted macromolecules.....46
6.2. Crystallization and structure determination of apo and holoenzyme.....46
6.3. Toward a molecular movie of enzymatic catalysis.....47
II. Material and methods.....48
1. Construct and strain.....49
2. Protein overexpression.....49
3. Purification of CCA-adding enzymes.....49
4. tRNA *in vitro* transcription.....51
5. Nano Differential Scanning Fluorimetry.....53
6. Crystallization.....53
7. Advanced methods.....55
8. Data collection.....55
9. Data processing.....56
10. Other softwares.....57
11. Electron microscopy.....57
III. Purification and crystallization of CCA-adding enzymes.....59
1. Purification of PhaCCA, RcuCC and AncCCA.....60
2. Preparative work on RcuCCA.....60
2.1. Stability of RcuCCA.....60
2.2. Crystallization of RcuCCA.....64
3. Crystallization of ancestral CCA-adding enzyme.....66
3.1. Identification of crystallization conditions.....69
3.2. Crystallization in the Xtal Controller XC900.....69
3.3. Further refinement by vapor diffusion.....72
3.4. In situ X-ray diffraction experiments at Proxima 2A.....73
3.5. Other cases similar to AncCCA.....74
3.6. “Salting in” behavior of AncCCA.....77

4. On-chip experiments at P14 T-REXX beamline.....77
4.1. Batch preparation.....78
4.2. Vapor diffusion preparation.....78
5. Preliminary cryo-TEM results.....80
IV. Adaptation of the *Romanomermis culicivorax* CCA-Adding Enzyme to Miniaturized Armless tRNA Substrates.....82
V. Monitoring the Production of High Diffraction-Quality Crystals of Two Enzymes in Real Time Using In Situ Dynamic Light Scattering.....108
VI. Crystallization and Structural Determination of an Enzyme:Substrate Complex by Serial Crystallography in a Versatile Microfluidic Chip.....128
VII. CCA-addition in the cold: Structural characterization of the psychrophilic CCA-adding enzyme from the permafrost bacterium *Planococcus halocryophilus*.....149
VIII. Conclusion & perspectives.....174
1. Exploring RcuCCA biochemical properties to improve its crystallization.....175
2. Combining advanced crystallization methods for a reproducible crystal production….176
3. Toward a molecular movie of enzymatic catalysis.....177
Bibliography.....179
Annexes.....193
1. Authors contribution.....195
1.1. Adaptation of the Romanomermis culicivorax CCA-Adding Enzyme to Miniaturized Armless tRNA Substrates.....195
1.2. Monitoring the Production of High Diffraction-Quality Crystals of Two Enzymes in Real Time Using In Situ Dynamic Light Scattering.....195
1.3. Crystallization and Structural Determination of an Enzyme:Substrate Complex by Serial Crystallography in a Versatile Microfluidic Chip.....195
1.4. CCA-addition in the cold: Structural characterization of the psychrophilic CCA-adding enzyme from the permafrost bacterium Planococcus halocryophilus.....196
2. Supplementary figures.....197
3. A simple and versatile microfluidic device for efficient biomacromolecule crystallization and structural analysis by serial crystallography.....199
4. Structural Analysis of RNA by Small-Angle X-ray Scattering.....211
5. Posters and presentations.....239
5.1. New microfluidic design for crystallization of biomolecules and their in situ

analysis by serial crystallography.....239
5.2. Monitoring the Production of High Diffraction-Quality Crystals in Real Time Using In Situ Dynamic Light Scattering.....239
5.3. Discovering the world of RNA: meeting and sharing science with the general public 239

Figures index

Figure I.1: Historical events in crystallography linked to this project.....3

Figure I.2: Major methods to explore 3D structure of macromolecules.....5

Figure I.3: Theoretical phase diagram of a macromolecule.....9

Figure I.4: One and two-step nucleation process.....11

Figure I.5: Conventional methods to grow crystals in lab.....15

Figure I.6: Effect of microseeding matrix screening (MMS).....18

Figure I.7: Effect of Tb-Xo4 nucleating agent.....19

Figure I.8: Usage of trace fluorescent labeling with carboxyrhodamine-succinimidyl ester. . .21

Figure I.9: Xtal controller picture and schematic view.....22

Figure I.10: Impact of different crystallization methods on the phase diagram.....22

Figure I.11: Dynamic light scattering principle.....26

Figure I.12: Plate screener system at Proxima 2A.....27

Figure I.13: Synchrotron and XFEL technology.....31

Figure I.14: Principle of serial time-resolved X-ray crystallography at an XFEL source.....33

Figure I.15: Time-scale reachable at XFEL and SSX beamlines.....34

Figure I.16: Canonical tRNA structure.....36

Figure I.17: tRNA maturation and partners.....37

Figure I.18: CCA-adding enzyme architecture.....39

Figure I.19: Minimal tRNAs found in nematodes and in Romanomermis culicivorax.....42

Figure I.20: phylogenetic tree used to reconstruct AncCCA.....44

Figure I.21: Pipeline adopted to characterize selected CCA-adding enzymes.....45

Figure III.1: Chromatograms from the different purification steps of AncCCA enzyme.....61

Figure III.2: DLS profiles of AncCCA and RcuCCA enzymes.....62

Figure III.3: RcuCCA degradation inspection by SDS-PAGE and mass spectrometry.....63

Figure III.4: Scattering signal of RcuCCA in nanoDSF buffer screening experiment.....64

Figure III.5: Potential crystallization hits for RcuCCA.....65

Figure III.6: Pipeline used to optimize crystal diffraction.....67

Figure III.7: Pictures of AncCCA initial crystallization hits obtained in commercial screens. 68

Figure III.8: AncCCA crystallization in the Xtal Controller.....71

Figure III.9: AncCCA crystals emerging from precipitate.....72

Figure III.10: *In situ* diffraction at room temperature with Plate Screener at PX2A.....76

Figure III.11: 3-11: Design of the HARE chip used on P14/T-REXX beamline.....79

Figure III.12: Crystallization of AncCCA from precipitate in hanging and sitting drop.....80

Figure III.13: Electron diffraction test on nanocrystalline material from AncCCA enzyme….81

Supplementary figure 1: Target CCA-adding enzymes sequences.....197

Supplementary figure 2: maps of plasmids used for CCA-adding enzymes overexepression 198

Supplementary figure 3: AncCCA and RcuCCA computational models obtained with AlphaFold2.....199

Abbreviations

ddH_2O: didistilled water

DTT: Dithiothreitol

mPEG: Polyethylene glycol monomethyl ether

IPTG: Isopropyl β-D-1-thiogalactopyranoside

LB: Lysogeny broth

MIP: metastable intermediate phase

MMS: Microseed Matrix Screening

nanoDSF: nano Differential Scanning Fluorimetry

PAA: Polyacrylamid

PBS: Phosphate buffer saline

PCR: Polymerase chain reaction

PEG: Polyethylene glycol

PX2A: Proxima 2A beamline

TB: Terrific broth

TBE: Tris-borate-EDTA

TEM: transmission elctron microscope

TIPP: Thermostable Inorganic Pyrophosphatase

Tm: melting temperature

Tris: Tris(hydroxymethyl)aminomethane

tRNA: transfer ribonucleic acid

XC900: Xtal Controller 900

I. Introduction

1. Crystallography in biology

1.1. History

Since Antique times mankind tried to understand the smallest units that compose universe. Leucippus (4th century BC) invented the concept of atomism, from ancient Greek atomos meaning ‘uncuttable’, to describe indivisible bodies at the basis of matter. It took more than two thousands years to see the emergence of a new scientific domain called crystallography, which allows to ‘see’ and describe the spatial arrangement of atoms in molecules. The word crystallography itself was introduced by Maurice Antoine Capeller in 1723. Then, in 1781 René-Just Haüy characterized calcite crystal as an orderly arrangement of “integrant molecules” or “nuclei”. Shortly after, in 1850, Auguste Bravais, a French physicist and mineralogist described the crystal as a repetition by translation in tridimensional space of an elementary motif. Thus, the 14 possible geometrical arrangements of nodes constituting the unit cell were were named after him as “Bravais lattices”. The discipline could suddenly jump from a purely theoretical to applied science thanks to Wilhelm Conrad Röntgen who discovered X-rays in 1895, thus allowing to see “beyond the matter”.

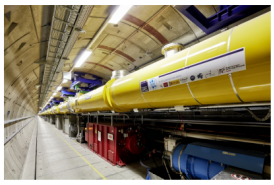
Seventeen years later, the first major experiment on crystals was made by Max von Laue who argued and proved that the short wavelength of X-rays should be on the same order of magnitude as inter atomic distances. Helped by his student Paul Peter Ewald, he showed that the interaction of X-rays with “resonators” from highly ordered crystals could give a diffraction pattern, reflecting the 3D structure of these “resonators”. The second major event happened in 1915 with the mathematical demonstration of the diffraction phenomenon on a crystal by father and son William Henry and William Lawrence Bragg. Their Bragg’s law is defined as following:

$$2d\sin(\theta)=n\cdot\lambda$$

With d corresponding to inter-reticular distance, θ half the angle formed by incident and reflected beam, n the diffraction order and λ the wavelength of incident beam.

2017: Opening of the euXFEL (Hamburg)

2009: First running XFEL (LCLS – University of Stanford)



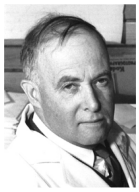
John Kendrew

& Max Perutz (1962)

1958: First protein structures solved by X-ray crystallography (myoglobin and hemoglobin)

George Porter (1967)

1945-1954: First structural time-resolved experiments on gaseous free-radicals formation



James Batcheller Sumner (1946)

1926: Isolation and crystallization of an enzyme (urease)

Max von Laue (1914)

& Paul Peter Ewald

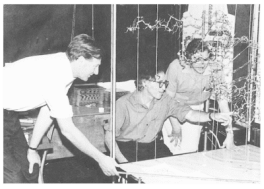
1912: Hypothesis and first X-ray diffraction experiment on a copper crystal

Auguste Bravais

1850: Description of crystal geometry and definition of lattices

Thomas Steitz, Ada Yonath, Venki Ramakrishnan (2009 - for ribosome structure)

2002: First CCA-adding enzyme structure



Alexander Rich & Aaron Klug

1973: First tRNA Structure

Francis Crick James D. Watson (1962) & Rosalind Franklin

1953: Helicoïdal structure of DNA

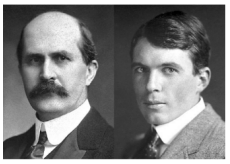
Edwin McMillan (1951)

& Vladimir Veksler

1949: First synchrotron built at University of Michigan

W. H. & W. L. Bragg (1915)

1912: Bragg's law



Wilhelm Conrad Röntgen

1895: Discovery and first usage of X-rays

René Just Haüy

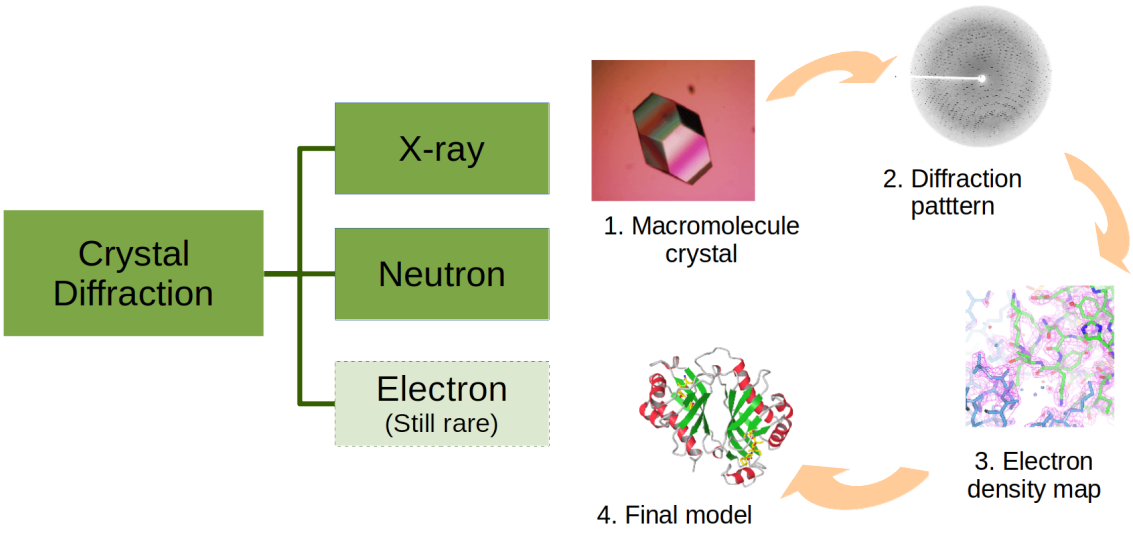
1781: Definition of a crystal

Figure I.1: Historical events in crystallography linked to this project

Under the impulse of such discoveries, modern crystallography gained first the chemistry field because of the ease to obtain crystals of small and simple molecules. In the late 1920's, James Batcheller Sumner isolated and crystallized the first enzyme, an urease from jack bean (James Sumner, 1926). The community had to wait until 1958 for the first structure of a protein solved by x-ray crystallography on myoglobin (Kendrew et al., 1958) quickly followed by hemoglobin (Perutz et al., 1960) by John Kendrew and Max Perutz. This breakthrough gave birth to a new field of research at the interface between biochemistry, physics and chemistry. For the last sixty years, X-ray crystallography has been and stays a powerful method to explore three-dimensional structures of macromolecules. As a matter of fact, macromolecular crystallography is one of the most fruitful field for prestigious awards (e.g., more than 25 Nobel prices)(Jaskolski et al., 2014), reflecting its massive contribution to life sciences community. From 1976 to November 2021, the Protein Data Bank (PDB) accumulated 160.871 crystal structures among the total 183.793 structures available. Nevertheless, it seems important to mention the great improvement made these last ten years with other methods. Indeed, dramatic advances in cryo electron microscopy (cryo-EM) could recently lead to solving the structure of apoferritin at the resolution of 1.22Å (Nakane et al., 2020). In addition, the size bottleneck in nuclear magnetic resonance (NMR) has also been pushed up, giving access to small protein complexes structure (Purslow et al., 2020) and dynamics (Kovermann et al., 2016) .

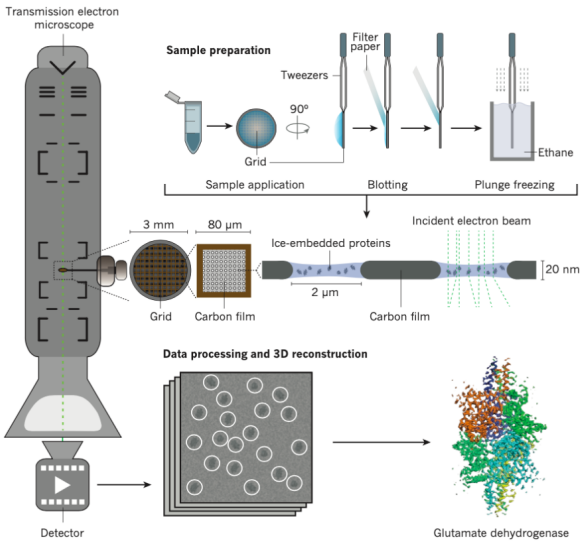
1.2. Structure and function of biological macromolecules

The word macromolecule was first used in the early 1920's to designate "big" molecules containing more than thousands of atoms found in the mineral, synthetic chemistry and biological world. The latter category encompasses polysaccharides, lipids, nucleic acids, ribonucleoproteins and proteins, which are found throughout all the living reign. Polysaccharids and lipids are in the vast majority composed of a repeated motifs with a variable length and have either a non-organized or organized 3D structure forming networks (e.g., cellulose and chitin). In the frame of this work, the word "macromolecule" directly refers to nucleic acids, ribonucleoproteins and proteins, which display specific 2D and 3D organization.



Cryo-electron microscopy (cryo-EM)

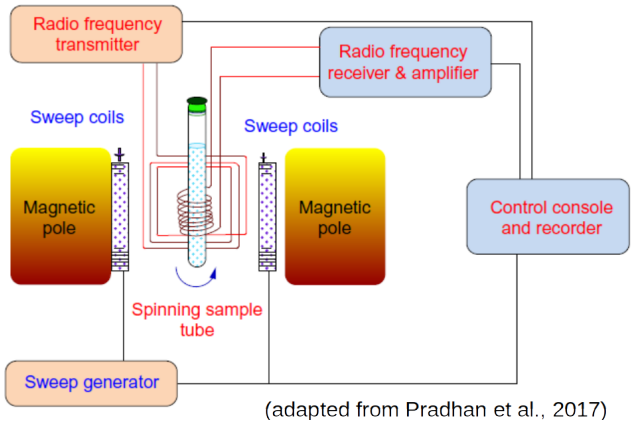
- Based on single particle imaging
- Uses all possible orientations of the macromolecule to reconstrcut a model



(adapted from Fernandez-Leiro and Scheres, 2016)

Nuclear magnetic resonance (NMR) spectroscopy

- Based on nucleus spin placed in a magnetic field
- Uses inter-atomic distances of the macromolecule to reconstruct a model



(adapted from Pradhan et al., 2017)

Figure I.2: Major methods to explore 3D structure of macromolecules

To begin with proteins, they consist of amino acids polymers linked by peptide bonds and flanked by a carboxyl (C-ter) and amino group (N-ter). Amide and carbonyl groups from the backbone can interact with adjacent amino-acid groups, forming hydrogen bonds and conferring particular sub-structure to the enzyme such as helices and ß-sheets. In addition, nature of amino-acids side chain confers both structural and functional properties to the protein. Depending on their charge or polarity, side chains can either create salt bridges between each others which may stabilize the protein architecture, but also interact with surrounding molecules (e.g., solvent, ligand).

An important point to consider for macromolecules stands in their native environment. When considering their location in the cell, proteins will display adapted composition and structure. On the one hand, amino-acids carrying alkyl or aromatic group show hydrophobic behavior and will commonly be found hidden inside the molecule for cytosolic proteins or, on the contrary, at the interface with hydrophobic tails from lipids for membrane proteins. On the other hand, charged and polar uncharged amino acids will favor interactions with the solvent molecules.

This statement is also true for nucleic acids, which basically corresponds to nucleotides oligomeres and are present in two distinct forms in nature: RNA (ribonucleic acid) or DNA (deoxyribonucleic acid). Both RNA and DNA molecules have variable length and can adopt single or double stranded architecture. Firstly considered as simple “encoding molecules”, further studies in the 1980’s highlighted the incredible versatile nature of nucleic acids. As a matter of fact, ribozymes are great example of RNAs adopting a 2D and 3D structure guiding a specific catalytic activity (Guerrier-Takada et al., 1983). Furthermore, such elements may associate with proteins (ribonucleoproteins) to combine their action such as in the ribosome (Ramakrishnan, 2002) or the vault complex, the biggest ribonucleoprotein in eukaryotic cells (13 Mda) (Kedersha et al., 1991).

Macromolecules’ functions are directly linked to their three-dimensional structure. Indeed, spatial organization of proteins and nucleic acids at the atomic scale, reveal important intra- and intermolecular mechanisms and interactions. In the late sixties, Max Perutz related probable structural abnormalities of hemoglobin to defects in respiratory function (Perutz & Lehmann, 1968). Quickly he laid the ground work to drug design field, especially in 1976 where he defined “tailored-made” inhibitors to fight diseases (Perutzt, 1976). For instance, structural biology highly contributed to understand diseases mechanistic and thus to design adapted drugs (so-called drug design). Following this idea, in 2004 the Structural Genomics

Consortium (SGC) was launched in North America and Sweden in order to explore structures of less well-studied areas of the human and pathogens genome (Williamson, 2000). With an active collaboration between academic and pharmaceutical worlds, the SGC had the ambition to accelerate research in drug-design field by making results available to the whole community. Finally, it is important to emphasize the strength of structural biology, especially in the actual context of the COVID-19 pandemic. As a matter of fact, dedicated structure databases have been established to help in immediate treatment research (Gowthaman et al., 2021).

Another contemporary field of application concerns ecology stakes, especially regarding plastics degradation. As a recent example, a French team solved the structure of an engineered Poly-ethylene-terephthalate (PET) hydrolase (Tournier et al., 2020), an enzyme capable of “digesting” one of most present plastic on Earth. In addition, the understanding of enzymes structure and dynamics is of great interests for industries. From food to pharmaceutical industry, enzyme engineering and design is used to improve yields at lower costs. Finally, structures inform on evolution and adaptation of biomolecule families to extreme conditions. Macromolecules from so-called extremophiles can resist to high or low temperatures, acidic or basic environments or heavy metal-containing medium. Such characteristics are of great interest for biotechnology and industry areas (Littlechild, 2015).

1.3. Crystallization process

1.3.1. The phase diagram

Every crystallographic study starts with the preparation of crystals (A. McPherson, 1976). Crystallization process refers to a phase transition from a soluble state of the macromolecule to an ordered solid state. Solubility of biological macromolecules in aqueous solution is influenced by a set of physical-chemical parameters such as their own concentration, precipitant concentration, temperature or pH. The precipitant or crystallant is defined as the compound that will drive the macromolecule/biomolecule beyond its solubility limit, i.e., to a non-equilibrium state called the supersaturation. The nature of the crystallant is very diverse and includes salts, organic compounds, a variety of polymers. Each macromolecule presents its specific solubility space which can be represented on a two-

dimensional phase diagram (Figure I.3) where the protein concentration on the Y-axis is plotted against the precipitant concentration on the X-axis. Below the solubility curve, molecules are moving freely in the solution and the system is said to be in undersaturated condition. As the concentration of either the protein or the precipitant is increased, the system crosses a first energy barrier (the solubility curve) to reach the supersaturated state. In this area, the thermodynamic driving force pushes the molecules outside of the solution and promotes the formation of a solid state (i.e., crystalline or aggregated). The supersaturated region is split in three main regions:

(1) The metastable zone which corresponds to a low-supersaturation state. There, macromolecules start to form small clusters but the energy is too low to allow the production of stable nuclei that can further grow into crystals. However, if nuclei are introduces in this zone (e.g., in a microseeding experiment), crystals will grow in this part of the diagram.

(2) The nucleation zone where the unstable system allows the formation of critical-sized nuclei. Once nuclei are formed, macromolecules will regularly pile onto these clusters to build up crystals, inducing a concentration decreasing and dragging the system back through the metastable zone to the thermodynamic equilibrium on the solubility curve.

(3) The precipitation zone located at the highest supersaturation part of the diagram. In this area, macromolecules can form plethora of nuclei, leading to crystalline precipitate, or also amorphous precipitate which is quite often irreversible.

In many cases, the limit between the nucleation and precipitation zones is quite thin, making the search for crystallization conditions quite challenging. “Traditional” crystallography tends to push the macromolecule to the nucleation zone, then drag the system (i.e., few big nuclei) back to the metastable area where crystals will grow until they reach the equilibrium on the solubility curve and stops to grow. In this regard, the main goal is to favor the growth of few nuclei fed by surrounding macromolecule to reach reasonable size for X-ray analysis, meaning from 40 to several hundreds micrometers.

1.3.2. One-step nucleation model

As previously described, the nucleation phenomenon takes place when the system is driven outside of equilibrium state where macromolecules are forced to interact and form small ordered clusters. Described as a stochastic event (Sleutel & Van Driessche, 2018),

nucleation occurs as the emergence of weak interactions (i.e., hydrogen and ionic bonds, or hydrophobic interactions) between macromolecules occurs and repeats around a primordial geometric pattern. This thermodynamic unstable state is in the vast majority obtained by a "salting out" effect, where macromolecule are pushed outside of the solution as the crystallant concentration is increased. However, in few cases, the inverse phenomenon called "salting

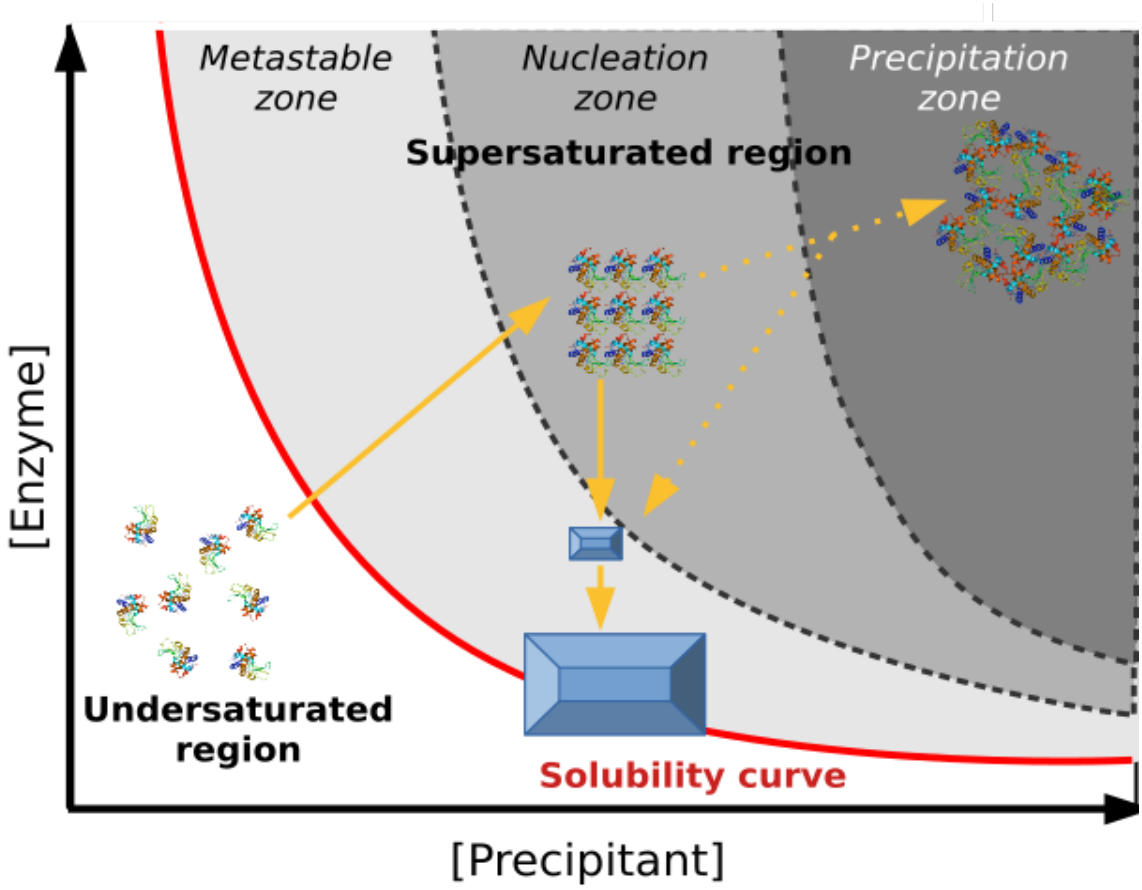


Figure I.3: Theoretical phase diagram of a macromolecule

Phase diagram describing the state of a macromolecule as a function of its concentration and that of the crystallant in the solution. Inside the undersaturated zone, macromolecules are moving freely into the solution. As the concentration of both the enzyme and/or the crystallant are increasing, the solubility curve is crossed and the system enters the supersaturated region. In the classical nucleation model (full yellow arrows), nuclei are spontaneously formed in the nucleation zone, then soluble macromolecules are integrated at surface of the nascent crystal (growth step) until reaching the solubility curve. In the two-step nucleation model, the system is pushed at the interface (or beginning) of the precipitation zone where nuclei are formed before coming back to the metastable zone and finally stop their growth while reaching the solubility curve.

in", can also promote the formation of nuclei by decreasing ionic strength of the media (e.g., by decreasing crystallant concentration) as it was described for T7 RNA polymerase (Jeruzalmi & Steitz, 1997).

Historically, two theories have been established to explain the nucleation process. The first and oldest one, introduced by Stranski and Kaischew in 1934 (Kaischew & Stranski, 1934; Stranski & Kaischew, 1934) , is consistent with low-supersaturation conditions. It stipulates that soluble macromolecules pushed in the supersaturation zone are firstly assembling as small clusters, considered as the transition to a solid phase, in equilibrium with the liquid phase, where molecules are moving freely (Figure I.4.A.). In this situation, nuclei are forming and dissolving unless they reach a critical radius (r^*) (García-Ruiz, 2003). This critical radius is defined as the minimal size of a cluster to reach the energetic barrier ΔG^* necessary to get a stable nucleus and grow a crystal. Considering nuclei as spherical objects, ΔG^* can be expressed as function of the radius of nuclei as follow (derived from Gibbs–Thompson relation):

$$\Delta G^* = \frac{16\pi v^2 \gamma^3}{3[kT\ln(S)]^2}$$

With v corresponding to the molar volume occupied by a single molecule, γ representing the surface energy on the forming nucleus, k the Boltzmann constant, T the temperature and S the supersaturation (ratio of the actual concentration divided by the concentration at the equilibrium).

Once the sufficient energy is reached, the following step corresponds to the regular packing of macromolecules coming from the surrounding medium also called the crystal growth.

1.3.3. Two-step nucleation

The second approach of crystal nucleation has been discussed for years as no direct physical methods were available to observe it. In the literature, the expression "fluid aggregate model" refers to the formation of a metastable intermediate phase (MIP) prior to the emergence of well-ordered nuclei in a high supersaturation state (Zhang et al., 2021). Concretely, macromolecules pushed rapidly in very high supersaturation are forming clusters with random orientations (Figure I.4.B.). While this phenomenon usually leads to aggregation and to more extent the formation of amorphous precipitate, a transition from random

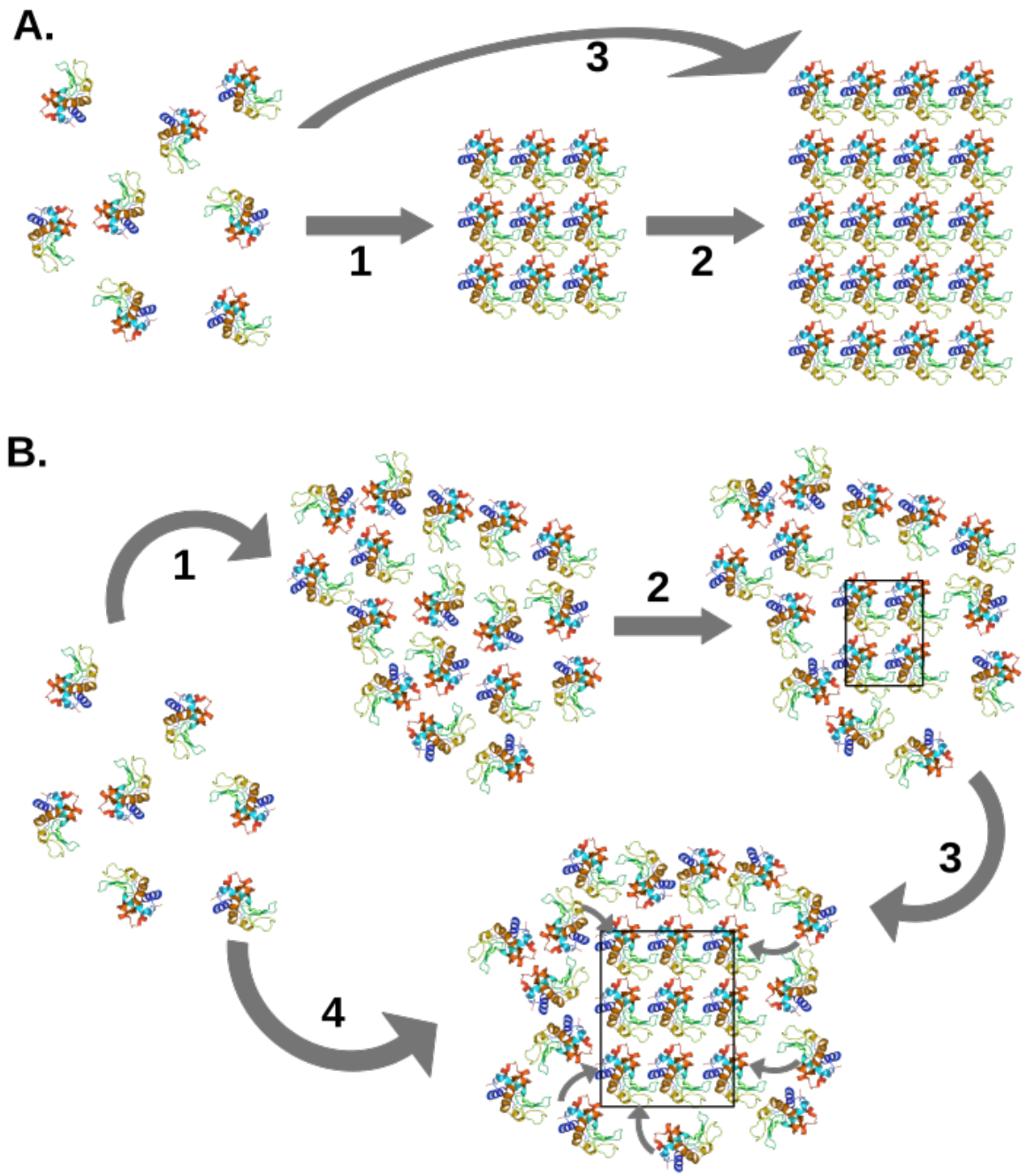


Figure I.4: One and two-step nucleation process

A. One-step nucleation model. (1) Spontaneous formation of a nucleus in the nucleation zone. (2) Growing crystal fed by (3) free macromolecules from the media. B. Two-steps nucleation model. (1) Macromolecules in the undersaturated region a pushed toward high supersaturation are of the phase diagram. (2) Formation of dense clusters which are reorganizing and give birth to a nucleus (crystalline state) in the bulk. (3)Growth of the nucleus into a crystal by the adsorption of surrounding macromolecules and (4) distant molecules moving freely in the medium

electrostatic bonds and hydrophobic interactions to geometric and periodic interactions between macromolecules can occur inside such dense molecular clusters. This primary geometric unit serves as a platform to drive the surrounding material to a higher ordered structure which will adsorb macromolecules from the medium and form a crystal. Recent work from Houben and co-workers used cryo-tomography to point out the early stages of ferritin crystal formation originating from amorphous precursor (Houben et al., 2020), thus reinforcing the two-step mechanism theory. In addition Frank Schreiber group highlighted crystallization process of human serum albumin in the presence of a metastable liquid–liquid Phase separation (LLPS) by combining small angle X-ray scattering (SAXS) and optical microscopy (Maier et al., 2020).

1.3.4. Crystal growth

Besides the central role of nucleation in crystal formation, the following growth step guides critical parameters defining, inter alia, crystal packing quality, size and resulting diffraction properties. Once the critical size of the nucleus is obtained and macromolecules start to pack onto the primordial ordered unit, the system is considered to enter the growth phase. As previously explained, crystallization is a transition from a liquid/soluble to a solid phase. Thus a crystal requires enough supersaturation to incorporate molecules from the surrounding and drives the system back to the solubility curve. In the past thirty years, *in situ* techniques like atomic force microscopy (AFM) and interferometry have been massively used to understand key mechanisms piloting the growth of macromolecular crystals (A. McPherson et al., 2000). To date two major mechanisms have been established to explain how complex molecules attach on the surface of a unique object, the crystal. The first and the most energetically favorable, called tangential growth, consists of the progressive addition of single macromolecules on a nascent layer at the surface of the crystal, causing lateral extension. As incoming macromolecule is interacting with partners from both nascent and underlying layers, its attachment to the crystal requires minimal energy. On the contrary, the second mechanism described as normal or face-growth, reflects the initiation nucleation at the surface of existing crystal. This configuration gives birth to small islands growing as steps on the top of crystal surface. Such a pathway needs higher energy than tangential growth and is usually a limited process leading to crystal growth abortion or crystal defects.

Another essential parameter guiding the quality stands in the speed of macromolecule incorporation. In the case of high supersaturation, one should consider that surrounding

macromolecules are either quickly integrated to crystal surface via one mechanism or the other. The faster these processes occur, the higher is the probability to accumulate defects (e.g., impurities, twinning) in the crystal (Alexander McPherson & Kuznetsov, 2014). This is a main reason for growth stop. This said, one can easily imagine that very small crystals (< 5 µm) would present less defects than larger ones with typical sizes of 100-500 µm. Otherwise, if crystals grow big enough, the limiting factor becomes the soluble macromolecule available in the surrounding solution as the system is driven back to an equilibrium between soluble macromolecule and the solid phase. This can be solved by bringing fresh macromolecule solution to the crystal. Finally, one should keep in mind that biological crystals are highly solvated, namely containing between 40 to 80% of solvent molecules (mostly water). This particularity makes such material fragile and subject to rapid desiccation if wrongly handled.

1.3.5. Ripening phenomenon

Coming back to the two-step nucleation process and crystal growth, it is important to mention the phenomenon of ripening, or re-crystallization, happening in some cases (Ng et al., 1996; Shtukenberg et al., 2021). As experimental conditions push the system to very high supersaturation state, the number of nucleation events dramatically increases. These single nuclei can evolve in an independent manner and grow into nanoscale crystals buried in the bulk. Ripening takes place when a nanocrystal becomes slightly bigger than its closest neighbors and starts to grow at the expense of the others which will redissolve. In the particular case of macromolecules, the corresponding mechanism is called Ostwald ripening (Ng et al., 1996). Concretely, it starts with the appearance of a precipitate in the experimental drop. Then, few crystals emerge from the mass (the mix of thousands of nanocrystals and amorphous precipitate) surrounded by a clear area corresponding to the depletion zone.

All crystallographers already experienced precipitation in crystallization trials. For decades, theses droplets have been often considered as amorphous precipitate. However, visible “precipitate” does not always reflect amorphous precipitate (Luft et al., 2011). Recent works highlighted the potential of such kind of hits to actually constitute microcristalline material (Schubert et al., 2015). Such samples would actually be great candidate for new-generation of serial crystallography (detailed in sections 4.2 and 4.3), thus rescuing old abandoned projects which could not lead to macrocrystal formation.

2. From standard to advanced crystallization methods

2.1. Conventional methods for macromolecules crystallization

2.1.1. Batch method

The oldest and maybe simplest way to trigger crystallization of a target molecule is called batch crystallization. It consists of mixing an undersaturated solution of macromolecule with a given crystallant in a proportion near the equilibrium of the system (Figure I.5.A.). By immediately reaching the low part of the precipitation zone or the nucleation zone, nuclei can quickly form and while the crystal grows the system comes back to the solubility curve. In this precise case, the thermodynamic equilibrium is reached on a short timescale and crystals might accumulate defaults if the process is too fast.

As the batch method often leads to high-amount of small crystals or even crystalline precipitate, it has slowly disappeared from a lot of laboratories. However, the recent evolution with new strategies for data collection allowing the usage of micro or nanocrystals (e.g., XFEL source or microfocus beamlines), makes batch great implement in condition screening and sample preparation for new generation X-ray analysis (Kupitz, Grotjohann, et al., 2014). As an example, a quick way to prepare micro-crystalline solution consists of mixing the stock tube of protein with an already known crystallant in the same proportion as for a vapor diffusion experiment. Then, using a centrifuge under vacuum, the mix is concentrated until the appearance of soft precipitate.

2.1.2. Vapor diffusion

Surely the most used method in laboratories, the vapor diffusion method allows to travel slowly from the undersaturated state to the nucleation zone. The experiment setup is usually designed as follows: (i) the macromolecule is mixed with the precipitant; (ii) the drop is placed in a closed compartment containing higher volume of the precipitant alone (called the reservoir solution) at the same concentration used before the mixing. The reservoir and the drop are separated by air; (iii) the system is slowly dragged to an equilibrium as water is evaporating from the drop, migrating to the reservoir, leading the drop to the same precipitant concentration as in the reservoir (Figure I.5.B). The equilibration speed is dependent on both the nature of the precipitant and of its concentration in the drop versus the reservoir (Vincent

Mikol et al., 1990).

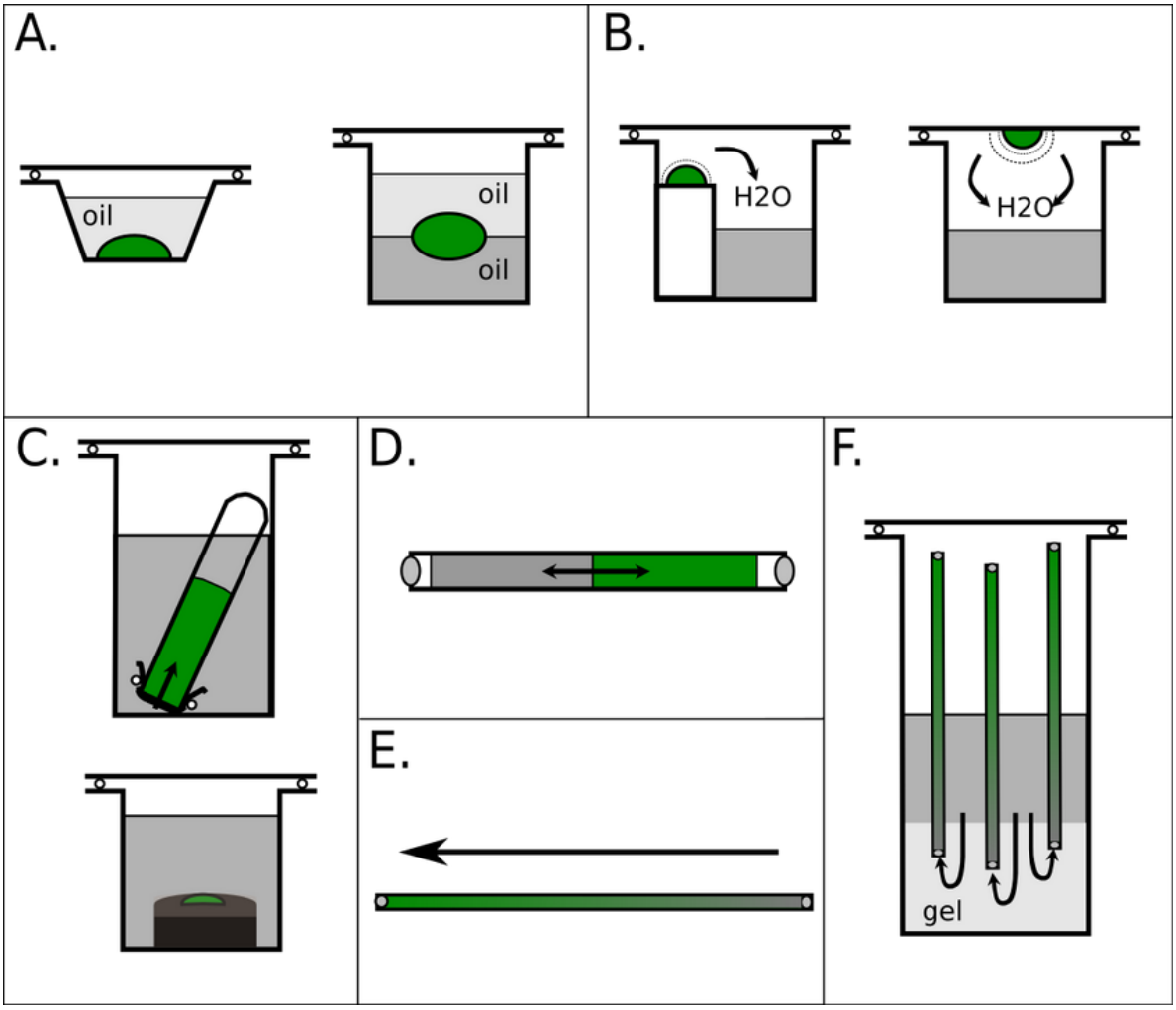


Figure I.5: Conventional methods to grow crystals in lab

The green and dark grey solution correspond to the macromolecule and crystallant solutions respectively. A. Batch under oil or at oil-oil interface. B. Vapor diffusion by sitting or hanging drop. C. Dialysis in sealed tube or in Cambridge button. D. Free-interface diffusion. E. Counter-diffusion in single microfluidic capillary. F. Counter-diffusion in capillaries combined with gel acupuncture. (adapted from Sauter et al., 2012)

There exists two main vapor diffusion configurations, the sitting and the hanging drop. The former technique consists of depositing the droplet on a pedestal surmounting the reservoir solution. This configuration is less dependent on gravity and widely used for high-throughput screening. The latter setup, places the drop in suspension over the reservoir. In this case the support type and the volume (weight) of the drop influence the drop shape. Thus, the

evaporation rate is usually faster compared to a classical sitting drop setup (DeTitta et al., 1995; Luft et al., 1996; Vincent Mikol et al., 1990). In addition, bigger elements formed in hanging droplets (i.e., crystals) will often converge to a single point at the bottom of the drop pointing toward the reservoir. Vapor diffusion benefited from the spread of crystallization robots, commercial screens and dedicated crystallization plates. On the one hand, robots reduce the experimental volumes down to hundred nanoliters and on the other hand, they permit mixing and dispense of droplets within few minutes for a typical 96-wells plates screening experiment.

2.2. Convection-less environments

Batch and vapor diffusion described above are surely the most widespread techniques and usually provide great results. However, these methods are systematically associated with a convection regime which should be avoided in crystallization experiments. In this condition it was shown that buoyancy driven fluid flow (induced by density difference) inside the crystallization drop is the dominant mechanism of solute transport during the crystal growth step (Savino & Monti, 1996). This regime dramatically affects crystal quality as the higher convection leads to the incorporation of impurities coming either from the protein source or the crystallant (Vekilov et al., 1998). On the contrary, if the flow strength is reduced in the medium, one can expect a slower growth rate and improved crystal quality. Different ways have been identified to favor diffusion over convection. For instance, in the late 1990's, low-gravity experiments in space-shuttle, highlighted the beneficial effect of growing individual crystals in suspension in the mother liquor (Ng et al., 2002; Snell & Helliwell, 2005). In the absence of sedimentation and convection, crystal growth is not disturbed by surrounding elements such as other crystals or experimental chamber walls. One way to mimic the micro-gravity effect on earth stands in using gels such as agarose or silica (B et al., 2009; Zhu et al., 2001). Their reticular nature, adjusted with their concentration in the crystallization medium, has displays several advantages. First, the gel matrix lowers nucleation rate by promoting slow diffusion of either the crystallant or the macromolecule. Second, it prevents growing crystals from falling to the bottom of the drop by gravity and assure better quality of crystals, as described in micro-gravity experiments. And finally, literature showed that gels have a protective effect on crystals, especially during soaking or cryo-cooling of samples. Taken together, these properties combined with low cost and easy handling of gels should confer the method a place of choice in crystallization attempts.

2.2.1. Free-interface and counter diffusion

Among convection-less techniques, the free-interface (or liquid-liquid) diffusion was introduced in the early 1970's by Salemme (Salemme, 1972) to avoid rapid-mixing of macromolecule and crystallant. It consists of using capillaries filled on one side (bottom) with the densest solution (usually high concentrated crystallant solution) and gently pour the less dense solution on the top (Figure I.5.D). At the interface of the two solutions, a supersaturation state appears thus promoting local nuclei formation. Both component will penetrate the opposite solution and so create a gradient. This situation allow either existing nuclei to reach lower crystallant concentration and grow or promote nuclei formation in lower parts of the phase diagram.

Based on the same concept, the counter-diffusion method also uses capillaries (Otálora et al., 2009) . Originally, it consists of a gel base (silica or agarose) at the bottom of a reservoir, which is covered by highly concentrated crystallizing agent (Figure 1.5.F). The glass capillary is filled with the target macromolecule and inserted into the gel layer. The crystallant migrates into the gel by diffusion, enter the capillary and forms a concentration gradient. The main advantage of the methods is the wide range of supersaturation state screened with minimal amount of material. Following this idea of miniaturization of counter-diffusion experiments, Dr. Claude Sauter developed a microfluidic chip, so called ChipX3, to screen a large variety of crystallant. Briefly, the ChipX3 is composed of one sample inlet on the left side, connected to eight microfluidic channels ending with individual reservoir on the right side. The channels are first filled with the macromolecule. Then the different crystallant solutions are added to the reservoirs, where they will diffuse into the channel and create a gradient. More detailed protocol and explanations stands in chapter VI.

2.2.2. From seeding to Microseed Matrix Screening

Finding reproducible nucleation conditions is a challenging operation. The microseeding method was popularized by Allan d'Arcy and collaborators to by-pass the step of spontaneous nuclei formation by inserting crystals "seeds" of a given macromolecule in an already supersaturated solution (Bergfors, 2007). By crushing low quality crystals, the user obtains a suspension of nano/micro-crystalline debris. These fragments are dispensed into new crystallization drops either by using a fiber (e.g., animal hair or synthetic material)

soaked into the seed stock, or by 'contaminating' crystallization drops with diluted seeds. Microseeding was shown to increase the number of hits obtained in screening, improve crystal diffraction properties and finally lower the number of growing crystals and lead to isolated bigger ones required for standard X-ray analysis (Figure I.6). A derived strategy, called microseed matrix screening (MMS), aims to use seeds obtained in one given condition and transfer them to unrelated crystallization conditions. Thanks to robotics implementation in actual laboratories, MMS has been widely applied to *de novo* crystallographic investigations. Such an approach may lead to different crystal packing, improvement of diffraction quality or increased of crystal size (D'Arcy et al., 2014).

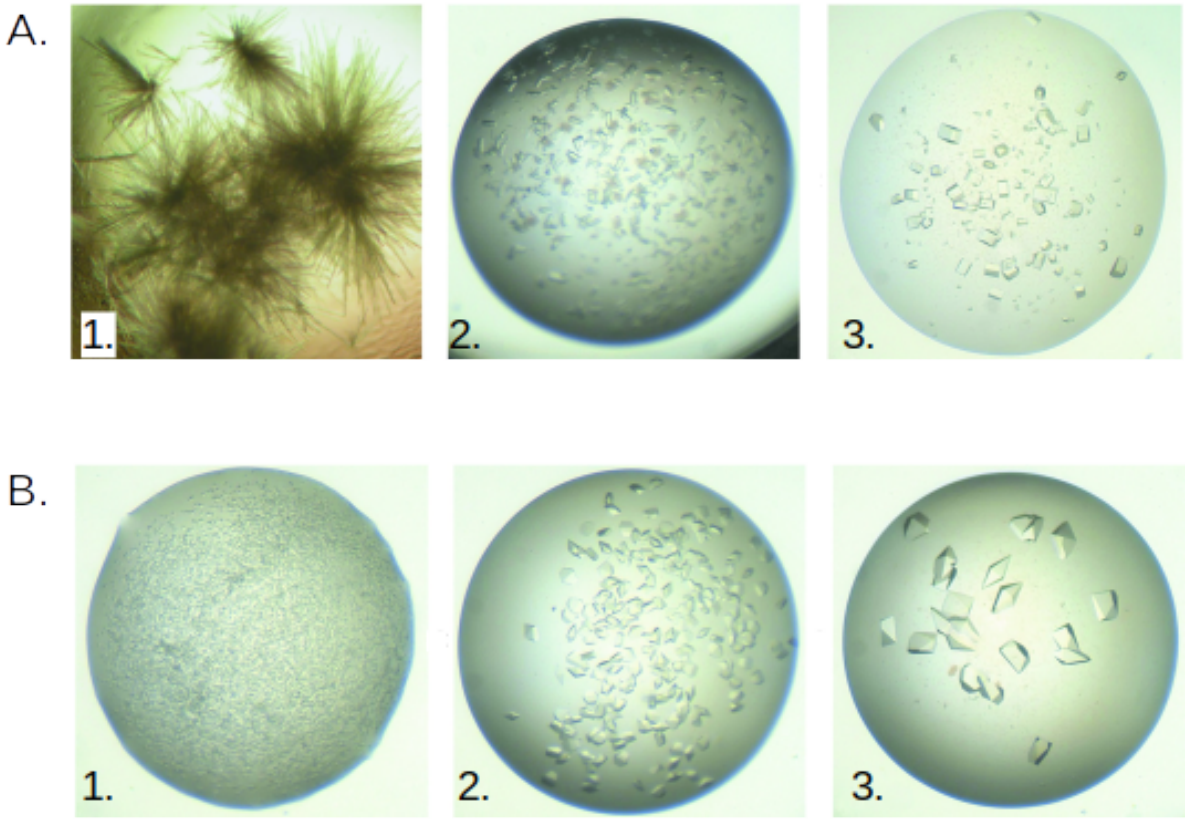


Figure I.6: Effect of microseeding matrix screening (MMS)

A. Improvement of crystal size, shape and number by iterative MMS with helicase crystals. Initial hit (1) followed by a first (2) and a second (3) round of MMS. B. Influence of the seed stock concentration with a tyrosine kinase. The drop were seeded with undiluted seed stock (1), dilution 1:100 (2) and dilution 1:1000 (3). (Adapted from D'Arcy et al. 2014)

To a larger extent, it is possible to use seeds from one macromolecule to enhance crystal growth of a closely related macromolecule (so-called cross-seeding). This way showed less positive results as the original method but should definitely be attempted for tricky

candidates. Finally, a variant of the seeding, called macroseeding, consists of transferring a crystal into a fresh macromolecule solution in supersaturated condition. Soluble macromolecules will then feed and increase the size of the crystal.

2.2.3. Nucleating agent – case of Tb-Xo4 crystallophore

Some tips and tricks have been developed to by-pass nucleation or lower the energy necessary for nucleation and crystal growth. When starting a project from scratch, the usage of compounds acting as "nucleating agents" or mimicking nuclei is considered as an alternative strategy. Such actors regroup natural nucleants such as animal hair (e.g., horse, rats, cats…) or dried seaweed (D'Arcy et al., 2003), but also engineered materials such as

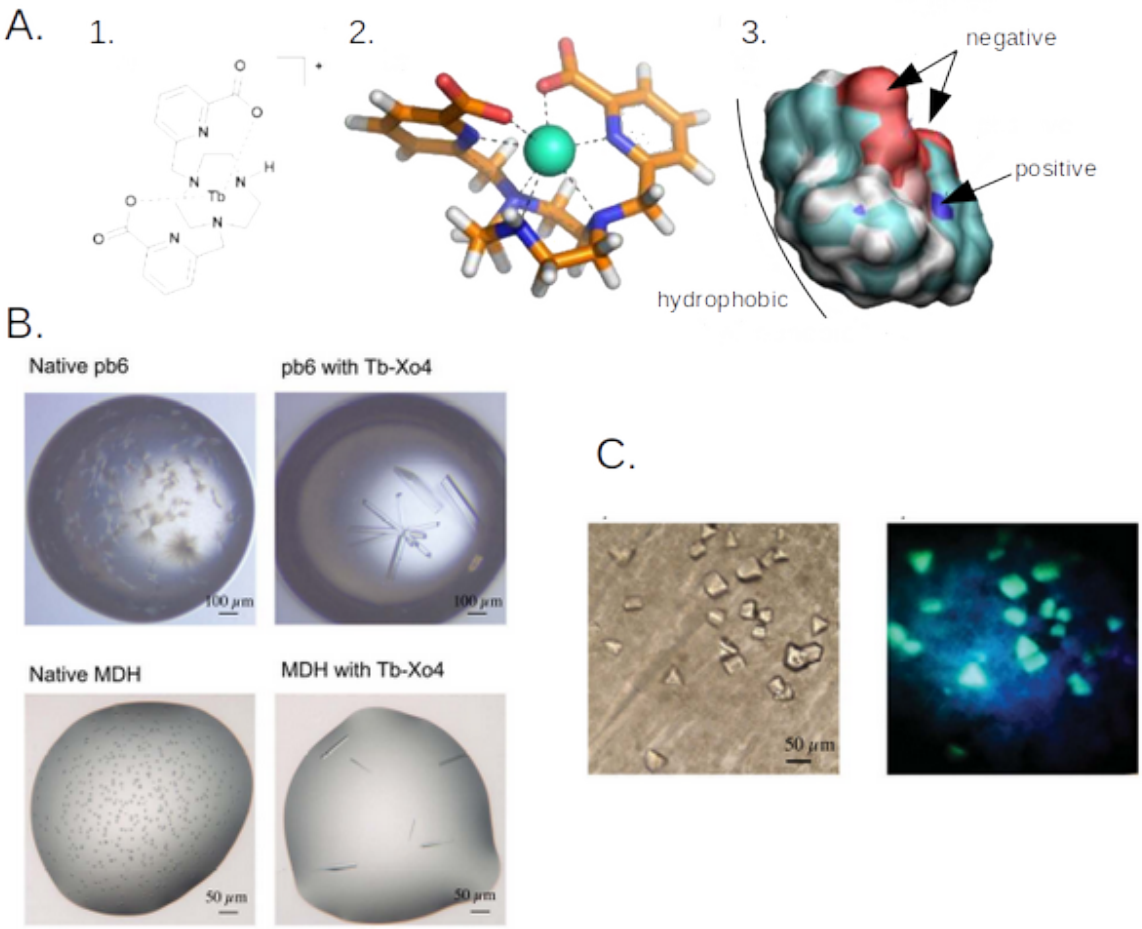


Figure I.7: Effect of Tb-Xo4 nucleating agent

A.(1) Chemical structure of Tb-Xo4. (2) Three-dimensional model of Tb-Xo4. (3) Surface properties of the compound. B. Nucleating effect of Tb-Xo4 tested on pb6 and MDH proteins. C. Fluorescence induced by Tb-Xo4 triggered under UV light. (adapted from Engilberge et al., 2017, 2018)

silicon-based porous nucleant (Chayen et al., 2001) or bioglass-based compound (e.g., Naomi's nucleant) (Saridakis & Chayen, 2009).

Recently, a French team developed a lanthanide-based crystallophore (Engilberge et al., 2017), which has (i) nucleating power by reinforcing molecular surface interactions (so-called "molecular glue") (Engilberge et al., 2018), (ii) phasing and (iii) luminescent properties under UV light thanks to the terbium (Tb-Xo4) ion coordinated with triazacyclononane macrocyclic scaffold (Figure I.7.A). Basically, the compound can be used following two different ways. The first one consists of mixing the compound with the macromolecule and directly use the solution for crystallization experiments. The second option uses soaking to incorporate the Tb-Xo4 into an already grown crystal and take advantage of the phasing property of the terbium atom for anomalous diffraction experiments.

2.2.4. Trace-fluorescent labeling

In the case of proteins, when screening plethora of crystallization conditions, a main bottleneck is the detection of tiny macromolecule crystals as well as their visual discrimination from salt crystals. The task is even harder when it comes to crystalline precipitate which, under white light, resemble amorphous precipitate. To deal with such a situation, it is possible to use the intrinsic fluorescence of proteins provided by aromatic residues (Trp, Tyr) under UV illumination. However, UV is not always convenient as it is dependent on tryptophane content, it can also interact with organic compounds present in the solution, leading to possible false positive results, or generating a blurry noise from the crystallization plate (Desbois et al., 2013).

An alternative method using fluorescent labels added as trace amount into stock protein solution has been introduced and is called "trace fluorescent labeling" (TFL). Among all available fluorophores, the accessible and cheap carboxyrhodamine-succinimidyl ester binds to surface lysines and was shown (Pusey et al., 2015) to highly facilitate identification of crystalline material (Figure I.8). The two-steps labeling process requires basic laboratory equipment and takes around 10 to 15 minutes to prepare usually enough material for all crystallization experiments of a project. Indeed, a "contamination" of around 0,5 to 1% of the unlabeled stock protein is usually far enough to get luminescent signal. Finally, observation of crystallization drops only requires a green light source at 520 nm and a low-pass filter at 550 nm to detect bright spots.

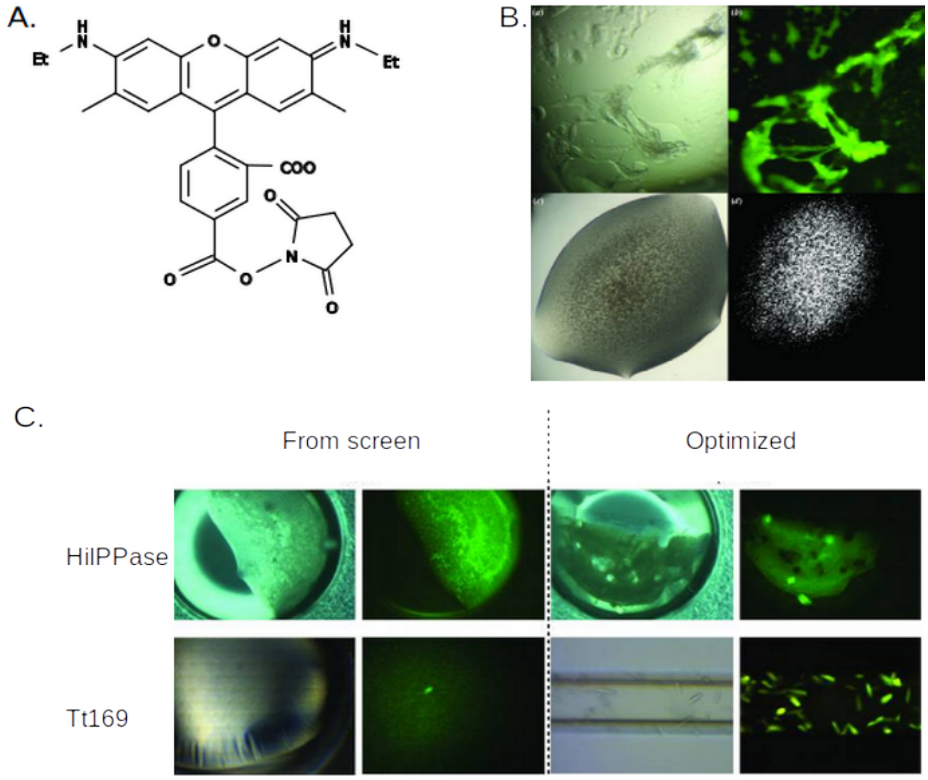


Figure I.8: Usage of trace fluorescent labeling with carboxyrhodamine-succinimidyl ester

A. Chemical drawing of 5-CR6G, SE [5-Carboxyrhodamine 6G, succinimidyl ester (www.aatbio.com). B. Identification of microcrystals clusters with TFL compared to SONICC analysis. C. Effects of trace fluorescent labelling on HilPPase and Tt169 protein before and after optimization. Labeling clearly highlights macromolecule cluster formation in initial screens. When the conditions are optimized, real macro crystals can be identifified under green light illumination. (Adapted from Pusey et al., 2015)

2.3. Phase diagram and Xtal Controller XC900

The dream of every crystallographer might be a single method allowing free navigation in the phase diagram during crystallization experiment. Resulting from the collaboration of the team of Pr. Dr. Christian Betzel (University of Hamburg), Pr. Dr. Rolf Hilgenfeld (University of Lübeck) and the company Xtal Concept (GmbH, Hamburg), the Xtal Controller 900 (XC900) was designed for such an application (Raphaël de Wijn et al., 2020) . Briefly, the instrument consists of a closed chamber where humidity and temperature are controlled (Figure I.9). At the center, the sample (i.e., the macromolecule solution) is deposited on a standard 22 mm siliconized glass coverslip which is placed on four pillars installed on an ultrasensitive balance. The exact weight of the experimental drop (typically

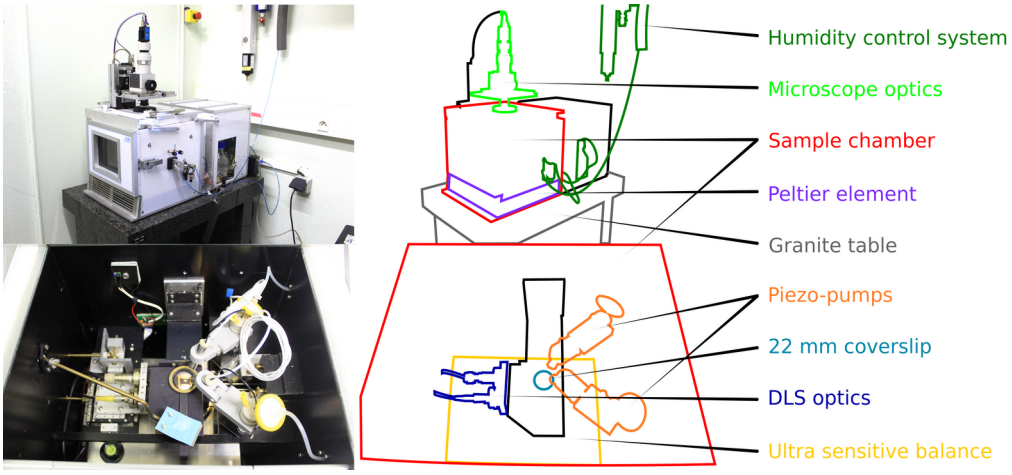


Figure I.9: Xtal controller picture and schematic view

The sample is put on the glass coverslip installed with a specialized holder on the top of the ultra sensitive balance . The DLS laser is oriented in the drop and adjusted via a screw system. The piezo-pumps are connected to 4mL glass vials containing either a crystallant solution or water. The whole injection system is mounted on a mobile metal holder adjusted with screws outside of the machine. The microscope optics is manually adjusted as well. The humidity control system is supplied by pressurized air and connected to a bottle filled with ultrapure water. (Adapted from Rollet *et al.*, 2020)

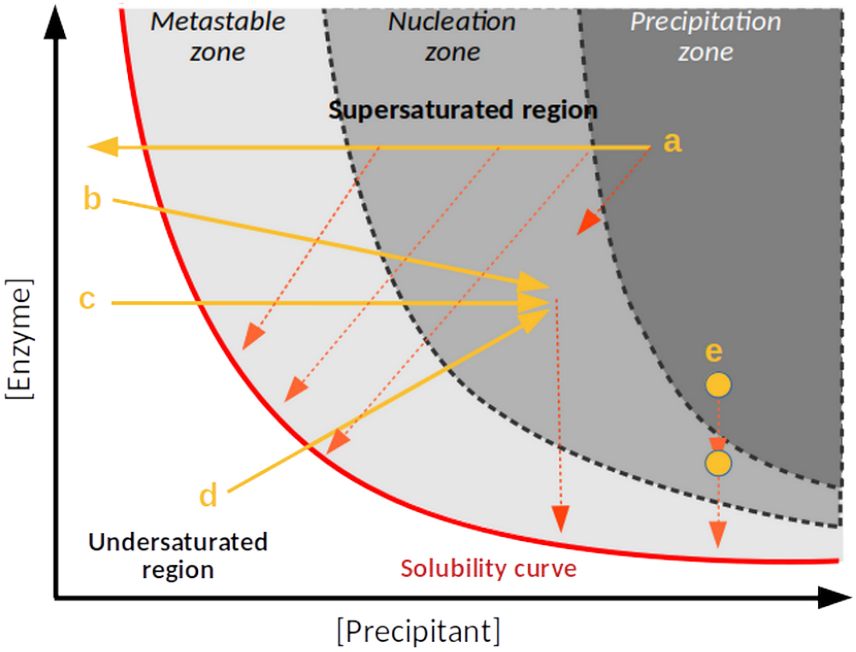


Figure I.10: Impact of different crystallization methods on the phase diagram

(a) Counter-diffusion. (b) The free-interface case. (c) Dialysis. (d) Vapor-diffusion. (e) Batch

from 5 to 10 µL of sample) is recorded to extrapolate the concentration of both themacromolecule and precipitant. On the two sides, piezoelectric pumps dispense (70 pL droplets) either desired crystallant to trigger crystallization or water to dilute or counter-balance evaporation of the crystallization drop. Placed below the glass coverlip, a DLS module provides the user a real-time monitoring of oligomerization and nucleation events happening in the drop. At last, user can visualize the crystallization drop thank to a microscope optic inserted on the top of the machine. A more detailed presentation of the XC900 can be found in chapter V**.** The XC900 can be used to modify the composition of the drop and hence to navigate in the phase diagram, while following the impact on the system's behavior in real time.

3. Sample quality control

A highly pure and homogeneous sample is a primordial prerequisite to go on for crystallization experiments (Lorber et al., 1987). A robust purification protocol should ensure the reproducibility of the sample and a strict quality control procedure is essential to prior any crystallization assay. The following methods provide a good basis for this task and are systematically by used in the team to compare and validate sample batches.

3.1. Analytic gel and Size Exclusion Chromatography

Polyacrylamide gel electrophoresis (PAGE) is a routine method to access sample homogeneity. Species present in the sample migrate through the gel, depending on their surface charge, size and weight. Native gel method informs on oligomeric state, on the presence of complexes and also of aggregates of the macromolecule of interest. Combined with denaturing sodium dodecyl sulphate–polyacrylamide gel (SDS-PAGE) which enables a separation only based on protein mass, it is possible to get a quick overview of sample quality right after the purification process (Lorber et al., 1987).

In addition, analytical size exclusion chromatography provides robust data on sample polydispersity. For instance, a sharp monomodal dispersion of the population would reflect the presence of clear single population, which is often wanted for crystallization attempts. On the opposite, satellite peaks would reveal polydispersity in the solution. Considering an homogeneous specie (examined by PAGE), the position and shape of the peaks inform on

either the presence of multimers, complexes, aggregates (on left side of monomer peak) or the potential degradation of the sample (on the right side of the monomer peak). This last decade saw a drastic improvement of separation power of size exclusion columns, so pushing the community to implement the tool upstream of other analytical methods, such as SAXS (small angle X-ray scattering)(Théobald-Dietrich et al., 2020; Yasushi Watanabe & Inoko, 2009) or MALS (multi-angle light scattering) (Some et al., 2019). This setups provide users (i) information on the quality of the sample and (ii) controlled dispensing and analysis of desired population from the solution.

3.2. Dynamic and static light scattering

Dynamic light scattering (DLS) is a non-invasive, low cost and fast method to explore sample stability and homogeneity (Schmitz, 1990). This technology is based on fluctuation of light scattered by particles (i.e., proteins, nucleic acids, complexes of both entities) moving freely in a solution. Knowing the viscosity of the solution, the concentration and approximative shape (e.g., globular or elongated) of the macromolecule, it enables a direct access to mono or polydispersity (spread of size of a distinct population), as well as size estimation of the molecule but also to presence of bigger particles (e.g aggregates or high-order complexes).

Scattering of particles smaller than wavelength of light was first established in 1871 by John William Strutt, better known as Lord Rayleigh, by explaining the sky-blue color is due to atmospheric particles that scatter light (Strutt, 1871) This particular kind of scattering was named after him as “Rayleigh scattering”. In 1906, right after describing the Brownian motion, Albert Einstein combined light scattering and diffusion behavior of molecules, described by Stokes in 1845 (Stokes, 2009) which linked the friction produced by moving particle to its radius and the surrounding solvent. This statement was summarized by the Einstein-Stokes equation (Stetefeld et al., 2016):

$$D_\tau = \frac{kT}{6\pi\eta R_h}$$

Where D_τ is the translational diffusion coefficient, k is Boltzmann coefficient (1.380 × 10−23 kg.m^2.s^{-2}.K^{-1}), T is an absolute temperature, η is the viscosity of medium and Rh the hydrodynamic radius. Here, it seems relevant that size, shape and mass of a given particle in a defined medium at a given temperature have a direct influence on D_τ. It was shown that objects with a size below λ/10 produce Rayleigh scattering.

Additionally, D_τ can also be used to calculate the frictional ratio of particles *f/f0* (Stokes radius over the volume of unsolvated molecule approximated to a sphere) which informs on solution conformation of macromolecules. *f* is defined as translational frictional coefficient obtained with following equation :

$$f = \frac{RT}{N_A D_\tau}$$

Where *R* is a gas constant (8,314 J.mol^{-1}.K^{-1}), *T* the absolute temperature and N_A the Avogadro's number (6,022 x 10^{23} mol^{-1}). *f/f0* increases as the shape deviates from a compact sphere.

With the development of laser optics and digital autocorrelator in the 1960's, modern light scattering can access with high reliability the characterization of macromolecules in solution and to a larger extent, the absolute quantification of present species (Stetefeld et al., 2016). Nowadays, vast majority of DLS equipment uses wavelength in the range of 600 to 660 nm which interaction with particles (scattering) can be recorded at a fixed angle by a detector (Figure I.11). The signal is processed by an autocorrelator, which serves to determine the normalized second order autocorrelation function (ACF). The latter describes the motion of particles in solution by correlating variations of scattered light intensity on a determined time delay (i.e., measurement of intensity at time t and time t+ τ).

Since the early 1990's, DLS has spread in laboratories as quality control prior to crystallization experiments, allowing the detection of mono or oligomers as well as high-order complexes, or even in the worst case the formation of aggregates. At the same period, the idea of using DLS to monitor nucleation events of biological macromolecule has been introduced (V Mikol et al., 1989; Wilson, 1990), using hen white egg lysozyme as model. While high energies of soft and hard X-rays could potentially disrupt forming nuclei, visible light lasers allow exploration of early crystallization stages in a non-destructive way. As a matter of fact, many studies illustrated how DLS can reveal the evolution of size distribution, ranging from few to hundreds nanometers, in crystallization experiments of model systems (e.g., trypsin or thaumatin).

Recent advances in X-ray Free Electron Lasers (XFEL) and Serial Synchrotron Crystallography brought DLS back in the spotlight to establish nucleation points and determine reproducible conditions to grow nano to micro-scale crystals (Boutet et al., 2018; Chapman et al., 2011). Following this idea, the Xtal Controller (described above) was

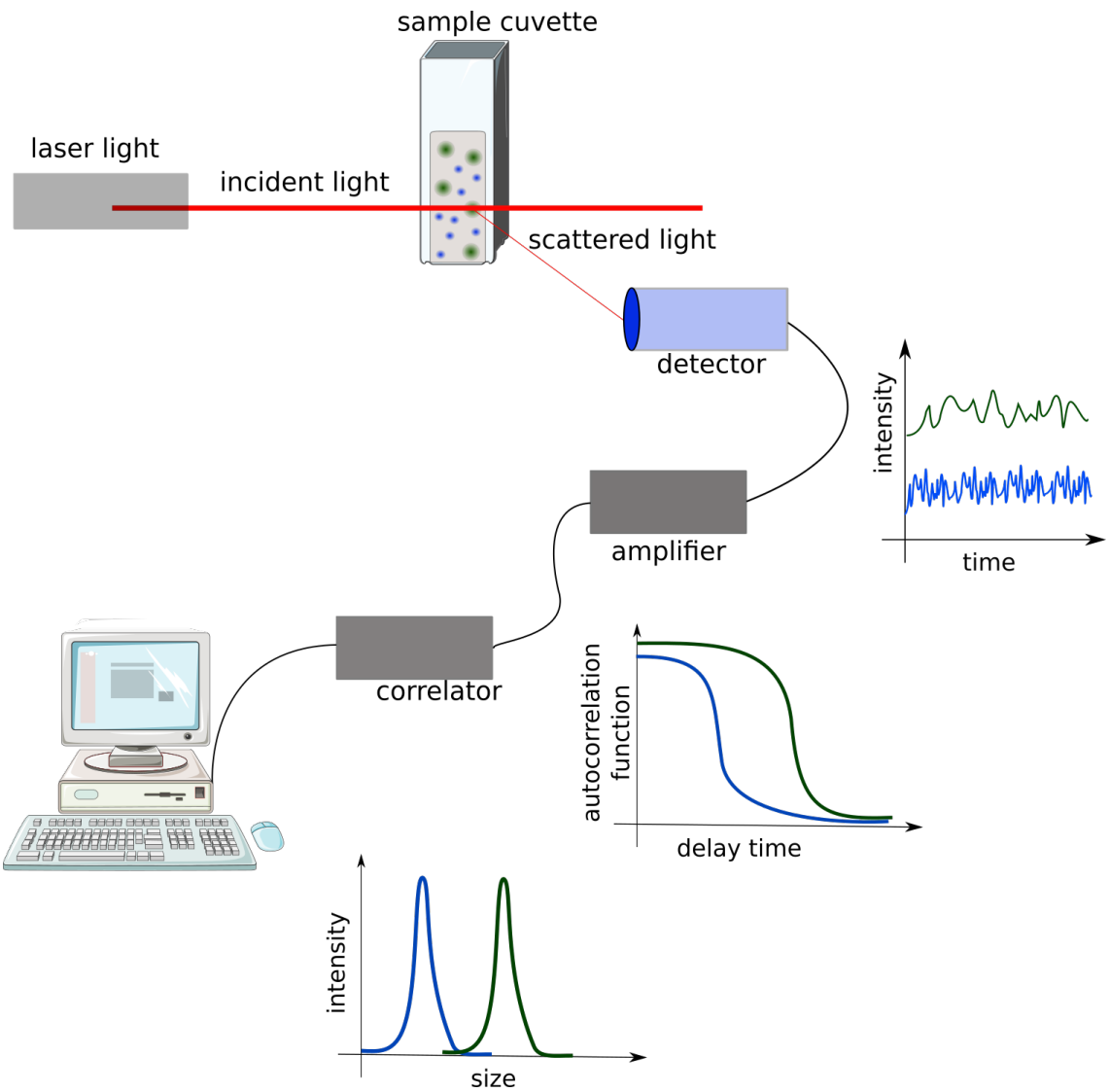


Figure I.11: Dynamic light scattering principle

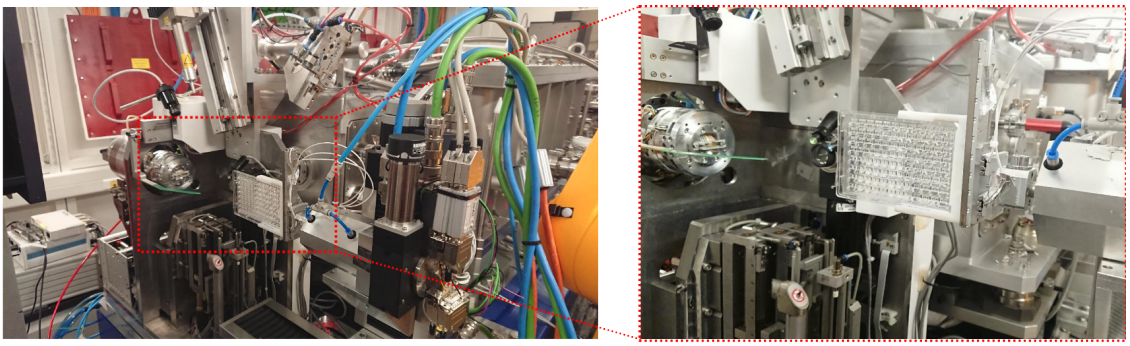
The sample to be analyzed is placed in a cuvette crossed by a red laser light. A part of the scattered light is detected and recorded at a fixed angle by the detector. After amplification, the signal is processed by a correlator which is essential to determine the autocorrelation function. Finally, data are processed by dedicated software and user have access to the size distribution in the sample. Blue and green colors corresponds to small and bigger particles respectively.

designed to control and follow nucleation in real time using DLS apparatus.

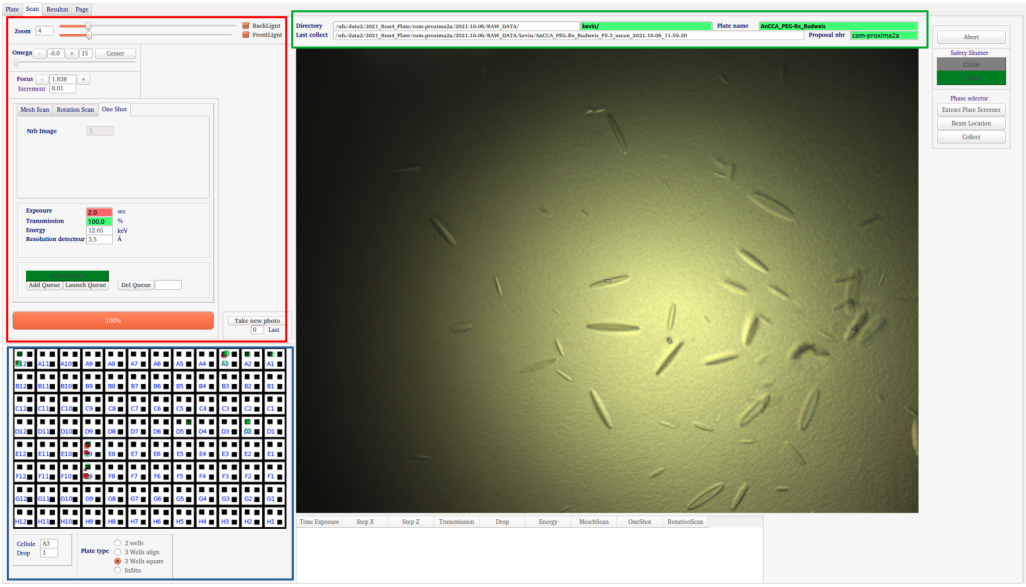
3.3. Detecting crystallization hits

Starting a crystallographic project *ex nihilo* means first to determine physico-chemical conditions allowing the birth and growth of crystals. To do so, the current and most common

A.



B.1.



2.

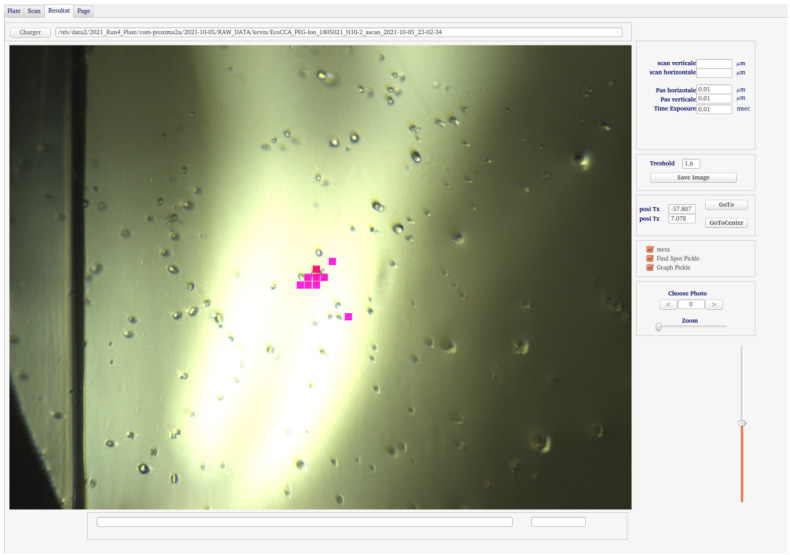


Figure I.12: Plate screener system at Proxima 2A

A. Overview of the plate screener setup. A 3D printed holder is fixed on motorised translation tables which allow horizontal, vertical and along the X-ray beam path movements to align the drop of interest. A supplementary table is dedicated to ω rotation axis with an angular range of +/-45°. A coaxial microscope completed with a back and frontlight allow the focusing on the drop from the experimental

hutch.The beamstop is placed behind the plate like in standard loop setup. B.1. Snapshot of the Python Qt4 interface controlling the data collection. The green box shows the name and access path of the loaded plate. The red box is dedicated to data collection parameters (e.g., ω rotation range, exposure time, resolution limit, transmission and energy). The blue frame shows the plate map which allows to pick desired position by simple one-click. B.2. Grid-scan analysis window to identify potential diffracting area. The threshold (arbitrary units) is manually applied.

strategy is based on screening of a wide range of different conditions. In this regard, a set of companies created and commercialized 96 conditions screens that can be dispensed in standard crystallization plates, facilitating the research process. In addition, the laboratory from Janett Newman established an online tool called C6 (Comparison of Crystallisation Conditions @ C3) (Newman et al., 2010), implemented on the C3 (Collaborative Crystallisation Centre) platform, which compares all commercially available screening kits. Combined with advanced crystallization methods described above, this swiss army knife website should be consulted prior to any experiment to target the largest set of distant conditions. Following this experimental scheme, only two to three 96-plates should be sufficient to obtain initial hits.

Once first screening plates are set up, users will usually observe various behaviors in droplets. Depending on the conditions, the appearance of precipitate (micro-crystalline or amorphous), small crystals with “nice” geometrical aspect or “ugly” shape or even spherulites can reflect different phenomenons (Bergfors, 2009). Determining the nature of such objects is still very empirical with regular laboratory tools. Traditionally, the usage of cross polarizers installed on stereo microscopes helps at least to discriminate salt from protein crystals. We showed above that fluorescent labeling as well as intrinsic UV fluorescence of protein are nice indicators to identify hits but may have some practical limitations (Desbois et al., 2013).

In the particular case of precipitates, several methods have been introduced to determine their crystalline or amorphous nature. As an example, SONICC (Second Order Nonlinear Imaging of Chiral Crystals) (Kissick et al., 2011) is based on the doubling of frequency of laser light shined on chiral crystalline material (e.g., protein crystal). Such equipment is still rare, quite expensive and does not always provide clear results on the nature of the sample. To date, the only real proof of getting crystalline material (e.g., protein crystal precipitate or few micrometers protein crystals), stands in diffraction experiments, either under an X-ray beam or by cryogenic electron microscopy (cryo-EM). The former technique should lead to powder diffraction-like patterns, as the X-ray-beam is interacting with a bulk of

stacked micro or nanocrystals. In contrast, micro cryo-EM diffraction (micro ED) gives access to single micro to nanocrystals analysis. Thus, like photons from X-rays, electrons are able to interact with matter and can be used for diffraction experiments on very small (< 1 µm) crystals.

For decades, the routine consisted of mounting small crystals on a microloop or collecting precipitate with micromeshes , prior to sample flash freezing and analysis under X-ray radiation. This time-consuming process requires (i) specific skill to avoid breaking tiny crystals during fishing, (ii) adapted cryo-cooling conditions and (iii) is also limited inquantity of sample to be analyzed. To speed up this search, few synchrotron beamlines such as Proxima 2A (Soleil, Saint-Aubin, France) recently implemented plate screener systems dedicated to *in situ* room temperature diffraction analysis (Figure I.12). The main advantage lays in the simplicity of the setup which does not require crystal handling or flash-cooling.

Concretely, standard 96-well plates are inserted on a dedicated plate holder installed on motorized stages allowing *x* and *y* translation for sample centering and ω rotation (with an angular range of +/-45°) for data collection. A camera, a front and a back lights serve for visualization and search for crystalline objects. The system is piloted from the hutch by a Python-based interface shown in Figure I.12.B. Alternatively, the nature of precipitates can also be tested by X-ray powder diffraction or micro-electron diffraction.

4. Strategies for crystal analysis by X-ray diffraction

When crystals are available, they are analyzed to explore the architecture of macromolecules requires methods using electromagnetic wavelengths able to interact with atomic-sized objects. X-rays are ranging from 0.01 to 10 nm and represents the best choice for such applications. When considering a single molecule illuminated by an X-ray radiation, electrons from atom outer shells scatter in all directions of space. Such a signal is too weak for isolated molecules and is, at the moment, not sufficient to reconstruct a detailed structure of the object. In solution, it possible to record and accumulate scattering signal from free molecules to obtain a global shape of the target macromolecule (i.e., SAXS).

However, the periodic arrangement of molecules into a crystal provides positive interference in directions dictated by the packing geometry. Indeed, the crystal acts as an amplifier of individual signals and resulting diffracted beams can be recorded on specialized

detector. As their phases are lost during the experiment (known as the phase problem), positions and intensities of the spots on diffraction patterns . Finally, dedicated softwares serve to calculate an electron density map corresponding to the unit cell content.

4.1. Standard crystal analysis

Prior to their analysis, crystals are usually fished with microloops and flash cooled in liquid nitrogen to extend their lifetime and reduce radiation damage induced by X-rays during data collection (Garman & Owen, 2005). Then, the analysis is traditionally performed either at laboratory X-ray home-source or at a synchrotron. Regarding home-sources, the radiation is generated by a rotating anode system or more recently by molten metal anode (Skarzynski, 2013). Despite drastic improvement of such compact instruments, synchrotron radiation stays the most powerful, stable and coherent source to easily and rapidly collect crystallographic datasets.

Since the early 1960's, several generations of synchrotrons have been built to always improve the power and coherence of light. We are currently observing the transition from the third (e.g., Soleil, Saint-Aubain France) to the fourth generation (e.g., ESRF, Grenoble, France) of synchrotron radiation. The principle of the synchrotron is summarized in Figure I.13. Briefly, electrons are generated in a central element called the linac (linear accelerator) where they approach the speed of light. Then, they are injected in the circular booster where they gain in energy. Once the maximum energy (the value depends on the synchrotron) is obtained, electrons are transferred to the storing ring (usually several hundreds meters of circumference) where they will be used by beamlines placed all around the ring. The synchrotron radiation is produced when the trajectory of electrons is modified by electromagnetic devices. Each beamline recovers part of the emitted radiation and thus uses a specific range of energy depending on the dedicated application. Beamlines dedicated to macromolecular crystallography usually are tunable with energies ranging from 6 to 20 keV.

For the last forty years, standard setup and data collection strategy undergone minor changes. Regarding sample preparation, microloops holding crystals are usually flash-cooled in liquid nitrogen or ethane. This trend became the norm in the 1990's to preserve crystals from radiation damage induced under X-ray exposure. Typically, the microloop is mounted on a goniometer head allowing rotations and translations in space, essential for sample centering and dataset collection. A first X-ray shot serves to determine the diffraction power of a crystal prior to the rotational collect. If diffraction at reasonable resolution is observed (e.g., < 3 µm),

then the goal of the experiment is to collect all reflections (or their symmetric) within the crystal lattice at least one time. Depending on the internal symmetry of the crystal, a full dataset can be recovered from a 45 to 180° rotation scan of the sample (Dauter, 2017) . This workflow highly benefited from robot implementation allowing automatic sample mounting and fast data collection. Nowadays, a full dataset can be collected in less than a minute.

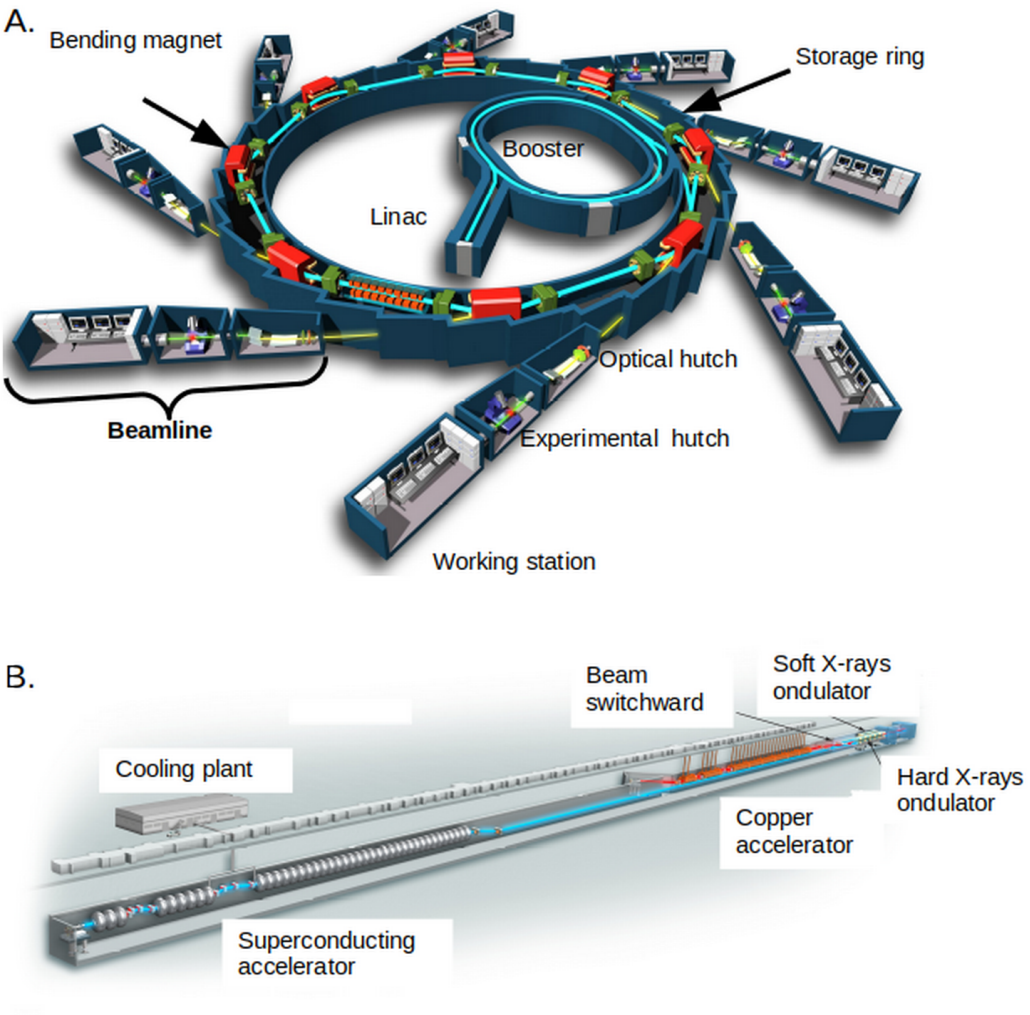


Figure I.13: Synchrotron and XFEL technology

A. In a synchtron, electrons are produced in the linac and accelerated in the booster. Once accelerated close to light speed, they are injected in the storage ring. The latter is composed of a succession of short straight parts containing magnets. Each time electrons are deviated, they emit photons which are collected in optical and experimental hutches (drawing adapted from www.wikimedia.org). B. In a XFEL, the electrons are oscillating at very high frequency in superconducting accelerators. The more the distance and acceleration are high, the more the light produced is powerful and coherent. (Drawing adapted from www.fnal.gov)

4.2. Serial synchrotron crystallography

The principle of serial crystallography emerged in the early 2000's with the improvement of synchrotrons radiation allowing the analysis of smaller crystals (less than 60µm). Typically, it consists of collecting X-ray diffraction images from several crystals grown with the same conditions and combine datasets afterwards. Small crystals are usually more sensitive to radiation damage, so that diffraction quality fades rapidly after limited time of exposition. However, it is possible to collect partial datasets of crystals in different spatial orientation which improves both the completeness and the redundancy of data (Diederichs & Wang, 2017). When considering a standard setup as described above, users can mount series of sample and proceed to X-ray analysis as long as crystals are diffracting at the same resolution. The main weakness of this pipeline relies on the time-consuming sample handling (fishing), preparation (cryo-cooling condition determination) and analysis (mounting and removing each and every single crystal).

Following this concept, specialized equipments such as the microfluidic ChipX (R. de Wijn et al., 2019) have been developed to perform fast *in situ* serial experiments on a unique device at room temperature. This last criteria is essential as it (i) fasten sample preparation and (ii) gives access to structural dynamics (developped later).

4.3. Time-resolved X-ray crystallography

As already mentioned, X-ray crystallography is a powerful tool to explore 3D structure of macromolecules in a fixed conformation. However, biological systems are dynamic and macromolecules are constantly undergoing spatial rearrangements, in particular during catalysis. Biochemical reactions happen from the femto (e.g., small chemical groups wiggling movement) to the microsecond range (e.g., proton release or uptake). In that respect, the idea of creating an ultra-fast imaging methods has been introduced in the early 2000's to capture snapshots from all intermediate steps of macromolecular mechanisms (Chapman et al., 2011). Thus, linear electron accelerators called X-ray Free Electron Lasers (XFEL) have been designed to produce highly powerful (10^9 times more powerful than actual synchrotrons), coherent and brilliant light allowing serial time resolved X-ray crystallography. Indeed, the frequency of photon bunches generated, for instance in the 3.4 km-long euXFEL accelerator (Hambourg, Germany), allows the beam to hit samples at very high repetition rate. To date, there exists only five working XFEL sources in the world as their cost and size is

much more superior than current synchrotrons.

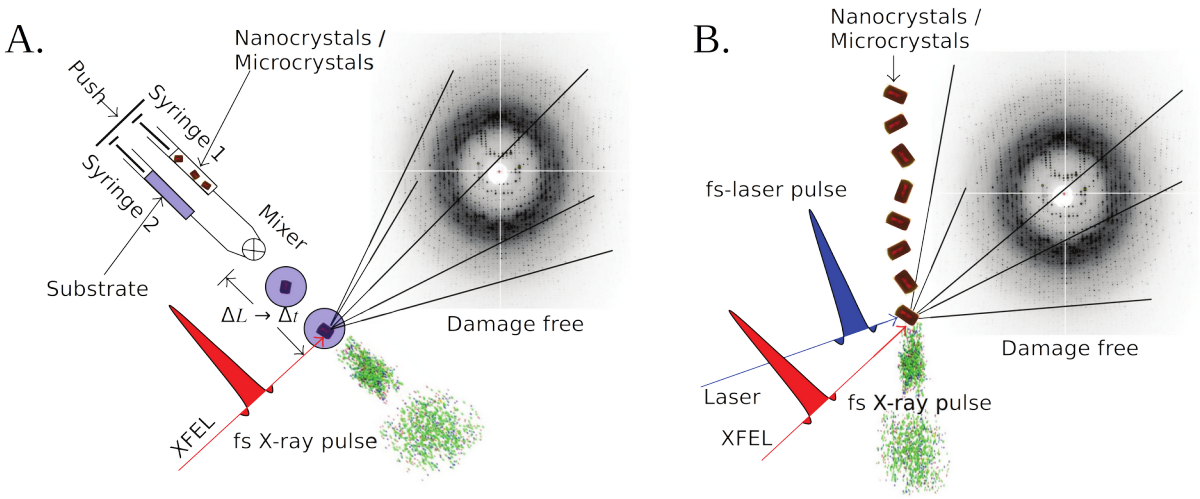


Figure I.14: Principle of serial time-resolved X-ray crystallography at an XFEL source

A. Schematic view of a mix and inject experiment at XFEL. The syringe 1 injects the nano or microcrystals solution while the syringe 2 brings the substrate of the macromolecule. Playing on the length of the incubation ramp or on mixing time, it is possible to catch different time points of the targeted reaction. When the X-ray beam hits a crystals, a single diffraction pattern is recorded before the complete destruction of the sample. B. Schematic view of a setup using photoactivable material. Here, crystals were either soaked with caged compounds or contain naturally photoactivable target. The reaction is triggered with a femtosecond (fs) laser just before the crystals cross the X-ray beam. (Modified from Schmidt, 2013)

Based on the concept of “diffraction before destruction” (Chapman et al., 2014; Doerr, 2011) , the samples (i.e., nano or microcrystals) are brought in a permanent flow in the front of the X-ray beam and are instantly disintegrated because of the intense energy. However, before complete destruction of the samples, a single diffraction pattern can be recorded. By collecting unique diffraction pictures from thousands of crystals in different spatial orientations, it is possible to (i) solve the structure of a macromolecule in its apo or holo-form, and (ii) to reconstruct a full movie of biological catalysis. The latter concept takes advantage of the small size of crystals produced in huge amount (milli to centiliters of crystal suspension required) and most commonly delivered by liquid jet systems (Figure I.14). Despite considerable developments to improve the hit rate of the X-ray beam on running crystals, this setup still induces the loss of more than a half of initial macromolecule stock. Furthermore, XFEL technology has mostly been designed to follow reactions occurring from the femto to

few nanoseconds scale (even if longer reaction are also accessible). At last but not least, such facilities are still hardly accessible for users without strong preliminary results.

Under the impetus of XFEL developments, few synchrotrons started to build and develop beamlines dedicated to serial time-resolved experiments (SSX). As a matter of fact, third and fourth generation light sources produce photon bunches with a sufficient frequency to explore macromolecular mechanisms from hundreds nano- to few seconds order (P. Mehrabi et al., 2021; Pearson & Mehrabi, 2020) (Figure I.15). Moreover, the continuous reduction of beam size but increased power allows the analysis of crystals with a size of a few micrometers. In addition, these beamlines benefit both from decades of synchrotron setups development and from more recent improvements at XFELs which can be easily adapted and implemented for time-resolved experiments. Besides liquid jets, fixed target methods on specialized chips were created to lower sample amount and improve hit rate by automated crystal research (Pedram Mehrabi, Müller-Werkmeister, et al., 2020). On a side note, hybrid systems have been designed, like rolling tapes holding crystallization drops and bringing sample in the front of the X-ray beam.

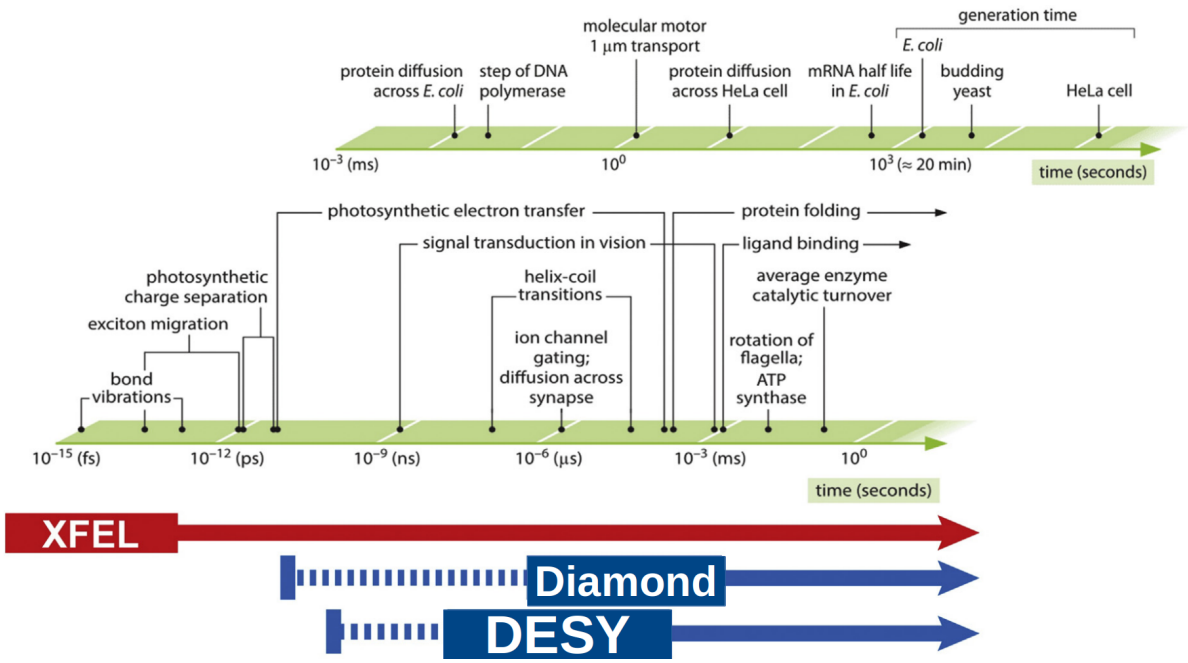


Figure I.15: Time-scale reachable at XFEL and SSX beamlines

(Adapted from Orville, 2020)

As explained above, time-resolved experiments serves to “observe” in real time dynamic interactions and conformational changes occurring during catalytic process. When

the latter uses small ligands (e.g., nucleotides), macromolecule crystals can be pre-soaked in an inactive state thanks to chemical groups called “caged compounds” which are sensitive to given light wavelength. Before crossing the X-ray beam, the reaction inside the crystals can be simply triggered by a synchronized laser.

Recently, the team of Pr. Dr. Arwen Pearson developed the P14 T-REXX (Time-Resolved X-ray crystallography) beamline at DESY:PETRA III (Hambourg, Germany) dedicated to such application and which presented its first results mid-2021. In the frame of a transnational collaboration between our laboratories, my project could benefit from the expertise and access to the facility.

5. Structural study of tRNA maturation enzymes

5.1. Transfer RNA

From the mid 1950’s to early 1960’s, the community put efforts on the understanding of genetic information transmission from DNA to a protein. In 1956, Zamecnik and Hoagland were the first to describe a “soluble RNA” capable of transferring a bonded radio-labeled amino-acid to the ribosome where the protein synthesis occurs (Kresge et al., 2005). In 1958, Paul Berg, in combination with previous observations from colleagues, described an “adaptor” RNA capable of carrying a specific amino-acid added by an amino-acyl synthetase (Barciszewska et al., 2015; Friedberg, 2014), thus using the word “transfer RNA” (tRNA). Folling this major discovery, Aaron Klug and Alexander Rich, solved and published the first $tRNA^{Phe}$ structure from yeast in 1973, revealing the typical L-shape of the molecule shown in Figure I.16. (Robertus et al., 1974).

Six decades of research were necessary to fully decipher the complex life cycle of tRNAs which are matured and regulated by a distinct set of enzymes shown in Figure I.17. After their transcription by an RNA polymerase (pol III in eukaryotes), they are first processed by RNase P and Z which remove the 5′ leader and trim the 3′ trailer, respectively (Hopper & Nostramo, 2019; Phizicky & Hopper, 2010). Then, the CCA-adding enzyme adds the essential 3’CCA tag, prior to its cognate amino-acid charging by aminoacyl-tRNA synthetase (Betat et al., 2010). Once matured and charged, tRNAs are carried by the elongation factor Tu (EF-Tu) to the A site of the ribosome and the amino-acid is transferred on

the C-terminus part of the nascent peptide in the P site (Ban et al., 2000).

As the most abundant non-coding RNA in the cell, tRNA can actually be found in every compartments. In eukaryotes, mitochondrial genome encodes a subset of tRNAs (22 in *H. sapiens*) (Suzuki et al., 2011). Interestingly, several mitochondrial tRNAs (mt tRNA) display a simplified structure, as they are lacking either or both D an T-arm, especially in nematodes where this evolutive adaptation is the most extreme (Okimoto & Wolstenholme, 1990; Wende et al., 2014; Wolstenholme et al., 1987).

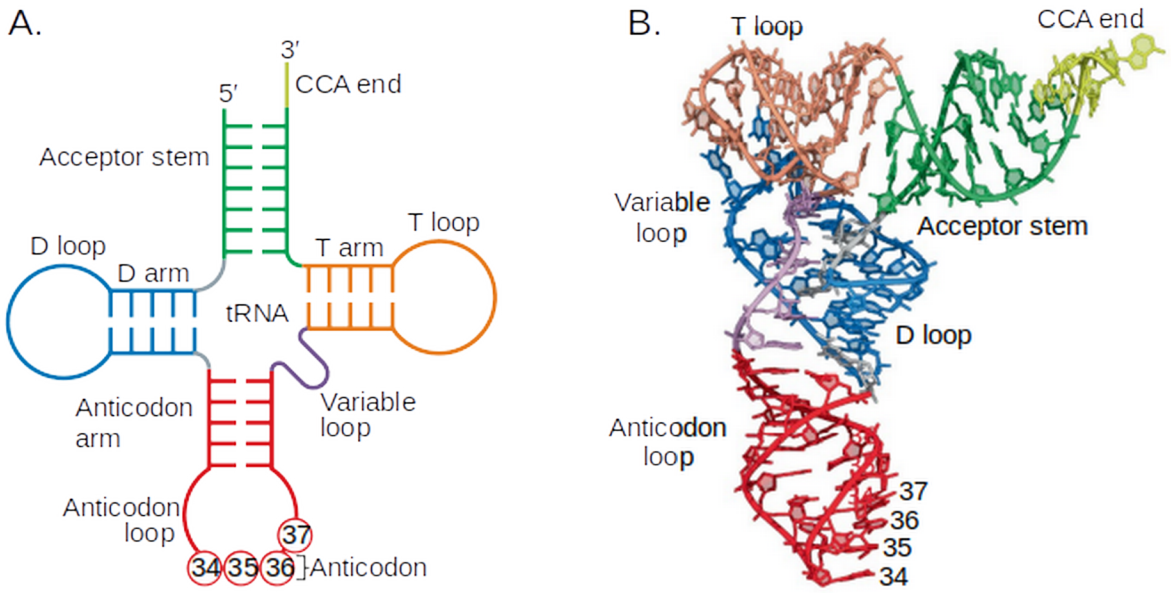


Figure I.16: Canonical tRNA structure

A. Cloverleaf 2D structure of tRNA composed of five main parts: the acceptor stem (green) carrying the 3'CCA trinucleotide, the D-arm and loop (blue), the anticodon arm and loop (red), the variable loop (purple) and the T-arm and loop (orange). B. L-shape 3D structure of yeast $tRNA^{Phe}$ (PDB ID: 1EHZ). The 3'CCA is not paired and points outside to allow interaction with the aminoacyl tRNA synthetase. (Adapted from Suzuki, 2021)

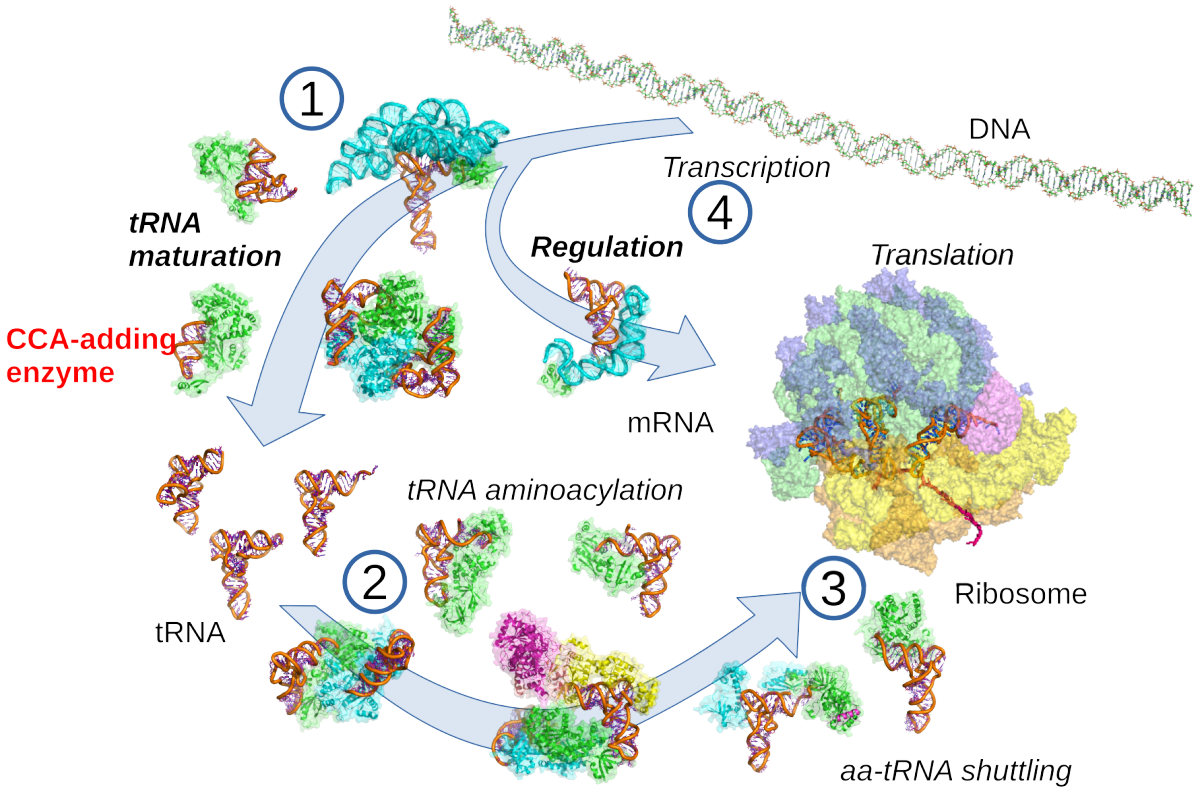


Figure I.17: tRNA maturation and partners

This schematic view of tRNA life shows major steps of tRNA interactions before entering the ribosome. (1) The maturation of tRNA by RNAse P , tRNAse Z and CCA-adding enzymes. (2) tRNA is charged with its cognate amino-acid by aminoacyl tRNA synthetase. (3) EF-TU elongation factor carries tRNA:aa to the ribosome. (4) tRNA can act as a regulatory element, as for instance shown here in complex with T-box riboswitch. (Adapted from Fernàndez-Millàn *et al.*, 2016)

5.2. CCA-adding enzymes

5.2.1. Classification

Throughout their metabolism pathway tRNAs undergo several maturation steps before amino-acid charging on their 3'CCA end. In eukaryotes, most bacteria and archaea, the essential 3'CCA end is not encoded in tRNA genes and needs to be post-transcriptionally added by specific enzymes called ATP(CTP):tRNA nucleotidyltransferases or CCA-adding enzymes. In addition, these specific transferases are involved in tRNA regulation by synthesizing a 3'CCACCA tag for cellular degradation pathway on regular tRNA and small RNAs (Kuhn et al., 2015; Wilusz et al., 2011).

These essential proteins belong to the polymerase ß superfamily (Aravind & Koonin, 1999; Yue et al., 1996) characterized by the common amino acid signature hG[GS]x(9,13)Dh[DE]h (x represents any amino acid, h represents a hydrophobic amino acid), which comprises poly(A) polymerase, DNA Polymerase β, terminalnucleotidyltransferases (TdT), antibiotic nucleotidyltransferases, proteinnucleotidyltransferases, 2′–5′-oligo(A) synthetases. The CCA-adding enzymes family is split into two classes (Aravind & Koonin, 1999; Yue et al., 1996). The class I enzyme is found in Archaea and comprises various enzymes such as polymerase ß and other members of the polymerase X family, protein nucleotidyltransferases, terminal nucleotidyltransferases (TdT), antibiotic nucleotidyltransferases, eukaryotic poly(A) polymerases, 20–50-oligo(A) synthetases, Trf4-like poly(A) polymerases, archaeal tRNA nucleotidyltransferases, and many others. This class does not present any sequence conservation from one species to another, even in catalytic parts where the nucleotide addition occurs.

On the opposite, the class II is restricted to bacterial and eukaryotic tRNA nucleotidyltransferases and bacterial poly(A) polymerases. This group also presents a highly conserved N-terminal domain of 25 kDa (detailed in the next section) (Li et al., 2002a). Furthermore, some bacteria and eukaryotes are characterized by the existence of distinct enzymes showing either a CC- or A-adding activity (Neuenfeldt et al., 2008; Tomita & Weiner, 2002) resulting from ancient evolutive divergence.

5.2.2. Structure and mode of action

Whereas the majority of polymerases needs a nucleic acid template, tRNA nucleotidyltransferases do not require any template, as their mode of action is only based on their structural properties. In the frame of this thesis work, we chose to focus on candidates from the class II family which are presented in the following sections. Both classes I and II display four main domains corresponding to a head and a neck forming the catalytic part of the enzyme, and a body and a tail (Figure I.18). On one hand, class I enzymes have a U-like shape and are close-related in their structure-function to eukaryotic poly A-polymerase. On the other hand, class II enzymes have a typical seahorse-like (Li et al., 2002b) or question mark shape resembling to bacterial PAP from *Escherichia coli* (Toh et al., 2011b). The head and the neck domain contain the catalytic site which recognize and binds top half of tRNAs (Tomita et al., 2004). The body and tail domains interacts with the sugar-phosphate backbone

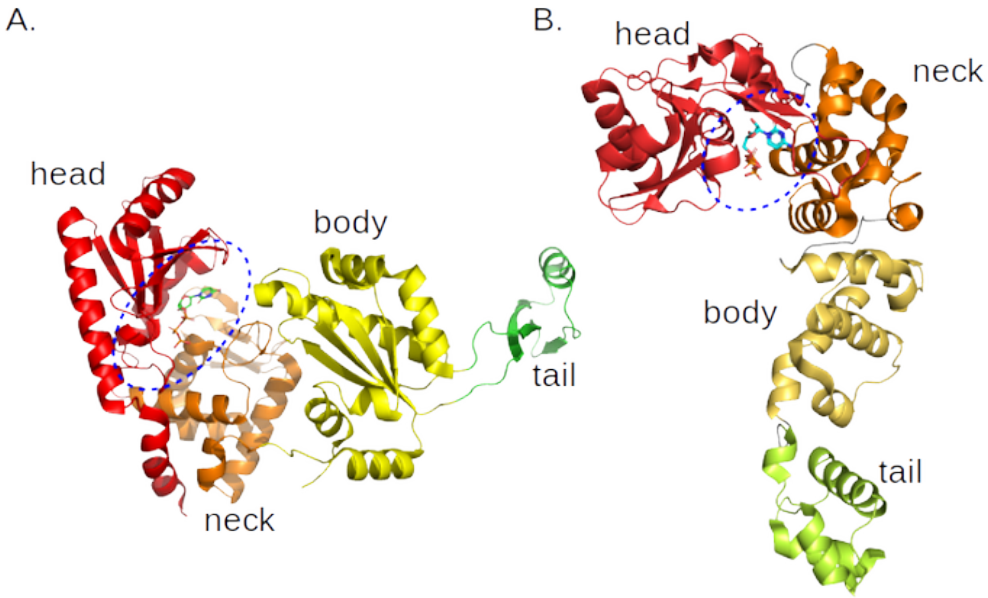


Figure I.18: CCA-adding enzyme architecture

A. Crystal structure of the class I CCA-adding enzyme from *Archaeoglobus fulgidus* in complex with UTP (PDB ID: 1R8C). The head the neck and body form a typical U-like shape. B. Crystal structure of the class II CCA-adding enzyme from *Planococcus halocryophilus* in complex with CTP (PDB ID: 6QXN). The whole enzyme has a typical sea-horse or question mark shape. The head, neck, body and tail are colored in red, orange, yellow and green, respectively. The blue dashed circle highlights the catalytic site.

on the acceptor and TΨC helices of tRNA.The head and neck carry highly conserved motifs that are essential in the catalytic process. The motif A carries the conserved sequence D*x*D (*x* is any amino-acid) (Holm & Sander, 1995) which coordinates two metal ions and catalyses the incoming nucleotide onto the tRNA (Sawaya et al., 1994). A following flexible loop, essential for the A-addition activity, presumably acts as a lever to accommodate the incoming ATP during nucleotide switching. For a proper A addition on tRNA, this loop is preceded by a B/A motif (Basic/Acidicresidues) which positions the 3'OH for the nucleophilic attack on the ATP. Indeed, when the flexible loop is deleted or shorten, the enzyme looses its A-adding activity. Then, a motif B helps for discrimination between ribose and deoxyribose, assuring specific enzymatic activity on RNA substrates (Li et al., 2002a). The motif C acts as a spring element to reorient the head and neck domain during the whole addition process (Ernst et al., 2015). While this polymerase does not require a nucleic acid template for its activity, it mimics Watson-Crick-like interactions between incoming nucleotide and highly conserved arginine and aspartate contained in motif D (Li et al., 2002a). Finally, motif E stabilizes the positioning of the tRNA acceptor stem in the catalytic core.

In the frame of this work, we used CCA-adding enzymes from three different origins as models to better understand and deciphering structural intermediates during the trinucleotide addition. The complexity of this mechanism based only on structural reorganization of specific domains make this enzyme a perfect candidate for time-resolved x-ray crystallography experiments. Furthermore, kinetic studies on class II CCA-adding enzymes indicate a quite slow catalytic process, which is an advantage for SSX technology. For instance, complete dissection of the process on *E. coli* enzyme revealed that single nucleotide addition and switching occur from the micro to millisecond range as long as the substrate used is adapted to the enzyme (Dupasquier et al., 2008; Kim et al., 2009). In that respect, efforts have been put in sample preparation suitable for the emerging technology of time-resolved crystallography. The main idea would be the real time visualization of the complete catalytic process.

From a different point of view, it seems important to notice the importance of the CCA-adding enzyme in human. As a matter of fact, in *Homo sapiens*, this unique polymerase is encoded by TRNT1 gene essential for cell survival. Various mutations have been reported to highly impact the functionality of the enzyme, thus causing a broad spectrum of disease (Slade et al., 2020). To date, all undergoing mechanisms are not well understood but future investigations should lead to adapted treatments and patient care.

5.2.3. *Planococcus halocryophilus* – adaptation to the cold

These last twenty years, the scientific community has put more interest in the understanding of biogeochemical cycles of melting arctic permafrost and the contribution of cold adapted microorganisms in this process. However, we still have partial information on psychrophilic organisms while climate change is accelerating. Among isolated organisms from Canadian permafrost, *Planococcus halocryophilus* showed normal growth and division at -15°C and 18% NaCl as well as an active metabolism at -25°C, pushing cold-adaptation mechanisms at farthest known limits (Mykytczuk et al., 2012, 2013). Such observation arise questions concerning both the evolution events and molecular mechanisms allowing life in extreme cold environments. Regarding enzymes, several studies highlighted the reduced amount of energy needed to fulfill their role in cold conditions which is directly correlated to higher protein flexibility. In the frame of this work we focused on functional and structural properties of the essential CCA-adding enzyme of *Planococcus halocryophilus* (PhaCCA). Previous investigations pinpointed the decreased fidelity of PhaCCA (Ernst et al., 2018) compared to its meso- or thermophilic analogs. To go further, we gathered both structural and biochemical data obtained for PhaCCA and compared them to the thermophilic homolog from *Geobacillus stearothermophilus* (GstCCA) in order to decipher undergoing cold-adaptation mechanisms.

5.2.4. *Romanomermis culicivorax* – adaptation to minimal tRNAs

In metazoan, tRNAs are encoded in nucleus but also in mitochondria which contains a subset of 22 tRNA genes, sufficent for mitochondrial proteins synthesis. While nuclear encoded tRNAs show a high sequence conservation, mitochondrial ones usually show evolutive divergence and shorter length. Both kind of tRNAs are successfully processed by a specialized set of enzymes (Figure I.17) such as the CCA-adding enzyme. The latter usually recognizes the typical L-shape of tRNAs and anchors the TΨC loop on its C-terminal part. Thus, small artifical tRNA-like substrates (i.e., micro or minihelices) are recognized but matured at lower efficiency by this polymerase (P. Y. Shi et al., 1998).

However, minimal tRNAs lacking either the T or D arm were identified in nematodes and proved to be efficiently processed by tRNA maturation machinery in these organisms (Tomari et al., 2002). In *Romanomermis culicivorax*, a roundworm parasite of mosquito

larvae, this particularity has come to an extreme case with a length down to 42 nucleotides, characterized by our teams from Strasbourg and Leipzig (Jühling et al., 2018b; Wende et al., 2014). Thus, the remaining question concerns the undergoing mechanism and adaptation to such reduced substrates.

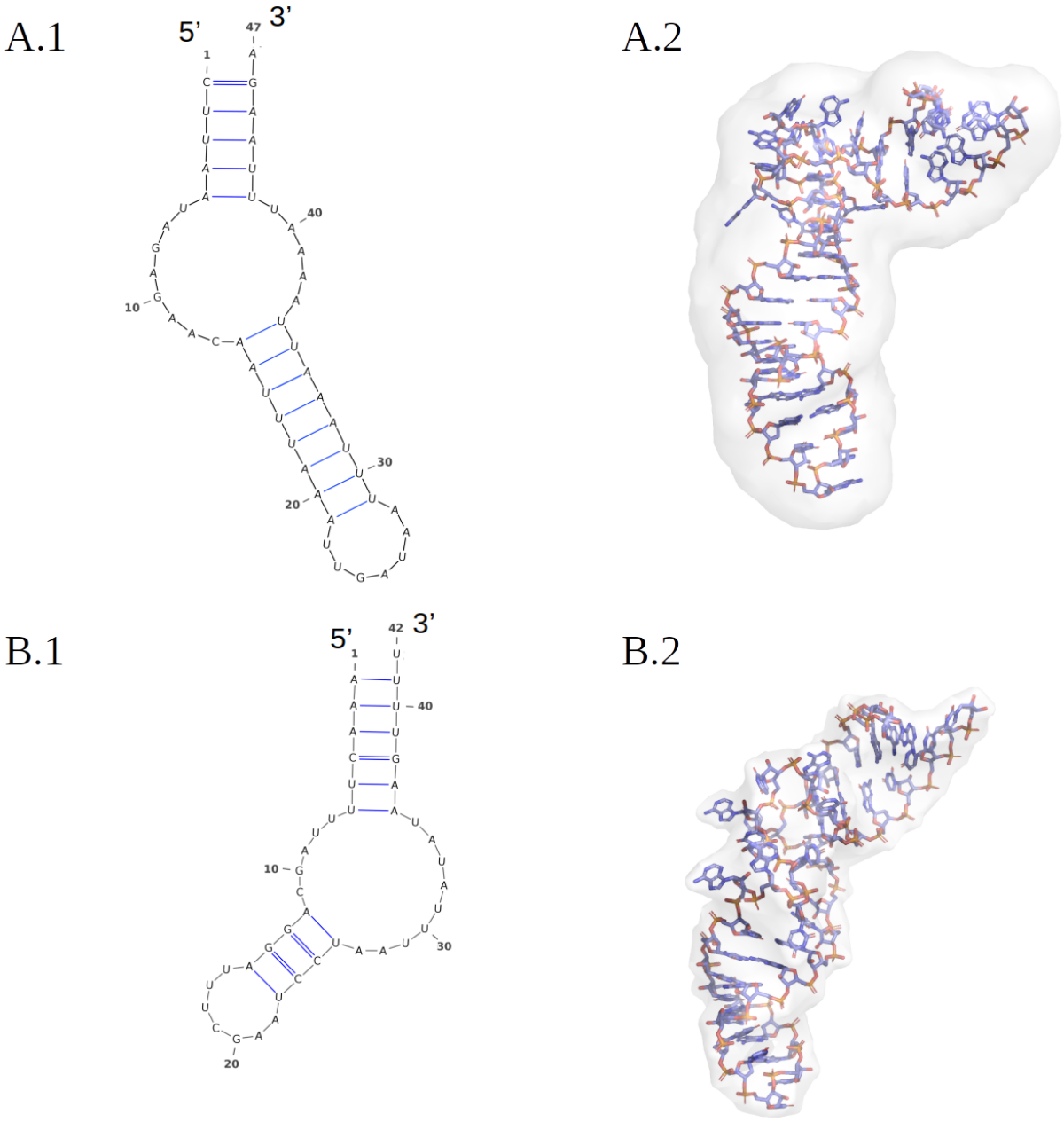


Figure I.19: Minimal tRNAs found in nematodes and in Romanomermis culicivorax

A 2D structure (A.1) and 3D model with SAXS envelope (A.2) of mitochondrial $tRNA^{Ile}$ from R. culicivorax (SASBDB ID: *SASDDF7)*. B. 2D structure (B.1) and 3D model with SAXS envelope (B.2) of mitochondrial $tRNA^{Arg}$ (SASBDB ID: *SASDDG7).*

5.2.5. Ancestral CCA-adding enzyme – Evolutive divergence of class II enzymes

As described above, class II tRNA nucleotidyltransferases regroup bacterial tRNA poly(A) polymerase as well as bacterial and eukaryotic CCA-adding enzyme. The evolutive diversity of this enzyme family highlights their essential role throughout all living organisms. Yet, the precise evolutive origin but also ancestral activity of class II nucleotidyltransferases remain unclear. To bring a first answer to this mystery, the group of Prof. Dr. Mario Mörl with Dr. Marie-Theres Pöhler (Institute for Biochemistry, University of Leipzig) in collaboration with Prof. Dr. Sonja Prohaska (Interdisciplinary Center for Bioinformatics, University of Leipzig) reconstructed the sequence of a Last Common Ancestor (LCA) for the CCA-adding enzyme using 102 sequences of recent gammaproteobacterial CCA-adding enzymes (Figure 1-20). This phylogenetic class presents only one gene encoding the CCA-adding enzyme which constitutes a solid basis for ancient catalyst reconstruction. Moreover, Gammaproteobacteria also have a unique gene encoding a poly(A) polymerase which ancestor can also be reconstructed by bioinformatics. Taken together, both ancestral enzymes would serve in future studies to determine which one emerged first and corresponds to primordial catalyst of class II tRNA nucleotidyltransferases. Finally, recent studies pinpointed that most of ancestral proteins, back from one to four billion years, have an increased thermostability and activity compared to their actual counterparts. This feature has been demonstrated for AncCCA (unpublished work from Dr. Marie-Theres Pöhler) and constitutes a real advantage for X-ray crystallography investigations.

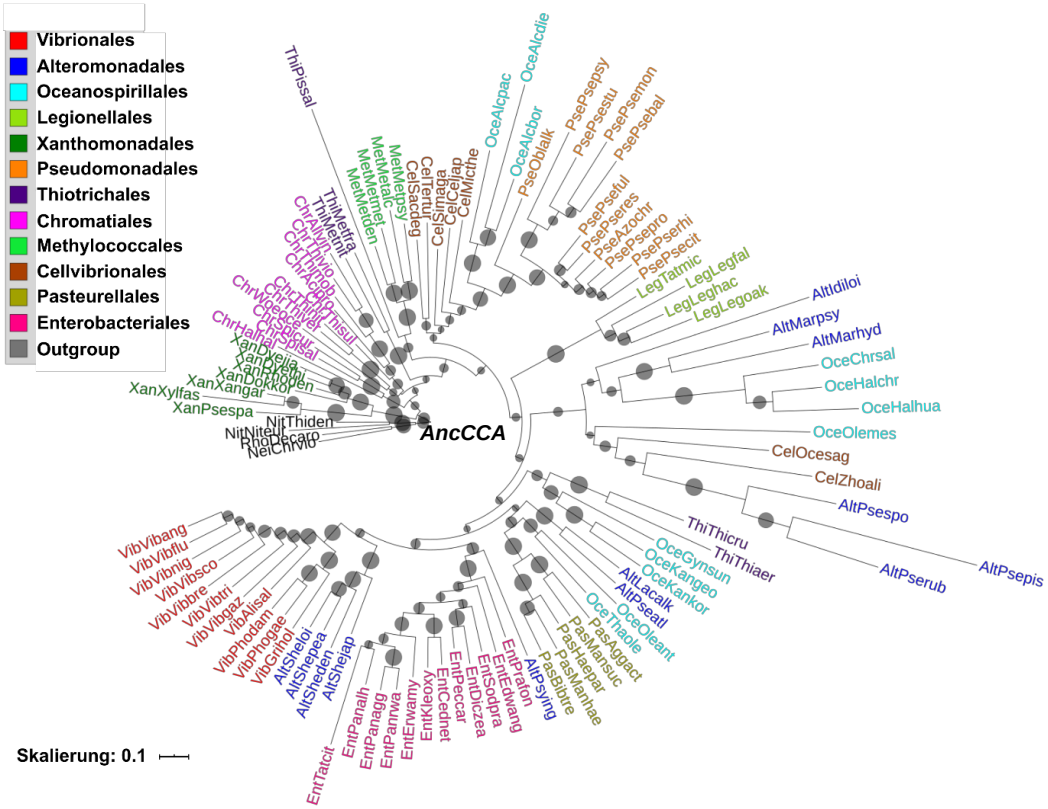


Figure I.20: phylogenetic tree used to reconstruct AncCCA

The ancestral class II CCA-adding enzyme AncCCA was reconstructed from the sequence alignment of 102 actual gammaproteobacteria CCA-adding enzymes. (Courtesy from Dr. Marie-Theres Pöhler)

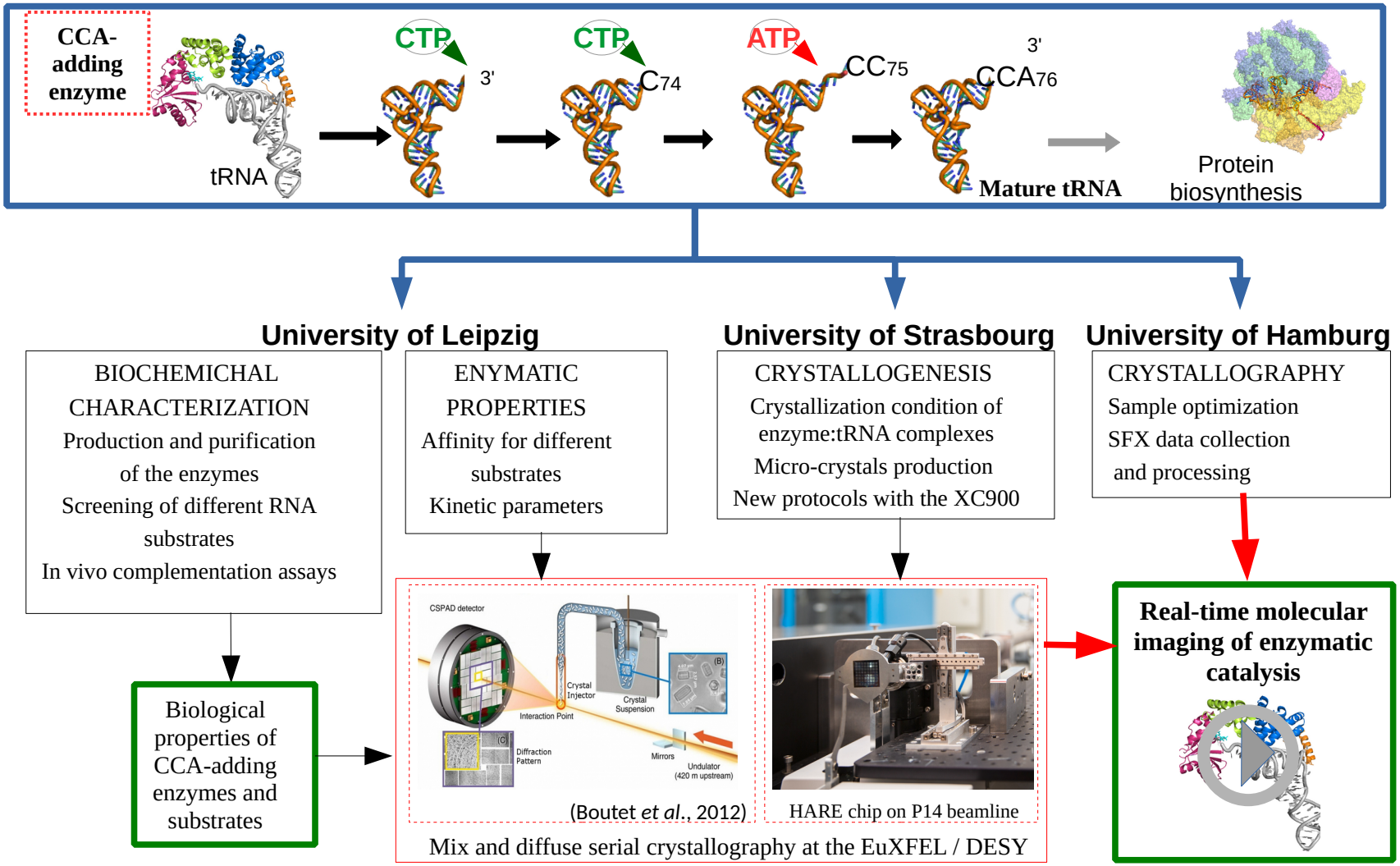


Figure I.21: Pipeline adopted to characterize selected CCA-adding enzymes

6. Goal of the thesis

My PhD work was carried out in the frame of a French-German cotutelle contract under the direction of Dr. Claude Sauter (IBMC, University of Strasbourg) and Prof. Dr. Mario Mörl (Institute for Biochemistry, Leipzig). The initial idea of the project was on the one hand to perform “biochemical” characterization of CCA-adding enzymes in Leipzig and on the other hand the crystallization of the selected candidates in their apo and holo-form.

6.1. Biochemical characterization of targeted macromolecules

The main task consisted of the production and purification not only of the enzymes but also of the different tRNA substrates. Besides the three selected enzymes, a set of eight tRNAs were *in vitro* transcribed and purified. This corresponds to canonical yeast $tRNA^{Phe}$, $tRNA^{Phe\text{-}CC}$, two $tRNA^{Phe}$ minihelices, *R. culicivorax* mitochondrial $tRNA^{Arg}$, $tRNA^{Ile}$, $tRNA^{Ile\text{-}C}$, $tRNA^{Ile\text{-}CC}$. In addition, I could determine the affinity for various substrate by electrophoretic mobility shift assay (EMSA) in the case of the RcuCCA and chimeric enzymes (see chapter IV). Getting information on the specificity and strength of an interaction between an enzyme and its substrate is also a prerequisite for crystallization experiments. Stronger interaction often indicates a more suitable complex to be crystallized. It’s even more important in the case of time-resolved X-ray crystallography experiments, where enzymatic reactions take place directly in the crystalline phase. A too weak interaction would simply lead to complex dissociation, thus triggering crystal dissolution during data collection.

As a complementary project, we wanted to explore the functionality of different CCA-adding enzymes integrated into an *in vivo* system developed in my German laboratory (Wellner et al., 2019). The system based on a modified *Escherichia coli* strain informs on the catalytic efficiency of selected enzymes in an *in cellulo* context. Combined with *in vitro* data, our results should reinforce chosen strategies to explore the CCA-addition mechanism in detail.

6.2. Crystallization and structure determination of apo and holoenzyme

In Strasbourg, my project was focused on the crystallization of enzymes in their apo form as well as in complex with tRNA substrate or nucleotide (i.e., CTP or ATP). This part was split on two main axis. The first one, was to obtain macrocrystals (from 80 to over

200µm) suitable for standard synchrotron analysis. The second part was oriented on the preparation of reproducible and stable samples to be used for X-ray time-resolved crystallography experiments at either dedicated synchrotron beamline (e.g., T-REXX) or XFEL. To do so, we combined all methods presented above from initial screening for crystallization condition to final optimization to improve the quality of samples.

6.3. Toward a molecular movie of enzymatic catalysis

While XFEL technology represents a major advance in the crystallography field, the vast majority of published studies still focuses on macromolecules described for years or decades. Such samples well characterized, easy to purify and to micro-crystallize in huge amount, were essential for an initial proof-of-concept and the rapid growth of the SFX approach. However, to tackle new molecular systems, it is crucial to develop a robust pipeline ensuring the reproducible production of micro-nano-crystalline samples for TR-SFX/SSX. This is a central motivation of my PhD project, together with setting the ground for the application of time-resolved enzymology to the fascinating family of CCA-adding enzymes and the obtention of a molecular movie describing the catalytic activity at atomic resolution.

II. Material and methods

1. Construct and strain

For their overexpression, the ancestral (AncCCA) and *R. culicivorax* (RcuCCA) CCA-adding enzyme genes were cloned in pET28a (Novagen, Bad Soden, Germany) vector. These two constructs carried a 6-histidine tag on the C-terminal side of the enzyme. Gene encoding for *P. halocryophilus* (PhaCCA) enzyme was cloned in a pET30 Ek/LIC (Novagen, Bad Soden, Germany) vector with a 43 amino acid expression tag, comprising the 6-histidine tag, on the N-terminal part of the enzyme. *Escherichia coli* BL21 (DE3) cca::cam (Novagen) were transformed with corresponding plasmid. We were using a knockout strain (Δ CCA), where the gene encoding the endogeneous *E. coli* CCA adding enzyme was knocked out to avoid contaminations with the EcoCCA in the enzyme preparations. Kanamycin and chloramphenicol antibiotics were both used for plasmidic and bacterial selection respectively. This preparative step was achieved by collaborators from my German host team in Leipzig.

2. Protein overexpression

Pre-culture of transformed *E. coli* BL21:DE3 were set overnight at 37°C with 200 rpm shaking (Edmund-Bühler SM-30). Then, 5 mL of pre-culture were transferred in 500 mL of lysogeny broth (LB; Carl Roth GmbH, Germany) medium for BL21-PhaCCA and BL21-AncCCA, or 500 mL of terrific broth (TB) medium for BL21-RcuCCA, both medium supplemented with 1 mM kanamycin and 1 mM chloramphenicol. The cultures were incubated at 30 °C and shaken at 200 rpm. After 4 hours (when the optical density at 600 nm reached 0.6), 500 mL of ice-cold medium supplemented with 2 mM IPTG and antibiotics was added to cultures which were transferred at 16 °C and shaken at 200 rpm. After 16 hours, cultures were harvested, centrifuged at 5000 xg for 20 min at 4 °C (Sorvall RC6 Plus, rotor F10S-6x500 Y). Finally, pellets were washed, resusupended in 20 mL of 1X phosphate buffer saline (PBS) and centrifuged at 4000 xg for 30 minutes at 4 °C (Sorvall Super T21). Supernatant was removed, pellets of approximately 4 g of cells were dried and stored at -80 or -20 °C for future purifications. All along my thesis period, about 250 g of pure cell pellets were necessary to purify the different enzymes.

3. Purification of CCA-adding enzymes

Purification of the different enzymes was carried out in my two host laboratories. Protocols, first designed for the Äkta™ purifier system (GE Healthcare, Freiburg) in Leipzig were then adapted to BioLogic DuoFlow™ HPLC system (Bio-Rad, USA) in Strasbourg. Considering that the procedure is similar for the different macromolecule targets and for more clarity only the protocol used for the AncCCA enzyme is described here.

Pellets were resuspended in 15 mL lysis buffer and cells were disrupted by sonication at 33 kHz four times 2 min on ice. Cell lysates were clarified by centrifugation at 200,000 xg for 45 min at 4 °C. Supernatant was filtered on 0.22 µm membrane prior to loading on 1 or 5 mL HisTrap™ Fast Flow column (GE Healthcare, Freiburg) equilibrated with HisTrap binding buffer. The enzyme was released/eluted from the column by applying an imidazole gradient with HisTrap elution buffer. Fractions of 1 mL containing the target macromolecule were identified with optical density (280 nm) record on the chromatogram and then were pooled

Buffer	**Composition**
Lysis	25 mM Tris/HCl pH 7.5 200 mM NaCl 10 mM $MgCl_2$ 10 mM imidazole 1 mM DTT 2 U DNase I (Sigma-Aldrich) 1/2 Complete™ protease inhibitor tab (Roche)
HisTrap binding	25 mM Tris/HCl pH 7.5 200 mM NaCl 10 mM imidazole
HisTrap wash = Heparin elution	25 mM Tris/HCl pH 7.5 1.5 M NaCl 10 mM imidazole
HisTrap elution	25 mM Tris/HCl pH 7.5 200 mM NaCl 500 mM imidazole
Heparin binding	25 mM Tris/HCl pH 7.5 150 mM NaCl 10 mM imidazole
Size exclusion (=storage buffer)	25 mM Tris/HCl pH 7.5 150 mM NaCl 5 mM $MgCl_2$ 5 % (w/v) Glycerol

Table II.1: Buffer used for AncCCA enzyme purification

(Figure 3-1.A). In order to get rid of nucleic acid contamination, pooled fractions were applied on a 5 mL HiTrap™ Heparin column (GE Healthcare, Freiburg) equilibrated with

heparin binding buffer. Pooled fractions were loaded onto the column and washed several times with binding buffer. Then, protein-matrix interaction was abolished with a sodium chloride gradient using HisTrap wash buffer. Fractions containing the enzyme were applied on a 27 mL Superdex™ 75 10/300 (GE Healthcare, Freiburg) column equilibrated with size exclusion buffer. This final step was necessary to remove potential aggregates, control the purity of the sample and finally exchange the buffer. In the end pure fractions were pooled (Figure III.1C.) and concentrated in 0.5 mL Amicon® Ultra centrifugal filter units (Merck Millipore, Molsheim) up to 5 $mg.mL^{-1}$. Samples were either directly used for crystallization assays or stored at 4 °C. Composition of size exclusion (i.e., storage) buffers for each CCA-adding are shown in Table II-2.

Before any crystallization experiment, potential aggregates or contaminant were removed from enzyme stocks by ultracentrifugation at 100,000 xg for 1 hour. Then, supernatant was recovered and systematically checked by dynamic light scattering.

Enzyme	**Isoelectric point**	**Molecular mass (kDa)**	**Storage buffer composition**
PhaCCA	6.13	48.4	20 mM Tris/HCl pH 7.5 200 mM NaCl 5 mM $MgCl_2$
RcuCCA	8.49	53.0	50 mM NaH_2PO_4 → final 50 mM Na_2HPO_4 → pH 7.0 200 mM NaCl 10 % (w/v) Glycerol
AncCCA	6.35	46.5	25 mM Tris/Hcl pH 7.5 150 mM NaCl 5 mM $MgCl_2$ 5 to 10 % (w/v) Glycerol

Table II.2: CCA-adding enzymes storage buffer composition

4. tRNA *in vitro* transcription

CCA-adding enzymes are able to process a variety of non-coding tRNA-like substrates which do not carry base modifications (Wende et al., 2015). Thus, *in vitro* transcribed tRNA are efficiently bound and recognized as substrates by this particular nucleotidyltransferase family. In attempts to obtain crystals of tRNA:CCA-adding enzyme complexes, four different tRNAs (Table II.3) were produced *in vitro* as described by Schürer et al. (2002), Jühling et al.

(2018) and Théobald-Dietrich et al. (2020). The mini-tRNA^{Phe} consists of upper part of the canonical yeast tRNA^{Phe}, that is acceptor stem and TψC arm/loop. Such a substrate was shown to be compatible for co-crystallization with class I (Pan et al., 2010) and class II (Toh et al., 2009; Yamashita & Tomita, 2016) CCA-adding enzymes .

As an example, tRNA^{Ile} and tRNA^{Arg} sequences from *R. culicivorax* were cloned in pCR2.1-TOPO (Invitrogen, Netherlands) with upstream T7 RNA polymerase (T7 RNA-pol) promoter and Hammerhead ribozyme (HH ribozyme) and downstream hepatitis delta virus ribozyme (HDV ribozyme). Sanger sequencing (Sanger et al., 1977) was used to check for right gene incorporation (MicroSynth Seqlab, Göttingen, Germany). Target sequences (T7/HH/tRNA/HDV) were amplified polymerase chain reaction (PCR).

In order to produce enough material for crystallization experiments, original transcription protocol volumes were scaledup to 30X or 50X. Thus, transcription mix was composed of 100 to 200 ng of PCR-products, 1X transcription buffer (80 mM Tris/HCl pH 8.8, 22 mM $MgCl_2$, 5 mM DTT, 1 mM Spermidin, 1.8 µM BSA), 200 mM DTT, 25 mM NTPs, 12 U TIPP, 30 µg T7 RNA-pol. Incubation of the mix at 37 °C for 8 to 10 hours was followed by a ribozyme-cycling (10X, 60 °C for 3 min, 25 °C for 3 min) to cleave off additional nucleotides. Xylene cyanol was added to the mix and tRNAs were separated on pre-heated 10% (tRNA^{Phe}) or 12.5% (tRNA^{Ile}, tRNA^{Arg}, mini-tRNA^{Phe}) 8 M polyacrylamide (PAA) gel at 30 W for 1 to 1.5 h. Visualization of tRNA was done under UV light, bands were cut off, split in 1.5 mL tubes with 400 µL ddH_2O and incubated overnight at 4 °C. Then, tubes were heated at 65 °C for 1 h and precipitated in ethanol as described in Sambrook et al., 1989. For about 500 pmol of tRNA, removal of 2',3'-cyclic phosphate (dephosphorylation) was done in dephosphorylation buffer (100 mM imidazole/HCl pH 6.0, 22 mM $MgCl_2$, 10 mM β-mercaptoethanol, 100 µM ATP, 300 nM BSA), 5 µL (home-made) T4 polynucleotide kinase (T4 PNK) at 37 °C for 1 h. Final tRNAs were extracted with phenol/chloroform procedure and precipitated as mentioned above.

Organism	**tRNA**	**Size (nucleotides)**	**Target enzyme**	**Protocol for synthesis**
S. cerevisiae	tRNA^{Phe}	73	AncCCA, PhaCCA	Schürer et al., 2002
	Mini-helix tRNA^{Phe}	20	AncCCA, PhaCCA, RcuCCA	Théobald-Dietrich et al., 2020
R. culicivorax	tRNA^{Ile}	47	RcuCCA	Schürer et al., 2002; Jühling et al., 2018
	tRNA^{Arg}	42	RcuCCA	

Table II.3: In vitro transcribed tRNAs for co-crystallization experiments

5. Nano Differential Scanning Fluorimetry

Stability of an enzyme can be assessed with its melting temperature (Tm), which corresponds to the temperature where half of the sample is in denatured (unfolded) state. Differential Scanning Fluorimetry (DSF) or Thermal Shift Assay (TSA) is a method based on progressive heating of a sample and the monitoring of denaturation of the macromolecule (Elgert *et al.*, 2020). In a DSF experiment, the target protein is associated with a fluorescent dye and placed in a reaction chamber where a temperature ramp is applied. Throughout the entire experiment the evolution of the fluorescence signal is recorded and plotted against temperature. As the protein undergoes physical changes (e.g., unfolding or aggregation), the signal drastically increases or decreases. At half of this curve jump, the Tm can be deduced.

The last ten years a labelled-free and miniaturized version of the method has been introduced (nanoDSF), using intrinsic protein fluorescence of tryptophans (and tyrosines to a lesser extend) which strongly depends on the 3D structure of their environment (Ghisaidoobe & Chung, 2014). The system has mostly been spread by the NanoTemper company and its Prometheus NT48 machine. Briefly, 10µL glass capillaries are filled with the sample and placed on a peltier holder in the machine. During the experiment, the samples are exposed under a light-emitting diode and fluorescence emitted at 330 and 350 nm is detected. Starting from a low (e.g 15 to 25°C) point, the temperature is progressively increased (from 1°C per minute) until 95°C in a typical experiment. At the end of the experiment, fluorescence values are plotted against temperature and four types of plots are cross-compared and used to determine the Tm. They consist of the first derivative of 330 nm and 350 nm evolution, the 350/330 ratio fluorescence and finally the scattering plot which informs on aggregation events.

NanoDSF system is also used to determine the (de)stabilizing effect of a ligand, different buffer compositions, or detergents on the macromolecule of interest. On more extend, the effects of mutations can also be assessed when compared to denaturation curve of the wild-type enzyme.

6. Crystallization

The different crystallization methods described in section I.2 were employed to obtain CCA-adding enzymes crystals suitable for X-ray data collection at synchrotron light sources.

Experiments were conducted at 20 °C (few attempts at 4 and 25 °C) using devices shown in Table II.4.

Method	Device (provider)	Number of experiments per device	Volume required (µL)	
			Macromolecule (drop)	Crystallant (reservoir)
Vapor diffusion hanging drop	EasyXtal®/ NeXtal® (NeXtal Biotech, Canada)	6 x 15 / 6 x 24	1 - 5	400 – 1.000
Vapor diffusion hanging drop	Linbro plates (Jena Bioscience, Germany)	15 / 24	1 - 10	400 – 1.000
Vapor diffusion sitting drop	CrystalQuick™ (Greiner Bio-One, Germany)	3 x 96	0.1 – 0.4	40 - 80
Vapor diffusion sitting drop	CrystalQuick™X (Greiner Bio-One, Germany)	2 x 96	0.1 – 0.4	40
Counter diffusion	ChipX3	8	5	5 x 8
Counter diffusion	GCB-Domino (Triana Sci & Tech, Spain)	6	0.3 – 0.5	2.500
Batch	Nunc™ MicroWell (Denmark)	1 x 60	0.5 - 5	none
Batch	Impact low profile (Greiner Bio-One, Germany)	1 x 96	0.5 - 5	none
Batch	0.5 / 1.5 mL Microtubes (Eppendorf™, Germany)	1	10 – 1.000	none
Dialysis	Dialysis Button™ (Hampton Research, USA)	1	5	700 – 1.000

Table II.4: Crystallization devices used in this project

Commercial crystallization screens consisted of JCSG++ (Jena Bioscience, Germany), Natrix HT, Crystal screen HT and PEG/Ion HT (Hampton Research, USA). Home-made screens were prepared in 96 Deep Well MASTERBLOCK® (Greiner Bio-One, Germany) with 1 to 2 mL per condition. Stock protein was used at 5 $mg.mL^{-1}$ in all experiments, except for batch crystallization where it was concentrated from 7 up to 12 $mg.mL^{-1}$ (if no aggregation instantly occurred) using 0.5 mL Amicon® Utra centrifugal filter units (Merck Millipore, Molsheim) pre-equilibrated with protein storage buffer. Crystallization droplets in 96-well plates were dispensed and mixed twice with a mosquito® pipetting robot (spt labtech, UK). Plates were sealed with Crystal Clear Sealing Film™ (Hampton Research, USA). Vapor diffusion and batch experiments were conducted following conventional procedures or complemented with low melting point 0.2% (w/v) agarose gel first heated at 90 °C and kept at 35 °C prior to dispensing. For experiments under cryogenic conditions, crystals were fished with MicroMounts™ and MicroLoops™ (MiTeGen, USA), transferred in cryo protecting solutions consisting of 1/3 of 50% (w/v) glycerol, 1/3 of mother liquor and 1/3 of crystallizing solution for 15 to 60 s. Eventually, crystals were briefly plunged in Paratone-N (Hampton Research, USA).

7. Advanced methods

Experiments with Tb-Xo4-01, Tb-Xo4-02 or Lu-Xo4 nucleants (Polyvalan SAS, France) were set following two different protocols from the provider. In the first one, the dried nucleant was resuspended in 10 µL of storage buffer of the target macromolecule to reach a concentration of 100 mM. Then, the solution was added to crystallization drops at a final concentration of 10 mM. In the second procedure, the dried compound was either resuspended in 25 µL of stock protein or, if the protein tends to precipitate, in 25 µL of 100 mM sodium hydrogencarbonate pH 7.1 and directly added to protein stock.

Streak seeding experiments were done with Seeding Tool (Hampton Research, USA) which was plunged in an initial drop with high concentration of nano or microcrystals and streaked in a series of drops with the enzyme prior to crystallant addition. For microseed matrix seeding, the seed stock was either prepared with Seed bead glass kit (Hampton research, USA) following provider recommendations or by crushing manually crystals in the drop with melted glass Pasteur pipette, transferring “crystal debris” in a 1.5 mL microtube and vortex three to four times for 30 s.

Trace fluorescent labeling of enzymes was done following Dr. Marc Pusey’s procedure (Pusey *et al.*, 2015) using 0.6 µL 5-Carboxyrhodamine 6G, Succinimidyl Ester (Invitrogen™, USA) to recover a final volume of about 80 µL of labeled enzyme at 1 to 1.5 $mg.mL^{-1}$.

8. Data collection

X-ray data collection were mostly performed at Soleil synchrotron (Saint-Aubin, France) on Proxima 1, and Proxima 2A beamlines equipped with Dectris Pilatus 6M and Eiger 9M detectors, respectively. Details concerning data collection on PhaCCA crystals are gathered in Table 1 from de Wijn *et al.*, (2021). *In situ* experiments at room temperature were conducted on Proxima 2A beamline with the precious help from Dr. William Shepard, Dr. Serena Sirigu and Damien Jeangerard who is developing the Plate Screener system (Cribleur). The complete setup and procedure are detailed in Figure 1-10 and in Jeangerard *et al.*, (2018). Initially, to determine whether crystals were diffracting or not, a first “one-shot” was tried on crystal extremities. Thus, the incident wavelength was set at 0.980 Å, the detector-sample distance was set at 0.216 m and an exposure time of 0.05 or 0.1 s at 10% of transmission energy was tried. If significant diffraction was detected (yielding enough diffraction spots to

index data), a series of rotation scans around ω angle were collected with +/- 20° amplitude, 100% transmission energy and 0.005 s exposure time per image.

P14/T-REXX beamline

The P14/T-REXX beamline (EMBL, Hamburg, Germany) is a brand new synchrotron beamline equipped with an Eiger 4M (Dectris, Switzerland) detector and designed for time-resolved X-ray crystallography experiments. Regarding the setup on the beamline, the 20.000 positions of the HARE chip system (see Figure 3-11 and Mehrabi *et al.*, 2020) are scanned within 10 min and X-ray data are automatically processed with CrystFEL (White *et al.*, 2012), a software initially developed for high-throughput experiments at XFEL sources. For our experiments, the wavelength was set at 0.97626 Å, the exposure time per image at 0.005 s, and the detector-sample distance at 0.120 m. Prior to chip loading, samples (i.e microcrystal solutions) were quickly checked under light microscope and size of crystals was evaluated inside hemocytomter C-Chip (NanoEnTek, South Korea).

9. Data processing

Dectris detectors used for instance on Proxima 1 and Proxima 2A beamlines compile X-ray diffraction data in Hierarchical Data Format 5 (hdf5) having a ".h5" extension. On the one hand, X-ray diffraction pictures (frames) are regrouped in blocks stored in a set of "data files" (usually 100 frames per "data.h5" file). On the other hand, a "master" file contains the list of all data files generated from a single collection as well as parameters used during the experiment (detector distance, exposure time, oscillation range…). Then, ADXV (A. Arvai, scripps.edu) and ALBULA (Dectris, Switzerland) program are used to visualize diffraction pictures.

For PhaCCA crystals, X-ray data processing and reduction were done as described in de Wijn et al. (2021). Concerning AncCCA crystals, the poor quality of data collected forced us to test the different available softwares gathered in Table 2-5 to exploit the more information available.

Softwares	Provider / Reference
XDS	MPI Heidelberg / Kabsch, 2010
DIALS (CCP4 package)	Diamond light source (UK) / (Winter *et al.*, 2022)
Xia2 (CCP4 package)	Diamond light source (UK) / (Winter, 2010)
Mosflm (CCP4 package)	CCP4 / Leslie, 2005
autoPROC	Global Phasing (UK) / Vonrhein et al., 2011

Tableau II.5: Softwares and packages used for X-ray data processing, reduction and scaling

10. Other softwares

All along my thesis, a selection of different softwares gathered in Table 2-6 were regularly used from the sample visualization to the final refinement and presentation of the model.

Program	Used for	Reference / Provider
Phenix suite	X-ray data and model refinement	Liebschner et al., 2019
CCP4 suite	X-ray data processing and model refinement	Winn *et al.*, 2011
ESPript 3.0	Amino acid sequence alignment to 2D structural elements	Robert & Gouet, 2014
ENDscript 2.0	Combines ESPript 3.0 results and 3D models from the PDB	Robert & Gouet, 2014
elNémo	Low frequency normal mode computation	Suhre & Sanejouand, 2004
Modeller	Protein 3D modelling	Webb & Sali, 2016
AlphaFold2	Protein 3D modelling	Jumper et al., 2021
Pymol	3D models visualization and animation	Schrödinger, LLC, 2015
Xtal Focus Acquire	Image acquisition from plate imager implemented on Axio Zeiss microscope	Explora Nova, France
Cell^{D}	Image acquisition on	Olympus, Germany
ImageJ	Microscopy images processing (TIFF, GIF, JPEG...)	Open source, NIH (USA)
Mendeley	Bibliography	Elsevier
Youtube Music	Overnight synchrotron shifts	Google Inc.

Table II.6: List of regularly used softwares in the frame of this work

11. Electron microscopy

Beside from conventional single particle analysis under cryogenic conditions (cryo-

EM), transmission electron microscopes (TEM) offer the possibility to observe nano/microcrystalline material issued, for instance, from precipitate-like of macromolecule suspensions. Furthermore, intrinsic nature of electrons allows them to interact with matter, more precisely with nuclei from atoms. Indeed, few studies from mid 1990's (Henderson, 1995; Kühlbrandt *et al.*, 1994; Shi *et al.*, 1998) proved that very thin nanocrystals of macromolecules could scatter electrons (as X-rays do by interacting with electrons) and thus give birth to a diffraction phenomenon. Until now, such experiments were hardly feasible due to material limitations. Very recently ThermoFisher company used the large spreading and developments of cryo-EM technology to implement new micro electron diffraction (micro-ED) systems on their own TEM. Thus, with the support from FRISBI (French Infrastructure for Integrated Structural Biology), we could get access to a Glacios™ cryo-TEM (ThermoFisher) from Instruct-ERIC infrastructure at IGBMC (Illkirch, France) to analyze nanocrystals of AncCCA enzyme.

Cryo-EM grids were cleaned with a plasma cleaner (Dubochet, 1982) for 2 min 30 s, at 35% intensity, 20% oxygen and 80% argon. Then, 3 µL of sample consisting of fresh (<15 h) AncCCA at 5 $mg.mL^{-1}$ mixed with PEG 3.350 (from 0.5 to 6 % w/v) and either 100 mM potassium fluoride or 100 mM sodium malonate, having a precipitate-like aspect, was deposited onto the grids inside the vitrobot apparatus (ThermoFisher, USA) . Finally, grids were blotted for 3.0 s, at force 5.0 and directly plunged into liquid nitrogen for storing.

III. Purification and crystallization of CCA-adding enzymes

1. Purification of PhaCCA, RcuCC and AncCCA

The three CCA-adding enzymes were purified as described in section II.3. For the three candidates, after HisTrap™ purification step, we notice a mid to high contamination of pooled fractions with nucleic acid (OD_{260nm}/OD_{280nm} ratio about 0.8 up to 1.2). Moreover, first size exclusion trials showed a clear double-peak profile in the same eluted fractions. Analysis of these fractions on 12% SDS-PAGE and 15% TBE-urea gel indicated that the first peak was actually composed of a mix of unspecific nucleic acids bound to the target enzyme whereas the second corresponded to pure protein. As described for a variety of RNA polymerases (Toh *et al.*, 2011) we used a heparin matrix based column to eliminate nucleic acid contaminants. Thus, one the one hand, we improved OD_{260nm}/OD_{280nm} ratio from above 1.0 down to 0.58 and, on the other hand, we recovered a higher quantity of protein. As an example, chromatograms from each AncCCA purification steps are shown in Figure III.1.

Then, prior to any experiment, homogeneity and monodispersity of purified enzymes were checked by DLS with a DynaPro NanoStar (Wyatt Technology, USA). While AncCCA (Figure III.2.A) and PhaCCA were pure and stable when stored at 4 °C, RcuCCA, on the contrary, was quickly showing signs of aggregation (Figure III.2.B).

2. Preparative work on RcuCCA

2.1. Stability of RcuCCA

The CCA-adding enzyme of *R. culicivorax* was purified as described in section II.3 Less than three days after purification, DLS highlighted the strong propensity of the sample to aggregate as shown in Figure III.2.B. In addition, a second lower band on 12% SDS PAGE could appear directly after purification of the enzyme, indicating a possible degradation of RcuCCA enzyme. Further investigations by mass spectrometry did not reveal specific cleavage sites for *E. coli* peptidase as shown in Figure III.3. Thus, degradation of the sample could be due to oxidation and non appropriate storage buffer. In order to limit aggregation, we stored the protein without any concentration step after purification. This way, the stability of enzyme stocks could be improved from few days up to 5 weeks of storage at 4 °C.

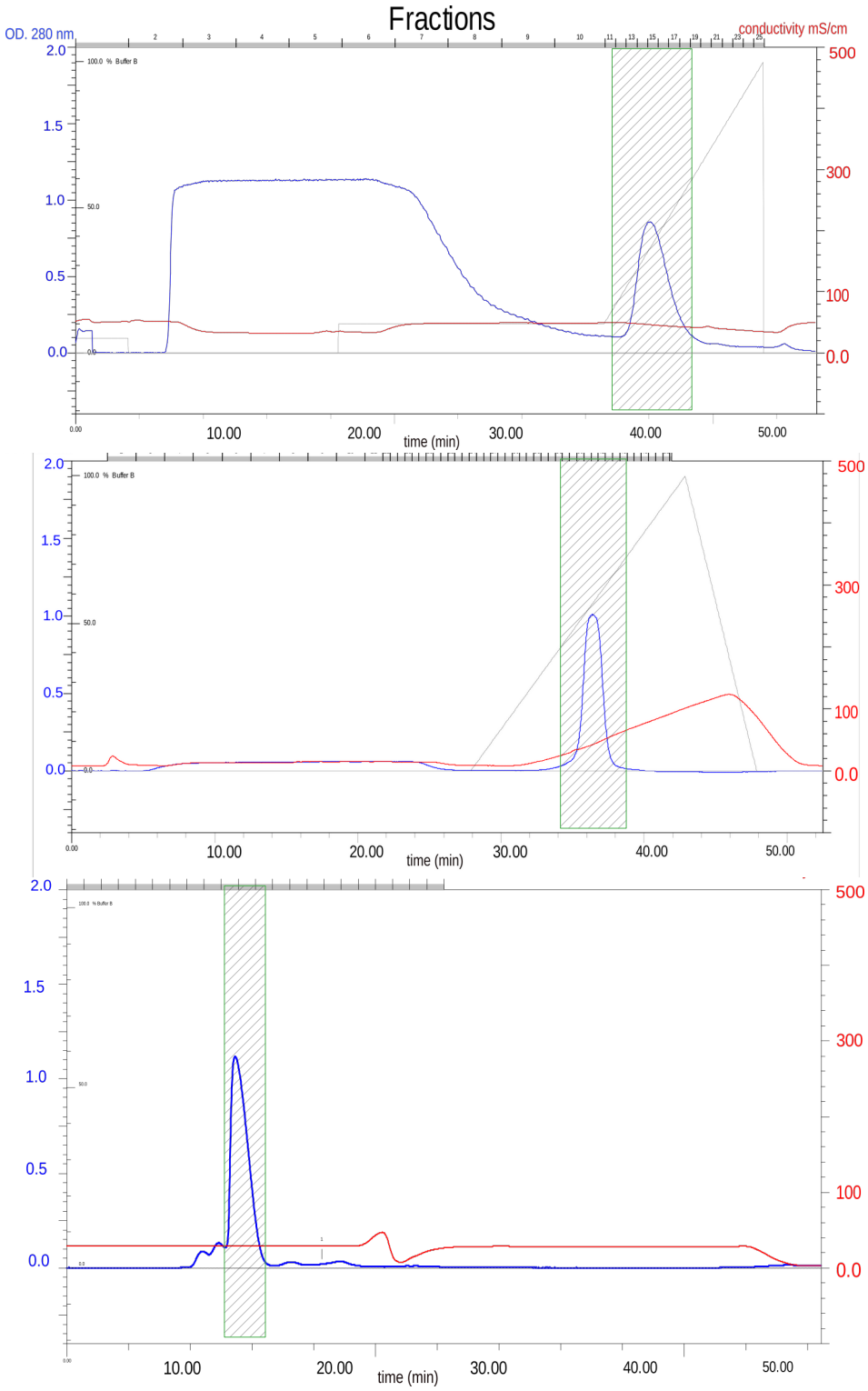


Figure III.1: Chromatograms from the different purification steps of AncCCA enzyme

A. Affinity purification (HisTrap) step. B. HiTrap heparin column. C. Size Exclusion Chromatography. Hatched rectangles corresponds to pooled fractions. The blue and red profiles corresponds to the absorbance at 280 nm and the conductivity (in mS/cm), respectively.

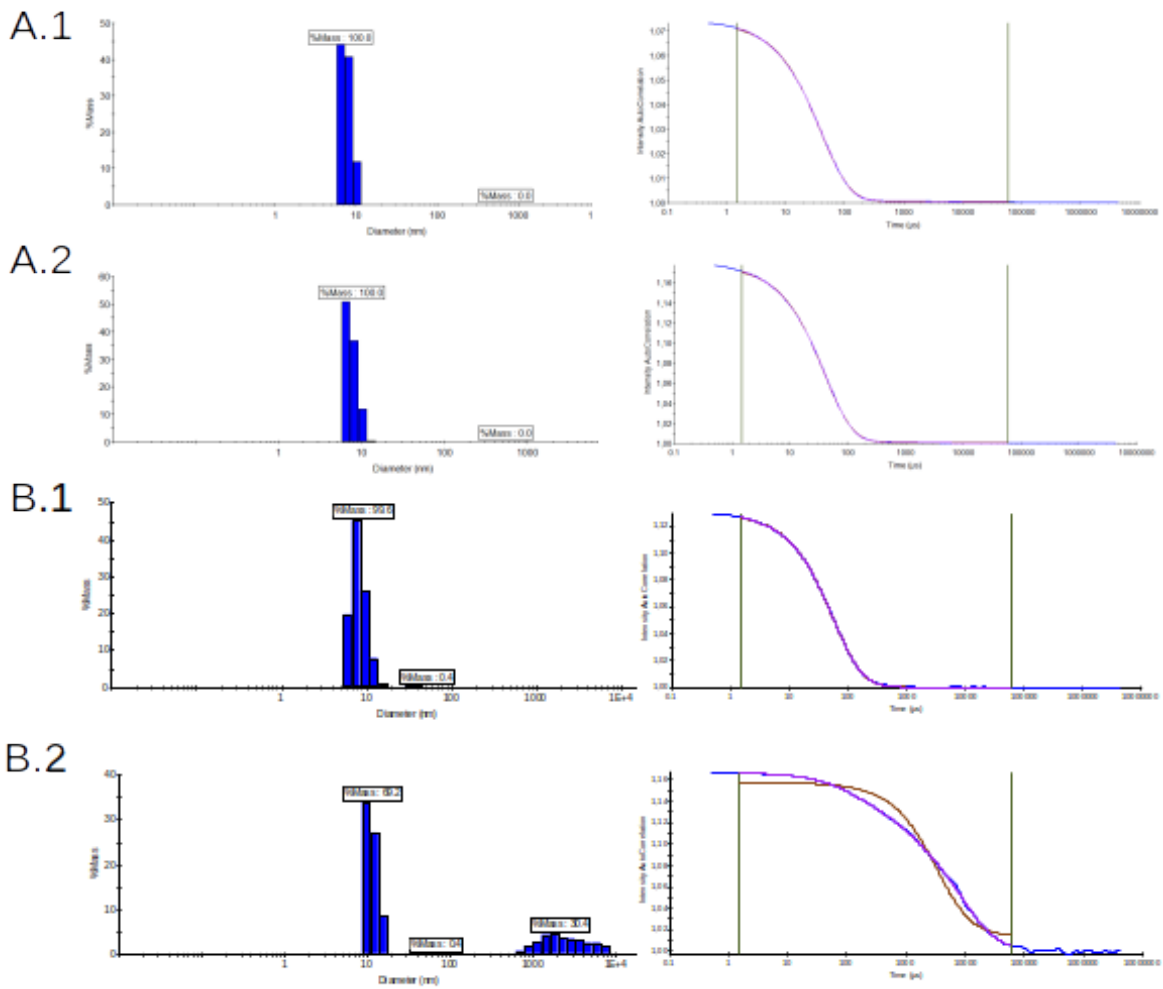


Figure III.2: DLS profiles of AncCCA and RcuCCA enzymes

The left graph plots the percentage of mass against particle diameter (in nm). The right graph corresponds to the autocorrelation function. A.1. AncCCA at 5 mg.mL-1 checked directly after ultra centrifugation (UC) at 100.000 xg for 1 hour. A.2. AncCCA stock checked after one week of storage at 4 °C. B.1. RcuCCA checked directly after UC at 100k G for 1 hour. B.2. RcuCCA stock checked after 2 days at 4 °C.

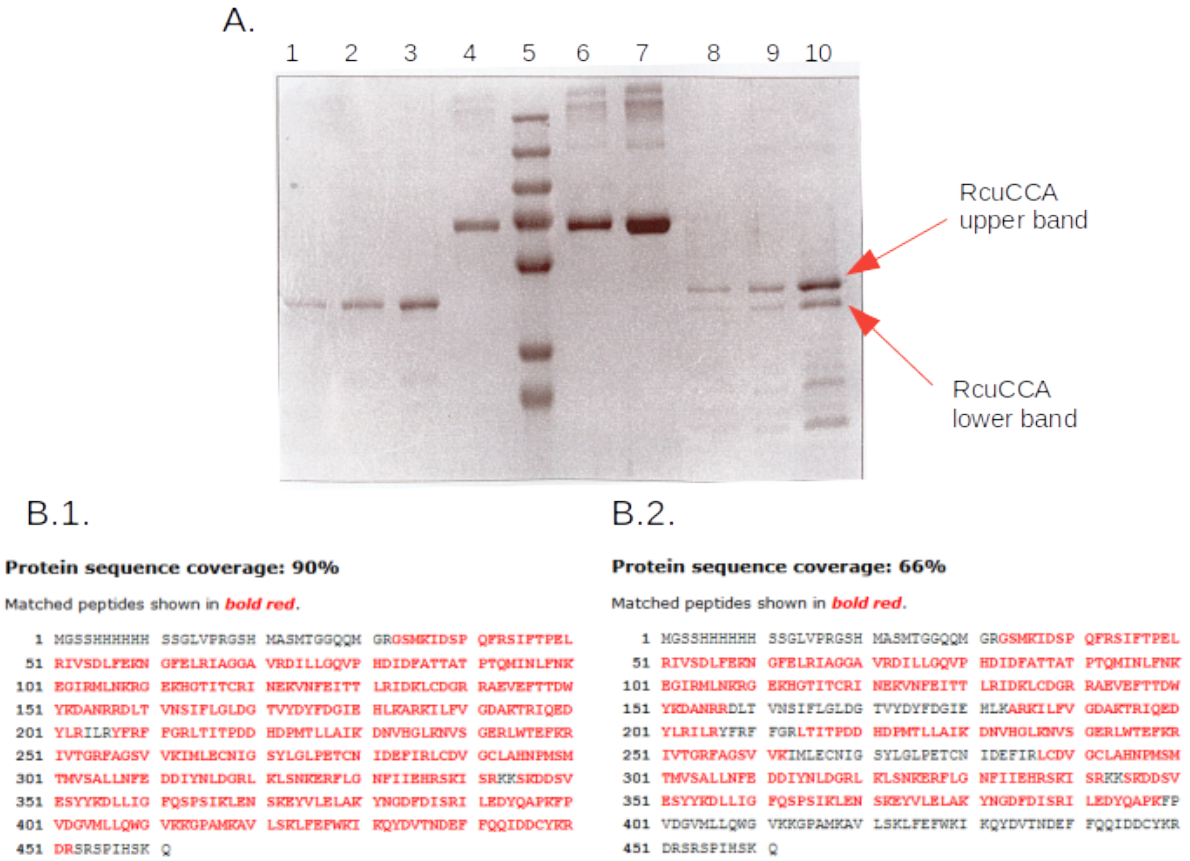


Figure III.3: RcuCCA degradation inspection by SDS-PAGE and mass spectrometry

A. SDS-PAGE 12% control on PhaCCA and RcuCCA stocks after four months at 4 °C. Lane 1: 0.2 µg PhaCCA stock. Lane 2: 0.4 µg PhaCCA stock. Lane 3: 0.8 µg PhaCCA stock. Lane 4: 2 µg control BSA (bovine serum albumin). Lane 5: Protein ladder. Lane 6: 4 µg control BSA. Lane 7: 4 µg control BSA. Lane 8: 0.2 µg RcuCCA stock. Lane 9: 0.4 µg RcuCCA stock. Lane 10: 0.8 µg RcuCCA stock. B. Mass spectrometry results for (B.1) RcuCCA “upper band” and (B.2) RcuCCA “lower band”.

Belatedly in the project, we could get access to nanoDSF equipment (Prometheus NT.48, Nanotemper) in order to screen a set of buffers with pH ranging from pH 4.0 up to pH 10.0. We used the light scattering data, corresponding to attenuation of the backreflected light intensity passing through the sample, to determine the onset temperature of aggregation (Kim *et al.*, 2021). At pH 7.0, corresponding to actual RcuCCA storage buffer, an onset temperature of 49.6 °C (Figure III.4.A) could be determined whereas pH 10.0 condition showed an onset at 65.3 °C (Figure III.4.B). These results clearly indicate a tendency for better stability at pH 10.0. Indeed, this behavior could be explained by a reduced proteolytic/gelatinolytic activity

far from neutral pH as already described in the literature (Sharma & Luthra-Guptasarma, 2009). Finally, pH 10.0 could be tested in new purification protocol of RcuCCA enzyme in order to improve stability and limit its degradation. However it seems important to keep in mind that such an artificial basic environment might be too far from *in vivo* conditions and so, could imply a different behavior of the enzyme than expected.

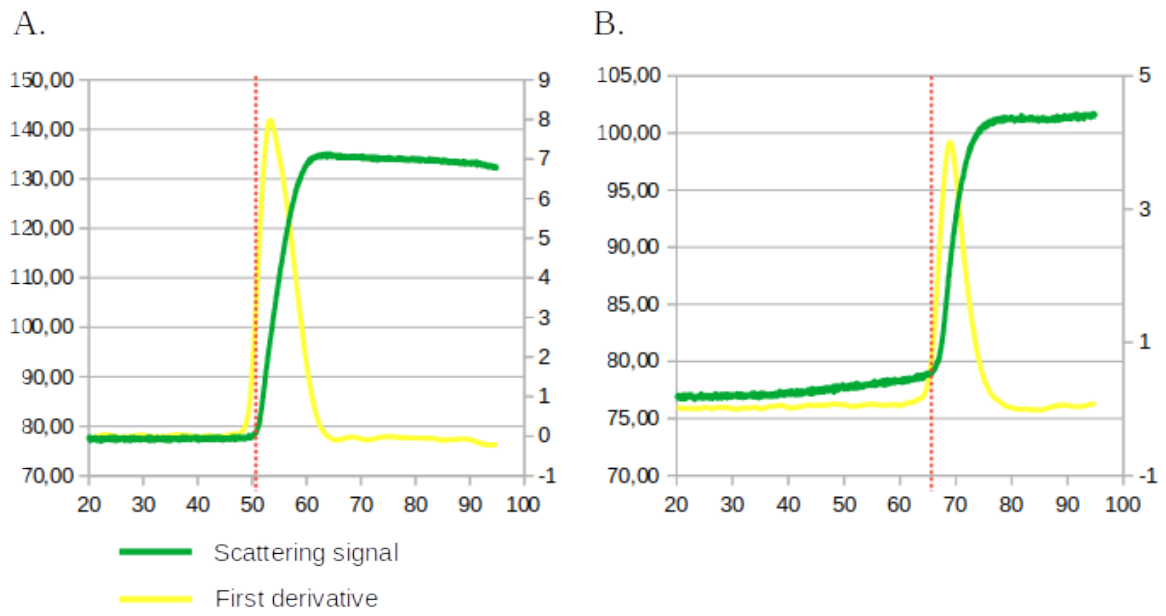


Figure III.4: Scattering signal of RcuCCA in nanoDSF buffer screening experiment

Here, the fluorescence (in arbitrary units) is plotted against the temperature (in °C). A. Scattering profile at pH 7.0 (pH of the original storage buffer) with an onset temperature of 49.6 °C (red dot line). B. Scattering profile at pH 10.0 with an onset temperature of 65.3 °C.

2.2. Crystallization of RcuCCA

For RcuCCA, a first series of crystallization attempts were performed by vapor diffusion with PEG/Ion HT, Crystal Screen HT, Natrix HT (Hampton Reasearch, USA) and JCSG++ HT (Jena Bioscience, Germany). I tried TFL labeling trials with carboxyrhodamine (Pusey et al., 2015) but it led to a mix irreversibly bound to the desalting column during the preparation process. Again, the non-adapted pH of protein and borate buffer used in labeling protocol, combined with concentration effect of the sample in the desalting column, may explain aggregation and the tendency to stick to the matrix. As an alternative, ultraviolet (UV) light was used to detect potential crystalline objects.

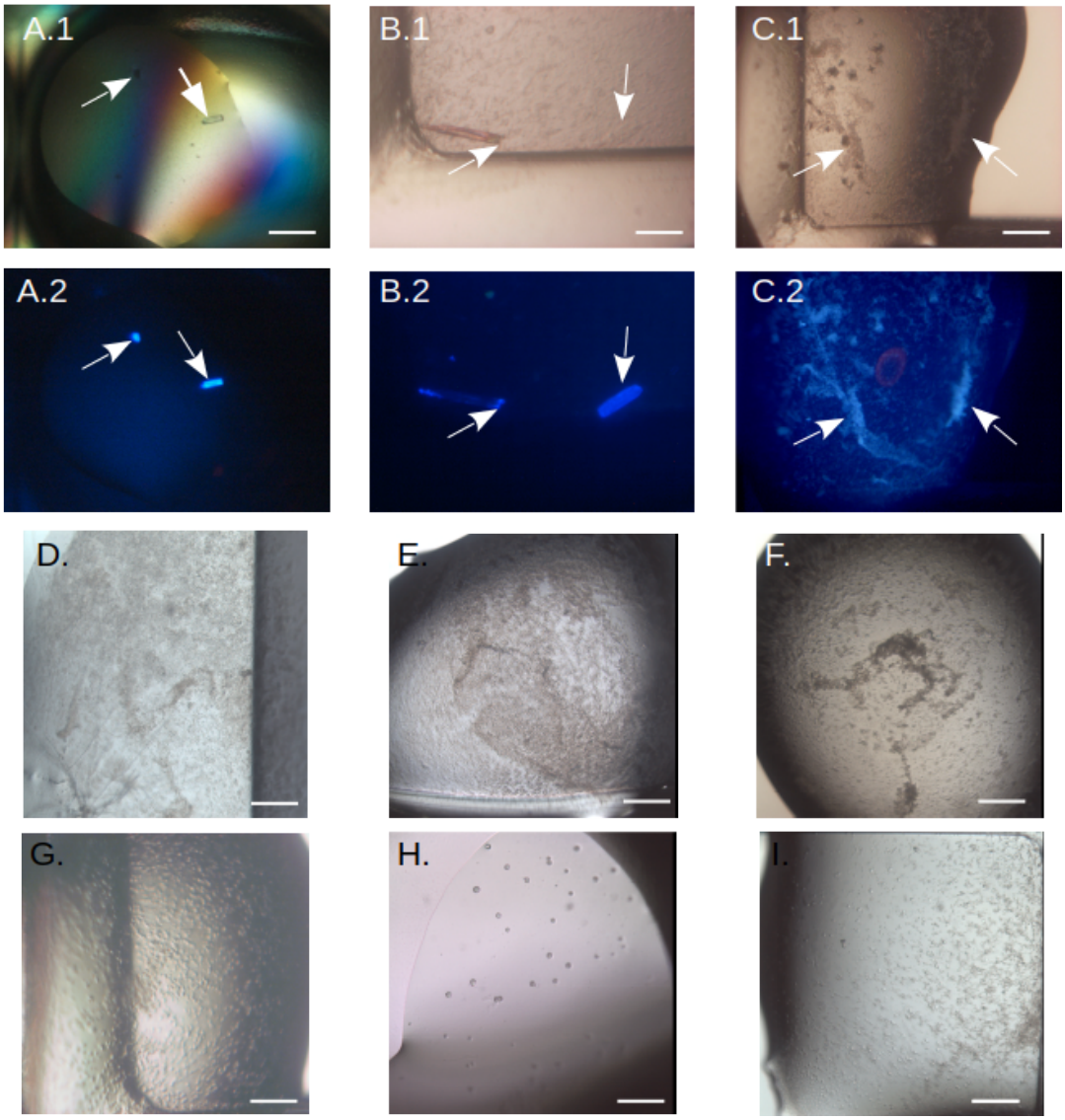


Figure III.5: Potential crystallization hits for RcuCCA

A.1 Index HT screen C4 (35% v/v tacsimate pH 7.0). A.2. Under UV illumination. B.1 JCSG++ H3.1 (25% w/v PEG 3.350, 100mM BIS-TRIS PH5.5). B.2. Under UV illumination. C.1. Index HT screen D12 (45% v/v MPD, 200 mM Calcium chloride dihydrate, 100 mM BIS-TRIS pH5.5). C.2. Under UV illumination .D. RcuCCA with Rcu-tRNAArg, 100X diluted AncCCA seeds, Index HT H12 (30% w/v mPEG 2.000 , 150 mM potassium bromide). E. 100X diluted AncCCA seeds, PEG/Ion HT E12 (20% w/v PEG 3.350, 8% v/v tacsimateTM pH5.0). F. RcuCCA with Rcu-tRNAArg, 100X diluted AncCCA seeds, Index HT F10 (25% w/v PEG 3.350, 200 mM sodium chloride, 100 mM BIS-TRIS pH5.5). G. PEG/Ion HT C1 (20% w/v PEG 3.350, 200 mM magnesium acetate tetrahydrate). H. RcuCCA with Rcu-tRNAIle, Natrix HT C5 (4.0 M Lithium chloride, 10 mM Magnesium chloride hexahydrate, 50 mM HEPES sodium pH7.0. I. Index HT B12 (2.8 M Sodium acetate trihydrate pH7.0). White scalebar corresponds to 150 µm. White arrows indicates possible crystals or microcrystalline material.

The enzyme was used in its apo form at 5.0 $mg.mL^{-1}$. Drops of 200 nL with protein:crystallant ratio of 1:1, 1:2 and 2:1 were deposited in crystallization plates. Experiments were regularly checked under light microscope but amorphous precipitate appeared in more than 95% of the conditions tested. Thus, considering the aggregation issue of the enzyme, we repeated the experiment with a concentration of RcuCCA lowered down to 3.5 $mg.mL^{-1}$. In order to increase our chances to obtain crystals, we applied the same optimization strategy as described for AncCCA (Figure III.6). Co-crystallization experiments with $tRNA^{Arg}$ and $tRNA^{Ile}$ from *R. culicivorax,* due to limited amount of material available, were conducted in three of the commercial screens (2 x 200 nL x 96 drops per screen) cited above, using a protein:tRNA ratio of 1:2.

Then few potential conditions could be identified as shown in Figure III.5. In the vast majority, crystallization drops had an amorphous precipitate aspect. However, when looking at the maximum magnification of the microscope (100X on Axio Zeiss microscope), we noticed that the precipitate was made out thousands small grains, especially in conditions containing PEG, or mPEG. Consistent with what has been observed for the AncCCA enzyme, it is very likely that nano and micro-crystals formed in identified conditions. Further refinement with fresh and optimized protein preparation should be eventually tried.

Access to *R. culicivorax* CCA-adding enzyme structure would constitute a key element in the understanding of miniaturized mitochondrial tRNA processing mechanisms. As exposed above, crystallization trials for the apo from or in complex with miniaturized tRNAs were not successful. However, my host team from Leipzig beautifully demonstrated (see chapter IV), that introduction of RcuCCA K74 residue at corresponding position in the β-turn of the human enzyme (HsaCCA) could confer the chimera a similar as wild type (wt) RcuCCA on these minimal substrates. The wt human CCA-adding enzyme has previously been crystallized in 100 mM sodium citrate (pH 5.6), 2.2 M ammonium sulfate, and its structure was also solved by my German team at a resolution of 3.4 Å (Augustin *et al.*, 2003). As a point mutation should not disturb crystal formation, our new strategy targets the crystallization of the chimera previously mentioned.

3. Crystallization of ancestral CCA-adding enzyme

This part of my thesis work combined all crystallization methods available in the laboratory to obtain and optimize crystals of the reconstituted ancestral CCA-adding enzyme of gammaproteobacteria group (AncCCA). The strategy used to improve diffraction quality of

crystals is summarized by the flowchart in Figure III.6. The particular behavior of the enzyme to crystallize by a “salting in” phenomenon made this candidate quite challenging to study.

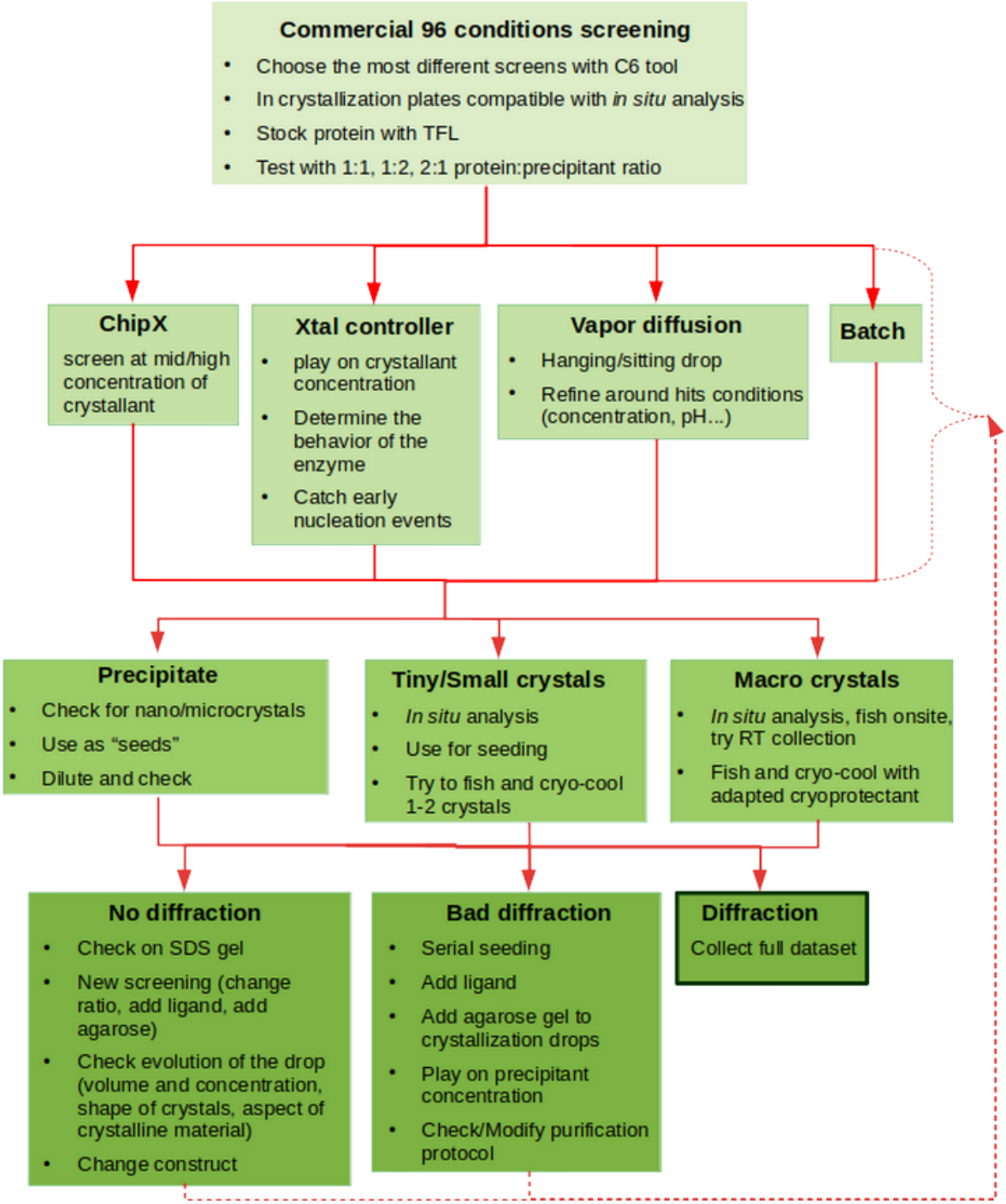


Figure III.6: Pipeline used to optimize crystal diffraction

Screen name	Position	Precipitant	Buffer
Crystal Screen (Hampton Research)	C12	8% w/v Polyethylene glycol 8,000	100 mM TRIS hydrochloride pH 8.5
	H2	20% w/v Polyethylene glycol 10,000	100 mM HEPES pH 7.5
PEG/Ion HT Screen (Hampton Research)	A2	20% w/v Polyethylene glycol 3,350	200 mM Potassium fluoride pH 7.3
	E5	12% w/v Polyethylene glycol 3,350	100 mM Sodium malonate pH 6.0
	G4	20% w/v Polyethylene glycol 3,350	200 mM Sodium formate pH 7.0
	G5	12% w/v Polyethylene glycol 3,350	100 mM Ammonium tartrate dibasic pH 7.0
	H3	20% w/v Polyethylene glycol 3,350	(40 mM Citric acid, 60 mM BIS-TRIS propane) / pH 6,4
	H4	20% w/v Polyethylene glycol 3,350	(30 mM Citric acid, 70 mM BIS-TRIS propane) / pH 7.6
	H9	15% w/v Polyethylene glycol 3,350	150 mM Cesium chloride
JCSG++ HTS (Jena)	E10	10% w/v PEG 6000	100 mM bicine pH 9.0

Tableau III.1: Initial crystallization hits for AncCCA in commercial screens

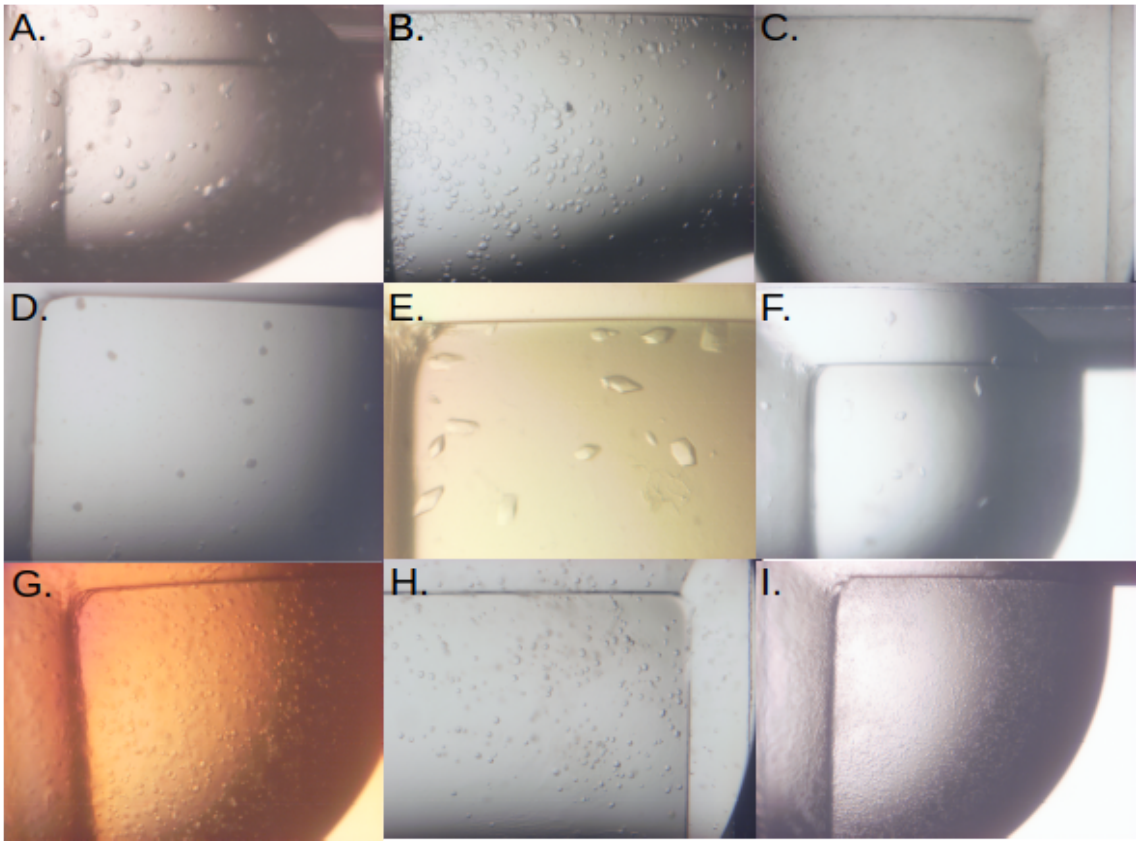


Figure III.7: Pictures of AncCCA initial crystallization hits obtained in commercial screens

A. PEG/Ion HT A2. B. PEG/Ion HT E5. C. PEG/Ion HT G4. D. PEG/Ion HT G5. E. PEG/Ion HT H3 F. PEG/Ion HT H4. G. PEG Ion HT H9. H. Crystal screen HT H2. I. JCSG ++ HT E10

Although no structure has been solved yet, new purification and crystallization approaches (discussed below) should considerably improve diffraction quality of crystals.

3.1. Identification of crystallization conditions

To determine initial crystallization conditions, the protein was first used used at 5 mg/mL in three different commercial screens: JCSG++ (Jena Bioscience, Germany), Crystal screen and PEG/Ion HT (Hampton Research, USA). Plates were stored at 20 °C and regularly checked. Crystals appeared within few days with an obvious tendency to form in the presence of polyethylene glycol as shown in Table 3-1.

Crystals obtained with commercial screens were fished with lithographic microloops (MiTeGen, USA), soaked in cryo-protecting solution consisting either of ½ mother liquor and ½ 25% glycerol, or paratone-N (Hampton Research, USA), and flash cooled in liquid nitrogen. Preliminary X-ray characterization experiments were done at Proxima 2A beamline (Saint-Aubin, France) and indicated that several hits corresponded to protein crystals yielding diffraction patterns at low resolution.

3.2. Crystallization in the Xtal Controller XC900

Considering diffraction results from initial screens, I selected three promising conditions from PEG/Ion HT screen (A2, E5, H3) to be tested in the Xtal Controller. Stock solutions of crystallant were all prepared at a concentration of 2.5% (w/v) PEG 3.350 for a viscosity compatible with piezoelectric pumps from the machine. Experiment was set as described in de Wijn *et al.* (2020), using an initial protein concentration of 5 $mg.mL^{-1}$. Briefly, a 10 µL drop of protein stock was deposited on a 22 mm siliconized glass coverslip and left in constant mode for at least 1 hour as a pre-incubation step. All along the experiment, DLS measurement were performed every 5 minutes to check for possible aggregation or nucleation events. Then, crystallant was progressively added for one hour until reaching twice the initial volume of the drop, which was incubated in constant mode while taking regular DLS measurements for one hour. The result is displayed as a “radius plot” which indicates for each measurement (as a function of time on the x-axis) the different particle populations represented as dots whose center and radius indicate their size (on the y-axis) and contribution

to the total scattering intensity, respectively. Finally, during one hour, controlled evaporation was used to come back to initial volume of the droplet.

As shown in Figure III.8.A.1, addition of 1 volume of crystallizing solution (here 2.5% (w/v) PEG 3.350, 100 mM sodium malonate pH 6.0) triggered appearance of bigger particles (from 100 to 1000 nm) into the drop, while the monomeric fraction (about 8 nm) was still present. After 1 hour at a constant regime, the drop was concentrated until reaching initial volume. After 14 hours, the coverslip was taken out of the machine and checked under microscope (Axio Observer, Zeiss, Germany) and light precipitate could be observed, as shown in Figure III.8.B. The coverslip was put back in the machine and, in an attempt to dissolve the precipitate, water was added to the drop until reaching a concentration of 1.5% (w/v) PEG 3.350, 60 mM sodium malonate pH 6.0. About 34 hours after crystallant addition, the drop was observed again under microscope and thousands of crystals ranging from 5 to 25 µm were observed. At this moment, the coverlip was transferred to standard Linbro plate and equilibrated against the same concentration of crystallant. The experiment was successfully repeated with solutions of 2.5% (w/v) PEG 3.350 buffered with either 200 mM potassium fluoride pH 7.3 (Figure III.8.C) or (30 mM citric acid, 70 mM BIS-TRIS propane) pH 7.6. In an other experiment, the same behavior regarding population size evolution on the radius plot was observed after the addition of 10 mM Tb-Xo4 (Figure III.8.A.2), but no crystals could be obtained at the end of the experiment.

We repeated several times these experiments and noticed that the addition of water was necessary for crystal growing. Moreover, when looking closer to microscope images from the Xtal controller, we could observe light precipitate at the water injection site, at very early stage of crystallant addition. In addition, radius plots indicated that bigger populations than the monomeric one could appear when the laser was close to water injection site. We then pushed this observation to an extreme point when, after addition of 1 volume of crystallant, we diluted the drop with water up to four times its original volume. This way, the precipitate was first disappearing from the drop and less than 20 hours after, crystals were appearing (Figure III.9). Such an observation was the first hint in favor of a “salting in” behavior of the protein instead of conventional “salting out” process. A series of crystals obtained with the different crystallant were fished, flash cooled and stored in liquid nitrogen. However, due to COVID-19 lockdown, crystals were lost after nitrogen evaporation from the storage deware.

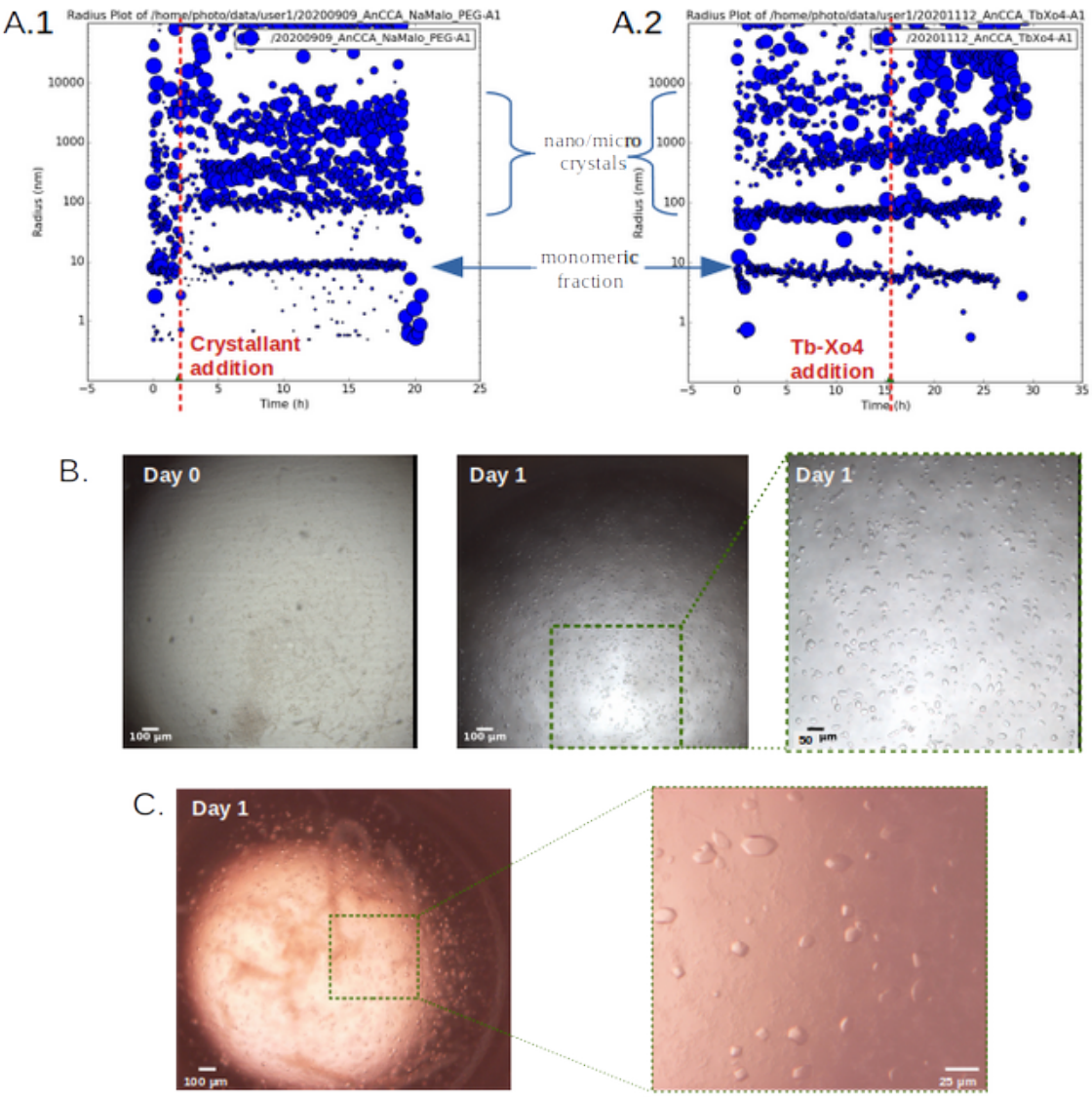


Figure III.8: AncCCA crystallization in the Xtal Controller

A.1. Radius plot showing the effect of progressive addition of 2.5% (w/v) PEG 3.350, 100 mM sodium malonate pH 6.0. Populations of bigger size (100 to 1000 nm) are appearing about 2 hours after injection. A.2. Radius plot with showing the effect of 10 mM Tb-Xo4 nucleant added after 15h incubation. B. Evolution of a drop after addition of crystallant to a final concentration of 1.5% (w/v) PEG 3.350, 60 mM sodium malonate pH 6.0. Day0 and Day1 corresponds to 14 and about 32 hours after crystallant addition respectively. C. Drop one days after addition of crystallant to a final concentration of 1% (w/v) PEG 3.350, 80 mM potassium fluoride pH 7.0.

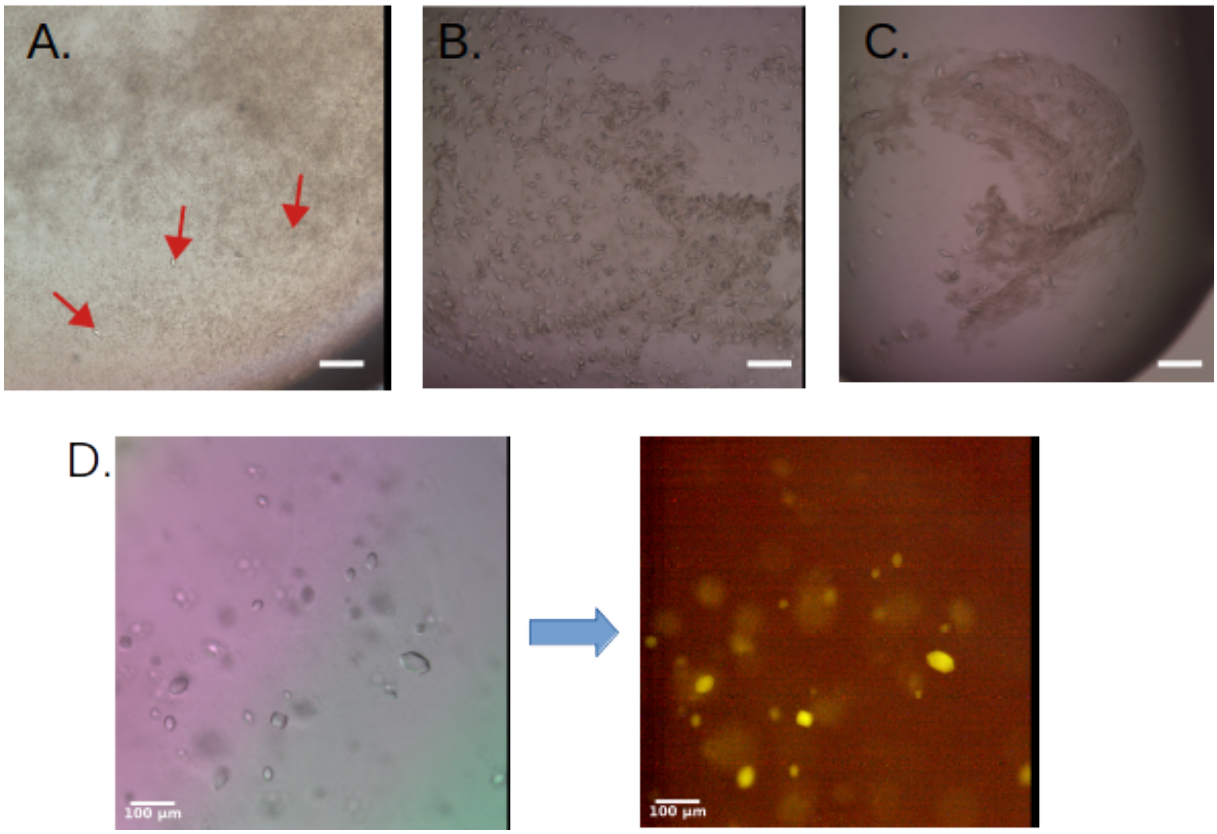


Figure III.9: AncCCA crystals emerging from precipitate

A. Drop with 1.0 % (w/v) PEG 3.350, 100 mM Na malonate pH 6.0, approximately 1.5 day after mixing. B. Drop with 3.0%(w/v) PEG 3.350, 200mM HEPES pH 6.5, 0.2% agarose, one week after mixing. C. Drop with 5.0% (w/v) PEG 3.350, 200mM HEPES pH 6.5, 0.2% agarose, one week after mixing. D. Drop with 1.2% (w/v) PEG 3.350, 50 mM potassium fluoride, 0.2% agarose, four days after mixing. The efficient trace fluorescent labelling with carboxyrhodamine serve to confirm the presence of AncCCA crystals. White scalebar corresponds to 100 µm.

3.3. Further refinement by vapor diffusion

While the Xtal Controller is a robust apparatus to navigate through the phase diagram of a macromolecule, it only functions with a single crystallization solution per experiment. In parallel, refinement around different PEG and salt concentrations as well as pH ranges were attempted with home-made 96 conditions screens. Sitting drop experiments were conducted in 96-well CrystalQuick™X crystallization plates (Greiner Bio One, USA) as they were suitable for direct *in situ* analysis at Proxima 2A beamline. Also, 24 wells EasyXtal plates (NeXtal Biotechnologies, Canada) were used for hanging drop trials. In total, six different 96 home-made screens with PEG 3.350, PEG 6.000 or PEG 10.000 concentrations ranging from 0.5 to 20 % (w/v) were tested. The buffers, used from 100 to 300 mM, consisted of potassium

fluoride pH7.3, sodium formate, sodium tartrate dibasic, bicine and pH ranging from pH 5 up to pH 8.5 were explored with sodium malonate or HEPES.

We first played on different ratios of protein versus crystallant and confirmed the tendency to form crystals at PEG concentrations below 3% (w/v). We observed the systematic formation of a precipitate at the mixing step with the crystallizing solution. Drops with very low PEG 3.350 concentrations (from 0.5 up to 3 %) were getting clear after 1.5 to 2 days and thousands of crystals were appearing after 2 to 4 days at 20 °C. On the contrary, drops with higher concentration of crystallant remained full of precipitate. Crystals from every working condition were manually fished (with sizes down to 10 µm), soaked in a series of different cryo protecting solutions, were flash cooled and tested at Proxima 1 or Proxima 2A beamlines. Despite several attempts, no relevant diffraction above 15 Å could be obtained.

3.4. In situ X-ray diffraction experiments at Proxima 2A

From 2021 on, we took benefit of availability of brand new Plate Screener called Cribleur (Jeangerard *et al.*, 2018) implement on Proxima 2A beamline (Figure I.12) to perform *in situ* X-ray diffraction characterization at room temperature and thus speeding up our research for ideal crystallization conditions. From then, we used agarose gel at 0.2 % (w/v) (Zhu *et al.*, 2001) (I) to limit nucleation events and thus push the system toward fewer and bigger crystals, and (II) to reduce convection in the mother liquor and so improve crystal growth and diffraction quality. In addition, we tested macro and microseeding at different dilutions and in series (up to four times). Cross-seeding was also attempted using crushed PhaCCA crystals with high diffraction quality properties (De Wijn *et al.*, 2018). We also used nucleating agents available from Polyvalan company (Lyon, France), namely Tb-Xo4, Tb-Xo4-2 (a more recent version of the nucleant), and Lu-Xo4 (an alternative form with a coordinated lutetium ion). Furthermore, 5 to 10 mM ATP, CTP or CMPcPP were added to crystallization drops or protein stocks in order to stabilize the flexible N-terminal part of the enzyme which might improve diffraction properties of crystals. Finally, we used *in situ* trace proteolysis (Wernimont & Edwards, 2009; Dong *et al.*, 2007) with trypsin and chemotrypsin to get rid of potential floppy parts of the enzyme which could disturb crystalline contacts and crystal packing.

In the end, more than thirty different plates containing crystals were screened at the

beamline. Unfortunately, the barrier of 6 Å could not be passed but few promising results are shown in Figure III.10. Low resolution data could be automatically indexed onsite with XDS package (Kabsch, 2010). Interestingly, indexing highlighted much larger unit cell parameters than expected. For instance, crystal shown in Figure III.10.B produced a dataset indexed in P422 space group with unit cell parameters of a=b=397.1 Å, c=129.1 Å and α=β=γ=90°. In comparison, for the same dataset, DIALS indexing tool (Diamond Light Source, UK) could also index in P4 space group with a=b=409.7 Å, c=131.5 Å and α=β=γ=90°. Another possibility obtained from the crystal in Figure III.10.A would be a=119.3 Å, b=198.1 Å, c=226.2 Å and α=115.4°, β=90.8°, γ=91.1°, determined by XDS package in space group P1.

It is important to notice the high fragility of AncCCA crystals under low-dose X-ray exposure. Indeed, after a "one shot" test at the lowest energy and exposure time required to detect a signal (0.05 seconds at 10% of maximum beam energy), the signal would quickly fade and disappear. In addition, even such a low energy dose was often destroying AncCCA crystals by leaving a clear radiation damage trace or cutting the crystal in two parts. Finally, this radiation damage effect seemed to propagate all along the crystal as the vast majority of spots above 50 Å resolution would completely disappear when changing shooting positing on a same crystal. To pursue on the quest for *in situ* room temperature diffraction experiments, crystallization trials were also set in microfluidic ChipX using a set of different PEG 3.350 concentrations and PEG/buffer couples. However, no crystalline material could be obtained so far in this device.

3.5. Other cases similar to AncCCA

My strategy to improve the size and the diffraction properties of AncCCA crystals was guided by three major relevant facts found in the literature. To begin with, one of the closest homologs of the ancestral enzyme in the Protein Data Bank (PDB) is the CCA-adding enzyme from *Geobacillus stearothermophilus* (GstCCA) (with 37.9% identity) which structure was solved in 2002 by the group of Thomas Steitz (Li *et al.*, 2002). The authors actually reported that "crystals (of GstCCA) appeared spontaneously when the protein at 8 mg/ml concentration was dialyzed in a buffer containing 50 mM TrisHCl (pH 7.5), 200 mM NaCl, and 10 mM $MgCl_2$" (sic), which is very close to AncCCA storage buffer (see Table II.2). Based on the same idea, we tested dialysis of AncCCA against its own storage buffer without glycerol and at decreasing salt concentrations in 5 µl dialysis button (Hampton Research, USA). However, as the equilibration rate was way faster than in vapor diffusion experiments, amorphous

precipitate appeared almost instantly and never evolved toward crystals. In our case, the few amount of PEG brought to the medium may probably act as a trigger for crystal nucleation. Nevertheless, it seems reasonable to consider the storage buffer of our candidate as a potential crystallization condition if well optimized.

As previously mentioned, we observed that AncCCA was crystallizing by vapor diffusion when drop volume was increasing after addition of very low quantity of PEG 3.350. In addition, we frequently obtained a microcrystalline state, i.e., thousands of objects ranging from 1 up to 10 µm in a single crystallization drop. Such samples are exactly the kind required for time-resolved X-ray crystallography experiments (Kupitz *et al.*, 2017; Kupitz, Grotjohann *et al.*, 2014; Boutet *et al.*, 2018). A similar example stands in the photosystem II from *Thermosynechococcus elongatus* which is first purified by several recrystallization steps (Kern *et al.*, 2005) as its solubility is inversely proportional to its purity. Then, in the presence of PEG 2.000, the pure enzyme is yielding millions of nanocrystals (Kupitz, Grotjohann *et al.*, 2014) which were succesfully used for time-resolved experiment at the Linac Coherent Light Source (LCLS) (Kupitz, Basu *et al.*, 2014). Hence, we notice that few drops, such as in Figure III.10.A, would first produce a mix of hundreds to thousands microccrystals in few days and which were not diffracting. Then, when traveling back and forth to Soleil synchrotron (Saint-Aubin, France), we had no real control on temperature and shaking effects on crystallization plates. Finally, three to four weeks after traveling, a single “tri-crystal” was formed into the drop. Next X-ray diffraction trip/test revealed spot up to about 15 to 12 Å.

Last but not least, several recent studies highlighted the power of ancestral enzymes to constitute a robust basis to improve crystallization chances and/or quality of certain candidates. As a matter of fact, the ancestral reconstruction technique usually leads to more thermostable macromolecules (Furukawa *et al.*, 2020), which are usually easier to purify, and can better undergo temperature variations during crystallization experiments (Koclega *et al.*, 2009). Furthermore, structural results from ancestral enzymes would put under spotlight the effects of the functionality of essential insertions/deletions occurring during evolution (Ross *et al.*, 2022) and giving birth to new classes of enzymes. In our case, AncCCA should provide information on structural elements that appeared or disappeared after evolutionary split from poly(A) polymerases in Gammaproteobacteria group, which presents a high sequence conservation between contemporary organisms.

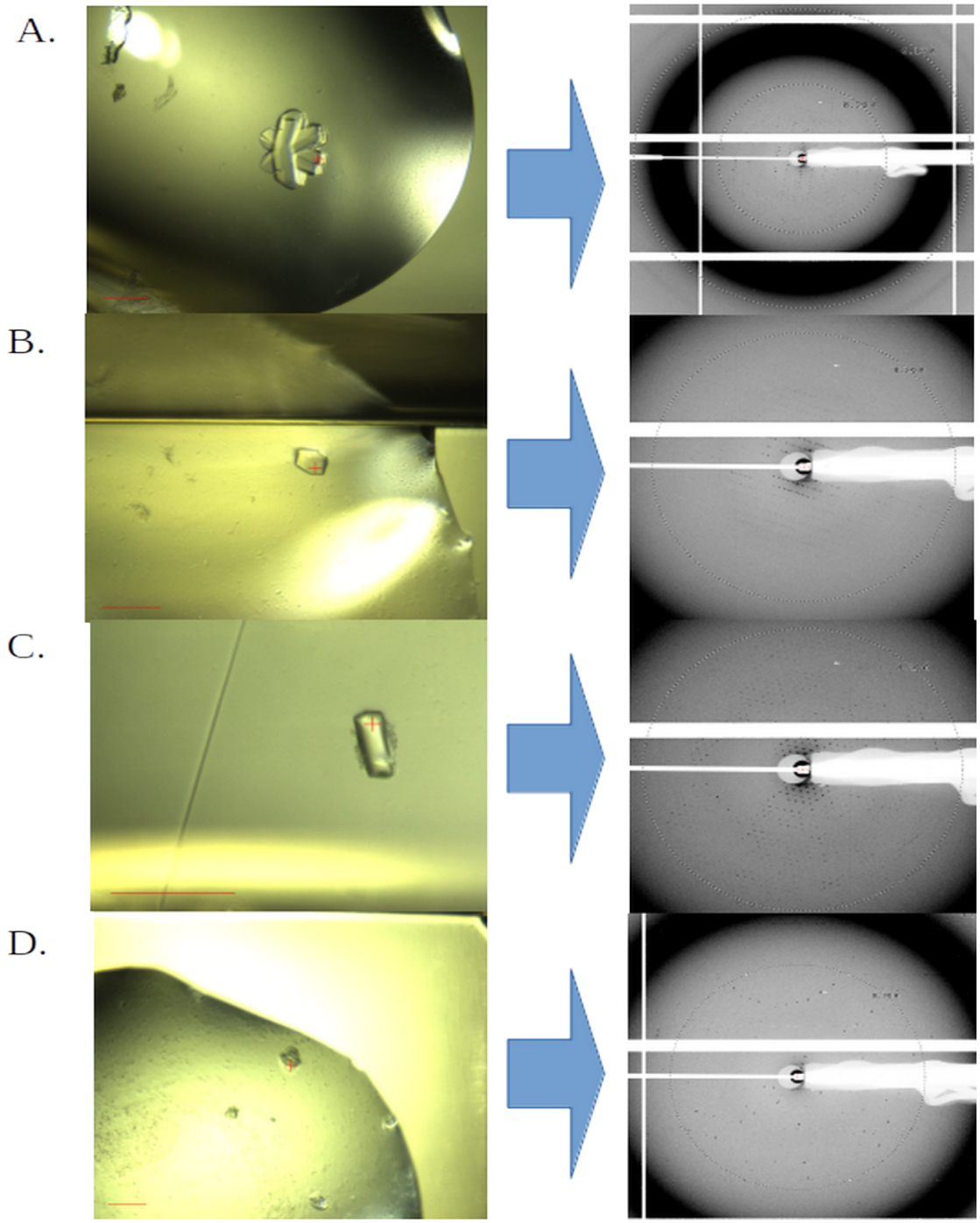


Figure III.10: *In situ* diffraction at room temperature with Plate Screener at PX2A

A. Recrystallized enzyme in 2% (w/v) PEG 3.350, 400 mM sodium formate. B. 10% (w/v) PEG 3.350, 100 mM sodium malonate pH 6.0. C. 1.5 % (w/v) PEG 3.350, 300 mM sodium formate. D. 10 % (w/v) PEG 3.350, 200 mM HEPES pH7.5. Few spots up to 6.7 Å. Red scalebar corresponds to 100 µm

Taken together, these features observed in different cases seem to be combined in the single case of AncCCA, making this particular target a quite unique and challenging candidate to study.

3.6. "Salting in" behavior of AncCCA

In very rare cases, macromolecules might crystallize by a phenomenon called "salting in". Typically, in a vapor diffusion experiment this is due to an inverse migration of water from the reservoir to the drop, thus resulting in the dilution of the latter (Giegé & McPherson, 2006). Throughout equilibration process, ionic force of the crystallization medium is diminishing upon drop dilution and macromolecules are pushed to form clusters until nucleation takes place, quickly followed by a growth phase. The literature reported few "inverse vapor diffusion" cases such as for the yeast transfer RNA^{Asp} (Giege *et al.*, 1977), Malate dehydrogenase from *Haloarcula marismortui* (Richard *et al.*, 2000) and the T7 RNA polymerase (Jeruzalmi & Steitz, 1997). In the latter situation, authors reported the exact same behavior as observed for AncCCA. Indeed, crystallization drops containing glycerol and equilibrated against deionized water triggered the formation of crystals in about two days. As such samples were not diffracting at a usable resolution, authors used a series of cosmotropic compounds (polyhydric alcohols, sugars, amino and methylamino acids) to stabilize the polymerase structure, thus leading to bigger and much better diffracting crystals. Inspired from this story, we are planing to use some of the cosmotropic compounds available in the laboratory to improve diffraction properties of AncCCA.

4. On-chip experiments at P14 T-REXX beamline

On December 2021, I had the opportunity to spend three weeks at Harbor Institute (DESY campus, Hamburg) in the team of Prof. Dr. Arwen Pearson. There, the main focus was put on the preparation of AncCCA microcrystalline suspensions suitable for analysis on P14/T-REXX beamline (size beam of 2 x 6 µm). As a matter a fact, the HARE silicon chip (Figure III.11) used for serial X-ray data collection is composed of 20.736 crystal wells with top size of 10 or 7 µm (Mehrabi *et al.*, 2020). The goal was to prepare, either by batch method or vapor diffusion, at least 100 µL of microcrystals suspension with a size ranging from 7 to

12 µm to load at least one chip.

4.1. Batch preparation

Batch experiments in 500 µL microtubes were based on results obtained from the Xtal controller as well as observations from vapor diffusion experiments to extrapolate final concentrations of crystallant and protein leading to AncCCA crystal formation. Thus, I selected the four following crystallizing solutions: 5 and 10% (w/v) PEG 3.350 buffered with either 200 mM of potassium fluoride pH 7.3 or 100 mM sodium malonate pH 6.0. In each tube, 10 µL of protein stock at 5.0 $mg.mL^{-1}$ were mixed with 5 µL of crystallizing solution. In order to maximize our chances, batch experiments were set at 4, 10 or 20 °C using peltier blocks, with or without AncCCA seeds (final dilution of 1/100). As expected, we observed precipitate appearing instantly in all tubes. After five days of incubation, as none of the precipitate would evolve into microcrystalline material, I mimicked “inverse vapor diffusion” by adding 5 µL of ddH_2O each two days and checked if mixes would evolve. In the end, two conditions (1% PEG 3.350, 20 mM sodium malonate, with and without seeds at 20 °C) led to suspensions with less than 3 µm crystals. However, resulting samples did not yield useful X-ray diffraction patterns.

4.2. Vapor diffusion preparation

In parallel, I used hanging drop in EasyXtal plates (NexTal, Canada) and sitting drop method in 4x6-wells XTalQuest crystallization plates (MiTeGen, USA) to increase drop volume compared to hanging drop experiments. Sitting drops of 5 µL of stock protein at 5 $mg.mL^{-1}$ mixed with 5 µL of reservoir solution consisting of the same crystallizing solutions used in batch experiments were prepared. In parallel, hanging drops had a final volume of 5 µL. Each line of the plate (6 wells) was dedicated to a single condition. All mixes were initially forming a precipitate. After about three days, hanging drops with initially 2.5% PEG (w/v) were already full of crystals ranging from 10 to 15 µm as shown in Figure III.12. In comparison, the different geometry and reduced air/liquid exchange surface of sitting drops seemed to slow the crystallization process. Thus, it took two to three additional days for sitting drops to reach the same aspect as hanging drops. In the end, each line was pooled individually to constitute aliquots of about 60 µL which were loaded on separated sections of the HARE chip.

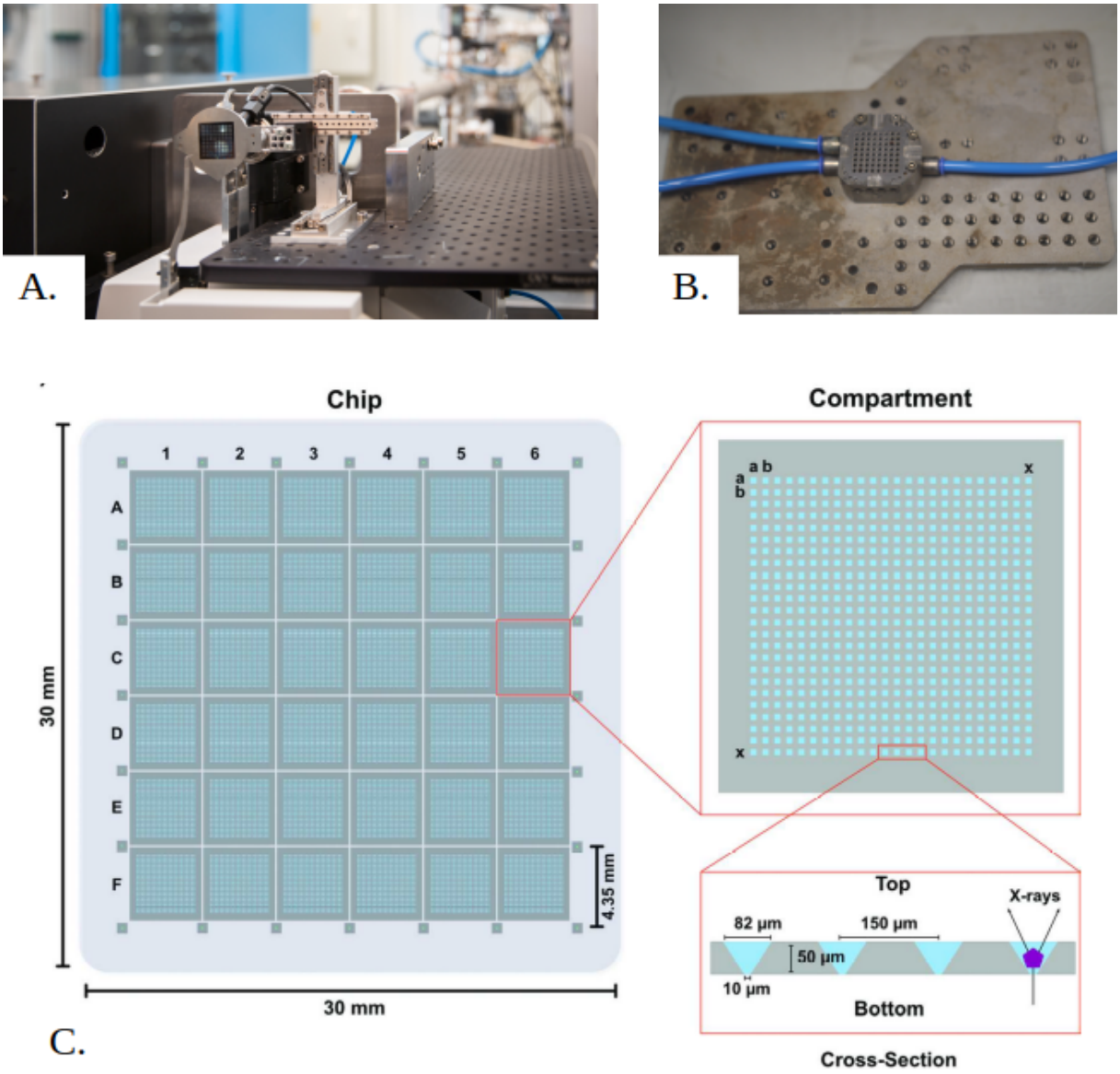


Figure III.11: 3-11: Design of the HARE chip used on P14/T-REXX beamline

A. HARE silicon chip mounted with adapted holder on P14 beamline (picture from https://www.embl-hamburg.de). B. Aluminum chip loading block in the humidity hood. The empty chip is place onto the device, then microcrystals solution is deposited in the compartments while vacuum is applied to sediment the crystals and remove the mother liquor. C. Schematic drawing of the chip (adapted from Mehrabi et al., 2020).

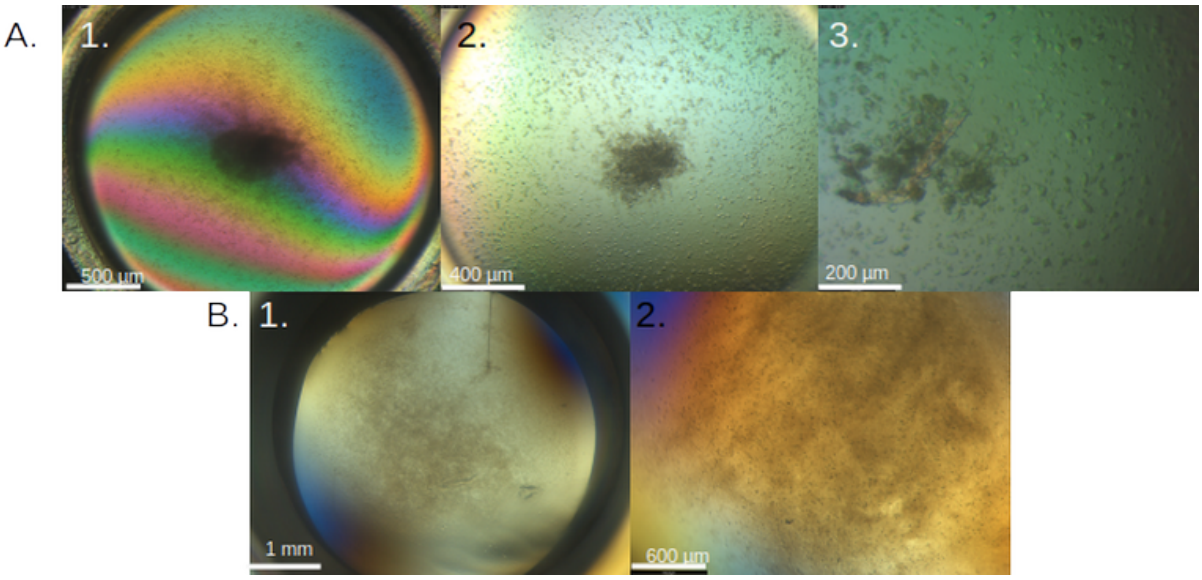


Figure III.12: Crystallization of AncCCA from precipitate in hanging and sitting drop

A. AncCCA crystallization in hanging drop with 5% (w/v) PEG 3.350, 100 mM sodium malonate pH6.0 at day 0 (A.1), day 3 (A.2) and day 5 (A.3). B. AncCCA crystallization in sitting drop with 5% (w/v) PEG 3.350, 200 mM potassium fluoride, at day 0 (B.1), at day 3 (B.2).

5. Preliminary cryo-TEM results

We exploited the potential of new electron microscope as described in section II.11 to explore the nature of early precipitates formed in AncCCA crystallization experiments. The aim of this experience was the detection and visualization of nano and microcrystals suspected to be present in the precipitate of AncCCA. Preparation of samples was conducted as described in the Material and Methods chapter. As we were the first users to conduct such experiments on microcrystalline material, we prepared a first series of EM-grids based on the single particle preparation protocol. The first observation was a “hydrophobic behavior” of samples with high (>10% w/v) PEG 3.350 concentration when deposited at the surface of the EM grid. We suspect that most of the sample was removed from the grid during the blotting as it was not completely plated on the surface.

The observation of samples consisting of low (0.5 to 2% w/v) PEG 3.350 concentration in the microscope did not reveal clear defined crystals. Yet, we could still distinguish geometrical particles, saying small objects with sharp edges, buried in dark clusters. Thus, we focused the electron beam on these potential crystals in order to check if a diffraction phenomenon could occur. As shown in Figure III.13, we observed a series of diffraction spots at the center of the detector. In addition, the distribution of the spots reveal a

crystal lattice linked to a macromolecule crystal. However, at the moment of the electron diffraction experiment, the new microscope was not calibrated for proper data collection and we could not estimate proper unit cell parameters. Further improvement of grid preparation and data collection protocols should provide more robust data on microcrystals characterization.

Still, these preliminary results consists an additional proof of the microcrystalline nature of the precipitate formed at early stage of a CCA-adding crystallization.

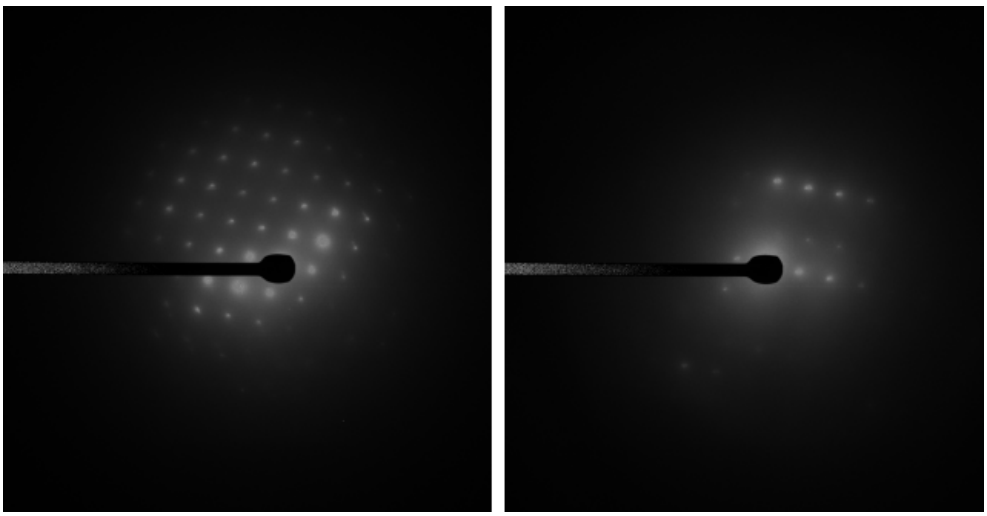


Figure III.13: Electron diffraction test on nanocrystalline material from AncCCA enzyme

The sample consists of a small geometric object obtained in batch-like experiment, using 1% (w/v) PEG 3.350, 100 mM KF and with AncCCA concentration set at 5.0 $mg.mL^{-1}$.

IV. Adaptation of the *Romanomermis culicivorax* CCA-Adding Enzyme to Miniaturized Armless tRNA Substrates

Protein synthesis is a central cellular process performed by the ribosome, which reads and decodes messenger RNAs to assemble polypeptide chains. In this process tRNAs act as decoders that recognize codons in the mRNA sequence and bring the corresponding amino acid bound to their universally conserved 3'-CCA tail. Hence, the activity of CCA-adding enzymes, i.e., the synthesis and maintenance of the CCA triplett, is an essential step in tRNA life cycle. In eukaryotes, it occurs in all compartments where tRNAs are required: cytoplasm, chloroplast or mitochondria. Interestingly, in metazoan mitochondrially encoded tRNAs frequently display shorter sequences than their cytosolic counterparts (Wolstenholme *et al.*, 1987; Okimoto & Wolstenholme, 1990; Watanabe *et al.*, 2014). This situation has reached an extreme case in the roundworm mosquito parasite *Romanomermis culicivorax* (Rcu) where these key molecules are still fully functional at a size down to 42 nucleotides (Rcu-tRNA^{Arg}) and thus, lacking both T- and D-arm (Wende *et al.*, 2014; Jühling *et al.*, 2018). While conventional tRNA:nucleotidyltransferase complexes show an interaction and anchoring on tRNA T-arm, our attention has been brought to the elements guiding this interaction in *R. culicivorax* CCA-adding enzyme (RcuCCA).

This chapter depicts the "divide to conquer" strategy employed to dissect the particular RcuCCA substrate specificity, using the human (HsaCCA) and *Escherichia coli* (EcoCCA) enzymes as controls. Our Rcu-tRNA candidates consisted of mitochondrial tRNA^{Arg} (42 nucleotides) and tRNA^{Ile} (47 nucleotides) compared to the canonical yeast tRNA^{Phe} (73 nucleotides). An initial experiment demonstrated that only RcuCCA was able to process miniaturized Rcu-tRNAs lacking both T- and D-arm and add the full 3'CCA trinucleotide. Then, knowing that CCA-adding enzymes carry conserved N-terminal motifs involved in tRNA anchoring and processing, we designed a set of chimeric enzymes (from A to E) composed of N-terminal fragments from RcuCCA transferred into the human enzyme. This work has been mainly conducted by Oliver Hennig and Susanne Philipp from my German host laboratory. I could take part in the purification of both enzymes and tRNAs, as well as affinity determination by electrophoretic mobility shift assay (EMSA).

Our results first showed that the integration of full RcuCCA N-terminal into HsaCCA makes the chimera capable of binding and fully processing the boomerang-like shape tRNAs from RcuCCA. Then, a progressive reduction of the chimeric part led us to an ultimate candidate (chimera E) where the simple substitution of the N-terminal ß-turn, located between motif A and the flexible loop, is sufficient to restore a wild type enzyme activity. Furthermore,

EMSA experiments revealed an equal affinity of chimera E and RcuCCA for small Rcu-$tRNA^{Ile}$. These observations motivated a deeper comparison of ß-turn composition between RcuCCA and HsaCCA, pointing the importance of two lysines (K74 and K89) which are essential for both mininal tRNA binding and processing. Finally, a computational 3D model of RcuCCA combined with our enzymatic assays highlighted the proximity of these two residues with tRNA binding site, thus suggesting a role in stable and correct orientation of 3' primer into the catalytic site.

Despite many attempts, no crystal structure of RcuCCA in its apo form of complex with tRNA could be obtained so far. However, based on the results described in the following article, a new strategy is envisaged using mutants of the human enzyme capable to process miniaturized tRNAs (see Perspectives).

La synthèse des protéines est un processus cellulaire central réalisé par le ribosome, qui lit et décode les ARN messagers pour assembler les chaînes polypeptidiques. Dans ce processus, les ARNt agissent comme des décodeurs qui reconnaissent les codons dans la séquence de l'ARNm et apportent l'acide aminé correspondant lié à leur queue 3'-CCA universellement conservée. Par conséquent, l'activité des enzymes d'addition du CCA, c'est-à-dire la synthèse et le maintien du 3'CCA terminal, est une étape essentielle du cycle de vie des ARNt. Chez les eucaryotes, elle se produit dans tous les compartiments où les ARNt sont nécessaires : cytoplasme, chloroplaste ou mitochondries. Il est intéressant de noter que chez les métazoaires, les ARNt codés par les mitochondries présentent fréquemment des séquences plus courtes que leurs homologues cytosoliques (Wolstenholme *et al.*, 1987 ; Okimoto & Wolstenholme, 1990 ; Watanabe *et al.*, 2014). Cette situation a atteint un cas extrême chez le parasite du moustique *Romanomermis culicivorax* (Rcu), un ver rond, où ces molécules clés sont encore pleinement fonctionnelles avec une taille de 42 nucléotides (Rcu-ARNt^{Arg}) et sont donc dépourvues de bras T et D (Wende *et al.* , 2014 ; Jühling *et al.* , 2018). Alors que les complexes classiques ARNt:nucléotidyltransférase montrent une interaction et un ancrage sur le bras T de l'ARNt, notre attention a été portée sur les éléments guidant cette interaction dans l'enzyme d'addition de l'ACC de *R. culicivorax* (RcuCCA).

Ce chapitre décrit la stratégie "diviser pour conquérir" employée pour disséquer la spécificité particulière du substrat RcuCCA, en utilisant les enzymes humaines (HsaCCA) et *Escherichia coli* (EcoCCA) comme contrôles. Les ARNt selectionnés du ver consistaient en l'ARNt^{Arg} (42 nucléotides) et d'ARNt^{Ile} (47 nucléotides) mitochondriaux mis en comparaison à l'ARNt^{Phe} canonique de la levure (73 nucléotides). Une première expérience a démontré que seule RcuCCA était capable de traiter les ARNt miniaturisés dépourvus de bras T et D et d'ajouter le trinucléotide 3'CCA complet. Ensuite, sachant que les enzymes d'addition de CCA portent des motifs N-terminaux conservés impliqués dans l'ancrage et le traitement des ARNt, nous avons conçu un ensemble d'enzymes chimériques (de A à E) composées de fragments N-terminaux de RcuCCA transférés dans l'enzyme humaine. Ce travail a été principalement mené par Oliver Hennig et Susanne Philipp de mon laboratoire d'accueil allemand. J'ai pu participer à la purification des enzymes et des ARNt, ainsi qu'à la détermination de l'affinité par retard sur gel (EMSA).

Nos résultats ont d'abord montré que l'intégration de la totalité du N-terminal de RcuCCA dans HsaCCA rend la chimère capable de lier et de traiter entièrement les ARNt en forme de boomerang du parasite. Ensuite, une réduction progressive de la partie chimérique nous a conduit à un candidat ultime (chimère E) où la simple substitution du coude ß en N-terminal, situé entre le motif A et la boucle flexible, est suffisante pour restaurer une activité enzymatique de type sauvage. En outre, les expériences EMSA ont révélé une affinité égale de la chimère E et de RcuCCA pour le petit $ARNt^{Ile}$. Ces observations ont motivé une comparaison plus approfondie de la composition du coude ß entre RcuCCA et HsaCCA, soulignant l'importance de deux lysines (K74 et K89) qui sont essentielles à la fois pour la liaison et la maturation de l'ARNt. Enfin, un modèle 3D computationnel de RcuCCA combiné à nos essais enzymatiques a mis en évidence la proximité de ces deux résidus avec le site de liaison de l'ARNt, suggérant ainsi un rôle dans l'orientation stable et correcte de l'amorce (primer) 3' dans le site catalytique.

Malgré de nombreuses tentatives, aucune structure cristallographique de RcuCCA dans sa forme apo ou en complexe avec unARNt n'a pu être obtenue jusqu'à présent. Cependant, sur la base des résultats décrits dans l'article suivant, une nouvelle stratégie est envisagée en utilisant des mutants de l'enzyme humaine capables de traiter des ARNt miniaturisés (voir Perspectives).

Die Proteinsynthese ist ein zentraler zellulärer Prozess, der vom Ribosom durchgeführt wird, das mRNAs liest und entschlüsselt, um Polypeptidketten zusammenzusetzen. Bei diesem Prozess fungieren tRNAs als Decoder, die Codons in der mRNA-Sequenz erkennen und die entsprechende Aminosäure an ihren universell konservierten 3'-CCA-Schwanz binden. Daher ist die Aktivität der CCA-addierenden Enzyme, d. h. die Synthese und Aufrechterhaltung dieses CCA-Tripletts, ein wesentlicher Schritt im tRNA-Lebenszyklus. In Eukaryonten findet er in allen Kompartimenten statt, in denen tRNAs benötigt werden: im Zytoplasma, in den Chloroplasten sowie in den Mitochondrien. Interessanterweise weisen mitochondrial kodierte tRNAs in Metazoen häufig kürzere Sequenzen auf als ihre zytosolischen Gegenstücke (Wolstenholme *et al.*, 1987; Okimoto & Wolstenholme, 1990; Watanabe *et al.*, 2014). Diese Situation hat einen Extremfall im Mückenparasiten *Romanomermis culicivorax* (Rcu) erreicht, wo diese Schlüsselmoleküle mit einer Größe von nur 42 Nukleotiden (Rcu-$tRNA^{Arg}$) noch voll funktionsfähig sind und somit sowohl der T- als auch der D-Arm fehlen (Wende *et al.* , 2014; Jühling *et al.* , 2018). Während herkömmliche tRNA:Nukleotidyltransferase-Komplexe eine Interaktion und Verankerung am T-Arm der tRNA zeigen, wurde unsere Aufmerksamkeit auf die Elemente gelenkt, die diese Interaktion im *R. culicivorax* CCA-addierenden Enzym (RcuCCA) steuern.

In diesem Kapitel wird die "divide-and-conquer"-Strategie beschrieben, die zur Entschlüsselung der besonderen RcuCCA-Substratspezifität angewandt wurde, wobei die Enzyme des Menschen (HsaCCA) und von *Escherichia coli* (EcoCCA) als Kontrolle dienten. Unsere tRNA-Kandidaten aus *R. culicivorax* bestanden aus mitochondrialer $tRNA^{Arg}$ (42 Nukleotide) und $tRNA^{Ile}$ (47 Nukleotide) im Vergleich zur kanonischen Hefe-$tRNA^{Phe}$ (73 Nukleotide). Ein erstes Experiment zeigte, dass nur RcuCCA in der Lage war, miniaturisierte Rcu-tRNAs, denen sowohl der T- als auch der D-Arm fehlte, zu verarbeiten und das vollständige 3'CCA-Trinukleotid hinzuzufügen. Da wir wissen, dass CCA-addierende Enzyme konservierte N-terminale Motive tragen, die an der tRNA-Verankerung und -Prozessierung beteiligt sind, haben wir eine Reihe von chimären Enzymen (von A bis E) entwickelt, in denen N-terminale Regionen von RcuCCA in das menschliche Enzym übertragen wurden. Diese Arbeit wurde hauptsächlich von Oliver Hennig und Susanne Philipp aus meinem deutschen Gastlabor durchgeführt. Ich konnte mich an der Reinigung der Enzyme und tRNAs sowie an der Affinitätsbestimmung mittels elektrophoretischem Mobilitätsverschiebungstest (EMSA) beteiligen.

Unsere Ergebnisse zeigten zunächst, dass die Integration des vollständigen N-Terminus von RcuCCA in HsaCCA die Chimäre in die Lage versetzt, die Bumerang-förmigen tRNAs von RcuCCA zu binden und vollständig zu verarbeiten. Eine schrittweise Verkleinerung des chimären Teils führte uns dann zu einem ultimativen Kandidaten (Chimäre E), bei dem die einfache Substitution der N-terminalen ß-Schleife, die sich zwischen Motiv A und einer flexiblen Region („flexible loop") befindet, ausreicht, um die Aktivität des Wildtyp-Enzyms wiederherzustellen. Darüber hinaus zeigten EMSA-Experimente eine gleiche Affinität von Chimäre E und RcuCCA für die armlose mt-tRNA^{Ile}. Diese Beobachtungen motivierten zu einem eingehenderen Vergleich der ß-Schleife-Zusammensetzung zwischen RcuCCA und HsaCCA, der die Bedeutung von zwei Lysinen (K74 und K89) aufzeigte, die sowohl für die Bindung als auch für die Verarbeitung der kleinen tRNA wesentlich sind. Ein 3D-Computermodell von RcuCCA in Kombination mit unseren enzymatischen Tests zeigte die Nähe dieser beiden Reste zur tRNA-Bindungsstelle, was auf eine Rolle bei der stabilen und korrekten Ausrichtung des 3'-Primers in der katalytischen Stelle hindeutet.

Trotz vieler Versuche konnte bisher keine Kristallstruktur von RcuCCA in seiner apo-Form des Komplexes mit tRNA erhalten werden. Auf der Grundlage der im folgenden Artikel beschriebenen Ergebnisse wird jedoch eine neue Strategie ins Auge gefasst, bei der Mutanten des menschlichen Enzyms verwendet werden, die in der Lage sind, miniaturisierte tRNAs zu verarbeiten (siehe Perspektiven).

International Journal of
Molecular Sciences

MDPI

Article

Adaptation of the *Romanomermis culicivorax* CCA-Adding Enzyme to Miniaturized Armless tRNA Substrates

Oliver Hennig [1,†], Susanne Philipp [1,†], Sonja Bonin [1], Kévin Rollet [1,2], Tim Kolberg [1], Tina Jühling [1,2], Heike Betat [1], Claude Sauter [2] and Mario Mörl [1,*]

1 Institute for Biochemistry, Leipzig University, Brüderstraße 34, 04103 Leipzig, Germany; oliver.hennig@uni-leipzig.de (O.H.); susanne.philipp@uni-leipzig.de (S.P.); sonja.bonin@uni-leipzig.de (S.B.); k.rollet@ibmc-cnrs.unistra.fr (K.R.); tim.kolberg@uni-leipzig.de (T.K.); tina.md@gmail.com (T.J.); heike.betat@uni-leipzig.de (H.B.)

2 Architecture et Réactivité de l'ARN, Université de Strasbourg, CNRS, IBMC, 67084 Strasbourg, France; c.sauter@ibmc-cnrs.unistra.fr

* Correspondence: mario.moerl@uni-leipzig.de; Tel.: +49-(0)-341-9736-911

† These authors contributed equally to this work.

Received: 29 October 2020; Accepted: 25 November 2020; Published: 28 November 2020

Abstract: The mitochondrial genome of the nematode *Romanomermis culicivorax* encodes for miniaturized hairpin-like tRNA molecules that lack D- as well as T-arms, strongly deviating from the consensus cloverleaf. The single tRNA nucleotidyltransferase of this organism is fully active on armless tRNAs, while the human counterpart is not able to add a complete CCA-end. Transplanting single regions of the *Romanomermis* enzyme into the human counterpart, we identified a beta-turn element of the catalytic core that—when inserted into the human enzyme—confers full CCA-adding activity on armless tRNAs. This region, originally identified to position the 3′-end of the tRNA primer in the catalytic core, dramatically increases the enzyme's substrate affinity. While conventional tRNA substrates bind to the enzyme by interactions with the T-arm, this is not possible in the case of armless tRNAs, and the strong contribution of the beta-turn compensates for an otherwise too weak interaction required for the addition of a complete CCA-terminus. This compensation demonstrates the remarkable evolutionary plasticity of the catalytic core elements of this enzyme to adapt to unconventional tRNA substrates.

Keywords: CCA-adding enzyme; co-evolution; evolutionary plasticity; minimalized armless tRNAs; tRNA nucleotidyltransferase

1. Introduction

tRNAs are the essential adaptor molecules which enable the decoding of the nucleic acid code into the amino acid sequence during the translational process [1]. To fulfill this function, they need to undergo several maturation steps and interact with the translational machinery [2–5]. In eukaryotes, this also includes the corresponding enzymes and proteins of mitochondria and chloroplasts [6,7]. For the efficient recognition by a wide range of processing enzymes, translation factors as well as ribosomes, tRNAs fold into a conserved cloverleaf-like secondary structure consisting of acceptor stem, anticodon arm as well as D- and T-arm that adopts an equally conserved three-dimensional L-shape [8–10]. The 3′-terminal CCA-triplet of the acceptor stem is a prerequisite for aminoacylation and the correct positioning of the charged tRNA in the ribosome [11,12]. In eukaryotes, this triplet is not genomically encoded but is added post-transcriptionally by tRNA nucleotidyltransferase (CCA-adding enzyme) [13–15].

tRNA nucleotidyltransferases represent essential enzymes and are ubiquitously found in all domains of life. Representing members of the polymerase β superfamily, they split up into two classes, based on the composition of their catalytic core [16]. Archaeal CCA-adding enzymes represent class I, while their bacterial and eukaryotic counterparts belong to class II [16]. The overall sequence identity among both tRNA nucleotidyltransferase classes is rather low, although they catalyze the same reaction [17]. The catalytic core motif common in both classes consists of two aspartate residues DxD (x, any amino acid) that coordinate the catalytically important metal ions [18–20]. In class II enzymes, the DxD sequence belongs to motif A, one of the five conserved motifs A to E located in the N-terminal part of this tRNA nucleotidyltransferase type (Figure 1) [21]. Motif A binds two Mg^{2+} ions required for nucleotide transfer onto the tRNA substrate via the general two metal ion mechanism of polymerases [19]. Motif B is involved in ribose binding [21], while motif C is a flexible element which coordinates interdomain movements, contributing to the proper orientation of the substrates within the active center [22,23]. Motif D represents an amino acid-based template, where arginine and aspartate residues form Watson–Crick-like hydrogen bonds with the incoming nucleotides, and their orientation in the catalytic core determines the specificity for CTP or ATP, respectively [21]. Lastly, motif E stabilizes a helix-turn structure in motif D and is discussed to interact with the tRNA primer strand in the catalytic core [21,24].

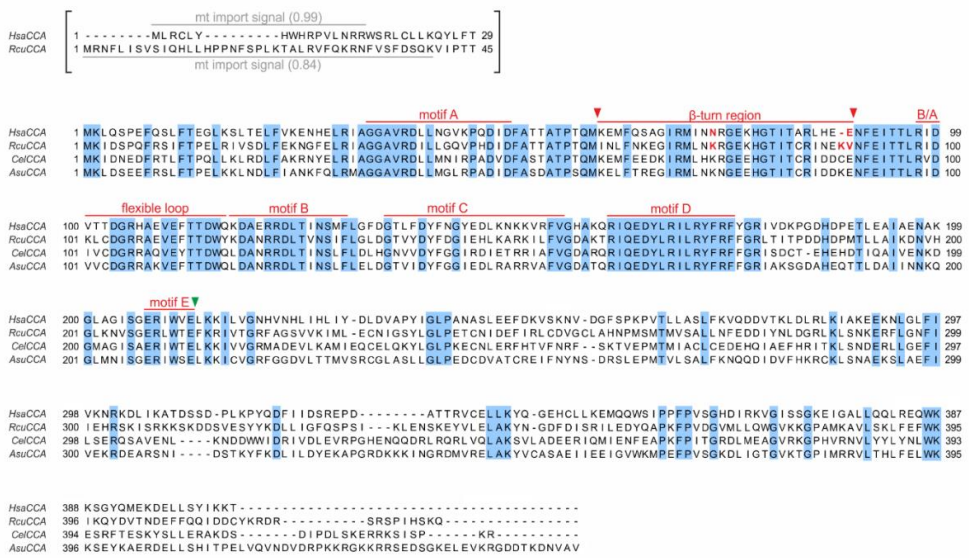


Figure 1. Sequence alignment of CCA-adding enzymes from *Homo sapiens*, *Romanomermis culicivorax*, *Caenorhabditis elegans*, and *Ascaris suum*. Light blue positions indicate identical residues. The predicted mitochondrial import signals for *H. sapiens* CCA (*Hsa*CCA) and the *R. culicivorax* CCA-adding enzyme (*Rcu*CCA) (grey bars, import probability is given in brackets) are shown in brackets and were excluded from the cloned open reading frames. Catalytically important elements are labeled in red. Fusion position of reciprocal chimeras A and B are indicated by a green arrowhead. Fusion positions of chimera E (β-turn element) are indicated by red arrowheads (K/I61–E/V90). Mutations K74N and K89Δ/V90E introduced in *Rcu*CCA and N74K and K89ins/E90V in *Hsa*CCA are indicated in red.

Another catalytically important region is a flexible loop consisting of 10–20 amino acids that is located immediately upstream of motif B. While it is not conserved at the sequence level [25], its interaction with the amino acid template of motif D is required for the specificity switch from C to A incorporation, where it acts as a lever to accommodate the ATP in the nucleotide binding pocket [22,26]. Between motif A and the flexible loop, a β-turn element was identified that is also involved in A-addition, as it binds and positions the priming 3′-end of the growing CCA terminus in the catalytic core [22]. Similar β-turn regions are present in many different polymerases, ranging from both classes of tRNA nucleotidyltransferases to poly(A) polymerases and DNA polymerases, underscoring

the catalytically important function of this element [27–30]. For a structural view of the reaction mechanism, we refer to Tomita and Yamashita, 2014 [31].

All these conserved motifs build up the active site in the N-terminal part of the enzyme. The C-terminus, in contrast, is much less conserved. Yet, it is of functional importance, as it is involved in tRNA binding, where it anchors the TΨC loop of the L-shaped tRNA substrate during nucleotide addition [24,32–35]. Correspondingly, artificial CCA-adding substrates like mini- or microhelices are recognized and accepted for CCA incorporation at a much lower efficiency [36,37]. Yet, metazoan mitochondria carry tRNA molecules that deviate from the cloverleaf structure, lacking either the D- or the T-arm [38–40]. As an example, the mammalian mt-tRNA^{Ser}(AGY) lacks the complete D-arm, so that D- and T-arm interactions do not exist [41,42]. In the mitochondrial genomes of nematodes, acari and arachnids, this situation comes to an extreme. These genomes are rich in genes for tRNAs that lack either the D- or the T-arm or even both [38,43–46]. In the mermithid *Romanomermis culicivorax*, mt-tRNA molecules with the most dramatic truncations were identified, resulting in miniaturized hairpin-like tRNAs with a length of down to 45 nts, in contrast to the standard average tRNA size of 76 nts [47,48]. Such extremely truncated tRNAs fold into a three-dimensional boomerang-like shape that deviates from the consensus L-form [47]. Yet, these organisms encode for a single CCA-adding enzyme that has to act on both cytosolic as well as mitochondrial tRNA pools [49–53], and it was shown for the corresponding enzyme of *Caenorhabditis elegans* that it recognizes mt-tRNAs lacking D- or T-arm [52]. Since in *R. culicivorax* nine mt-tRNAs lack both arms, representing the strongest deviation from the consensus structure [45], we investigated the co-evolution and substrate adaptation of its CCA-adding enzyme. In a comparative analysis, we identified the β-turn element as a major adaptation to the hairpin-like tRNA substrates. This adaptation is based on an increased substrate affinity of the enzyme. Hence, while the conventional substrate binding based on interactions between the enzyme's C-terminus and the TΨC loop/T-arm is not possible with such tRNA hairpins, a different part of the enzyme took over this function to assure a sufficiently strong substrate binding for CCA-addition, demonstrating a surprising evolutionary plasticity of these enzymes.

2. Results

2.1. RcuCCA Adds a Complete CCA-Triplet to Armless and Canonical tRNAs

To investigate the catalytic activity and substrate specificity of the *R. culicivorax* CCA-adding enzyme (*Rcu*CCA), we identified a corresponding singular open reading frame in the *R. culicivorax* genome assembly available at https://parasite.wormbase.org/Romanomermis_culicivorax_prjeb1358/Info/Index/ [54]. The sequence conservation was investigated in an alignment with the corresponding human enzyme (*Hsa*CCA). As it was reported that the CCA-adding enzyme of *Caenorhabditis elegans* is adapted to the bizarre mitochondrial tRNAs of this organism [52], the sequence of this enzyme as well as that of *Ascaris suum* (another nematode with mt-tRNAs lacking D- or T-arm [38,39,43]) were included (Figure 1). At the amino acid level, the overall sequence identity and similarity between *Rcu*CCA and *Hsa*CCA is 48% and 66%, respectively. Carrying the complete active site as well as a putative mitochondrial import sequence as predicted [55], the N-terminus shows a comparatively higher conservation (78% sequence similarity), as it is true for the corresponding enzymes of *C. elegans* and *A. suum* (Figure 1). In the less conserved C-terminal part, *Rcu*CCA and the other nematode enzymes carry a short insertion and a terminal extension.

A conserved methionine residue downstream of the mt import signal was chosen as the N-terminus of the expressed open reading frame (labeled as position 1 in Figure 1). In previous experiments on the human enzyme, this position was successfully used and the absence of the mt import signal had no effect on its catalytic activity [37,50,56,57]. The open reading frame was synthesized as a codon-optimized DNA sequence and recombinantly expressed in *Escherichia coli*. Together with the corresponding enzymes from *E. coli* (*Eco*CCA; an organism exclusively carrying conventional cloverleaf-like tRNAs) and *Homo sapiens* (*Hsa*CCA; an organism carrying conventional cytosolic

as well as moderately reduced mt-tRNAs), the purified enzyme was tested in vitro for activity. As substrates, three different radioactively labeled tRNA transcripts were generated by in vitro transcription (Figure 2A), as it is well established that tRNA nucleotidyltransferases from all kingdoms readily accept in vitro transcripts lacking base modifications [25,58–60]. tRNA^{Phe} from *Saccharomyces cerevisiae* is one of the best characterized tRNAs and represents a standard substrate for many tRNA-interacting enzymes [58,61,62], since the unmodified in vitro transcript folds into a structure almost identical to the native tRNA [63,64]. Furthermore, two armless mt-tRNAs from *R. culicivorax* were generated. With a length of 42 nts and 47 nts, respectively, mt-tRNA^{Arg} and mt-tRNA^{Ile} represent the shortest tRNAs identified so far, and the in vitro transcripts fold into hairpin-like structures with two single-stranded connector elements replacing D- and T-arm (Figure 2A) [47]. On the standard tRNA^{Phe}, all enzymes added a complete CCA-triplet, indicating highly active enzyme preparations (Figure 2B). On the armless mt-tRNA substrates, however, the bacterial enzyme *Eco*CCA was completely inactive and did not add any nucleotides. The human enzyme *Hsa*CCA that has to recognize the human D-arm-lacking mt-tRNA^{Ser}(AGY) (Figure S1) catalyzed a moderate incorporation of two residues on the armless *Rcu* mt-tRNA^{Ile}, but was almost inactive on the even smaller *Rcu* mt-tRNA^{Arg}. In contrast, *Rcu*CCA added complete CCA-triplets, regardless whether the substrate represented a conventional cloverleaf-structured tRNA or an armless hairpin-like tRNA, indicating an efficient adaptation to these miniaturized substrates (Figure 2B).

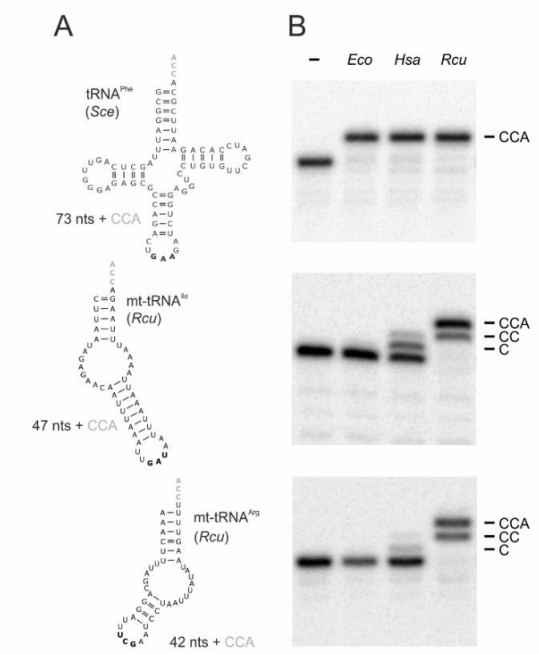


Figure 2. CCA-addition on conventional and hairpin-like tRNA substrates. (**A**) tRNA^{Phe} from *Saccharomyces cerevisiae* (*Sce*) represents a conventionally structured tRNA substrate of standard size (73 nts without CCA), while the mitochondrial tRNAs for isoleucine and arginine from *R. culicivorax* (*Rcu*) considerably deviate in size (47 and 42 nts, respectively; both without 3′-terminal CCA-triplet) and structure, lacking both D- and T-arms. Anticodons are indicated in bold. (**B**) CCA-addition on radioactively labeled tRNA transcripts catalyzed by the corresponding enzymes (20 ng each) of *Escherichia coli* (*Eco*), *H. sapiens* (*Hsa*), and *R. culicivorax* (*Rcu*). Incubation without enzymes represent negative controls (–). All enzymes completely convert the canonical tRNA^{Phe} from *S. cerevisiae* into a mature transcript with CCA-end. On armless mt-tRNAs, the *E. coli* enzyme shows no activity at all, while the human enzyme adds only one or two C residues to mt-tRNA^{Arg} and mt-tRNA^{Ile}, respectively. In contrast, the enzyme of *R. culicivorax* readily synthesizes a complete CCA-end on both transcripts, although the time of incubation was not sufficient for 100% A-addition. The experiment was done in three independent replicates. The panel shows a representative autoradiogram.

To investigate the substrate preferences of *Hsa*CCA and *Rcu*CCA in more detail, a Michaelis–Menten kinetics analysis was performed. Due to the limited RNA solubility, excessive saturating amounts of tRNA cannot be used, and the obtained parameters represent apparent values [24,25,30,58,65]. As *Eco*CCA showed no activity on armless tRNAs, this enzyme was excluded from further analysis. To discriminate between C- and A-addition, assays were performed on tRNAs lacking the CCA-end in the presence of either α-^{32}P-CTP or α-^{32}P-ATP and unlabeled NTPs [23,37,66]. As shown in Table 1, the kinetic parameters of both enzymes on tRNA^{Phe} are rather similar for CC-incorporation, and *Rcu*CCA is somewhat less efficient (0.5-fold) in adding the complete CCA terminus (CCA*). On the armless mt-tRNAs for isoleucine and arginine, *Hsa*CCA and *Rcu*CCA are almost equally active in CC-addition, with somewhat higher values for *Rcu*CCA (1.2 to 1.7x). However, when the complete CCA-incorporation is investigated (CCA*), the *Romanomermis* enzyme is much more efficient on mt-tRNA^{Ile} and mt-tRNA^{Arg} (3.3 to 10x). As the obtained k_{cat} values for *Hsa*CCA in this reaction are at the detection limit, the actual efficiency of *Rcu*CCA relative to *Hsa*CCA is probably much higher.

Table 1. Kinetic parameters of *Hsa*CCA and *Rcu*CCA for CC- and A-addition. As indicated by the relative activity of *Rcu*CCA compared to *Hsa*CCA (change), CC-addition on standard and armless tRNA substrates is similar. In terminal A-addition on the armless tRNAs, *Rcu*CCA shows a 3.3 to 10-fold increase, a clear indication of its adaptation to these substrates. *Hsa*CCA is strongly affected in this reaction, and the obtained low values likely represent an overestimation, as they are close to the detection limit. The actual values are probably much lower. For each analysis, three independent experiments were performed.

Substrate	*Hsa*CCA			*Rcu*CCA			Change
	k_{cat} [s^{-1}]	K_M [μM]	k_{cat}/K_M	k_{cat} [s^{-1}]	K_M [μM]	k_{cat}/K_M	(*Rcu*CCA)
tRNA^{Phe}							
CCA*	0.091 ± 0.012	4.28 ± 1.22	0.02	0.041 ± 0.007	4.66 ± 1.67	0.01	0.5↓
C*C*	0.214 ± 0.034	4.12 ± 1.51	0.05	0.166 ± 0.025	3.01 ± 1.18	0.06	1.2↑
mt-tRNA^{Ile}							
CCA*	0.006 ± 0.001	2.00 ± 0.95	0.003	0.052 ± 0.011	7.84 ± 2.06	0.01	3.3↑
C*C*	0.165 ± 0.016	5.58 ± 1.08	0.03	0.224 ± 0.042	5.43 ± 2.08	0.04	1.3↑
mt-tRNA^{Arg}							
CCA*	0.003 ± 0.000	4.77 ± 1.61	0.001	0.012 ± 0.001	1.30 ± 0.56	0.01	10↑
C*C*	0.041 ± 0.004	1.30 ± 0.44	0.03	0.081 ± 0.011	1.69 ± 0.76	0.05	1.7↑

Taken together, both enzymes prefer a conventionally structured tRNA as substrate. On the hairpin-like tRNAs, *Rcu*CCA still adds complete CCA-ends, although at a lower efficiency for mt-tRNA^{Arg}. The human enzyme, however, strongly prefers the conventional tRNA and is severely affected in A-addition on armless tRNAs, resulting in incomplete and hence non-functional tRNA molecules as already seen in CCA-addition assay (Figure 2B).

2.2. In the Romanomermis Enzyme, Especially the Catalytic Core is Adapted to Armless tRNA Substrates

To identify the contribution of individual enzyme regions to the recognition of armless tRNAs as substrates for CCA-addition, we followed a strategy that we successfully applied to investigate several CCA-adding enzymes concerning their enzymatic features [33,58,67]. We reciprocally exchanged N- and C-termini of *Hsa*CCA and *Rcu*CCA, carrying the complete catalytic core and the region involved in tRNA binding, respectively. Based on the sequence alignment shown in Figure 1, we selected a glutamate at position 212 (*Hsa*CCA) and 213 (*Rcu*CCA), representing the last invariant residue of motif E [68] as fusion position (Figure 3A). To allow for a direct comparison of enzymatic activities of the resulting proteins, we adjusted the efficiency of CCA-addition on the canonically structured tRNA^{Phe} substrate and defined an arbitrary unit as the amount of enzyme required for 50% substrate turnover, ranging between 0.3 ng for wild type (wt) enzymes and 0.5 to 1.3 ng for chimeras. Incubation of the standard substrate tRNA^{Phe} with 1 to 50 arbitrary units of both wt enzymes as well as chimera A

(N-terminal catalytic core of *Rcu*CCA, C-terminus of *Hsa*CCA) and the reciprocal chimera B indicate that all enzymes are fully active and efficiently synthesize a complete CCA-terminus (Figure 3B, left gel panel). On the minimalized mt-tRNA^{Ile} (Figure 3B, right gel panel), even 50 units of the human wt enzyme added only two C residues. In contrast, the same amount of *Romanomermis* enzyme added a complete CCA-end. Surprisingly, chimera B (with the catalytic core of *Rcu*CCA) synthesized a complete CCA-end at considerable efficiency, while identical units of chimera A, carrying the tRNA-binding C-terminus of *Rcu*CCA, catalyzed the full CCA-addition to a lesser extent.

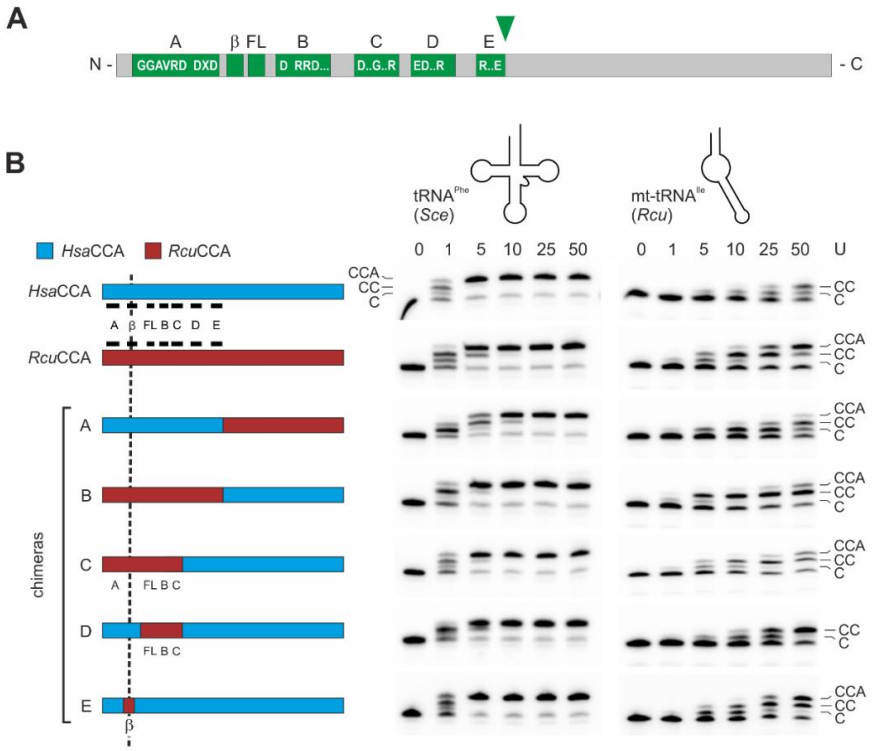


Figure 3. Catalytic activity of wild-type and chimeric enzymes on tRNA^{Phe} and mt-tRNA^{Ile}. (**A**) Bar representation of class II CCA-adding enzymes. The N-terminal region contains the catalytic core consisting of five motifs A to E, a β-turn element (β) and a flexible loop (FL). The green arrowhead indicates the fusion position for enzyme chimeras A and B. (**B**) Left: Bar representation of tested enzyme chimeras consisting of *Hsa*CCA (cyan) and *Rcu*CCA (red) regions. Catalytic core elements are indicated in black. The reciprocal chimeras A and B are fused after position E213, downstream of motif E. The replaced β-turn element (β) in chimera E is located between motif A and the flexible loop (FL). **Gel panels:** CCA-addition on the conventional tRNA^{Phe} from yeast (*Sce*, left) and the armless mt-tRNA^{Ile} from *R. culicivorax* (*Rcu*, right) with increasing amounts of enzymes indicated as arbitrary units (U). All enzymes catalyze an efficient CCA-incorporation on the conventional tRNA. On the armless tRNA^{Ile}, the *Romanomermis* wt enzyme synthesizes a complete CCA-end, while the corresponding human enzyme adds only two C-residues. Chimeras A and B also add the terminal A, but at a somewhat reduced level. On this tRNA substrate, chimera B, carrying the catalytic core of *Rcu*CCA, is more efficient than chimera A, where A-incorporation is only visible at the highest enzyme concentration. Chimera C still shows full CCA-addition, whereas no terminal A-incorporation is observed for chimera D. Chimera E shows an efficiency comparable to that of the *Rcu*CCA wt enzyme, indicating the importance of the β-turn region in the reaction on armless tRNAs. The fact that chimera E is more active than chimera B likely reflects differences in the compatibility of the chosen fusion positions in these chimeras, resulting in different protein folding and/or catalytic efficiency. For each construct, up to four independent experiments were performed. For each tRNA substrate, a representative gel is shown.

While these reactions indicate that both enzyme parts participate in the adaptation to the armless tRNA substrates, the contribution of the catalytic core seems to have a greater impact on the acceptance of these substrates. Based on these results, we generated a series of chimeric enzymes carrying parts of the *Rcu*CCA catalytic core in the context of the human enzyme (Figure 3B, Table S1). For chimera C, the human C-terminal part was extended to include motifs D and E. Comparable to chimera B, this construct added the full CCA-end to *Rcu* mt-tRNA^{Ile}. As the flexible loop (FL) and motif C are involved in the enzymes' specificity switch to A-addition [22,23,26], chimera D was constructed, consisting of *Hsa*CCA where the region spanning these two elements was replaced by the corresponding part from *Rcu*CCA. While this enzyme was fully active on tRNA^{Phe}, the loop and motif C replacement did not result in A-addition on mt-tRNA^{Ile}. These results led to the conclusion that an element upstream of the flexible loop must be responsible for A-addition on armless tRNA substrates. As this region contains a small β-turn element that was described to contribute to tRNA-CC primer positioning and, as a consequence, A-addition [22], chimera E was constructed, where a stretch of 30 amino acid residues (positions 61–91) carrying this element was replaced in the backbone of the human enzyme. The resulting enzyme catalyzed CCA-addition on both tRNA^{Phe} and mt-tRNA^{Ile} at efficiencies comparable to the wt *Rcu*CCA enzyme (Figure 3B). Hence, this divide-and-conquer approach allowed us to identify the β-turn as an important element involved in the adaptation of *Rcu*CCA to the armless tRNA substrates.

2.3. The β-Turn of the R. culicivorax CCA-Adding Enzyme Strongly Contributes to Substrate Binding and CCA Incorporation on Armless tRNAs

To investigate whether the β-turn of *Rcu*CCA contributes to an increased efficiency in binding of armless tRNA substrates, we performed electrophoretic mobility shift experiments on chimera E and the parental enzymes *Hsa*CCA and *Rcu*CCA with an armless tRNA transcript. Radioactively labeled mt-tRNA^{Ile} lacking the CCA terminus was incubated with increasing amounts of recombinantly expressed enzymes and separated by native polyacrylamide gel electrophoresis (Figure 4). Enzyme-bound and free substrates were visualized and binding parameters were determined by nonlinear regression. The *Romanomermis* enzyme showed an efficient and robust binding to mt-tRNA^{Ile}, resulting in a K_d value of 1.4 μM, while for the human enzyme, no significant binding could be detected, as previously reported for this enzyme class [23,33,69]. Interestingly, and concurrent with its wt-like activity on mt-tRNA^{Ile} (Figure 3B), chimera E also showed a high affinity for mt-tRNA^{Ile} at a K_d of 1.8 μM.

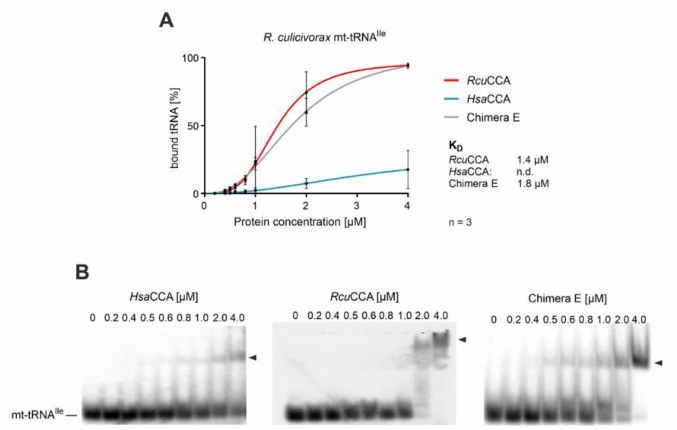


Figure 4. Binding of wt and chimeric CCA-adding enzymes to an armless tRNA. Quantitative analysis of enzyme binding to the armless mt-tRNA^{Ile} determined by electrophoretic mobility shifts. (**A**) While the tRNA interaction of *Hsa*CCA over the whole concentration range (0–4 μM) is too weak to calculate dissociation constants, *Rcu*CCA as well as chimera E exhibit a strong affinity to this substrate, resulting in dissociation constants of 1.4 and 1.8 μM, respectively. Data are means ± SD; $n = 3$. (**B**) Images of representative gel shift assays on *Hsa*CCA, *Rcu*CCA, and chimera E with mt-tRNA^{Ile} as substrate.

When transcripts of mt-tRNA^{Ile} with different 3′-ends (mt-tRNA^{Ile}, mt-tRNA^{Ile}-C, mt-tRNA^{Ile}-CC, mt-tRNA^{Ile}-CCA) were offered to *Rcu*CCA, the enzyme bound all substrates with similar affinity, indicating that the composition of the tRNA 3′-end does not affect this interaction (Figure S2A). The human enzyme, in contrast, did not show efficient binding to any of these substrates. Gel shift experiments on the conventional substrate tRNA^{Phe} indicate that this high affinity of the *Rcu*CCA enzyme is not restricted to armless tRNA substrates (Figure S2B). With a binding constant of 1.3 μM, the enzyme exhibits a similar binding behavior to the cloverleaf-shaped tRNA as it does on the armless mt-tRNA^{Ile} (K_d = 1.4 μM), while the human enzyme again shows almost no interaction. Hence, the adaptation of the *Romanomermis culicivorax* CCA-adding enzyme to process both cloverleaf-structured cytosolic as well as armless mitochondrial tRNAs is obviously achieved by a general exceptional tight interaction with its substrates, regardless whether they represent canonical or minimalized tRNAs.

While the β-turn itself (GEKH) is identical in *Rcu*CCA and *Hsa*CCA, the flanking sequences in the *Rcu*CCA enzyme differ in 13 positions from the counterpart, and some of them likely contribute to the extended substrate specificity of this enzyme. In these regions, the *Romanomermis* enzyme carries several lysine residues that are not present in *Hsa*CCA (Figure 5). K74 is also found in the corresponding enzymes of *C. elegans* and *A. suum*, while the human enzyme carries an asparagine at this position (Figure 1). As the basic lysine might enhance the enzymes' binding to negatively charged tRNA substrates, it represents a top candidate for such a substrate adaptation. At positions 89 and 90, *Rcu*CCA carries a further lysine residue followed by a valine position that might also be involved in tRNA primer binding, although these residues are not present in *Cel*CCA and *Asu*CCA. Hence, we replaced these positions in *Rcu*CCA by the corresponding residues of the human enzyme, resulting in *Rcu*CCA K74N, K89Δ/V90E, and the combination thereof (K74N/K89Δ/V90E). The CCA-adding activity of the recombinantly expressed enzyme variants was determined on yeast tRNA^{Phe} and adjusted to arbitrary units (1 U corresponding to 0.3 ng protein; Figure 5D,E). On this conventional tRNA, all enzymes showed a comparable activity, with a slight reduction in activity of *Rcu*CCA K74N (Figure 5D, upper panel). On mt-tRNA^{Ile}, however, the variants carrying the K74N replacement were considerably affected in incorporating the terminal A residue. *Rcu*CCA K89Δ/V90E, in contrast, added a full CCA-end on this tRNA, although it seems to be somewhat less efficient than the wt enzyme. The reciprocal amino acid exchanges and insertions introduced in the human enzyme support these results (Figure 5E). *Hsa*CCA N74K (for reasons of clarity, numbering is according to *Rcu*CCA positions) shows an efficient addition of a complete CCA-end on mt-tRNA^{Ile}. Intriguingly, also *Hsa*CCA 89Kins/E90V and the combination N74K/K89ins/E90V exhibit improved CCA-addition on this armless substrate. As a conclusion, K74, and to a certain extent also K89/V90, represent important positions to accept armless tRNA substrates and probably contribute to a stable interaction with the tRNA and the correct positioning of its 3′-end as a primer.

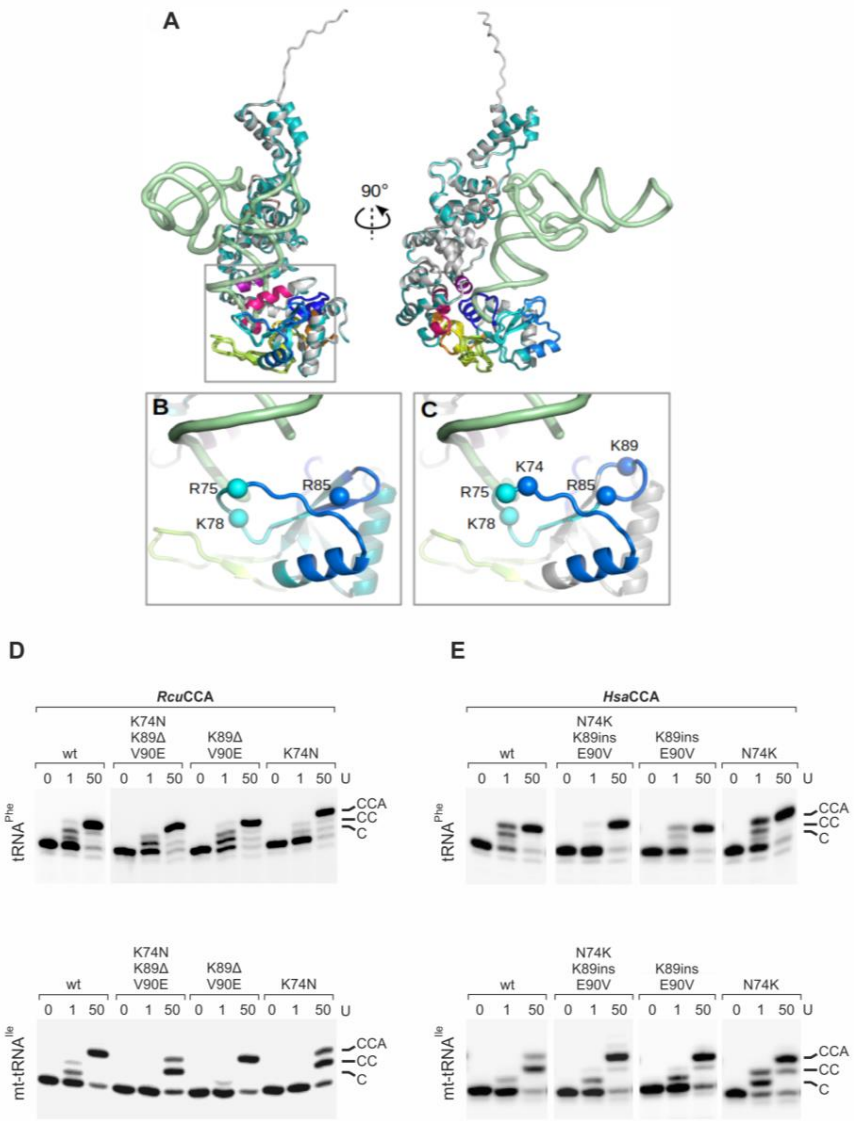


Figure 5. The β-turn in *Hsa*CCA and *Rcu*CCA enzymes. (**A**) Superimposed full-length models of *Hsa*CCA (cyan) and *Rcu*CCA (light gray) with the backbone of a bound tRNA (green) in two perpendicular views. The tRNA position was obtained by superimposing the A-adding enzyme:tRNA-CC complex from *Aquifex aeolicus* onto the human enzyme [34,35,70]. Motif A (dark blue), β-turn region (medium blue), β-turn (light blue), B/A motif and flexible loop (green), motif B (yellow), motif C (orange), motif D (red) and motif E (violet) are indicated. (**B**,**C**). Zoom into the β-turn region and the tRNA 3′-end (corresponding to the squared region in (**A**) of *Hsa*CCA (**B**) and RcuCCA (**C**). Spheres represent the Cα positions of positively charged residues (K and R). *Rcu*CCA carries two additional lysines at positions 74 and 89 that might contribute to tRNA binding and primer positioning. Model is based on the crystal structure of the A-adding enzyme from *Aquifex aeolicus* [34]. (**D**) Enzymatic activity of *Rcu*CCA carrying mutations K74N, K89Δ/V90E and K74N/K89Δ/V90E. 0, 1, and 50 arbitrary units of enzyme variants were incubated with yeast $tRNA^{Phe}$ and the armless mt-$tRNA^{Ile}$. *Rcu*CCA wt accepts both tRNAs for CCA-addition, while *Rcu*CCA K74N is less active on mt-$tRNA^{Ile}$, resulting in a considerably reduced A-addition. In contrast, *Rcu*CCA K89Δ/V90E catalyzes full CCA-addition on the conventional (comparable to wt activity) as well as on the armless tRNA. The triple variant *Rcu*CCA K74N/K89Δ/V90E shows the same activity as *Rcu*CCA K74N. (**E**) The introduction of the corresponding amino acids of *Rcu*CCA into *Hsa*CCA enables this enzyme to add a complete CCA-end on mt-$tRNA^{Ile}$, in contrast to the wildtype situation. These results indicate that especially position K74 of the β-turn, but to a certain extent also K89/V90, contribute to the substrate adaptation of *Rcu*CCA.

3. Discussion

3.1. A Specific Adaptation within the Catalytic Core Enables CCA-Addition to Minimalized tRNA Substrates

As the tRNA genes of most organisms do not encode the 3′-terminal CCA-triplet, this essential feature has to be added post-transcriptionally by the CCA-adding enzyme. In eukaryotes, a single enzyme is responsible for this maturation step in both cytosolic as well as organellar tRNA pools [50,51,71]. A second function of CCA-adding enzyme is to monitor the intactness of its tRNA substrate, so that only undamaged molecules are accepted for CCA-incorporation [72–76]. The CCA-adding enzyme does not recognize a specific sequence or base pair in its substrates, but relies on common elements of the overall tRNA cloverleaf structure, as it is also observed for RNase P, tRNase Z, and some tRNA modifying enzymes [77–81]. Metazoan mitochondria, however, encode for tRNAs with deviations from the cloverleaf, where D- or T-arms are reduced or lacking [38,41,48]. As these tRNAs still carry a conventional acceptor stem, they are correctly processed by RNase P, since this enzyme predominantly recognizes this structural feature [78,79]. The CCA-adding enzyme, in contrast, specifically interacts with the tRNA elbow region, and especially with the T-loop region [32,69,82–84]. Hence, a bacterial enzyme like the *E. coli* version, evolved for conventional tRNAs, is strongly impaired on truncated tRNAs [52]. The most extreme truncations are observed in nematodes like *R. culicivorax*, in acari and in arachnids, where hairpin-like tRNAs are found that require a specific co-evolution of the corresponding enzymes. On such substrates, the *E. coli* enzyme is completely inactive (Figure 2B). However, the human enzyme co-evolved to accept the D-armless $tRNA^{Ser}{}_{AGY}$ found in human mitochondria (Figure S1) [41]. Hence, this enzyme accepts the armless tRNAs to a certain extent and adds two C-residues, but not the terminal A, indicating that A incorporation requires a specific adaptation to such extreme substrates (Figure 2B, Figure 3, Table 1). In contrast, the *R. culicivorax* enzyme is adapted to these tRNA structures and efficiently adds the complete CCA-triplet. For *Hsa*CCA and *Rcu*CCA, the kinetic parameters for CC-incorporation on standard as well as armless tRNAs are quite similar and correspond to published values [37]. However, when it comes to A-addition, only *Rcu*CCA is adapted to the hairpin-like tRNAs (Figure 2B, Figure 3), although this reaction step is less efficient than C-addition (Table 1).

In the transition from C- towards A-addition, CCA-adding enzymes undergo domain rearrangements to accommodate the growing tRNA 3′-end in the catalytic core and to switch the specificity of the amino acid template in the nucleotide binding pocket from CTP towards ATP recognition [21–23,85]. During this structural rearrangement, the tRNA substrate has to remain bound to the enzyme, and this is usually accomplished by specific interactions of the tRNA 3′-end in the catalytic core and of the elbow region (T- and D-loop) with the C-terminal region of the CCA-adding enzyme [22,24]. For CCA-adding enzymes adapted to conventional tRNAs or tRNAs lacking only one arm, this interaction is not very tight, as no K_d values could be determined [23,33,69] — yet it is sufficient for a complete synthesis of the CCA-end. However, for armless tRNAs, these interactions seem to be insufficient, and the observed high substrate affinity of *Rcu*CCA (in contrast to the human enzyme; Figure 4, Figure S2) corroborates this hypothesis. Obviously, the *Romanomermis* enzyme is able to bind its substrate very tightly, and since this is also the case for armless tRNAs, this interaction cannot involve the conventional contacts between C-terminus of the enzyme and T-loop of the tRNA but must be located elsewhere, although the C-terminus contributes to the terminal A-addition to a certain extent, as shown by the *Rcu*/*Hsa*CCA chimera A (Figure 4).

In the detailed analysis of enzyme chimeras between *Rcu*CCA and *Hsa*CCA, an adaptation of both enzymes to different tRNA substrates is obvious. In *Hsa*CCA and chimeras B, C, D, and E, the tRNA-binding C-terminus helps to confer an efficient CCA-addition on the conventional $tRNA^{Phe}$ (Figure 3B). As *Hsa*CCA shows a rather weak interaction with tRNA substrates (Figure 4, Figure S2), this C-terminal part obviously contributes to an efficient CCA-synthesis by enabling the final product release [85]. In contrast, in *Rcu*CCA, an adaptation to armless tRNA substrates in the N-terminus is obvious. Here, two elements of the catalytic core are described to play an important role in the specificity switch from C- to A-addition. The flexible loop acts as a lever that adjusts the templating

amino acids for correct ATP binding [22,25,26,66], and motif C represents a springy hinge that supports the domain rearrangements involved in this reaction step [22,23]. Accordingly, one could expect that both elements are adapted to the CCA-addition on armless tRNAs. Yet, chimera D (human enzyme with flexible loop and motif C of *Rcu*CCA) does not show any A-addition on the armless tRNA, while it is fully active on a standard substrate (Figure 3), excluding a specific adaptation of these two elements.

A detailed inspection of the residual N-terminal *Rcu*CCA components in the chimeras revealed that constructs carrying a β-turn element located between strands 3 and 4 of the β-sheet in the catalytic core are able to add the terminal A, indicating that this element represents a major adaptation to armless tRNAs (Figure 3B, chimeras B, C, and E)

3.2. The ß-Turn Element Impacts the Substrate Affinity of R. culicivorax CCA-Adding Enzyme

In chimera E, the inserted β-turn element consists of 30 amino acids located between motif A and a conserved stretch upstream of the flexible loop (Figure 1). Toh et al., showed that in CCA-adding enzymes for conventional tRNAs, this region plays an important role in binding the growing CC-end of the tRNA and adjusting it in the catalytic core for the nucleophilic attack initiating A-addition [22]. Similar β-turn elements are found for the archaeal class I CCA-adding enzymes [29,30,86], poly(A) polymerases as well as for DNA- and RNA-polymerases, where they are described to position the 3′-hydroxyl of the primer 3′-end in close vicinity of the catalytically important metal ions located in the active site [68]. In *Rcu*CCA, however, the β-turn element has an additional function, as it dramatically contributes to tRNA binding. Here, especially the lysine residue at position 74 (K74) seems to be involved, as its replacement by asparagine, the corresponding position in *Hsa*CCA, affects A-addition on an armless tRNA (Figure 5D). The introduction of K74 into *Hsa*CCA supports the importance of this lysine residue, as it enables a strongly improved CCA-addition on mt-tRNA^{Ile} by this enzyme. Intriguingly, also the *Hsa*CCA K89ins/E90V variant shows an enhanced CCA-addition on mt-tRNA^{Ile}. This could be a general effect of the positively charged lysine side chain that keeps the tRNA in close proximity during polymerization, enabling a more efficient nucleotide transfer. In *Rcu*CCA, these positions seem to be less important, as their replacement by the human residues have only very subtle effects on CCA-addition. It is conceivable that other positions not directly involved in tRNA binding might structurally support the substrate adaptation of *Rcu*CCA by folding the loop region into a conformation that optimally positions K74 (and other interacting residues) for primer binding during polymerization. Corresponding crystal structures of *Rcu*CCA in complex with an armless tRNA are needed to clarify this point.

The presented results on *Rcu*CCA, *Hsa*CCA, and the corresponding chimeras support the following hypothesis. As described above, the binding of the tRNA's T-loop to its C-terminus represents one of the major substrate interactions of the CCA-adding enzyme, ensuring that the tRNA primer remains correctly located for A-addition during the domain rearrangements inducing the specificity switch [34,35]. As a consequence, product release seems to be a limiting factor in the reaction [85]. As this interaction is not possible with armless tRNAs, the human enzyme loses contact after adding the C residues, while a conventional tRNA remains correctly bound in the T-loop/C-terminus interaction and gets elongated by a complete CCA-terminus. In *R. culicivorax*, the CCA-adding enzyme had to adapt to armless tRNAs and evolved a different mode of tRNA interaction, as a binding of T-loop and C-terminus is not possible anymore. Here, the function of the β-turn region evolved from simple primer positioning for A-incorporation into an enhanced high-affinity substrate binding compensating for the loss of the conventional C-terminal tRNA interaction. As this strong interaction is visible at all stages of CCA-addition on canonical as well as armless tRNAs (Figure S2A), the function of the β-turn region in *Rcu*CCA is not limited to bind a tRNA-CC intermediate but generally contributes to an efficient substrate binding during the complete CCA-incorporation reaction. This ensures that the enzyme's interaction with armless tRNAs is sufficiently strong to survive the structural rearrangements during polymerization. A drawback of this tight binding of *Rcu*CCA might be a reduced product

release, and the slightly less efficient CCA-addition on the conventional $tRNA^{Phe}$ could be an indication of this (Table 1).

3.3. An Orthogonal Translation System?

Swapping domains between proteins is a useful method to generate proteins with new functions—an approach that is frequently used by nature as well as researchers [87–89]. The fact that the replacement of a small β-turn element converts the human tRNA nucleotidyltransferase into an enzyme that accepts armless tRNAs as substrates raises the question whether it is possible to generate an orthogonal system for the incorporation of non-natural amino acids in a host cell system. Representing a fascinating idea, this is highly unlikely, as the armless tRNAs are so different to their conventional counterparts. It is known that many mt-tRNAs are not recognized and charged by cytosolic aminoacyl-tRNA synthetases but require specifically adapted mitochondrial enzymes [90]. Similarly, mitochondrial ribosomes have undergone a specific co-evolution to compensate for the unusual structural features of mt-tRNAs [91]. While metazoan mt-rRNAs are usually shorter than their cytosolic counterparts, the mitochondrial ribosomal proteins are generally enlarged [92–95]. Again, nematodes represent the extreme case, having ribosomes strongly enriched in protein content but with reduced rRNA components [96–98]. Due to this intricate co-evolution of the mitochondrial protein synthesis components, it is probably not feasible to use armless tRNAs as an orthogonal tool in synthetic biology.

Taken together, the CCA-adding enzyme of *R. culicivorax* shows a remarkable adaptation to hairpin-like tRNA where the loss of substrate interactions with the C-terminus is compensated by enhanced tRNA binding of a different enzyme region. The evolutionary plasticity of enzymes is described for the composition of active site residues, where amino acids with identical catalytic roles are located at different positions in the primary sequence [99]. The catalytic core, however, remains unchanged and structurally almost identical. In contrast, the β-turn element of *Rcu*CCA is not a mimicry of the T-loop/C-terminus interaction but recognizes a very different region of the tRNA (the 3′-end) and is still part of the catalytic core. Hence, its specific adaptation to the miniaturized tRNAs add a new layer of evolutionary plasticity. The structural resolution of this enzyme in complex with its tRNA substrate is expected to shed more light into this unusual and fascinating substrate adaptation.

4. Materials and Methods

4.1. Construction of Recombinant Enzymes

Open reading frames of CCA-adding enzymes from *Escherichia coli* and *Homo sapiens* were cloned into pET30 Ek/LIC plasmid with an N-terminal His_6-Tag. The mt target signals were not included, and in both enzymes, the cloned coding regions started at the following conserved methionine residue as described [37,50,56,57]. For the CCA-adding enzyme of *Romanomermis culicivorax*, the coding sequence was identified in the *R. culicivorax* genome assembly available at https://parasite.wormbase.org/Romanomermis_culicivorax_prjeb1358/Info/Index/, codon-optimized for expression in *E. coli* and synthesized in pET28a by GenScript®® (Piscataway, NJ, USA). All alignments were done using Jalview 2 [100]. Point mutations were introduced as described [23].

4.2. Cloning of Chimeric Enzymes

Chimeric enzymes of CCA-adding enzymes from *Homo sapiens* and *Romanomermis culicivorax* were generated via site-directed mutagenesis in pET30-Ek/LIC or pET28a plasmids, respectively. All chimeras were cloned with an N-terminal His_6-tag. The fusion positions of all chimeras are shown in Table S1 and Figure S3.

4.3. Expression and Purification of Recombinant Enzymes

E. coli BL21 (DE3) cca::cam lacking the endogenous CCA-adding enzyme were transformed with plasmids encoding the CCA-adding enzymes from *H. sapiens* (*Hsa*CCA), *R. culicivorax* (*Rcu*CCA), or chimeric enzymes. For CCA-adding enzyme from *E. coli* (*Eco*CCA), *E. coli* BL21 (DE3) was used. Cells were grown in 400 ml LB or TB with 50 µg/ml kanamycin and 35 µg/ml chloramphenicol (only for cca::cam strains) at 30 °C. Expression was induced at OD_{600} = 1.5 by adding 400 ml ice-cold LB or TB containing both antibiotics and IPTG to a final concentration of 1 mM. Cultures were incubated over night at 16 °C and then harvested at 6340 g for 15 min.

Pellets were resuspended in 8 ml ice-cold lysis-buffer (25 mM Tris/HCl pH 7.6, 500 mM NaCl, 1 mM DTT for *Eco*CCA and *Hsa*CCA, 100 mM phosphate buffer pH 7.0, 500 mM NaCl, 10% glycerol, 0.2% NP-40, 1 mM DTT for *Rcu*CCA) and disrupted with 5 g Zirconia beads and Fastprep-24 homogenizer (6 m/s, 30 s). Cell lysates were centrifuged at 30,600 g, 30 min, 4 °C, sterile filtrated and loaded onto a HisTrap FF 1 ml or 5 ml column (GE Healthcare). Column wash was performed with 4–10 column volumes of binding buffer (25 mM Tris/HCl pH 7.6, 500 mM NaCl for *Eco*CCA and *Hsa*CCA, 100 mM phosphate buffer pH 7.0, 500 mM NaCl, 10% glycerol for *Rcu*CCA) with 50 mM imidazole. His-tagged proteins were eluted with 3–8 column volumes of elution buffer (binding buffer with 500 mM imidazole). If necessary, protein containing fractions were further purified by size exclusion chromatography on a HiLoad 16/60 Superdex 75 pg column in binding-buffer containing 200 mM NaCl. Protein-containing fractions were combined and concentrated on Vivaspin 6 columns (15–30 kDa MWCO, GE Healthcare). Proteins were stored in 40% glycerol (*v*/*v*) at −80 °C. Protein concentration was determined according to Bradford [101].

4.4. tRNA Preparation

Armless mitochondrial tRNAs for isoleucine and arginine from *R. culicivorax* [48] and canonical cytosolic $tRNA^{Phe}$ from *Saccharomyces cerevisiae* were generated as in vitro transcripts lacking the CCA-end in the presence of α^{32}P-ATP (3000 Ci/mmol). Homogeneous 5′- and 3′-ends of the transcripts were generated as described [102]. For kinetic analyses, tRNAs were transcribed without α^{32}P-ATP.

4.5. Electrophoretic Mobility Shift Assay (EMSA)

0.5 pmol α^{32}P-ATP-labeled tRNA substrates were heated for 2 min at 90 °C, cooled to room temperature and incubated with 0 to 4 µM of enzyme in HEPES/KOH (pH 7.6), 30 mM KCl and 6 mM $MgCl_2$ at 20 °C for 10 min. After addition of glycerol to a final concentration of 18.5%, tRNAs were separated by 5% native polyacrylamide gel electrophoresis. For visualization of enzyme-bound and free substrates, a Typhoon 9410 scanner (Cytiva, Freiburg, Germany) was used. Dissociation constants were determined in three independent experiments by nonlinear regression using GraphPad Prism 7.

4.6. Activity Test and Determination of Arbitrary Units

Initial activity tests for CCA-addition were performed in 30 mM HEPES/KOH pH 7.6, 30 mM KCl, 6 mM $MgCl_2$, 2 mM DTT, 0.5 mM NTPs, 5 pmol tRNA, and 20 ng of enzyme in a reaction volume of 20 µl at 20 °C.

For comparative analysis of CCA-addition on armless tRNAs (Figure 3), activity of all enzyme preparations was normalized using canonical $tRNA^{Phe}$ as a substrate. CCA-addition was performed in in the same buffer as described above. For calculation of arbitrary units, 5 pmol of tRNA were incubated with increasing amounts of enzymes for 30 min at 20 °C. Reactions were ethanol-precipitated and analyzed on 10% or 12.5% polyacrylamide gels by autoradiography. Enzyme amounts leading to 50% substrate-turnover were defined as 1 arbitrary unit. Nucleotide incorporation assays were performed as mentioned above. 5 pmol of tRNA were incubated with 1, 5, 10, 25, and 50 arbitrary units of each enzyme preparation for 30 min at 20 °C and analyzed as described [25,58].

4.7. Kinetic Analysis

Steady-state Michaelis–Menten kinetics were performed as described [103]. Each reaction contained 1 mM NTPs and 3 µCi of α^{32}P-CTP or α^{32}P-ATP (3000 Ci/mmol) and 30–35 ng *Hsa*CCA or 30–75 ng *Rcu*CCA. Non-labeled tRNA transcripts without CCA-end were titrated from 1-10 µM and incubated for 15–20 min at 20 °C. Determination of incorporated radioactivity was performed as described [25,66] and kinetic parameters of three independent experiments were calculated by non-linear regression Michaelis–Menten kinetic (GraphPad Prism). As the tRNA transcripts are not soluble at excessive saturating conditions, the calculated kinetic parameters represent apparent values, as frequently used for this type of enzymes [24,25,30,58,65].

4.8. Enzyme Modeling

A homology model of *Rcu*CCA enzyme was built using Modeller [104] based on a sequence alignment with the human enzyme and on the human structure (PDB id: 4x4w) [70]. A model of the human enzyme was also built to include all loops that are not visible in the crystal structure. Models were superimposed to the structure of A-adding enzyme:tRNA complex for (PDB id: 4wc2) [34] in PyMOL (v2.4.0, Schrödinger) to position a tRNA in the active site and visualize loops in the vicinity of the tRNA 3′-tail.

Supplementary Materials: The following are available online at http://www.mdpi.com/1422-0067/21/23/9047/s1.

Author Contributions: H.B., C.S. and M.M. designed the experiments; S.P., O.H., T.K., S.B., K.R., and T.J. performed the experiments; H.B., C.S., M.M., S.P., K.R., and O.H. analyzed the data; O.H., S.P., H.B., and M.M. wrote the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Deutsche Forschungsgemeinschaft, grant number Mo 634/10-1. We acknowledge support from the Deutsche Forschungsgemeinschaft (DFG) and Leipzig University within the program of Open Access Publishing.

Acknowledgments: We thank Philipp Schiffer (University of Cologne) for help with the annotation of the CCA-adding enzyme from *Romanomermis culicivorax* and Barbara Klüver for expert technical assistance.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Rak, R.; Dahan, O.; Pilpel, Y. Repertoires of tRNAs: The Couplers of Genomics and Proteomics. *Annu. Rev. Cell Dev. Biol.* **2018**, *34*, 239–264. [CrossRef] [PubMed]
2. Wolin, S.L.; Matera, A.G. The trials and travels of tRNA. *Genes Dev.* **1999**, *13*, 1–10. [CrossRef] [PubMed]
3. Hopper, A.K. Transfer RNA post-transcriptional processing, turnover, and subcellular dynamics in the yeast Saccharomyces cerevisiae. *Genetics* **2013**, *194*, 43–67. [CrossRef]
4. Caetano-Anollés, G.; Sun, F.-J. The natural history of transfer RNA and its interactions with the ribosome. *Front. Genet.* **2014**, *5*, 127. [CrossRef] [PubMed]
5. Hopper, A.K.; Nostramo, R.T. tRNA Processing and Subcellular Trafficking Proteins Multitask in Pathways for Other RNAs. *Front. Genet.* **2019**, *10*, 96. [CrossRef] [PubMed]
6. D'Souza, A.R.; Minczuk, M. Mitochondrial transcription and translation: Overview. *Essays Biochem.* **2018**, *62*, 309–320. [CrossRef]
7. Zoschke, R.; Bock, R. Chloroplast Translation: Structural and Functional Organization, Operational Control, and Regulation. *Plant Cell* **2018**, *30*, 745–770. [CrossRef]
8. Kim, S.H.; Suddath, F.L.; Quigley, G.J.; McPherson, A.; Sussman, J.L.; Wang, A.H.; Seeman, N.C.; Rich, A. Three-dimensional tertiary structure of yeast phenylalanine transfer RNA. *Science* **1974**, *185*, 435–440. [CrossRef]
9. Jühling, F.; Mörl, M.; Hartmann, R.K.; Sprinzl, M.; Stadler, P.F.; Pütz, J. tRNAdb 2009: Compilation of tRNA sequences and tRNA genes. *Nucleic Acids Res.* **2009**, *37*, D159–D162. [CrossRef]

10. Giegé, R.; Jühling, F.; Pütz, J.; Stadler, P.; Sauter, C.; Florentz, C. Structure of transfer RNAs: Similarity and variability. *Wiley Interdiscip. Rev. RNA* **2012**, *3*, 37–61. [CrossRef]
11. Sprinzl, M.; Cramer, F. The -C-C-A end of tRNA and its role in protein biosynthesis. *Prog. Nucleic Acid Res. Mol. Biol.* **1979**, *22*, 1–69. [PubMed]
12. Green, R.; Noller, H.F. Ribosomes and translation. *Annu. Rev. Biochem.* **1997**, *66*, 679–716. [CrossRef] [PubMed]
13. Weiner, A.M. tRNA maturation: RNA polymerization without a nucleic acid template. *Curr. Biol.* **2004**, *14*, R883–R885. [CrossRef] [PubMed]
14. Xiong, Y.; Steitz, T.A. A story with a good ending: tRNA 3′-end maturation by CCA-adding enzymes. *Curr. Opin. Struct. Biol.* **2006**, *16*, 12–17. [CrossRef]
15. Betat, H.; Rammelt, C.; Mörl, M. tRNA nucleotidyltransferases: Ancient catalysts with an unusual mechanism of polymerization. *Cell. Mol. Life Sci.* **2010**, *67*, 1447–1463. [CrossRef]
16. Yue, D.; Maizels, N.; Weiner, A.M. CCA-adding enzymes and poly(A) polymerases are all members of the same nucleotidyltransferase superfamily: Characterization of the CCA-adding enzyme from the archaeal hyperthermophile Sulfolobus shibatae. *RNA* **1996**, *2*, 895–908.
17. Xiong, Y.; Li, F.; Wang, J.; Weiner, A.M.; Steitz, T.A. Crystal structures of an archaeal class I CCA-adding enzyme and its nucleotide complexes. *Mol. Cell* **2003**, *12*, 1165–1172. [CrossRef]
18. Holm, L.; Sander, C. DNA polymerase beta belongs to an ancient nucleotidyltransferase superfamily. *Trends Biochem. Sci.* **1995**, *20*, 345–347. [CrossRef]
19. Steitz, T.A. A mechanism for all polymerases. *Nature* **1998**, *391*, 231–232. [CrossRef]
20. Aravind, L.; Koonin, E.V. DNA polymerase beta-like nucleotidyltransferase superfamily: Identification of three new families, classification and evolutionary history. *Nucleic Acids Res.* **1999**, *27*, 1609–1618. [CrossRef]
21. Li, F.; Xiong, Y.; Wang, J.; Cho, H.D.; Tomita, K.; Weiner, A.M.; Steitz, T.A. Crystal structures of the Bacillus stearothermophilus CCA-adding enzyme and its complexes with ATP or CTP. *Cell* **2002**, *111*, 815–824. [CrossRef]
22. Toh, Y.; Takeshita, D.; Numata, T.; Fukai, S.; Nureki, O.; Tomita, K. Mechanism for the definition of elongation and termination by the class II CCA-adding enzyme. *EMBO J.* **2009**, *28*, 3353–3365. [CrossRef] [PubMed]
23. Ernst, F.G.M.; Rickert, C.; Bluschke, A.; Betat, H.; Steinhoff, H.-J.; Mörl, M. Domain movements during CCA-addition: A new function for motif C in the catalytic core of the human tRNA nucleotidyltransferases. *RNA Biol.* **2015**, *12*, 435–446. [CrossRef] [PubMed]
24. Tomita, K.; Fukai, S.; Ishitani, R.; Ueda, T.; Takeuchi, N.; Vassylyev, D.G.; Nureki, O. Structural basis for template-independent RNA polymerization. *Nature* **2004**, *430*, 700–704. [CrossRef]
25. Hoffmeier, A.; Betat, H.; Bluschke, A.; Gunther, R.; Junghanns, S.; Hofmann, H.-J.; Mörl, M. Unusual evolution of a catalytic core element in CCA-adding enzymes. *Nucleic Acids Res.* **2010**, *38*, 4436–4447. [CrossRef]
26. Neuenfeldt, A.; Just, A.; Betat, H.; Mörl, M. Evolution of tRNA nucleotidyltransferases: A small deletion generated CC-adding enzymes. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 7953–7958. [CrossRef]
27. Davies, J.F.; Almassy, R.J.; Hostomska, Z.; Ferre, R.A.; Hostomsky, Z. 2.3 Å crystal structure of the catalytic domain of DNA polymerase β. *Cell* **1994**, *76*, 1123–1133. [CrossRef]
28. Sawaya, M.R.; Pelletier, H.; Kumar, A.; Wilson, S.H.; Kraut, J. Crystal structure of rat DNA polymerase beta: Evidence for a common polymerase mechanism. *Science* **1994**, *264*, 1930–1935. [CrossRef]
29. Cho, H.D.; Verlinde, C.L.; Weiner, A.M. Archaeal CCA-adding enzymes: Central role of a highly conserved beta-turn motif in RNA polymerization without translocation. *J. Biol. Chem.* **2005**, *280*, 9555–9566. [CrossRef]
30. Tomita, K.; Ishitani, R.; Fukai, S.; Nureki, O. Complete crystallographic analysis of the dynamics of CCA sequence addition. *Nature* **2006**, *443*, 956–960. [CrossRef]
31. Tomita, K.; Yamashita, S. Molecular mechanisms of template-independent RNA polymerization by tRNA nucleotidyltransferases. *Front. Genet.* **2014**, *5*, 36. [CrossRef]
32. Li, Z.; Gillis, K.A.; Hegg, L.A.; Zhang, J.; Thurlow, D.L. Effects of nucleotide substitutions within the T-loop of precursor tRNAs on interaction with ATP/CTP:tRNA nucleotidyltransferases from Escherichia coli and yeast. *Biochem. J.* **1996**, *314 Pt 1*, 49–53. [CrossRef]
33. Tretbar, S.; Neuenfeldt, A.; Betat, H.; Mörl, M. An inhibitory C-terminal region dictates the specificity of A-adding enzymes. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 21040–21045. [CrossRef] [PubMed]
34. Yamashita, S.; Martinez, A.; Tomita, K. Measurement of Acceptor-TΨC Helix Length of tRNA for Terminal A76-Addition by A-Adding Enzyme. *Structure* **2015**, *23*, 830–842. [CrossRef]

35. Yamashita, S.; Takeshita, D.; Tomita, K. Translocation and rotation of tRNA during template-independent RNA polymerization by tRNA nucleotidyltransferase. *Structure* **2014**, *22*, 315–325. [CrossRef] [PubMed]
36. Li, Z.; Sun, Y.; Thurlow, D.L. RNA minihelices as model substrates for ATP/CTP:tRNA nucleotidyltransferase. *Biochem. J.* **1997**, *327 Pt 3*, 847–851. [CrossRef]
37. Lizano, E.; Scheibe, M.; Rammelt, C.; Betat, H.; Mörl, M. A comparative analysis of CCA-adding enzymes from human and E. coli: Differences in CCA addition and tRNA 3′-end repair. *Biochimie* **2008**, *90*, 762–772. [CrossRef] [PubMed]
38. Wolstenholme, D.R.; Macfarlane, J.L.; Okimoto, R.; Clary, D.O.; Wahleithner, J.A. Bizarre tRNAs inferred from DNA sequences of mitochondrial genomes of nematode worms. *Proc. Natl. Acad. Sci. USA* **1987**, *84*, 1324–1328. [CrossRef]
39. Okimoto, R.; Wolstenholme, D.R. A set of tRNAs that lack either the T psi C arm or the dihydrouridine arm: Towards a minimal tRNA adaptor. *EMBO J.* **1990**, *9*, 3405–3411. [CrossRef]
40. Watanabe, Y.-I.; Suematsu, T.; Ohtsuki, T. Losing the stem-loop structure from metazoan mitochondrial tRNAs and co-evolution of interacting factors. *Front. Genet.* **2014**, *5*, 109. [CrossRef]
41. de Bruijn, M.H.; Schreier, P.H.; Eperon, I.C.; Barrell, B.G.; Chen, E.Y.; Armstrong, P.W.; Wong, J.F.; Roe, B.A. A mammalian mitochondrial serine transfer RNA lacking the "dihydrouridine" loop and stem. *Nucleic Acids Res.* **1980**, *8*, 5213–5222. [CrossRef] [PubMed]
42. Hunter, S.E.; Spremulli, L.L. Interaction of mitochondrial elongation factor Tu with aminoacyl-tRNAs. *Mitochondrion* **2004**, *4*, 21–29. [CrossRef] [PubMed]
43. Watanabe, Y.; Tsurui, H.; Ueda, T.; Furushima, R.; Takamiya, S.; Kita, K.; Nishikawa, K.; Watanabe, K. Primary and higher order structures of nematode (Ascaris suum) mitochondrial tRNAs lacking either the T or D stem. *J. Biol. Chem.* **1994**, *269*, 22902–22906. [PubMed]
44. Klimov, P.B.; Oconnor, B.M. Improved tRNA prediction in the American house dust mite reveals widespread occurrence of extremely short minimal tRNAs in acariform mites. *BMC Genom.* **2009**, *10*, 598. [CrossRef]
45. Jühling, F.; Pütz, J.; Florentz, C.; Stadler, P.F. Armless mitochondrial tRNAs in Enoplea (Nematoda). *RNA Biol.* **2012**, *9*, 1161–1166. [CrossRef] [PubMed]
46. Pons, J.; Bover, P.; Bidegaray-Batista, L.; Arnedo, M.A. Arm-less mitochondrial tRNAs conserved for over 30 millions of years in spiders. *BMC Genom.* **2019**, *20*, 665. [CrossRef] [PubMed]
47. Jühling, T.; Duchardt-Ferner, E.; Bonin, S.; Wöhnert, J.; Pütz, J.; Florentz, C.; Betat, H.; Sauter, C.; Mörl, M. Small but large enough: Structural properties of armless mitochondrial tRNAs from the nematode Romanomermis culicivorax. *Nucleic Acids Res.* **2018**, *46*, 9170–9180. [CrossRef]
48. Wende, S.; Platzer, E.G.; Jühling, F.; Pütz, J.; Florentz, C.; Stadler, P.F.; Mörl, M. Biological evidence for the world's smallest tRNAs. *Biochimie* **2014**, *100*, 151–158. [CrossRef] [PubMed]
49. Wolfe, C.L.; Lou, Y.C.; Hopper, A.K.; Martin, N.C. Interplay of heterogeneous transcriptional start sites and translational selection of AUGs dictate the production of mitochondrial and cytosolic/nuclear tRNA nucleotidyltransferase from the same gene in yeast. *J. Biol. Chem.* **1994**, *269*, 13361–13366.
50. Reichert, A.S.; Thurlow, D.L.; Mörl, M. A eubacterial origin for the human tRNA nucleotidyltransferase? *Biol. Chem.* **2001**, *382*, 1431–1438. [CrossRef]
51. Nagaike, T.; Suzuki, T.; Tomari, Y.; Takemoto-Hori, C.; Negayama, F.; Watanabe, K.; Ueda, T. Identification and characterization of mammalian mitochondrial tRNA nucleotidyltransferases. *J. Biol. Chem.* **2001**, *276*, 40041–40049. [CrossRef] [PubMed]
52. Tomari, Y.; Suzuki, T.; Ueda, T. tRNA recognition by CCA-adding enzyme. *Nucleic Acids Res. Suppl.* **2002**, *2*, 77–78. [CrossRef] [PubMed]
53. Shikha, S.; Schneider, A. The single CCA-adding enzyme of T. brucei has distinct functions in the cytosol and in mitochondria. *J. Biol. Chem.* **2020**, *295*, 6138–6150. [CrossRef] [PubMed]
54. Schiffer, P.H.; Kroiher, M.; Kraus, C.; Koutsovoulos, G.D.; Kumar, S.; Camps, J.I.R.; Nsah, N.A.; Stappert, D.; Morris, K.; Heger, P.; et al. The genome of Romanomermis culicivorax: Revealing fundamental changes in the core developmental genetic toolkit in Nematoda. *BMC Genom.* **2013**, *14*, 923. [CrossRef]
55. Claros, M.G.; Vincens, P. Computational method to predict mitochondrially imported proteins and their targeting sequences. *Eur. J. Biochem.* **1996**, *241*, 779–786. [CrossRef]
56. Augustin, M.A.; Reichert, A.S.; Betat, H.; Huber, R.; Mörl, M.; Steegborn, C. Crystal structure of the human CCA-adding enzyme: Insights into template-independent polymerization. *J. Mol. Biol.* **2003**, *328*, 985–994. [CrossRef]

57. Lizano, E.; Schuster, J.; Müller, M.; Kelso, J.; Mörl, M. A splice variant of the human CCA-adding enzyme with modified activity. *J. Mol. Biol.* **2007**, *366*, 1258–1265. [CrossRef]
58. Ernst, F.G.M.; Erber, L.; Sammler, J.; Jühling, F.; Betat, H.; Mörl, M. Cold adaptation of tRNA nucleotidyltransferases: A tradeoff in activity, stability and fidelity. *RNA Biol.* **2018**, *15*, 144–155. [CrossRef]
59. Okabe, M.; Tomita, K.; Ishitani, R.; Ishii, R.; Takeuchi, N.; Arisaka, F.; Nureki, O.; Yokoyama, S. Divergent evolutions of trinucleotide polymerization revealed by an archaeal CCA-adding enzyme structure. *EMBO J.* **2003**, *22*, 5918–5927. [CrossRef]
60. Erber, L.; Franz, P.; Betat, H.; Prohaska, S.; Mörl, M. Divergent Evolution of Eukaryotic CC- and A-Adding Enzymes. *Int. J. Mol. Sci.* **2020**, *21*, 462. [CrossRef]
61. Oommen, A.; Li, X.Q.; Gegenheimer, P. Cleavage specificity of chloroplast and nuclear tRNA 3′-processing nucleases. *Mol. Cell. Biol.* **1992**, *12*, 865–875. [CrossRef] [PubMed]
62. Loria, A.; Pan, T. The 3′ substrate determinants for the catalytic efficiency of the Bacillus subtilis RNase P holoenzyme suggest autolytic processing of the RNase P RNA in vivo. *RNA* **2000**, *6*, 1413–1422. [CrossRef]
63. Shi, H.; Moore, P.B. The crystal structure of yeast phenylalanine tRNA at 1.93 A resolution: A classic structure revisited. *RNA* **2000**, *6*, 1091–1105. [CrossRef]
64. Byrne, R.T.; Konevega, A.L.; Rodnina, M.V.; Antson, A.A. The crystal structure of unmodified tRNAPhe from Escherichia coli. *Nucleic Acids Res.* **2010**, *38*, 4154–4162. [CrossRef] [PubMed]
65. Wende, S.; Bonin, S.; Götze, O.; Betat, H.; Mörl, M. The identity of the discriminator base has an impact on CCA addition. *Nucleic Acids Res.* **2015**, *43*, 5617–5629. [CrossRef] [PubMed]
66. Just, A.; Butter, F.; Trenkmann, M.; Heitkam, T.; Mörl, M.; Betat, H. A comparative analysis of two conserved motifs in bacterial poly(A) polymerase and CCA-adding enzyme. *Nucleic Acids Res.* **2008**, *36*, 5212–5220. [CrossRef] [PubMed]
67. Betat, H.; Rammelt, C.; Martin, G.; Mörl, M. Exchange of regions between bacterial poly(A) polymerase and the CCA-adding enzyme generates altered specificities. *Mol. Cell* **2004**, *15*, 389–398. [CrossRef]
68. Martin, G.; Keller, W. Sequence motifs that distinguish ATP(CTP): tRNA nucleotidyl transferases from eubacterial poly(A) polymerases. *RNA* **2004**, *10*, 899–906. [CrossRef]
69. Shi, P.Y.; Maizels, N.; Weiner, A.M. CCA addition by tRNA nucleotidyltransferase: Polymerization without translocation? *EMBO J.* **1998**, *17*, 3197–3206. [CrossRef]
70. Kuhn, C.-D.; Wilusz, J.E.; Zheng, Y.; Beal, P.A.; Joshua-Tor, L. On-Enzyme Refolding Permits Small RNA and tRNA Surveillance by the CCA-Adding Enzyme. *Cell* **2015**, *160*, 644–658. [CrossRef]
71. Chen, J.Y.; Joyce, P.B.; Wolfe, C.L.; Steffen, M.C.; Martin, N.C. Cytoplasmic and mitochondrial tRNA nucleotidyltransferase activities are derived from the same gene in the yeast Saccharomyces cerevisiae. *J. Biol. Chem.* **1992**, *267*, 14879–14883. [PubMed]
72. Dupasquier, M.; Kim, S.; Halkidis, K.; Gamper, H.; Hou, Y.-M. tRNA integrity is a prerequisite for rapid CCA addition: Implication for quality control. *J. Mol. Biol* **2008**, *379*, 579–588. [CrossRef] [PubMed]
73. Hou, Y.-M. CCA addition to tRNA: Implications for tRNA quality control. *IUBMB life* **2010**, *62*, 251–260. [CrossRef] [PubMed]
74. Wellner, K.; Betat, H.; Mörl, M. A tRNA's fate is decided at its 3′ end: Collaborative actions of CCA-adding enzyme and RNases involved in tRNA processing and degradation. *Biochim. Biophys. Acta Gene Regul. Mech.* **2018**, *1861*, 433–441. [CrossRef] [PubMed]
75. Wellner, K.; Czech, A.; Ignatova, Z.; Betat, H.; Mörl, M. Examining tRNA 3′-ends in Escherichia coli: Teamwork between CCA-adding enzyme, RNase T, and RNase R. *RNA* **2018**, *24*, 361–370. [CrossRef] [PubMed]
76. Wilusz, J.E.; Whipple, J.M.; Phizicky, E.M.; Sharp, P.A. tRNAs marked with CCACCA are targeted for degradation. *Science* **2011**, *334*, 817–821. [CrossRef] [PubMed]
77. Levinger, L.; Bourne, R.; Kolla, S.; Cylin, E.; Russell, K.; Wang, X.; Mohan, A. Matrices of paired substitutions show the effects of tRNA D/T loop sequence on Drosophila RNase P and 3′-tRNase processing. *J. Biol. Chem.* **1998**, *273*, 1015–1025. [CrossRef]
78. Zahler, N.H.; Christian, E.L.; Harris, M.E. Recognition of the 5′ leader of pre-tRNA substrates by the active site of ribonuclease P. *RNA* **2003**, *9*, 734–745. [CrossRef]
79. Kirsebom, L.A. RNase P RNA mediated cleavage: Substrate recognition and catalysis. *Biochimie* **2007**, *89*, 1183–1194. [CrossRef]

80. Levinger, L.; Hopkinson, A.; Desetty, R.; Wilson, C. Effect of changes in the flexible arm on tRNase Z processing kinetics. *J. Biol. Chem.* **2009**, *284*, 15685–15691. [CrossRef]
81. McKenney, K.M.; Rubio, M.A.T.; Alfonzo, J.D. The Evolution of Substrate Specificity by tRNA Modification Enzymes. *Enzymes* **2017**, *41*, 51–88. [CrossRef] [PubMed]
82. Spacciapoli, P.; Doviken, L.; Mulero, J.J.; Thurlow, D.L. Recognition of tRNA by the enzyme ATP/CTP:tRNA nucleotidyltransferase. Interference by nucleotides modified with diethyl pyrocarbonate or hydrazine. *J. Biol. Chem.* **1989**, *264*, 3799–3805. [PubMed]
83. Spacciapoli, P.; Thurlow, D.L. Purines in tRNAs required for recognition by ATP/CTP:tRNA nucleotidyltransferase from rabbit liver. *J. Mol. Recognit.* **1990**, *3*, 149–155. [CrossRef] [PubMed]
84. Hegg, L.A.; Thurlow, D.L. Cytidines in tRNAs that are required intact by ATP/CTP:tRNA nucleotidyltransferases from Escherichia coli and Saccharomyces cerevisiae. *Nucleic Acids Res.* **1990**, *18*, 5975–5979. [CrossRef] [PubMed]
85. Kim, S.; Liu, C.; Halkidis, K.; Gamper, H.B.; Hou, Y.-M. Distinct kinetic determinants for the stepwise CCA addition to tRNA. *RNA* **2009**, *15*, 1827–1836. [CrossRef] [PubMed]
86. Xiong, Y.; Steitz, T.A. Mechanism of transfer RNA maturation by CCA-adding enzyme without using an oligonucleotide template. *Nature* **2004**, *430*, 640–645. [CrossRef] [PubMed]
87. Kries, H.; Niquille, D.L.; Hilvert, D. A subdomain swap strategy for reengineering nonribosomal peptides. *Chem. Biol.* **2015**, *22*, 640–648. [CrossRef] [PubMed]
88. Ostermeier, M.; Benkovic, S.J. Evolution of protein function by Domain swapping. In *Evolutionary Protein Design*; Arnold, F.H., Ed.; Academic Press: San Diego, CA, USA, 2001; pp. 29–77. ISBN 9780120342556.
89. Truman, A.W.; Dias, M.V.B.; Wu, S.; Blundell, T.L.; Huang, F.; Spencer, J.B. Chimeric glycosyltransferases for the generation of hybrid glycopeptides. *Chem. Biol.* **2009**, *16*, 676–685. [CrossRef]
90. Tzagoloff, A.; Gatti, D.; Gampel, A. Mitochondrial aminoacyl-tRNA synthetases. *Prog. Nucleic Acid Res. Mol. Biol* **1990**, *39*, 129–158.
91. Salinas-Giegé, T.; Giegé, R.; Giegé, P. tRNA biology in mitochondria. *Int. J. Mol. Sci.* **2015**, *16*, 4518–4559. [CrossRef]
92. Greber, B.J.; Bieri, P.; Leibundgut, M.; Leitner, A.; Aebersold, R.; Boehringer, D.; Ban, N. Ribosome. The complete structure of the 55S mammalian mitochondrial ribosome. *Science* **2015**, *348*, 303–308. [CrossRef] [PubMed]
93. Koc, E.C.; Burkhart, W.; Blackburn, K.; Moyer, M.B.; Schlatzer, D.M.; Moseley, A.; Spremulli, L.L. The large subunit of the mammalian mitochondrial ribosome. Analysis of the complement of ribosomal proteins present. *J. Biol. Chem.* **2001**, *276*, 43958–43969. [CrossRef] [PubMed]
94. Sharma, M.R.; Koc, E.C.; Datta, P.P.; Booth, T.M.; Spremulli, L.L.; Agrawal, R.K. Structure of the Mammalian Mitochondrial Ribosome Reveals an Expanded Functional Role for Its Component Proteins. *Cell* **2003**, *115*, 97–108. [CrossRef]
95. Suzuki, T.; Terasaki, M.; Takemoto-Hori, C.; Hanada, T.; Ueda, T.; Wada, A.; Watanabe, K. Structural compensation for the deficit of rRNA with proteins in the mammalian mitochondrial ribosome. Systematic analysis of protein components of the large ribosomal subunit from mammalian mitochondria. *J. Biol. Chem.* **2001**, *276*, 21724–21736. [CrossRef]
96. Okimoto, R.; Macfarlane, J.L.; Wolstenholme, D.R. The mitochondrial ribosomal RNA genes of the nematodes Caenorhabditis elegans and Ascaris suum: Consensus secondary-structure models and conserved nucleotide sets for phylogenetic analysis. *J. Mol. Evol.* **1994**, *39*, 598–613. [CrossRef]
97. Zhao, F.; Ohtsuki, T.; Yamada, K.; Yoshinari, S.; Kita, K.; Watanabe, Y.-I.; Watanabe, K. Isolation and physiochemical properties of protein-rich nematode mitochondrial ribosomes. *Biochemistry* **2005**, *44*, 9232–9237. [CrossRef]
98. Mears, J.A.; Cannone, J.J.; Stagg, S.M.; Gutell, R.R.; Agrawal, R.K.; Harvey, S.C. Modeling a Minimal Ribosome Based on Comparative Sequence Analysis. *J. Mol. Biol* **2002**, *321*, 215–234. [CrossRef]
99. Todd, A.E.; Orengo, C.A.; Thornton, J.M. Plasticity of enzyme active sites. *Trends Biochem. Sci.* **2002**, *27*, 419–426. [CrossRef]
100. Waterhouse, A.M.; Procter, J.B.; Martin, D.M.A.; Clamp, M.; Barton, G.J. Jalview Version 2—A multiple sequence alignment editor and analysis workbench. *Bioinformatics* **2009**, *25*, 1189–1191. [CrossRef]
101. Bradford, M.M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **1976**, *72*, 248–254. [CrossRef]

102. Schürer, H.; Lang, K.; Schuster, J.; Mörl, M. A universal method to produce in vitro transcripts with homogeneous 3′ ends. *Nucleic Acids Res.* **2002**, *30*, e56. [CrossRef] [PubMed]
103. Scheibe, M.; Bonin, S.; Hajnsdorf, E.; Betat, H.; Mörl, M. Hfq stimulates the activity of the CCA-adding enzyme. *BMC Mol. Biol.* **2007**, *8*, 92. [CrossRef] [PubMed]
104. Webb, B.; Sali, A. Comparative Protein Structure Modeling Using MODELLER. *Curr. Protoc. Bioinform.* **2016**, *54*, 5.6.1–5.6.37. [CrossRef] [PubMed]

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).

V. Monitoring the Production of High Diffraction-Quality Crystals of Two Enzymes in Real Time Using In Situ Dynamic Light Scattering

A crystallographic investigation always starts with the determination of crystallization conditions. Once identified, these do not necessarily lead to perfect crystals diffracting to high resolution. We have previously shown that plethora of methods exist to fine-tune the conditions leading to the best crystal quality. However, taken together, the full refinement process is not only time-consuming but also based on “semi-blind” strategy. Indeed, conditions (e.g., crystallant concentration, pH, temperature) are set at the beginning of the experiment, then users must wait for spontaneous crystal formation and growth without having full access or control on early crystallization events. In this regard, the company Xtal Concepts (Hamburg, Germany), in collaboration with the teams from Pr. Dr. Christian Betzel and Pr. Dr. Rolf Hilgenfeld, developed the Xtal Controller XC900 (Figure I.9) to monitor and control the crystallization process in a single experiment using dynamic light scattering (DLS) and videomicroscopy (Meyer *et al.*, 2012; Baitan *et al.*, 2018). Its design described in part I.2.3 and Figure I.9, allows the control over a single drop (from 5 to 10 µL) placed on a glass coverslip atop of an ultrasensitive balance. On each side, two piezoelectric pumps are dispensing 70 pL droplets of either crystallant solution to trigger nucleation/growth or water to counter-balance evaporation in the crystallization drop. In order to visually track nascent crystals, a microscope coupled with a CCD camera is inserted on the top lid of the machine. Last but not least, the DLS module gives access in real time to early nuclei formation and particle size increase toward crystalline objects. While regular optical microscope serves to identify crystalline objects with usually a minimal size of 5µm, DLS informs on early stages of crystallization, said the nucleation step. Assisted by a Python based software, user have the possibility to record DLS profiles, by plotting particle radius against time, take pictures and modulate physicochemical parameters (i.e., solution concentration, temperature, humidity) all along the experiment.

The present study was published in Crystals journal (MDPI). It exposes recent results obtained using the XC900 to study the crystallization of the psychrophilic CCA-adding enzyme from *Planococcus halocryophilus* (PhaCCA) and the hen white egg lysozyme combined with Tb-Xo4 nucleant. In both cases we could precisely determine the nucleation point of macromolecules following a progressive size increase in DLS profiles. Regarding PhaCCA, a succession of crystallization/dissolution cycles was applied to the system, leading to improved diffraction power of final crystals. In parallel, we observed a tendency to form a reversible crystalline precipitate prior to visible crystal formation. Taking advantage of this

versatile nature, we established a reproducible workflow to obtain nano to microcrystalline solution of the enzyme withing few hours.

In the frame of an active collaboration with Polyvalan company (SAS, Lyon, France) and one of its founders, Dr Eric Girard (IBS, Grenoble), we also demonstrated the immediate nucleating effect of Tb-Xo4 compound on lysozyme. The enzyme was first resuspended in either pure water or in 10 mM sodium acetate pH 4.5, 40 mM NaCl, and then deposited inside the instrument. Prior to Tb-Xo4 injection, a phase during which the protein concentration was maintained constant under regular DLS measurements showed a monodisperse sample. Rapidly after nucleant injection (from 15 to 30 minutes), the size distribution of particles, initially around few nanometers, jumps to 100 up to 1000 nm. No visible object could be directly observed. However, after few weeks of incubation outside of the machine, crystals in the range of several hundreds of micrometers were obtained and the enzyme structure including the Tb-Xo4 could be solved.

On the one hand, we demonstrate here the power of the Xtal Controller which provides real-time monitoring of early nucleation events of a given macromolecule. On the other hand, we highlight its role as a joystick to navigate through the phase diagram of an enzyme. Regarding the increased need for reproducible and high-quality sample suitable for serial time-resolved crystallography at either XFEL (Kupitz *et al.*, 2014, 2017) or synchrotron light source (Pearson & Mehrabi, 2020), we can easily imagine how useful this instrument will be in the future development of the field of TR-SFX/SSX.

Une étude cristallographique commence toujours par la détermination des conditions de cristallisation. Une fois identifiées, celles-ci ne conduisent pas nécessairement à des cristaux parfaits diffractant à haute résolution. Nous avons montré précédemment qu'il existe une pléthore de méthodes permettant d'affiner les conditions conduisant à la meilleure qualité de cristal. Cependant, pris dans son ensemble, le processus d'affinement complet est non seulement long mais aussi basé sur une stratégie "semi-aveugle". En effet, les conditions (par exemple, la concentration en agent crystallisant, le pH, la température) sont fixées au début de l'expérience, puis les utilisateurs doivent attendre la formation et la croissance spontanées des cristaux sans avoir un accès ou un contrôle total sur les premiers événements de cristallisation. À cet égard, la société Xtal Concepts (Hambourg, Allemagne), en collaboration avec les équipes du Pr. Dr Christian Betzel et du Pr. Dr. Rolf Hilgenfeld, a développé le Xtal Controller XC900 (Figure 1-9) pour surveiller et contrôler le processus de cristallisation dans une seule expérience en utilisant la diffusion dynamique de la lumière (DLS) et la vidéomicroscopie (Meyer et al. , 2012 ; Baitan et al. , 2018). Sa conception décrite dans la partie I.2.3 et la figure 1-9, permet le contrôle d'une seule goutte (de 5 à 10 µL) placée sur une lamelle en verre au sommet d'une balance ultrasensible. De chaque côté, deux pompes piézoélectriques distribuent des gouttelettes de 70 pL soit de solution de crystallization pour déclencher la nucléation/croissance, soit d'eau pour contrebalancer l'évaporation dans la goutte de cristallisation. Afin de suivre visuellement les cristaux naissants, un microscope couplé à une caméra CCD est inséré sur le couvercle supérieur de la machine. Enfin, le module DLS donne accès en temps réel à la formation précoce de noyaux et à l'augmentation de la taille des particules vers les objets cristallins. Alors que le microscope optique ordinaire sert à identifier les objets cristallins dont la taille minimale est généralement de 5 µm, le module DLS fournit des informations sur les premiers stades de la cristallisation, notamment l'étape de nucléation. Grâce à un logiciel encodé en Python, l'utilisateur a la possibilité d'enregistrer des profils de DLS, en traçant le rayon des particules en fonction du temps, de prendre des photos et de moduler les paramètres physico-chimiques (c'est-à-dire la concentration de la solution, la température, l'humidité) tout au long de l'expérience.

La présente étude a été publiée dans la revue Crystals (MDPI). Elle expose les résultats récents obtenus en utilisant le XC900 pour étudier la cristallisation de l'enzyme d'addition du CCA sur les tRNAs de Planococcus halocryophilus (PhaCCA) et du lysozyme

d'œuf blanc de poule combiné avec le nucléant Tb-Xo4. Dans les deux cas, nous avons pu déterminer avec précision le point de nucléation des macromolécules en suivant une augmentatio progressive de la taille dans les profils de DLS. Concernant le PhaCCA, une succession de cycles de cristallisation/dissolution a été appliquée au système, conduisant à une amélioration du pouvoir de diffraction des cristaux finaux. En parallèle, nous avons observé une tendance à former un précipité cristallin réversible avant la formation de cristaux visibles. En tirant profit de cette nature versatile, nous avons établi une procédure reproductible pour obtenir une solution nano à microcristalline de l'enzyme en quelques heures.

Dans le cadre d'une collaboration active avec la société Polyvalan (SAS, Lyon, France) et l'un de ses fondateurs, le Dr Eric Girard (IBS, Grenoble), nous avons également démontré l'effet nucléant immédiat du composé Tb-Xo4 sur le lysozyme. L'enzyme a d'abord été resuspendue soit dans de l'eau pure, soit dans un tampon d'acétate de sodium 10 mM pH 4.5, 40 mM NaCl, puis déposé à l'intérieur de l'instrument. Avant l'injection de Tb-Xo4, une phase pendant laquelle la concentration de la protéine était maintenue constante sous des mesures régulières de DLS a montré un échantillon monodisperse. Rapidement après l'injection du nucléant (de 15 à 30 minutes), la distribution de taille des particules, initialement de quelques nanomètres, passe à 100 jusqu'à 1000 nm. Aucun objet visible n'a pu être observé directement. Cependant, après quelques semaines d'incubation à l'extérieur de la machine, des cristaux de l'ordre de plusieurs centaines de micromètres ont été obtenus et la structure de l'enzyme incluant le Tb-Xo4 a pu être résolue.

D'une part, nous démontrons ici la puissance du Xtal controller qui permet de suivre en temps réel les événements de nucléation précoce d'une macromolécule donnée. D'autre part, nous soulignons son rôle de « joystick » pour naviguer dans le diagramme de phase d'une enzyme. Compte tenu de la nécessité croissante de disposer d'échantillons reproductibles et de haute qualité pour la cristallographie sérielle résolue en temps, que ce soit au XFEL (Kupitz et al., 2014, 2017) ou pour une source de lumière synchrotron (Pearson & Mehrabi, 2020), nous pouvons facilement imaginer l'utilité de cet instrument pour de futurs développements du domaine du TR-SFX/SSX.

Eine kristallografische Untersuchung beginnt immer mit der Bestimmung der Kristallisationsbedingungen. Sind diese erst einmal festgelegt, führen sie nicht unbedingt zu perfekten Kristallen mit hoher Auflösung. Wir haben bereits gezeigt, dass es eine Fülle von Methoden zur Feinabstimmung der Bedingungen gibt, die zur besten Kristallqualität führen. Zusammengenommen ist der gesamte Verfeinerungsprozess jedoch nicht nur zeitaufwändig, sondern basiert auch auf einer "halbblinden" Strategie. Die Bedingungen (z. B. Kristallkonzentration, pH-Wert, Temperatur) werden nämlich zu Beginn des Experiments festgelegt, dann müssen die Forscher auf die spontane Kristallbildung und das Wachstum warten, ohne vollen Zugang oder Kontrolle über die frühen Kristallisationsereignisse zu haben. In diesem Zusammenhang hat die Firma Xtal Concepts (Hamburg, Deutschland) in Zusammenarbeit mit den Teams von Prof. Dr. Christian Betzel und Prof. Dr. Rolf Hilgenfeld den Xtal Controller XC900 (Abbildung I. 9) zur Überwachung und Steuerung des Kristallisationsprozesses in einem einzigen Experiment mittels dynamischer Lichtstreuung (DLS) und Videomikroskopie entwickelt (Meyer et al. , 2012; Baitan et al. , 2018). Das in Teil I.2.3 und Abbildung I. 9 beschriebene Design ermöglicht die Kontrolle eines einzelnen Tropfens (5 bis 10 µl), der auf einem Glasdeckglas über einer ultrasensiblen Waage platziert wird. Auf jeder Seite geben zwei piezoelektrische Pumpen 70 pl Tröpfchen entweder mit Kristallisationslösung zum Auslösen der Keimbildung/des Wachstums oder mit Wasser zum Ausgleich der Verdampfung im Kristallisationstropfen ab. Um die entstehenden Kristalle visuell zu verfolgen, ist ein Mikroskop mit einer CCD-Kamera auf dem oberen Deckel der Maschine angebracht. Nicht zuletzt gibt das DLS-Modul in Echtzeit Aufschluss über die frühe Keimbildung und die Zunahme der Partikelgröße zu kristallinen Objekten. Während ein normales optisches Mikroskop dazu dient, kristalline Objekte mit einer Mindestgröße von 5 µm zu identifizieren, gibt das DLS-Modul Aufschluss über die frühen Stadien der Kristallisation, d. h. die Keimbildung. Mit Hilfe einer auf Python basierenden Software kann der Benutzer DLS-Profile aufzeichnen, indem er den Partikelradius gegen die Zeit aufträgt, Bilder macht und physikalisch-chemische Parameter (z.B. Lösungskonzentration, Temperatur, Feuchtigkeit) während des gesamten Experiments moduliert.

Die vorliegende Studie wurde in der Zeitschrift Crystals (MDPI) veröffentlicht. Sie stellt aktuelle Ergebnisse vor, die mit dem XC900 bei der Untersuchung der Kristallisation des psychrophilen CCA-addierenden Enzyms aus Planococcus halocryophilus (PhaCCA) und

des Hühnerei-Lysozyms in Kombination mit dem Nukleierungsmittel Tb-Xo4 erzielt wurden. In beiden Fällen konnten wir den Nukleationspunkt der Makromoleküle aufgrund einer progressiven Größenzunahme in den DLS-Profilen genau bestimmen. Bei PhaCCA wurde das System einer Reihe von Kristallisations-/Auflösungszyklen unterzogen, was zu einer verbesserten Beugungsleistung der endgültigen Kristalle führte. Gleichzeitig beobachteten wir eine Tendenz zur Bildung eines reversiblen kristallinen Präzipitats vor der sichtbaren Kristallbildung. Unter Ausnutzung dieser vielseitigen Eigenschaften haben wir einen reproduzierbaren Arbeitsablauf entwickelt, um innerhalb weniger Stunden eine nano- bis mikrokristalline Lösung des Enzyms zu erhalten.

Im Rahmen einer aktiven Zusammenarbeit mit der Firma Polyvalan (SAS, Lyon, Frankreich) und einem ihrer Gründer, Dr. Eric Girard (IBS, Grenoble), haben wir auch die sofortige keimbildende Wirkung der Tb-Xo4-Verbindung auf Lysozym nachgewiesen. Das Enzym wurde zunächst entweder in reinem Wasser oder in 10 mM Natriumacetat pH 4,5, 40 mM NaCl resuspendiert und dann im Gerät deponiert. Vor der Tb-Xo4-Injektion zeigte eine Phase, in der die Proteinkonzentration unter regelmäßigen DLS-Messungen konstant gehalten wurde, eine monodisperse Probe. Schnell nach der Injektion des Nukleierungsmittels (15 bis 30 Minuten) springt die Größenverteilung der Partikel, die zunächst bei wenigen Nanometern lag, auf 100 bis 1000 nm. Es konnte kein sichtbares Objekt direkt beobachtet werden. Nach einigen Wochen der Inkubation außerhalb der Maschine wurden jedoch Kristalle im Bereich von mehreren hundert Mikrometern erhalten, und die Enzymstruktur einschließlich Tb-Xo4 konnte gelöst werden.

Einerseits demonstrieren wir hier die Leistungsfähigkeit des Xtal Controllers, der eine Echtzeitüberwachung der frühen Nukleationsereignisse eines bestimmten Makromoleküls ermöglicht. Andererseits betonen wir seine Rolle als Joystick, um durch das Phasendiagramm eines Enzyms zu navigieren. Angesichts des steigenden Bedarfs an reproduzierbaren und qualitativ hochwertigen Proben, die für die serielle zeitaufgelöste Kristallographie am XFEL (Kupitz et al., 2014, 2017) oder an Synchrotronlichtquellen (Pearson & Mehrabi, 2020) geeignet sind, können wir uns leicht vorstellen, wie nützlich dieses Instrument für die zukünftige Entwicklung des Bereichs TR-SFX/SSX sein wird.

crystals

MDPI

Article

Monitoring the Production of High Diffraction-Quality Crystals of Two Enzymes in Real Time Using In Situ Dynamic Light Scattering

Raphaël de Wijn [1,†], Kévin Rollet [1,2,†], Sylvain Engilberge [3], Alastair G. McEwen [4], Oliver Hennig [2], Heike Betat [2], Mario Mörl [2], François Riobé [5], Olivier Maury [5], Eric Girard [3], Philippe Bénas [1], Bernard Lorber [1] and Claude Sauter [1,*]

1 Institut de Biologie Moléculaire et Cellulaire, CNRS, UPR 9002, Architecture et Réactivité de l'ARN, Université de Strasbourg, F-67084 Strasbourg, France; raphael.de.wijn@xfel.eu (R.d.W.); k.rollet@ibmc-cnrs.unistra.fr (K.R.); p.benas@ibmc-cnrs.unistra.fr (P.B.); b.lorber@ibmc-cnrs.unistra.fr (B.L.)
2 Institute for Biochemistry, Leipzig University, D-04103 Leipzig, German; oliver.hennig@uni-leipzig.de (O.H.); heike.betat@uni-leipzig.de (H.B.); mario.moerl@uni-leipzig.de (M.M.)
3 Institut de Biologie Structurale, CNRS, CEA, Université Grenoble Alpes, F-38044 Grenoble, France; sylvain.engilberge@psi.ch (S.E.); eric.girard@ibs.fr (E.G.)
4 Plateforme de Biologie et Génomique Structurale, Institut de Génétique et de Biologie Moléculaire et Cellulaire, Université de Strasbourg, CNRS UMR 7104, INSERM U 1258, F-67404 Illkirch, France; alastair@igbmc.fr
5 Univ Lyon, ENS Lyon, CNRS-UMR 5182, Université Lyon 1, F-69342 Lyon, France; francois.riobe@ens-lyon.fr (F.R.); olivier.maury@ens-lyon.fr (O.M.)
* Correspondence: c.sauter@ibmc-cnrs.unistra.fr; Tel.: +33-388-417-102
† These authors contributed equally to the work.

Received: 25 December 2019; Accepted: 14 January 2020; Published: 23 January 2020

Abstract: The reproducible preparation of well-diffracting crystals is a prerequisite for every structural study based on crystallography. An instrument called XtalController has recently been designed that allows the monitoring of crystallization assays using dynamic light scattering and microscopy, and integrates piezo pumps to alter the composition of the mother liquor during the experiment. We have applied this technology to study the crystallization of two enzymes, the CCA-adding enzyme of the psychrophilic bacterium *Planococcus halocryophilus*, and the lysozyme from hen egg white in the presence of a synthetic chemical nucleant. We were able to (i) detect early nucleation events and (ii) drive the crystallization system (through cycles of dissolution/crystallization) toward growth conditions yielding crystals with excellent diffraction properties. This technology opens a way to the rational production of samples for crystallography, ranging from nanocrystals for electron diffraction, microcrystals for serial or conventional X-ray diffraction, to larger crystals for neutron diffraction.

Keywords: enzyme; crystallization; dynamic light scattering; nucleation; nucleant; Tb-Xo4 crystallophore; microcrystals; nanocrystals; X-ray diffraction; XtalController

1. Introduction

Since its birth in the 1960s, biocrystallography has been a primary source of structural information, contributing more than 90% of 3D structures accessible in the Protein Data Bank [1] and remains a central player in structural biology, alongside NMR and CryoEM. Over the past decade, new experimental setups have been introduced that widen its applicability and transform the daily practice of crystal growers and crystallographers. The recent advent of X-ray free electron lasers (X-FEL) has enabled the serial femtosecond diffraction (SFX) analysis of micro- and nano-crystals, and offers unprecedented

possibilities for time-resolved experiments [2–4]. At the same time, electron microscopes have been hijacked to perform micro-electron diffraction (μED), opening the way to the characterization of nanocrystals using laboratory-based instruments [5–8]. Though, like conventional ones relying on synchrotron or neutron sources, these new crystallographic approaches require crystalline material and call for the development of means facilitating the production of calibrated samples (i.e. nano-, micro-, or macrocrystals) with a size adapted to the radiation (electrons, X-rays, or neutrons) and the experimental setup.

Growing crystals of a new biomolecule (protein, DNA, RNA, and their complexes) is often a time-consuming task that involves a trial-and-error screening step to find solvent conditions generating promising crystalline or microcrystalline phases. It is followed by an optimization step to improve the quality of one or several crystalline forms and make them suitable for diffraction analysis [9]. However, before the first diffraction test, the evaluation of this two-step process mainly relies on optical microscopy observations. As a consequence, early crystal growth events, including nucleation, nano-crystal, or nano-cluster formation that directly impact the final crystallization outcome, remain hidden to the crystal grower. For this reason there is a clear need for a system enabling the preparation and the optimization of crystals under well-defined and controlled conditions, and ensuring reproducible crystalline properties and quality. The concept of such a system emerged in the nineties in the frame of a European research consortium on crystal growth (European Bio-crystallogenesis Initiative, 1998–2000) and was developed in the context of the OptiCryst European consortium (2006–2010) [10]. The current implementation called XtalController (or XC900; Xtal Concepts GmbH, Hamburg, Germany) is composed of a crystallization chamber for a single experiment with precise temperature and humidity control [11]. The composition of an initial drop (volume 5–10 μl) of a solution of the target biomolecule can be modified by the injection of various solutions (such as water, buffer or crystallant) using two piezo injectors spraying 70 picoliter droplets (Figure 1A,A'). The sample drop can also be concentrated by evaporation of volatile chemicals (generally water) and its composition (i.e., the concentrations of components) is continuously calculated from its weight recorded to ±1 μg by an ultra-sensitive balance. The instrument provides diagnostic means to track the drop content along the experiment; the early occurrence of association events leading to nucleation, as well as nanocrystals can be detected in real time by dynamic light scattering (DLS) and the growth of crystals by video microscopy as soon as they reach a size exceeding a few microns. The XtalController also provides means to navigate in the phase diagram, from an undersaturated solution to a supersaturated state leading to crystal growth or precipitation. Hence, a specific phase may be stabilized or the system may be driven in the phase diagram toward another phase by varying experimental conditions in real time using the piezo injectors.

For more than three decades DLS has proven to be instrumental to study nucleation, to perform quality control of biological samples, to predict the propensity of the latter to crystallize, and, more recently, to follow their behavior in crystallization assays [12–16]. With the unique and innovative combination of piezo injectors to modify the experimental conditions and DLS to track in real time the effect of various physicochemical parameters (chemical composition, biomolecule concentration, temperature) on biomolecules in solution, the XtalController opens a wealth of possibilities for basic and applied crystallogenesis. First examples included the observation of liquid dense clusters formed during nucleation [17] and the preparation of crystals with well-defined size [18].

Here we used this technology to study the crystallization of two enzymes, the CCA-adding enzyme from the psychrophilic bacterium *Planococcus halocryophilus* (PhaCCA) and the hen egg-white lysozyme (HEWL). In the first case, classical vapor-diffusion assays produced numerous small crystals or precipitates. The XtalController helped better define the appropriate crystallant concentration to nucleate and grow large crystals of PhaCCA. In the second case, we used the XtalController to highlight the nucleating effect of a lanthanide complex, the crystallophore Tb-Xo4 [19,20] on HEWL in the absence of crystallant. Both examples illustrate the potential of this technology in

crystallogenesis and for the design of protocols to produce calibrated crystalline samples for a variety of crystallographic applications.

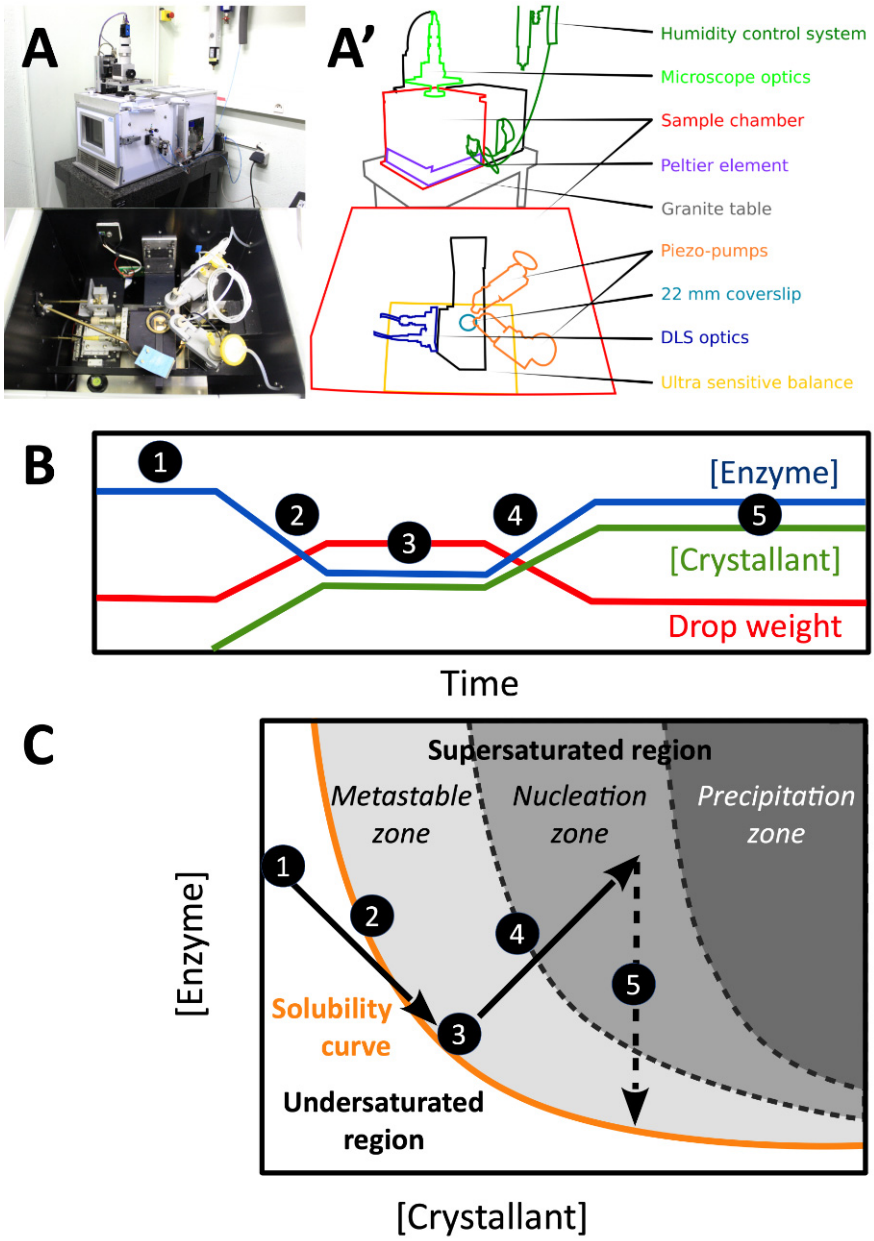


Figure 1. XtalController setup. (**A**) Side and top pictures of the instrument and (**A'**) schematic view highlighting its major components. (**B**) Example of experimental schedule with curves indicating the variation of drop weight (red curve) and associated variations of enzyme and crystallant concentrations (blue and green curves, respectively). The same color code is used in all figures. (**C**) Corresponding trajectory in a theoretical phase diagram. Step 1: incubation of the enzyme solution at constant concentration. Step 2: addition of crystallant (increase of crystallant concentration, decrease of biomolecule concentration in the drop). Step 3: incubation, constant drop in weight (compensation of evaporation by water injection) to keep biomolecule and crystallant concentrations constant. Step 4: controlled evaporation of the drop leading to increased concentrations of biomolecule and crystallant to reach the nucleation zone. Step 5: incubation until crystals start to grow, consume part of the soluble enzyme stock and bring the system back to equilibrium on the solubility curve.

2. Materials and Methods

2.1. Chemicals and Enzyme Samples

Chemicals used for the preparation of buffers and crystallization solutions were of highest purity grade. Solutions were filtered on 0.22 µm porosity Ultrafree-MC membranes (Millipore, Molsheim, France). PhaCCA (monomer of 420 amino acids, 48 kDa) was produced in *Escherichia coli* cells, purified, concentrated to 5 mg/mL, and stored at 4 °C in 50 mM HEPES-NaOH pH 7.5, 100 mM NaCl as described previously [21]. HEWL (monomer of 129 amino acids, 14 kDa) was purchased from Seikagaku Corp. (Tokyo, Japan, Cat. N° 100940), Roche (Basel, Switzerland, Cat. N° 10153516103) and Sigma-Aldrich (St. Louis, MO, USA, Fluka Cat. N° 62970-5G-F). It was used without any further treatment and dissolved in milliQ water (Roche) or in 10 mM sodium acetate pH 4.5, 40 mM NaCl (Seikagaku and Sigma) at concentrations ranging from 25 to 71 mg/mL. Lysozyme stock solutions were filtered prior to concentration measurement. The crystallophore Tb-Xo4 used as nucleant for HEWL was synthesized and purified as described [19]. It was dissolved in water to prepare a 100 mM stock solution.

2.2. Crystallization in the XtalController

The humidity and temperature of the crystallization chamber of the instrument were set to 99.5% and 20.0 °C, respectively, one hour before starting an experiment to ensure the stability of experimental conditions. A drop of 10 µl of enzyme stock solution was deposited on a siliconized glass coverslip (Ø 22 mm) placed on the balance. One pump was loaded with the appropriate crystallant solution and the second with pure water. The crystallization chamber was closed and the protocol started with a 5 min step to monitor drop evaporation and compensate the loss of weight by injection of water to keep the drop weight constant. Several steps of crystallant addition, drop evaporation or dilution were scheduled to explore the phase diagram. The shooting frequency of the pumps was adjusted to vary the slope of concentration variations. DLS measurements and drop image capture were scheduled at regular time intervals (e.g., every 5 to 30 min) to follow nucleation, aggregation, or crystal growth events. At the end of the experiment, the coverslip was transferred onto a 24-well Linbro plate (HR3-110, Hampton Research, Aliso Viejo, CA, USA) for incubation.

2.3. Standard DLS Measurements

In parallel to DLS measurements performed inside the XtalController, a benchtop Nanostar light scattering instrument (Wyatt Technology, Inc., Santa Barbara, CA, USA) was used to record the effect of Tb-Xo4 on lysozyme. Total of 10 µl of Sigma lysozyme (71 mg/mL) were transferred into a quartz cell for DLS measurements at 20 °C. The drop was covered with 10 µl paraffin oil and the cuvette was sealed with Parafilm™ foil to prevent evaporation. Subsequently, 1 µl of a 100 mM Tb-Xo4 stock solution in 10 mM sodium acetate pH 4.5 was added. In control experiments the Tb-Xo4 solution was replaced by buffer. Data were corrected for solvent viscosity and refractive index.

2.4. Crystal Analysis

Crystals of PhaCCA were analyzed in cryogenic conditions at FIP/BM30A beamline (ESRF, Grenoble, France) using an ADSC Quantum 315r detector. A crystal was soaked for a few seconds in the mother liquor supplemented with 20% (w/v) glycerol, mounted in a cryoloop and flash-frozen in liquid nitrogen. 240 images were collected with a rotation of 0.5° and an exposure time of 1 s per frame, yielding a dataset at a resolution of 2.28 Å. Crystals of HEWL were analyzed at ambient temperature using a Rigaku FR-X diffractometer at the FRISBI platform (IGBMC, Illkirch, France) with an EIGER R 4M detector (DECTRIS) with a 2θ offset of 10°. Several crystals grown with Roche, Sigma, and Seikagaku lysozymes were tested and diffracted to up to 1.5 Å resolution but showed rapid decay because of radiation damage at ambient temperature. The exposure time and rotation speed were adapted accordingly to collect a full dataset from a crystal of Seikagaku lysozyme plunged in viscous Parabar 10312 (Hampton Research), mounted in a cryoloop and protected from dehydration using the

MicroRT room temperature kit from MiTeGen. 720 images were collected with a rotation of 0.25° and an exposure time of 2 s per frame. Data were processed with the XDS package [22].

2.5. Structure Determination

The structures of PhaCCA and HEWL were refined in PHENIX [23] using PDB entries 6QY6 and 6F2I (cleared of solvent molecules and ligands), respectively, as starting models for initial rigid body adjustment. Several rounds of refinement and manual inspection in COOT [24] were performed until convergence of R_{free} (calculated using 5% of reflections). The model of PhaCCA was refined using one TLS group and includes 365 residues (the N-terminal expression tag and the flexible loop encompassing residues 83–94 are not visible), two phosphate ions, two acetate ions, and three glycerol molecules present in the mother liquor and the cryoprotection solution. The model of HEWL was refined using anisotropic atomic displacement parameters (ADPs) and includes the full enzyme sequence, one sodium, two chloride ions, a full Tb-Xo4 complex and two additional Tb^{3+} positions clearly identified in the anomalous density map but for which the ligand was not observed because of low occupancy.

3. Results and Discussion

3.1. Triggering the Growth of Large Crystals of PhaCCA

We recently reported the crystallization of a new tRNA nucleotidyltransferase or CCA-adding enzyme from the cold-adapted bacterium *P. halocryophilus* living in the permafrost [21]. During the process of optimization we observed that the protein had a tendency to produce either clear drops or drops full of small crystals or light precipitate, indicating a narrow nucleation zone. To increase the reproducibility of crystallization, microseeding was systematically used, both in vapor-diffusion or in counter-diffusion assays [21,25]. In the present work the goal is to exploit the XtalController technology to determine appropriate crystallant concentration sufficient for the nucleation of a minimal number of crystals.

Figure 2 shows a typical crystallization assay with PhaCCA. The experiment starts with a drop of enzyme at 5 mg/mL in its storage buffer. The crystallant (stock solution: 1 M diammonium phosphate, 0.1 M ammonium acetate pH 4.5) is gradually added into the enzyme solution using a piezo pump. The DLS signal clearly indicates the start of protein association around 0.2 M diammonium phosphate with the appearance of objects larger than the initial PhaCCA monomer and increasing in size. This phenomenon is amplified during the step of drop evaporation when both the concentration of crystallant and enzyme simultaneously increase, hence increasing the supersaturation of the enzyme solution. The monomeric form of PhaCCA remains present throughout the experiment, in equilibrium with nano/microcrystalline objects.

After 48 h, the monitoring was stopped and the cover slip was transferred from the instrument to a classical 24-well Linbro plate to be stored in equilibrium with a reservoir containing a crystallant solution at the same concentration as that reached at the completion of the protocol (i.e., 1 M diammonium phosphate, 0.1 M ammonium acetate pH 4.5). Small crystals that were already visible after 28 h in the periphery of the drop, grew slowly to a useful size (100–200 μm) for X-ray diffraction (see Section 3.3) over a period of several weeks.

In an attempt to increase the size of PhaCCA crystals, protocols including several cycles of drop concentration/dilution were tested. The experimental idea was to trigger protein interaction, then redissolve partially the population of nuclei to select a few for growth in the next drop concentration step. Figure 3 illustrates this strategy and displays the resulting crystals. The largest reached a size of 0.5 mm after 11 days of incubation. This strategy may be exploited in the future to grow large crystals for neutron diffraction. A similar concept was applied using temperature variation coupled with dialysis to adjust the supersaturation and reduce the number of nuclei to promote the growth of large mm-sized crystals [26]. The great advantage of the XtalController is the possibility to monitor the fluctuation of particle populations in real time. The DLS signal shows that the monomer of PhaCCA

is converted into larger objects during concentration steps while the drop remains optically clear. If the process is too fast and evolves toward precipitation, the whole system can be driven back to lower concentrations favoring the dissolution of precipitates and the re-solubilization of monomers. A series of cycles of decreasing concentration variation allows a slow and controlled convergence to the target condition, similar to that used in classical vapor diffusion assays but resulting in many fewer nucleation events and, thus, in larger crystals.

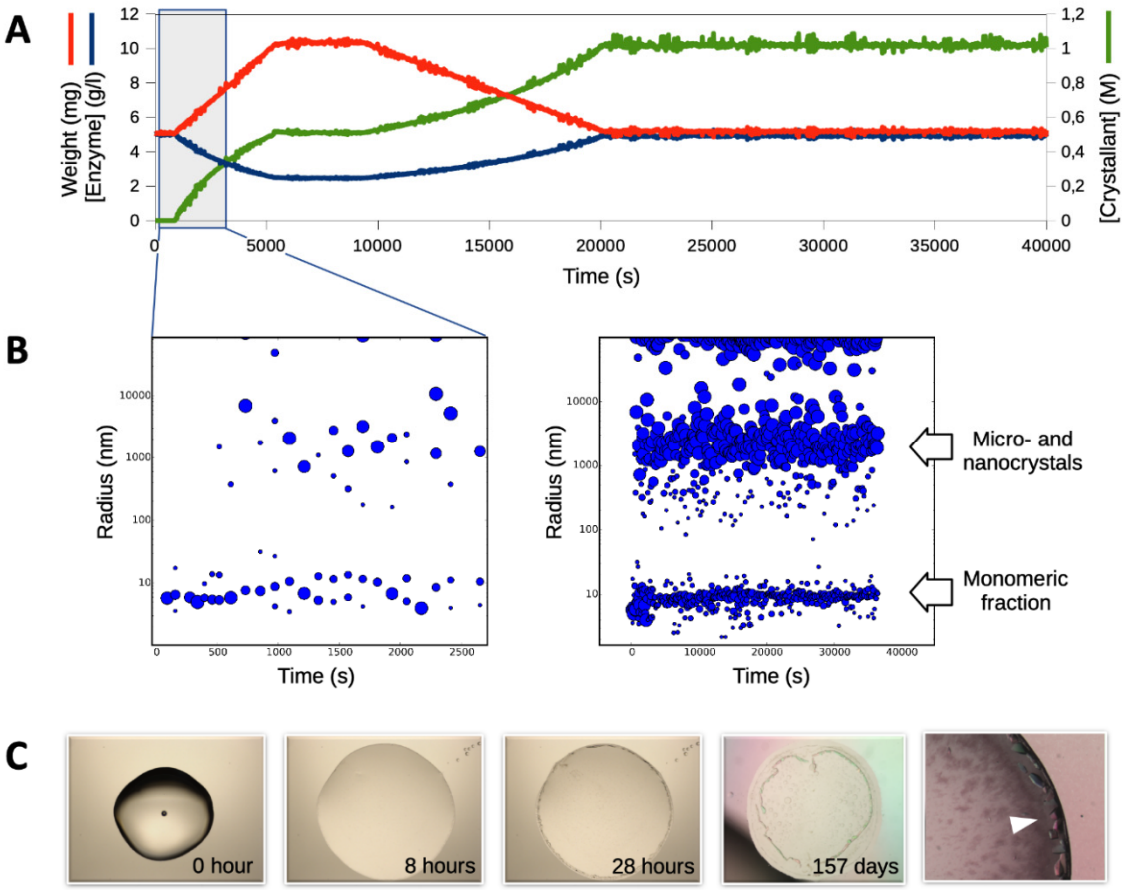


Figure 2. Crystallization of PhaCCA enzyme in the XtalController. Panels (A) and (B) show the experimental records over the first 40,000 sec (~11 h). No visible change occurred between t = 11 and 48 h, i.e., when the experiment was transferred to a Linbro plate. (**A**) Experimental parameters are recorded and curves display the variation of drop weight and corresponding variation of enzyme and crystallant concentrations over time. (**B**) The evolution of particle size in solution is monitored by dynamic light scattering (DLS) over time. At each time point, different populations detected in solution are represented as blue dots. The center of the dot indicates the hydrodynamic radius R_h of the population and the surface its relative contribution to the overall scattered signal. The DLS distribution on the left highlights early events (blue window in A) and shows that the enzyme starts to react at concentrations of crystallant as low as 0.2 M. Particle size distribution on the right corresponds to the monitoring of the complete experiment. The signal of monomeric PhaCCA (R_h = 4 nm) decreases while larger objects, likely nano-crystals or nano-clusters, appear upon crystallant addition. (**C**) A selection of drop micrographs taken along the experiment shows the appearance of small objects after 28 h, leading to useful samples for diffraction analysis within several weeks of incubation over a reservoir in a Linbro plate. The close-up view on the right-hand side shows typical PhaCCA crystals (white arrow) used for data collection (Table 1).

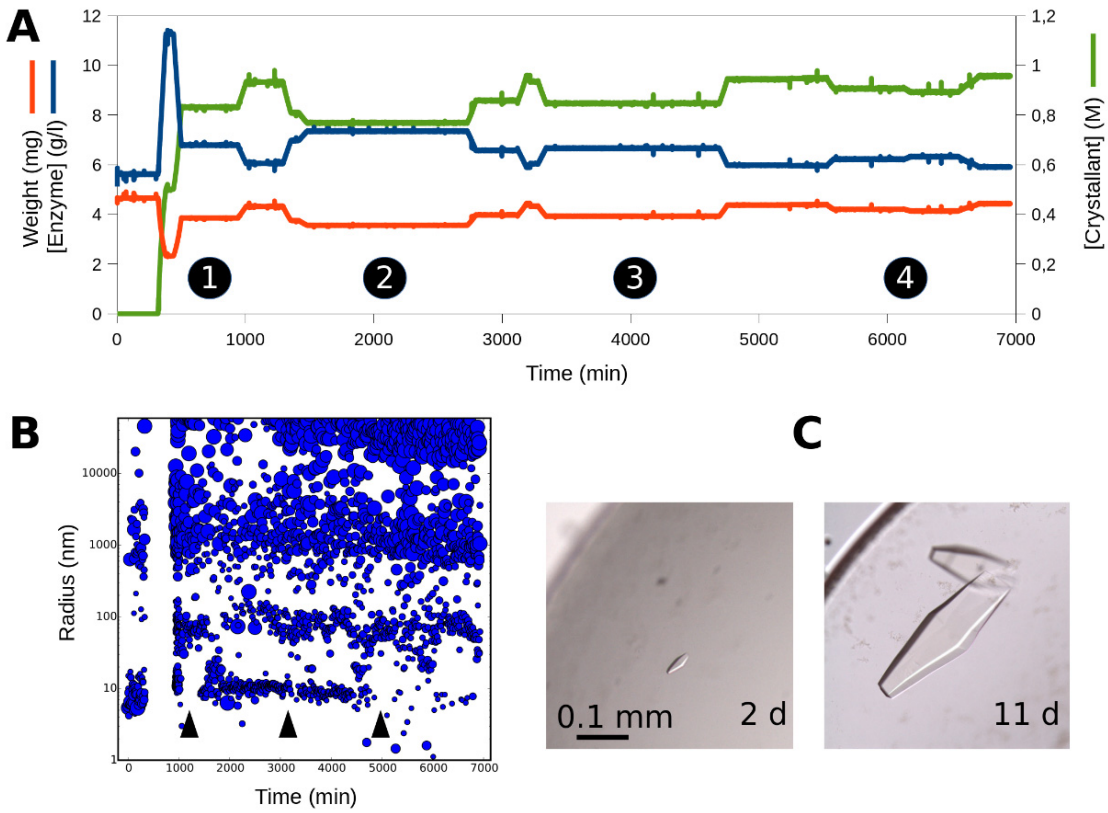


Figure 3. Production of large PhaCCA crystals. (**A**) The crystallant was injected in the sample drop following the same protocol as in Figure 2. Four cycles (see numbers) of drop concentration/dilution were performed. (**B**) The DLS monitoring shows that each time the drop is concentrated, the signal for the population of monomeric PhaCCA decreases (see arrows) in favor of larger particles. At the end of this protocol, the clear drop was transferred from the XtalController to a Linbro plate for incubation. (**C**) First crystals appeared after 2 days and the biggest reached a size 0.5 mm in 11 days.

3.2. Tracking the Nucleant Effect of the Crystallophore Tb-Xo4on HEWL

A terbium complex, called crystallophore N°1 or Tb-Xo4, with phasing and nucleating properties was recently described [19,20,27]. When preparing lysozyme solutions in pure milliQ water including 10 mM Tb-Xo4 for screening experiments, we observed that the addition of the compound led to the formation of a pellet composed of single lysozyme crystals at the bottom of Eppendorf tubes after 20 to 30 days in the cold room. To confirm the ability of the nucleant to promote lysozyme self-association in conditions containing very low salt concentration, the XtalController was used to monitor the behavior of the enzyme in solution upon injection of 10 mM Tb-Xo4 dissolved in water. The control experiment consisted in injecting an equivalent volume of pure water. After incubating 10 µl of a fresh HEWL solution (25 mg/mL) at 20 °C to check its stability by DLS (Figure 4), the crystallization chamber was opened for a few seconds to inject 1 µl of a 100 mM Tb-Xo4 solution (or water) directly into the drop using a 1 µl Hamilton syringe. Accordingly, the drop weight increased to 11 mg and rapidly went back to 10 mg upon water evaporation controlled by the instrument scheduler. The drop was incubated for hours and DLS measurements were recorded before its transfer to a Linbro plate for long-term storage.

Figure 4 shows that the addition of Tb-Xo4 triggered the instantaneous formation of particles of larger size and this phenomenon of association, which was not observed upon water addition, led to the growth of large crystals after weeks of incubation. Tb-Xo4 decreases the solubility of HEWL and promotes interactions between enzyme particles: the monomer seems to be converted in a slightly larger entity (Figure 5), possibly bridged by the nucleant (Figure 6B,C) and large assemblies grow up to a size of 1 µm. Again, the XtalController was instrumental to highlight a situation leading

to nucleation in a drop that remained clear for hours or days in videomicroscopy. Indeed, when a supersaturated state has been created, the system is thermodynamically set for either precipitation or nucleation. If nuclei are favored over precipitates, then crystal growth is a matter of time (i.e., kinetics). The "activated" drop is then simply transferred from the crystallization chamber to a Linbro plate for incubation until crystals grow.

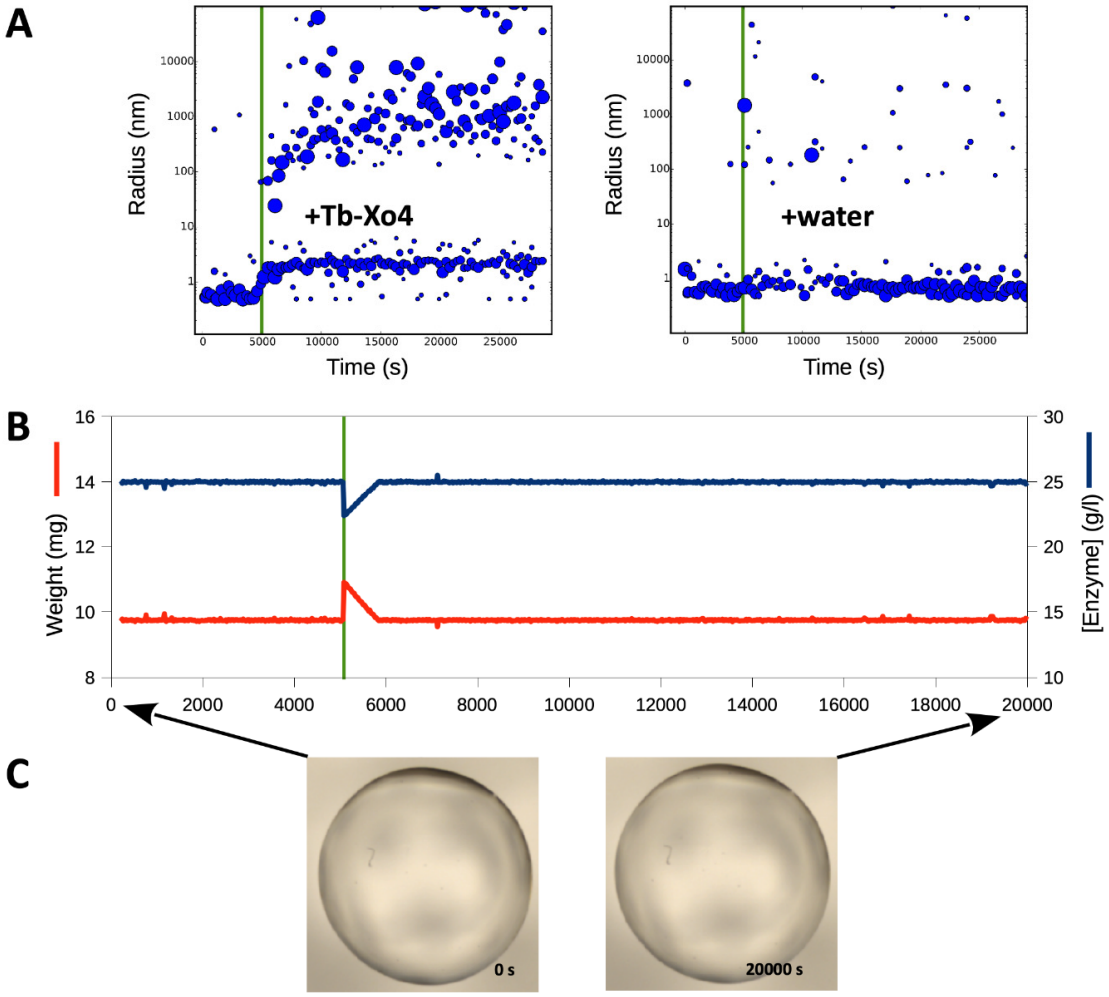


Figure 4. Monitoring the nucleant effect of Tb-Xo4 on HEWL. In this simple experiment, a Roche lysozyme solution (25 mg/mL in water) was incubated in the XtalController. (**A**) Upon injection of 1 µl of nucleant (100 mM Tb-Xo4 solution), symbolized by the green line and highlighted by a 1 mg jump of the weight (**B**), the enzyme immediately reacted and formed larger objects that grew up to size of ~1 µm (**A**, left). The drop was then maintained at a constant weight as illustrated by weight and concentration curves (**B**). Records (**A**) and (**B**) are displayed for t = 0 to 20,000 sec (~ 5.6 h). No visible change was observed afterwards and, as the drop remained clear (**C**), it was transferred after 8 h to a Linbro plate for incubation. The control injection of 1 µl of water does not change the DLS signal (A, right).

We further tested the nucleating property of Tb-Xo4 on the crystallization of HEWL from other suppliers (Sigma and Seikagaku) to check for potential enzyme preparation and batch effect. Following the protocol described in Figure 4, the concentration of enzyme was increased to 50 and 71 mg/mL to promote the growth of larger crystals. As for the Roche HEWL, the addition of Tb-Xo4 triggered the rapid formation (i.e., within 15 to 30 min) of populations of particles with radii ranging from 100 to 1000 nm in the XtalController (Figure 5A). The same phenomenon was reproduced in the quartz cuvette of a Nanostar DLS instrument using the same lysozyme batch. Although they contributed significantly to the scattering signal, the percentage in mass of the larger populations was always low,

in the range of a few percent (Figure 5B). Drops prepared with Sigma and Seikagaku lysozyme batches in the XtalController and incubated for several months against a buffer reservoir produced crystals of several hundreds of µm suitable for crystallographic analyses (Figure 5C).

Table 1. Statistics of crystal analysis and structure refinement.

Enzyme	***PhaCCA***	***HEWL***
Data collection		
X-ray source	FIP/BM30A – ESRF	Rigaku Fr-X – IGBMC
Wavelength (Å)	0.9799	1.5418
Detector	ADSC Q315r	EIGER R 4M
Temperature (K)	100	293
Space group	$P4_32_12$	$P4_32_12$
Cell parameters (Å)	70.53, 70.53, 291.48	78.81, 78.81, 38.33
Mosaicity (°)	0.31	0.12
Solvent content	67.4	40.8
Resolution range (Å)	47 – 2.28 (2.42 – 2.28)	35 – 1.51 (1.60 – 1.51)
Number of reflections	237664 (37915)	164648 (14601)
Number of unique reflections	34470 (5385)	19535 (3017)
Multiplicity	6.9 (7.0)	8.4 (4.8)
Completeness (%)	99.5 (98.5)	99.2 (96.5)
Mean I/sigma(I)	15.5 (1.5)	29.7 (1.9)
R_{meas} (%)[1]	8.6 (126)	3.6 (72.6)
R_{pim} (%)[2]	3.4 (44.8)	1.1 (31.7)
SigAno	–	2.1 (0.6)
$CC_{1/2}$	99.9 (76.0)	100 (78.1)
Structure refinement		
R_{work}, R_{free}	0.213, 0.257	0.141, 0.175
Number of non-H atoms enzyme, ligands, solvent	2989, 36, 105	1001, 35, 54
RMSD on bonds (Å) and angles (°)	0.009, 0.96	0.004, 0.69
Average ADPs (Å^2) overall, enzyme, ligands, solvent	51.0, 50.8, 62.9, 51.6	30.1, 28.7, 51.7, 41.9
Ramachandran plot: % of residues in favored, allowed, unfavored regions	96.4, 3.3, 0.3	99.2, 0.8, 0
PDBid	6TVZ	6TVY

[1] $R_{meas} = \Sigma_{hkl} \sqrt{(n/n-1)} \Sigma_i |I_i(hkl) - \langle I(hkl)\rangle| / \Sigma_{hkl} \Sigma_i I_i(hkl)$; [2] $R_{pim} = \Sigma_{hkl} \sqrt{(1/n-1)} \Sigma_i |I_i(hkl) - \langle I(hkl)\rangle| / \Sigma_{hkl} \Sigma_i I_i(hkl)$.

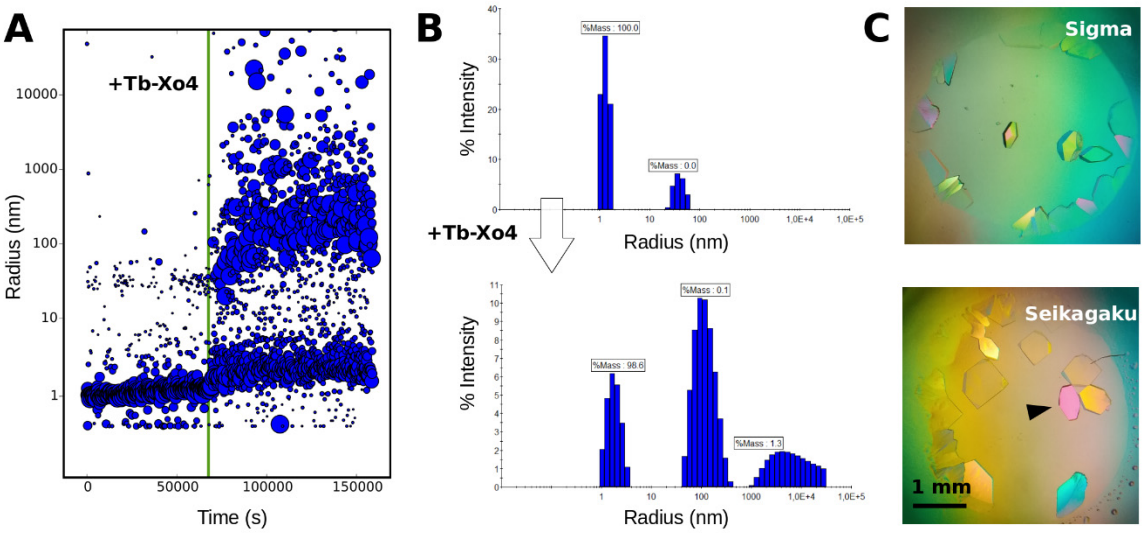


Figure 5. Production of Tb-Xo4-HEWL crystals for structural analyses. The experimental procedure described in Figure 4 was applied to Sigma and Seikagaku lysozyme samples at a higher concentration (50 and 71 mg/mL, respectively) to promote the growth of larger crystals. (**A**) the DLS monitoring of Sigma lysozyme in the XtalController upon Tb-Xo4 injection (green line) shows same features as in Figure 4A, yet amplified by the higher enzyme concentration. (**B**) The phenomenon was followed on the same batch in a Nanostar DLS instrument. The top distribution detects almost exclusively the monomeric lysozyme before the injection, whereas populations with a R_h of ~100 nm and >1 µm appear in the bottom distribution already 2.5 min after Tb-Xo4 injection. Crystallization assays performed with both batches were transferred from the XtalController to a Linbro plate for incubation. (**C**) Images showing resulting crystals after 9 months before harvesting for data collection. The crystal marked by an arrow was used to collect the dataset presented in Table 1.

3.3. Characterization of Crystals Grown in the XtalController

The quality of PhaCCA and HEWL crystals grown using the XtalController was assessed by X-ray diffraction. Crystals of PhaCCA were analyzed under cryogenic conditions at the FIP-BM30A synchrotron beamline (ESRF, Grenoble, France). Table 1 shows excellent statistics for such a crystal leading to a complete diffraction dataset at 2.28 Å resolution. The signal extended isotropically up to 1.8 Å (not shown) as for our best data collected on a much stronger beamline installed on an insertion device (PROXIMA2A, SOLEIL, Saint-Aubin, France). Here, data were truncated because of the presence of two diffraction rings at resolutions of about 1.9 and 2.15 Å. They were probably due to suboptimal cryocooling of the solvent and significantly altered the statistics in the high resolution shell (<2.2 Å). However, the resulting electron density maps are extremely clear and allowed us to observe two phosphate ions (Figure 6A) of the crystallant bound to the surface of this RNA binding enzyme.

Lysozyme crystals were analyzed at room temperature using the X-ray lab source from the INSTRUCT-FRISBI platform at the IGBMC (Illkirch, France). A fast data collection was applied using an EIGER pixel detector to collect complete anomalous data before crystal decay. The Tb^{3+} ions gave a strong anomalous signal at 1.5418 Å (f″ of 9 electrons at the Cu Kα wavelength). The dataset collected at 1.51 Å resolution (Table 1) allowed the identification of three anomalous peaks corresponding to Tb^{3+} sites, a major site with an occupancy of q = 0.8 for which the ligand could be built, and a minor site with two alternate positions (Figure 6C). These sites are consistent with those described in PDBid 6F2I determined with an HEWL crystal grown in 0.8 M NaCl and 100 mM Tb-Xo4 [20]. The two Tb^{3+} positions corresponding to the minor site have a low occupancy (q = 0.2 and 0.4) and no visible ligand, which is also consistent with the low concentration of Tb-Xo4 (10 mM) providing a ratio of two Tb-Xo4 complexes for one lysozyme molecule. This concentration seems to be sufficient to trigger the

association of lysozyme particles, nucleation and subsequent crystal growth, but not to guarantee a full occupancy in the crystal lattice.

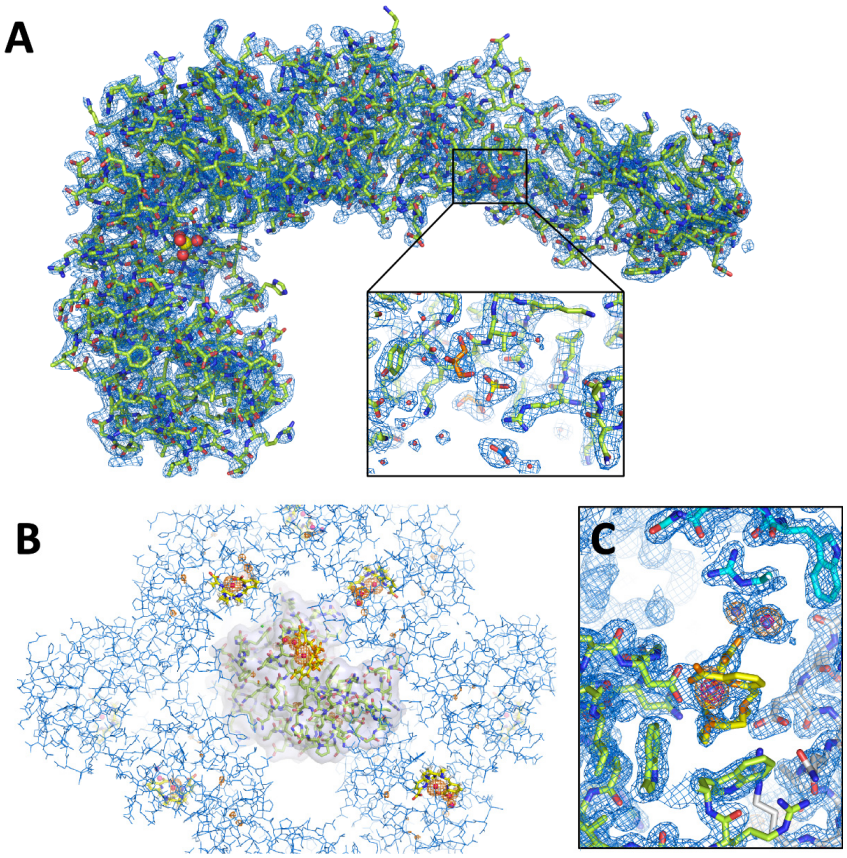


Figure 6. 3D structures of PhaCCA and HEWL crystallized in the XtalController. (**A**) Overall structure of PhaCCA and close-up on a phosphate, an acetate and two glycerol molecules bound to the enzyme surface. The blue *2mFo-DFc* map is contoured at 1.2 σ. (**B**) View of tetragonal crystal packing of HEWL seen down the four-fold axis. One lysozyme monomer is represented with a Tb-Xo4 complex bound to its surface. Symmetry related monomers are depicted in line mode. The anomalous difference map shown in orange is contoured at 5 σ and indicates the presence of Tb^{3+} ions (red spheres) as well as sulfur atoms. (**C**) Zoom on the Tb-Xo4 binding sites. The side chain of Asp101 directly coordinates the Tb^{3+} of the major site for which the organic ligand (in yellow) is partially visible in the blue *2mFo-DFc* map contoured at 1 σ. The anomalous difference map depicted in orange and red is contoured at 5 and 20 σ, respectively, and highlights the strong anomalous signal of Tb^{3+} ions. Two alternate positions are observed but their ligand is not visible due to low occupancy. The two adjacent Tb-Xo4 sites constitute a bridge (molecular glue) between the green and blue lysozyme monomers.

4. Conclusions

In the perspective of emerging time-resolved studies of enzyme:substrate systems using SFX and XFEL facilities, it becomes increasingly important to gain more control over sample production, quality and reproducibility. These two examples of enzyme crystallization highlight the usefulness of the XtalController to monitor the evolution of crystallization assays and to act on the process. Beyond helping define and optimize crystallization conditions, the XtalController and its integrated DLS module may also be an ideal tool to:

- Explore phase diagrams of biomolecules with a direct feedback on nucleation events;
- Study the stability of biomolecules in solution with respect to various parameters such as temperature, pH, ligands, etc.;

- Determine the optimum conditions for introducing a cryoprotectant;
- Ensure the reproducibility of crystals in the context of structural biology investigations, rational drug design, and fragment-based screening.
- Produce calibrated nanocrystals on demand (difficult to monitor and control otherwise) for diffraction analyses using X-ray free electron lasers and μED, or, conversely, to promote the selective growth of large crystals for neutron diffraction.

More generally, this type of versatile instruments provides a more rational approach to crystallization and a great alternative to extensive blind screening. We do believe that this technology has a bright future.

Author Contributions: Formal analysis, R.D.W., K.R., E.G., P.B., B.L. and C.S.; funding acquisition, M.M., O.M., E.G. and C.S.; investigation, R.D.W., K.R., S.E., A.G.M., O.H., O.M., E.G., P.B., B.L. and C.S.; resources, F.R. and O.M.; supervision, M.M., O.M., E.G. and C.S.; writing—original draft, C.S.; writing—review and editing, R.D.W., K.R., S.E., A.G.M., O.H., H.B., M.M., F.R., O.M., E.G., P.B., B.L. and C.S. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the French Centre National de la Recherche Scientifique (CNRS), the University of Strasbourg, the project Ln23 (ANR-13-BS07-0007-01), the LabEx consortia "NetRNA" (ANR-10-LABX-0036_NETRNA) and "INRT" (ANR-10-LABX-0036_INRT), a French State fund managed by the Agence National de la Recherche under the frame program Investissements d'Avenir ANR-10-IDEX-0002-02, a PhD funding to R.dW from the Excellence initiative (IdEx) of the University of Strasbourg, a PhD funding to K.R from the French-German University (UFA/DFH), the Deutsche Forschungsgemeinschaft (grant no. Mo 634/10-1). The authors benefitted from the PROCOPE Hubert Curien cooperation program (French Ministry of Foreign Affair and Deutscher Akademischer Austauschdienst). The XtalController was acquired with the financial support of the LabEx consortium "MitoCross" (ANR-11-LABX-0057_MITOCROSS).

Acknowledgments: The authors thank the team of FIP beamline at the European Synchrotron Radiation Facility (ESRF, Grenoble, France) for beamtime allocation; Alexandra Bluhm and Léna Coudray for their contribution to crystallization experiments and structure refinement; Robin Schubert and Christian Betzel (University of Hamburg), as well as Karsten Dierks, Arne Mayer and Annette Eckhardt (Xtal-Concepts GmbH, Hamburg) for fruitful discussions and exchanges about the XtalController. The authors acknowledge the support and use of the resources of the French Infrastructure for Integrated Structural Biology (FRISBI, ANR-10-INBS-05) and of Instruct-ERIC.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Giegé, R.; Sauter, C. Biocrystallography: Past, present, future. *HFSP J.* **2010**, *4*, 109. [CrossRef] [PubMed]
2. Chapman, H.N.; Fromme, P.; Barty, A.; White, T.A.; Kirian, R.A.; Aquila, A.; Hunter, M.S.; Schulz, J.; DePonte, D.P.; Weierstall, U.; et al. Femtosecond X-ray protein nanocrystallography. *Nature* **2011**, *470*, 73–77. [CrossRef] [PubMed]
3. Spence, J.C.H.; Chapman, H.N. The birth of a new field. *Philos. T. Roy. Soc. B* **2014**, *369*, 20130309. [CrossRef] [PubMed]
4. Johansson, L.C.; Stauch, B.; Ishchenko, A.; Cherezov, V. A Bright Future for Serial Femtosecond Crystallography with XFELs. *Trends Biochem. Sci.* **2017**, *42*, 749–762. [CrossRef]
5. Shi, D.; Nannenga, B.L.; Iadanza, M.G.; Gonen, T. Three-dimensional electron crystallography of protein microcrystals. *Elife* **2013**, *2*, e01345. [CrossRef]
6. Nannenga, B.L.; Shi, D.; Leslie, A.G.W.; Gonen, T. High-resolution structure determination by continuous-rotation data collection in MicroED. *Nat. Methods* **2014**, *11*, 927–930. [CrossRef]
7. Clabbers, M.T.B.; van Genderen, E.; Wan, W.; Wiegers, E.L.; Gruene, T.; Abrahams, J.P. Protein structure determination by electron diffraction using a single three-dimensional nanocrystal. *Acta Cryst. D* **2017**, *73*, 738–748. [CrossRef]
8. Nederlof, I.; Li, Y.W.; van Heel, M.; Abrahams, J.P. Imaging protein three-dimensional nanocrystals with cryo-EM. *Acta. Crystallogr. D. Biol Crystallogr.* **2013**, *69*, 852–859. [CrossRef]
9. Sauter, C.; Lorber, B.; McPherson, A. Crystallization–General Methods. In *International Tables of Crystallography, Vol. F, Crystallography of Biological Macromolecules*, 2nd ed.; Arnold, E., Himmel, D.M., Rossmann, M.G., Eds.; John Wiley and Sons: Chichester, UK, 2012; pp. 99–120.

10. Garcia-Caballero, A.; Gavira, J.A.; Pineda-Molina, E.; Chayen, N.E.; Govada, L.; Khurshid, S.; Saridakis, E.; Boudjemline, A.; Swann, M.J.; Shaw Stewart, P.; et al. Optimization of Protein Crystallization: The OptiCryst Project. *Cryst. Growth Des.* **2011**, *11*, 2112–2121. [CrossRef]
11. Meyer, A.; Dierks, K.; Hilterhaus, D.; Klupsch, T.; Mühlig, P.; Kleesiek, J.; Schöpflin, R.; Einspahr, H.; Hilgenfeld, R.; Betzel, C. Single-drop optimization of protein crystallization. *Acta Crystallog. F* **2012**, *68*, 994–998. [CrossRef]
12. Mikol, V.; Hirsch, E.; Giegé, R. Diagnostic of precipitant for biomacromolecule crystallization by quasi-elastic light-scattering. *J. Mol. Biol.* **1990**, *213*, 187–195. [CrossRef]
13. Kadima, W.; McPherson, A.; Dunn, M.F.; Jurnak, F.A. Characterization of precrystallization aggregation of canavalin by dynamic light scattering. *Biophys. J.* **1990**, *57*, 125–132. [CrossRef]
14. George, A.; Wilson, W.W. Predicting protein crystallization from a dilute solution property. *Acta Cryst. D Biol Cryst.* **1994**, *50*, 361–365. [CrossRef]
15. Zulauf, M.; D'Arcy, A. Light scattering of proteins as a criterion for crystallization. *J. Cryst. Growth.* **1992**, *122*, 102–106. [CrossRef]
16. Dierks, K.; Meyer, A.; Einspahr, H.; Betzel, C. Dynamic Light Scattering in Protein Crystallization Droplets: Adaptations for Analysis and Optimization of Crystallization Processes. *Cryst. Growth Des.* **2008**, *8*, 1628–1634. [CrossRef]
17. Schubert, R.; Meyer, A.; Baitan, D.; Dierks, K.; Perbandt, M.; Betzel, C. Real-Time Observation of Protein Dense Liquid Cluster Evolution during Nucleation in Protein Crystallization. *Cryst. Growth Des.* **2017**, *17*, 954–958. [CrossRef]
18. Baitan, D.; Schubert, R.; Meyer, A.; Dierks, K.; Perbandt, M.; Betzel, C. Growing Protein Crystals with Distinct Dimensions using Automated Crystallization Coupled with in Situ Dynamic Light Scattering. *JoVE.* **2018**, e57070. [CrossRef]
19. Engilberge, S.; Riobé, F.; Di Pietro, S.; Lassalle, L.; Coquelle, N.; Arnaud, C.-A.; Pitrat, D.; Mulatier, J.-C.; Madern, D.; Breyton, C.; et al. Crystallophore: A versatile lanthanide complex for protein crystallography combining nucleating effects, phasing properties, and luminescence. *Chem. Sci.* **2017**, *8*, 5909–5917. [CrossRef]
20. Engilberge, S.; Riobé, F.; Wagner, T.; Pietro, S.D.; Breyton, C.; Franzetti, B.; Shima, S.; Girard, E.; Dumont, E.; Maury, O. Unveiling the Binding Modes of the Crystallophore, a Terbium-based Nucleating and Phasing Molecular Agent for Protein Crystallography. *Chemistry* **2018**, *24*, 9739–9746. [CrossRef]
21. De Wijn, R.; Hennig, O.; Ernst, F.G.M.; Lorber, B.; Betat, H.; Mörl, M.; Sauter, C. Combining crystallogenesis methods to produce diffraction-quality crystals of a psychrophilic tRNA-maturation enzyme. *Acta Crystallog. F* **2018**, *74*, 747–753. [CrossRef]
22. Kabsch, W. XDS. Acta Crystallogr. *D. Biol. Crystallogr.* **2010**, *66*, 125–132. [CrossRef] [PubMed]
23. Adams, P.D.; Afonine, P.V.; Bunkóczi, G.; Chen, V.B.; Davis, I.W.; Echols, N.; Headd, J.J.; Hung, L.-W.; Kapral, G.J.; Grosse-Kunstleve, R.W.; et al. PHENIX: A comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. *D. Biol. Cryst.* **2010**, *66*, 213–221. [CrossRef] [PubMed]
24. Emsley, P.; Cowtan, K. Coot: Model-building tools for molecular graphics. *Acta Crystallogr. D. Biol. Cryst.* **2004**, *60*, 2126–2132. [CrossRef] [PubMed]
25. De Wijn, R.; Hennig, O.; Roche, J.; Engilberge, S.; Rollet, K.; Fernandez-Millan, P.; Brillet, K.; Betat, H.; Mörl, M.; Roussel, A.; et al. A simple and versatile microfluidic device for efficient biomacromolecule crystallization and structural analysis by serial crystallography. *IUCrJ* **2019**, *6*, 454–464. [CrossRef]
26. Junius, N.; Oksanen, E.; Terrien, M.; Berzin, C.; Ferrer, J.-L.; Budayova-Spano, M. A crystallization apparatus for temperature-controlled flow-cell dialysis with real-time visualization. *J. Appl. Cryst.* **2016**, *49*, 806–813. [CrossRef]
27. Engilberge, S.; Wagner, T.; Santoni, G.; Breyton, C.; Shima, S.; Franzetti, B.; Riobé, F.; Maury, O.; Girard, E. Protein crystal structure determination with the crystallophore, a nucleating and phasing agent. *J. Appl. Cryst.* **2019**, *52*, 722–731. [CrossRef]

© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).

VI. Crystallization and Structural Determination of an Enzyme:Substrate Complex by Serial Crystallography in a Versatile Microfluidic Chip

X-ray crystallography stays the most common and powerful method to characterize 3D structure of macromolecules. Evolution of light sources toward more powerful, coherent and microfocused beams requires adapted samples and setups for fast and efficient data collection. In the mean time, the strategy to determine crystallization conditions has evolved over the last two decades as described in section I.2. To date, vapor diffusion (VD) is surely the most common and successful method to grow crystals in laboratories. However, this technique is without doubt time and sample-consuming, as it is based on trial and error process in single crystallization drops. Moreover, once crystals are obtained, users would fish them with a microloop and flash cool the sample in liquid nitrogen, propane or ethane. The latter constitutes a critical and risky step during which users may destroy their precious sample just before any diffraction experiment. Among other current crystallization methods, counter-diffusion (CD) may be a useful alternative to rescue a crystallization project (Otálora *et al.*, 2009). First, CD offers the possibility to screen a wide range of crystallant concentrations in a single experiment. Second, CD capillaries can usually be mounted directly onto beamline goniometer without requiring direct crystal handling for room temperature analysis.

Taking advantage of microfluidics spreading, Dr. Claude Sauter developed a crystallization chip, so-called ChipX, which combines advantages of standard CD with very low sample consumption (from 5-6 µL per chip) and the possibility to screen eight different crystallant solutions on a single device suitable for direct *in situ* X-ray analysis. Briefly, the macromolecule solution is injected in the eight channels via a common sample inlet. Then, at the other extremity of the channels, crystallant solutions are deposited in individual reservoirs from where they will diffuse into the channels, thus creating a concentration gradient. The chip is closed with regular crystallization tape and stored at desired temperature. Finally, once crystals are identified, the ChipX can be mounted on a 3D printed holder and directly installed onto regular beamline goniometers. Labels along channels serve to localize potential crystalline objects.

In this chapter, we present how simple, fast and efficient is the set up of a chipX. In addition, this paper highlights results we obtained with CCA-adding enzyme from psychrophilic bacterium *Planococcus halocryophilus*. This publication also completes results published in 2019 in IUCr journal (de Wijn *et al.*, 2019), showing a set of different macromolecule structures solved after crystallization within ChipX.

La cristallographie aux rayons X reste la méthode la plus courante et la plus puissante pour caractériser la structure 3D des macromolécules. L'évolution des sources lumineuses vers des faisceaux plus puissants, cohérents et microfocalisés nécessite des échantillons et des montages adaptés pour une collecte de données rapide et efficace. Dans le même temps, la stratégie pour déterminer les conditions de cristallisation a évolué au cours des deux dernières décennies, comme décrit dans la section I.2. À ce jour, la diffusion de vapeur (VD) est certainement la méthode la plus courante et la plus efficace pour faire croître des cristaux en laboratoire. Cependant, cette technique est sans aucun doute gourmande en temps et en échantillons, car elle repose sur un processus d'essais et d'erreurs dans des gouttes de cristallisation uniques. De plus, une fois les cristaux obtenus, les utilisateurs doivent les pêcher à l'aide d'une micro-boucle et refroidir rapidement l'échantillon dans de l'azote liquide, du propane ou de l'éthane. Ce dernier point constitue une étape critique et risquée au cours de laquelle les utilisateurs peuvent détruire leur précieux échantillon juste avant une expérience de diffraction. Parmi les autres méthodes de cristallisation actuelles, la contre-diffusion (CD) peut être une alternative utile pour sauver un projet de cristallisation (Otálora et al., 2009). Premièrement, la CD offre la possibilité de cribler une large gamme de concentrations de cristallisants en une seule expérience. Deuxièmement, les capillaires de CD peuvent généralement être montés directement sur le goniomètre de la ligne de faisceau sans nécessiter la manipulation directe des cristaux pour l'analyse à température ambiante.

Profitant de la propagation de la microfluidique, le Dr Claude Sauter a développé une puce de cristallisation, appelée ChipX, qui combine les avantages de la CD standard avec une très faible consommation d'échantillons (de 5 à 6 µL par puce) et la possibilité de cribler huit solutions de cristallisation différentes sur un seul dispositif adapté à l'analyse directe *in situ* aux rayons X. En bref, la solution de macromolécules est injectée dans les huit canaux via une entrée dédiée commune. Ensuite, à l'autre extrémité des canaux, les solutions cristallines sont déposées dans des réservoirs individuels d'où elles diffuseront dans les canaux, créant ainsi un gradient de concentration. La puce est fermée avec un ruban de cristallisation ordinaire et stockée à la température souhaitée. Enfin, une fois que les cristaux sont identifiés, la ChipX peut être montée sur un support imprimé en 3D et directement installée sur des goniomètres de ligne de lumière ordinaires. Les étiquettes le long des canaux servent à localiser les objets cristallins potentiels.

Dans ce chapitre, nous présentons la simplicité, la rapidité et l'efficacité de la mise en place d'une ChipX. En outre, cet article met en évidence les résultats que nous avons obtenus avec l'enzyme d'addition du CCA provenant de la bactérie psychrophile *Planococcus halocryophilus*. Cette publication complète également les résultats publiés en 2019 dans la revue IUCr (de Wijn et al., 2019), montrant un ensemble de différentes structures de macromolécules résolues après cristallisation au sein de ChipX.

Die Röntgenkristallographie ist nach wie vor die gängigste und leistungsfähigste Methode zur Charakterisierung der 3D-Struktur von Makromolekülen. Die Entwicklung von Lichtquellen hin zu leistungsfähigeren, kohärenteren und mikrofokussierten Strahlen erfordert angepasste Proben und Aufbauten für eine schnelle und effiziente Datenerfassung. In der Zwischenzeit hat sich die Strategie zur Bestimmung der Kristallisationsbedingungen in den letzten zwei Jahrzehnten weiterentwickelt, wie in Abschnitt I.2 beschrieben. Bis heute ist die Dampfdiffusion (VD) sicherlich die am weitesten verbreitete und erfolgreichste Methode zur Kristallzüchtung in Labors. Allerdings ist diese Technik zweifellos zeit- und probenaufwändig, da sie auf dem Versuch-und-Irrtum-Verfahren in einzelnen Kristallisationstropfen beruht. Außerdem müssen die gewonnenen Kristalle mit einer Mikroschleife isoliert und die Probe in flüssigem Stickstoff, Propan oder Ethan abgekühlt werden. Letzteres ist ein kritischer und riskanter Schritt, bei dem der Benutzer seine wertvolle Probe kurz vor einem Beugungsexperiment zerstören kann. Neben anderen aktuellen Kristallisationsmethoden kann die Gegendiffusion (CD) eine nützliche Alternative sein, um ein Kristallisationsprojekt zu retten (Otálora et al., 2009). Erstens bietet die CD-Methode die Möglichkeit, in einem einzigen Experiment einen breiten Bereich von Kristallisationskonzentrationen zu untersuchen. Zweitens können CD-Kapillaren in der Regel direkt auf ein Strahlengang-Goniometer montiert werden, ohne dass eine direkte Handhabung der Kristalle für die Analyse bei Raumtemperatur erforderlich ist.

Unter Ausnutzung der Vorteile der Mikrofluidik hat Dr. Claude Sauter einen Kristallisationschip, den so genannten ChipX, entwickelt, der die Vorteile der Standard-CD mit einem sehr geringen Probenverbrauch (5-6 µL pro Chip) und der Möglichkeit kombiniert, acht verschiedene kristalline Lösungen auf einem einzigen Gerät zu screenen, das für die direkte In-situ-Röntgenanalyse geeignet ist. Kurz gesagt, die Makromoleküllösung wird über einen gemeinsamen Probeneinlass in die acht Kanäle injiziert. Am anderen Ende der Kanäle werden die kristallinen Lösungen in einzelnen Reservoirs deponiert, von wo aus sie in die Kanäle diffundieren und so einen Konzentrationsgradienten erzeugen. Der Chip wird mit normalem Kristallisationsband verschlossen und bei der gewünschten Temperatur gelagert. Sobald die Kristalle identifiziert sind, kann der ChipX auf eine 3D-gedruckte Halterung montiert und direkt auf regulären Strahlengang-Goniometern installiert werden. Etiketten entlang der Kanäle dienen zur Lokalisierung potenzieller kristalliner Objekte.

In diesem Kapitel stellen wir vor, wie einfach, schnell und effizient der Aufbau des ChipX ist. Außerdem werden in diesem Beitrag Ergebnisse vorgestellt, die wir mit dem CCA-addierenden Enzym des psychrophilen Bakteriums Planococcus halocryophilus erzielt haben. Diese Publikation vervollständigt auch die 2019 in der Zeitschrift IUCr veröffentlichten Ergebnisse (de Wijn et al., 2019), die eine Reihe verschiedener Makromolekülstrukturen zeigen, die nach der Kristallisation in ChipX gelöst wurden.

Crystallization and Structural Determination of an Enzyme:Substrate Complex by Serial Crystallography in a Versatile Microfluidic Chip

Raphaël de Wijn[1,4], **Kévin Rollet**[1,2], **Vincent Olieric**[3], **Oliver Hennig**[2], **Nicola Thome**[1], **Camille Noûs**[1], **Caroline Paulus**[1], **Bernard Lorber**[1], **Heike Betat**[2], **Mario Mörl**[2], **Claude Sauter**[1]

[1] Université de Strasbourg, Architecture et Réactivité de l'ARN, UPR 9002, CNRS, Institut de Biologie Moléculaire et Cellulaire [2] Biochemistry and Molecular Biology, Institute for Biochemistry, Leipzig University [3] Paul Scherrer Institute, Swiss Light Source [4] European XFEL GmbH

Corresponding Author

Claude Sauter
c.sauter@ibmc-cnrs.unistra.fr

Citation

de Wijn, R., Rollet, K., Olieric, V., Hennig, O., Thome, N., Noûs, C., Paulus, C., Lorber, B., Betat, H., Mörl, M., Sauter, C. Crystallization and Structural Determination of an Enzyme:Substrate Complex by Serial Crystallography in a Versatile Microfluidic Chip. *J. Vis. Exp.* (169), e61972, doi:10.3791/61972 (2021).

Date Published

March 20, 2021

DOI

10.3791/61972

URL

jove.com/video/61972

Abstract

The preparation of well diffracting crystals and their handling before their X-ray analysis are two critical steps of biocrystallographic studies. We describe a versatile microfluidic chip that enables the production of crystals by the efficient method of counter-diffusion. The convection-free environment provided by the microfluidic channels is ideal for crystal growth and useful to diffuse a substrate into the active site of the crystalline enzyme. Here we applied this approach to the CCA-adding enzyme of the psychrophilic bacterium *Planococcus halocryophilus* in the presented example. After crystallization and substrate diffusion/soaking, the crystal structure of the enzyme:substrate complex was determined at room temperature by serial crystallography and the analysis of multiple crystals directly inside the chip. The whole procedure preserves the genuine diffraction properties of the samples because it requires no crystal handling.

Introduction

Crystallography is a method to decipher the 3D architecture of biological macromolecules. The latter is important to understand how an enzyme selects and processes its substrates. The determination of a crystal structure requires the crystallization of the target macromolecule and the conditioning of the crystals for their analysis by X-ray diffraction[1]. Both crystal preparation and handling are crucial but delicate steps that can affect crystal quality and diffraction properties, and, thus, the resolution (i.e., the accuracy) of the resulting 3D structure. To facilitate the preparation of

Copyright © 2021 JoVE Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported License

high-quality crystals and eliminate unnecessary handling to preserve their diffraction properties, we designed a user-friendly and versatile microfluidic device called ChipX[2,3,4].

In this article, we will demonstrate how to load the protein solution into ChipX channels using conventional laboratory material to prepare crystals by counter-diffusion. This crystallization method provides an efficient screening of supersaturation and of potential nucleation conditions along the microfluidic channels containing the enzyme solution due to the concentration gradient generated by the diffusion of the crystallizing agent[5,6].

The chip setup is simple, it uses only standard laboratory pipets and does not require any costly equipment. When crystals have grown in ChipX, ligands of the enzyme can be introduced by diffusion. Diffraction data are then collected at room temperature on a series of crystals contained in the channels of the chip using a synchrotron X-ray source. The structural study described here led to the determination of structures of a tRNA maturation enzyme in its apo form and in complex with an analog of its CTP substrate introduced by soaking. This protein called CCA-adding enzyme polymerizes the CCA trinucleotide tail at the 3' end of tRNAs. The comparison of the two 3D images obtained by serial crystallography reveals the local conformational changes related to the binding of the ligand in conditions that are more physiological than those used in cryo-crystallography. The protocol described in this video is generally applicable to any biomolecule, be it a protein, a nucleic acid or a multi-component complex.

Protocol

1. Setting up crystallization assays in ChipX

NOTE: The ChipX microfluidic device can be obtained from the authors. A description of the chip is given in **Figure 1**. Solutions containing the crystallant (or crystallizing agent) used to trigger crystallization may be of commercial origin or prepared by the experimenter.

1. Loading the biomolecule sample

   NOTE: The sample volume actually required to perform an individual counter-diffusion assay in the straight section of each channel of ChipX is 300 nL. However for convenience we suggest to load 5 µL to completely fill the eight channels taking into account the variable length of their curved section and inlet dead volumes.

   1. Pipet 5 µL of enzyme solutions using standard 10 µL pipet and tip.
   2. Introduce the tip vertically in the sample inlet and inject the solution until the eight channels are filled up to their opposite end (entry of the crystallant reservoir).
   3. Inject 1 µL of paraffin oil in the sample inlet in order to disconnect the channels from each other.
   4. Recover the extra solution in the crystallant reservoir at the extremity of each channel using a standard 10 µL pipet.
   5. Seal the sample inlet with a 1 cm x 1 cm piece of tape.

2. Loading the crystallization solutions

   1. Pipet 5 µL of crystallization solution using standard 10 µL pipet and tip. The reservoir volume is 10

Copyright © 2021 JoVE Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported License

µL, but loading only half of it avoids overflow when sealing with tape and facilitates further addition of ligand for soaking experiments. If initial crystallization conditions were obtained by vapor diffusion, increase the crystallant concentration by a factor of 1.5 - 2. Solutions can be different in every reservoir (in the presented case, 1 M diammonium hydrogen phosphate, 100 mM sodium acetate, pH 4.5 was used throughout).

2. Orient the pipet tip towards the entry of the channel in the funnel shaped part of the reservoir to avoid the formation of an air bubble upon solution deposition. It would prevent the contact between the two solutions and crystallant diffusion into the channel.
3. Inject the crystallant solution into the reservoir.
4. Seal the reservoirs with a 2.5 cm x 1 cm piece of tape.
5. Incubate the chip at 20 °C (temperature can be adjusted depending on the target, typically between 4 and 37 °C [4]).

2. Protein labeling with carboxyrhodamine for fluorescence detection

NOTE: This step is optional. It must be performed prior to sample loading to facilitate the detection of crystals in the chip using fluorescence. The detailed method of trace fluorescent labeling was described by Pusey and coworkers[7]. All steps are carried out at room temperature.

1. Dissolve 5 mg of carboxyrhodamine ester powder in 1 mL anhydrous dimethyl-formamide, split the solution in 0.6 µL aliquots to be stored at -20 °C.
2. Prepare a 1 M Na-borate pH 8.75 stock solution.
3. Dilute the stock to prepare the reaction buffer at 0.05 M Na-borate pH 8.75.
4. Rinse a desalting column (7 kDa MWCO, 0.5 mL) with 800 µL of reaction buffer.
5. Centrifuge the column for 1 min at 1400 x g, remove the filtrate.
6. Repeat this operation twice (steps 2.4-2.5) to wash the column.
7. Deposit 80 µL of protein in its storage buffer on the column (the protein can be diluted down to 1 mg/mL to increase the volume if needed).
8. Centrifuge the column for 1 min at 1400 x g. This step is intended to transfer the protein from its storage buffer to the reaction buffer.
9. Recover the flow-through (containing the protein in the reaction buffer) and mix it with 0.6 µL of carboxyrhodamine solution.
10. Incubate 5 min at room temperature.
11. Meanwhile, rinse the column 3 x with the storage buffer, centrifuge the column for 1 min at 1400 x g and discard the filtrate.
12. Deposit the reaction solution on the column.
13. Centrifuge the column for 1 min at 1400 x g and recover the flow-through (i.e., solution of labeled protein in its storage buffer).
14. Supplement the stock protein solution with 0.1-1 % (w/w) of labeled protein.
15. Setup the ChipX crystallization assays as described in section 1.

Copyright © 2021 JoVE Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported License

16. Check for the presence of protein crystals in the assays by exciting the fluorescent probe with a 520 nm wavelength light source.

3. Crystal observation

NOTE: The ChipX device may be handled without special care, even with crystals inside, except if the temperature needs to be controlled.

1. Use any stereomicroscope to check the outcome of crystallization assays in ChipX. Its footprint has the standard dimensions of microscope slides and is compatible with any system and slide holder.
2. Check the content of microfluidic channels starting from the reservoir where the crystallant concentration is the highest to the sample inlet where the crystallant concentration is the lowest. The ChipX material is transparent to visible light, compatible with the use of polarizers, as well as with UV illumination for protein crystal identification by intrinsic tryptophane fluorescence[8].
3. Record crystal positions using the labels embossed along the channels or mark crystal locations with a permanent marker by drawing color dots next to them on the chip surface.

4. Crystal soaking with ligands

NOTE: This procedure is optional. It is used to introduce ligands, enzyme substrates or heavy atoms into the crystals and should be carried out at least 24-48 h before X-ray analysis to allow the compound diffusion along the channels and into the crystals.

1. Gently remove the sealing tape from the reservoirs.
2. Add up to 5 µL of ligand solution in one or several reservoirs using a 10 µL micropipet (in the example, 3 µL of 10 mM cytidine-5'-[(α,β)-methyleno] triphosphate (CMPcPP) solution was added to achieve a final concentration of 3.75 mM). CMPcPP is a non-hydrolizable analog of CTP, a natural substrate of the enzyme.
3. Seal the reservoirs with a 2.5 cm x 1 cm piece of tape.
4. Incubate the chip under controlled temperature for 24-48 h to allow ligand diffusion along the channels of the chip.

5. Crystal analysis by serial crystallography

NOTE: This part of the protocol needs to be adapted depending on the beamline setup and the diffraction properties of the crystals. Only general indications are given for the crystallographic analysis based on experiments performed at X06DA beamline (SLS, Villigen, Switzerland).

1. ChipX mounting on the beamline goniometer

   NOTE: The file for 3D printing the ChipX holder is provided in ref[4].

   1. Turn off the cryo-jet of the beamline. The analysis here is carried out at room temperature.
   2. Mount the ChipX on a dedicated holder with the channel containing the crystals to be analyzed positioned at the center of the holder. The ChipX holder[4] does not require any screw or additional part, as it was designed to provide a perfect fit for ChipX.
   3. Attach the holder to the goniometer.
2. Data collection

Copyright © 2021 JoVE Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported License

1. Orient the thickest layer (top layer, **Figure 1**) of ChipX (in this orientation labels along the channels are directly readable using the centering camera of the beamline), towards the direct beam and the thinnest face behind the crystal to minimize the attenuation of the diffracted signal as described in ref[3].
2. To avoid collision of ChipX with the surrounding material (beamstop, collimator), restrict goniometer movements in the range ±30° (0° corresponding to the channels being perpendicular to the X-ray beam).
3. Find crystals position with the help of the labels embossed along the channels.
4. Select a crystal position.
5. Center the crystal either by standard low dose grid/ raster screening or 1-click procedure (the video shows an example of grid screening).
6. Collect diffraction data within the range -30° / +30°.
7. Restart the procedure at steps 5.2.4-5.2.6 on another crystal in the same channel after translation of the chip.
8. Manually realign another ChipX channel at the center of the holder and carry on data collection on crystals present in this channel.
9. Use standard crystallographic packages and procedures to process and merge the data, then to solve and refine the structure.

Representative Results

The microfluidic chip described here was designed to enable easy setup of crystallization assays and crystal analysis at room temperature. The procedure described above and in the video was applied in the frame of the structural characterization of the CCA-adding enzyme from the cold-adapted bacterium *Planococcus halocryophilus.* This enzyme belongs to an essential polymerase family that catalyses the sequential addition of the 3' CCA sequence on tRNAs using CTP and ATP[9,10].

The chip was first used to prepare crystals of the enzyme for structural analysis by the method of counter-diffusion. To this end, the enzyme solution was loaded in the eight microfluidic channels (crystallization chambers) by a single injection in the sample inlet of the chip (see **Figure 1**). The enzyme was used at 5.5 mg/mL in its storage buffer containing 20 mM Tris/HCl pH 7.5, 200 mM NaCl and 5 mM $MgCl_2$. This step was performed manually with a standard 10 µL micropipet. Crystallization solutions (100 mM sodium acetate pH 4.5,1 M diammonium hydrogen phosphate) were then deposited in the reservoirs at the other extremity of the channels.

The loading procedure is straightforward and does not take longer than five minutes (**Figure 2**). The crystallant then diffuses into the channels, creates a gradient of concentration that triggers crystal nucleation and growth. This gradient evolves dynamically and explores a continuum of supersaturation states[5,6] until reaching an equilibrium of crystallant concentration between the channels and the reservoir. Crystallization assays are typically checked under the micoscope over a period of 2 - 4 weeks to track the growth of crystals. Bipyramidal crystals of CCA-adding enzyme appeared throughout the channels after a few days of incubation at 20 °C (**Figure 3**). The optional fluorescent labeling[7] of the protein greatly facilitates the identification of protein crystals and their discrimination from salt crystals (**Figure 4**).

Copyright © 2021 JoVE Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported License

We exploited the diffusive environment in chip channels to deliver a substrate to the enzyme that builds up the crystals. In the present case, CMPcPP, a CTP analog, was added to the reservoir solutions at a final concentration of 3.75 mM (**Figure 5**). This addition was performed two days before the crystallographic analysis to allow CMPcPP to reach and occupy the catalytic site of the enzyme, as later confirmed by the crystal structure (see below).

We manufactured a chip holder (**Figure 6**) in polylactic acid using a 3D printer. The holder enables chip mounting on goniometers using standard magnetic heads. Hence the chip can be easily positioned and translated in the X-ray beam to bring the crystals in diffraction position. The data collection strategy needs to be adapted depending on beamline characteristics and on crystal properties. In the case of the CCA-adding enzyme, data were collected at X06DA and X10SA beamlines, Swiss Light Source (SLS), with an X-ray wavelength of 1.0 Å and Pilatus 2M-F and 6M pixel detectors, respectively. 30-60° of rotation were collected on each crystal at room temperature with images of 0.1° or 0.2° and 0.1 s exposure (see **Table 1**). Partial datasets were processed individually and cut when the resolution of diffraction patterns started to decay due to radiation damage (detected by the decrease of signal-to-noise ratio $\langle I/\sigma(I)\rangle$ and $CC_{1/2}$, and an increase of R_{meas} in the high resolution shell). Full datasets were reconstituted by merging data from 5 crystals (**Table 1**). Crystal structures were derived by molecular replacement using standard crystallographic packages and procedures for data processing[11] and refinement[12]. The comparison of the structures of the enzyme and of its complex with CMPcPP reveals the local conformational adaptation that accompanies substrate binding in the active site of the CCA-adding enzyme (**Figure 7**).

Copyright © 2021 JoVE Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported License

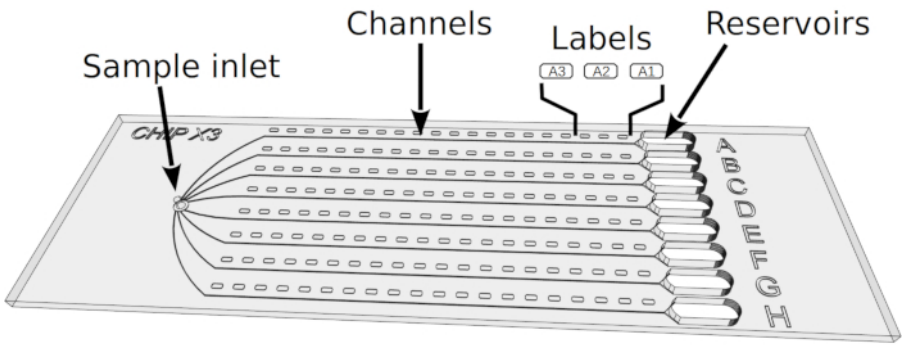


Figure 1: ChipX design. The chip consists of a top layer made of COC (thickness: 1 mm) in which eight microfluidic channels and reservoirs are imprinted. The entire chip is sealed with a layer of COC (thickness: 0.1 mm). All channels are connected to a single inlet on the left hand side for simultaneous sample injection and to individual reservoirs on the right hand side in which crystallization solutions are deposited. The channels, which constitute the actual crystallization chambers of the chip, are 4 cm long and have a cross section of 80 µm x 80 µm. Labels (A1, A2, A3, etc.) embossed along the channels facilitate crystal positioning under the microscope and the preparation of a sample list for data collection. ChipX has the size of a standard microscope slide (7.5 cm x 2.5 cm). Please click here to view a larger version of this figure.

Copyright © 2021 JoVE Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported License

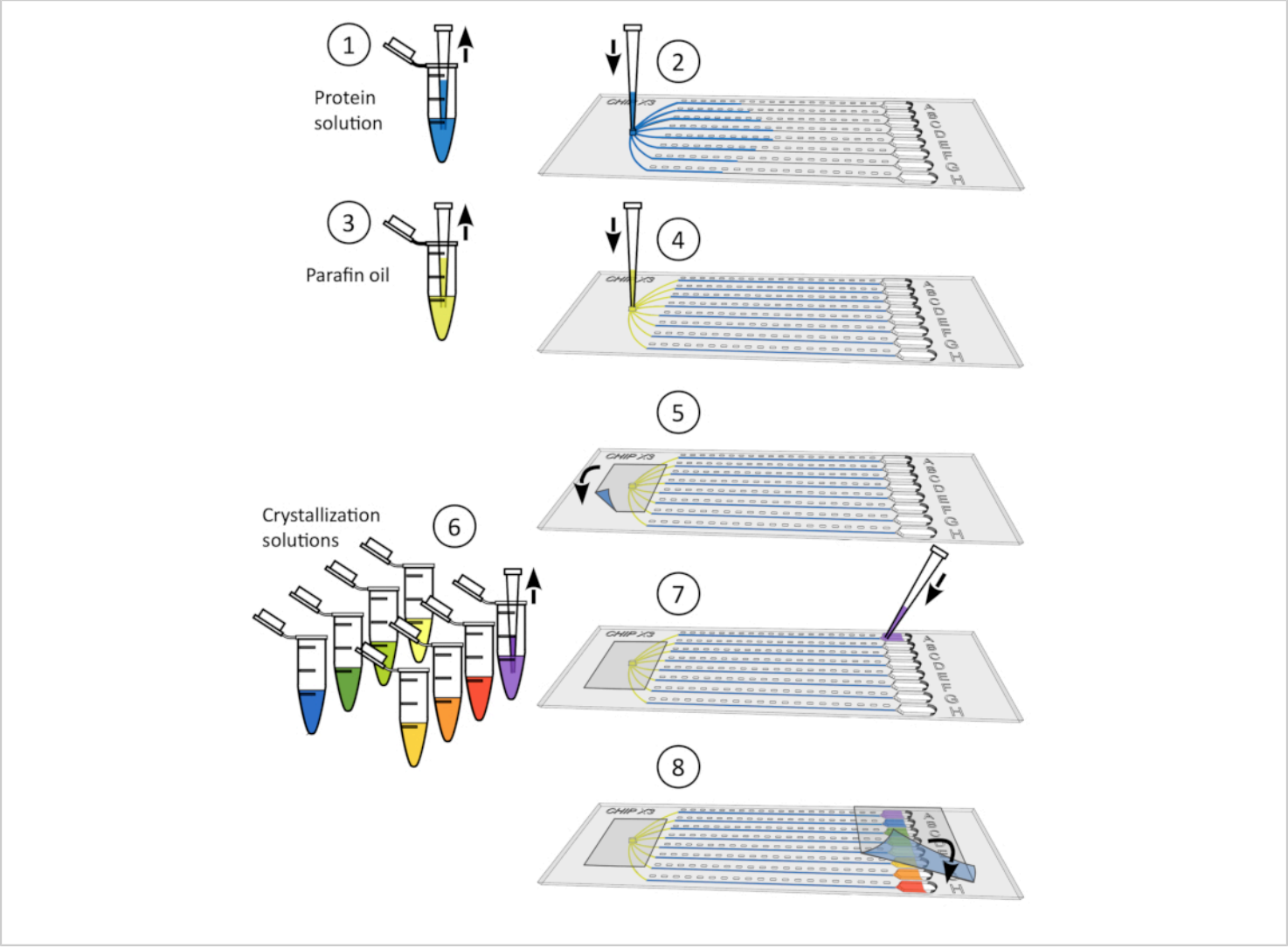


Figure 2: Setting up crystallization assays in ChipX. 1) Deposit 5-6 µL of enzyme solutions using standard 10 µL pipet and tip. 2) Introduce the tip vertically in the sample inlet and inject the solution in the eight channels. 3) Pipet 1 µL of paraffin oil. 4) Introduce the tip vertically in the sample inlet and inject the oil in order to disconnect the channels from each other. 5) Seal the inlet with a piece of tape. 6) Pipet 5 µL of crystallization solution using standard 10 µL pipet and tip. Solutions can be different in every reservoir (e.g., from a screening kit). 7) Orient the pipet tip towards the entry of the channel in the funnel shaped part of the reservoir (to avoid the formation of an air bubble upon solution deposition) and inject the crystallant solution in the reservoir. 8) Seal the reservoirs with a piece of tape and incubate the chip at controlled temperature. Please click here to view a larger version of this figure.

Copyright © 2021 JoVE Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported License

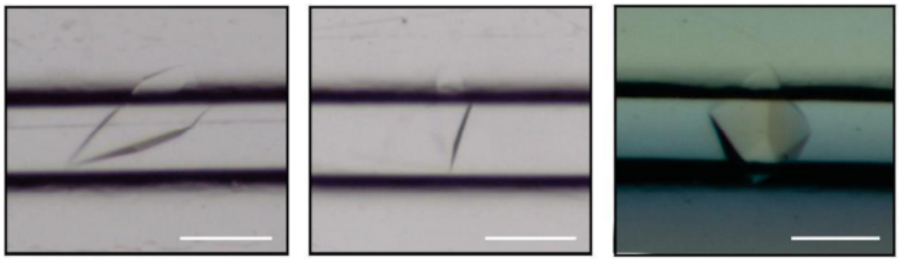


Figure 3: Crystals of CCA-adding enzyme grown by counter-diffusion in the microfluidic channels of ChipX. Scale bar is 0.1 mm. Please click here to view a larger version of this figure.

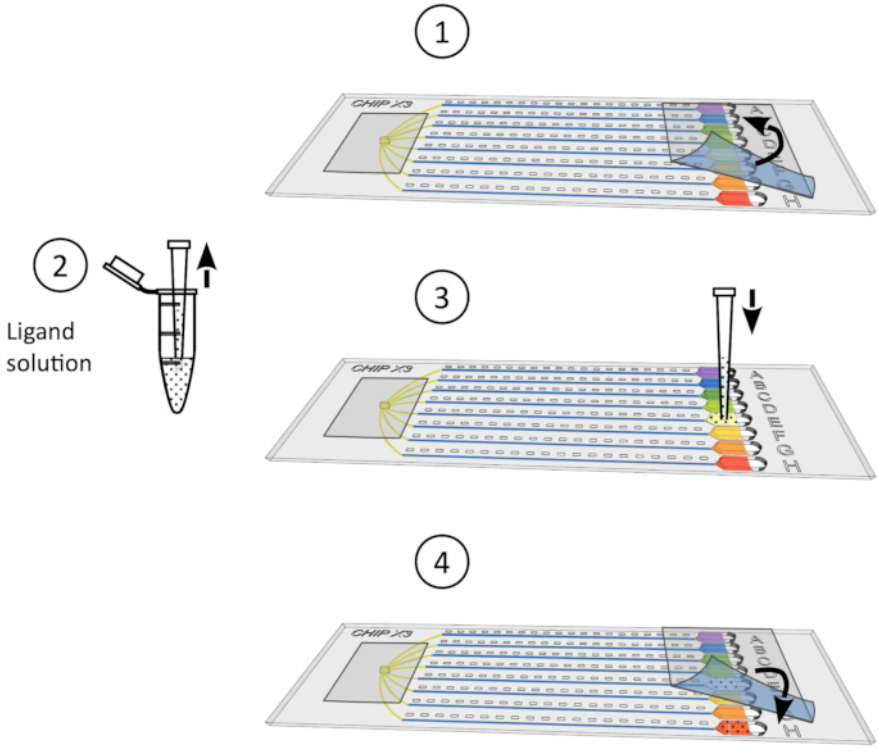


Figure 4: Crystal soaking procedure. 1) Gently remove the tape from the reservoirs. 2) Deposit up to 5 µL of ligand solution using a 10 µL micropipet. 3) Add the ligand to one or several reservoirs. 4) Seal again the reservoirs with a piece of tape and incubate the chip under controlled temperature for 24-48 h before data collection. Please click here to view a larger version of this figure.

Copyright © 2021 JoVE Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported License

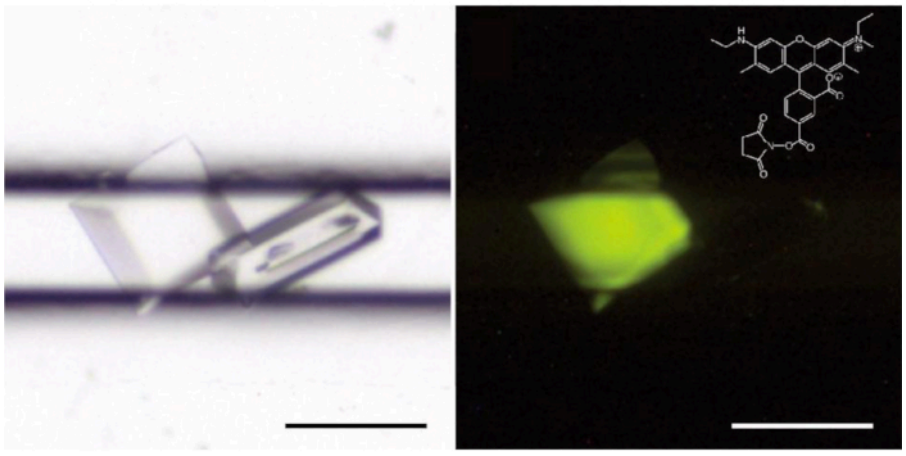


Figure 5: Trace fluorescent labeling discriminates protein (left) from salt (right) crystals. The CCA-adding enzyme solution contained 0.4 % (w/w) of protein labeled with carboxyrhodamine. On the right, crystals are illuminated with a 520 nm wavelength light source and the image is taken with a low pass filter at 550 nm (LP550); (inset) structure of carboxyrhodamine-succinimidyl ester. Please click here to view a larger version of this figure.

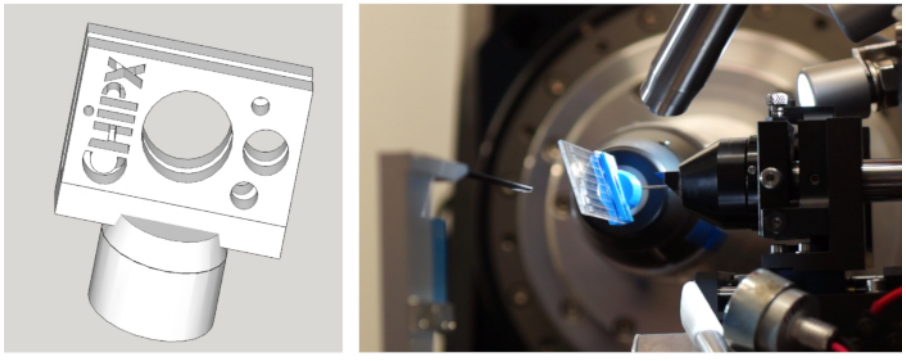


Figure 6: (Left) Drawing of the ChipX holder and (Right) ChipX mounted on the goniometer of beamline X06DA at SLS (Villigen, Switzerland) for serial crystal analysis. Please click here to view a larger version of this figure.

Copyright © 2021 JoVE Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported License

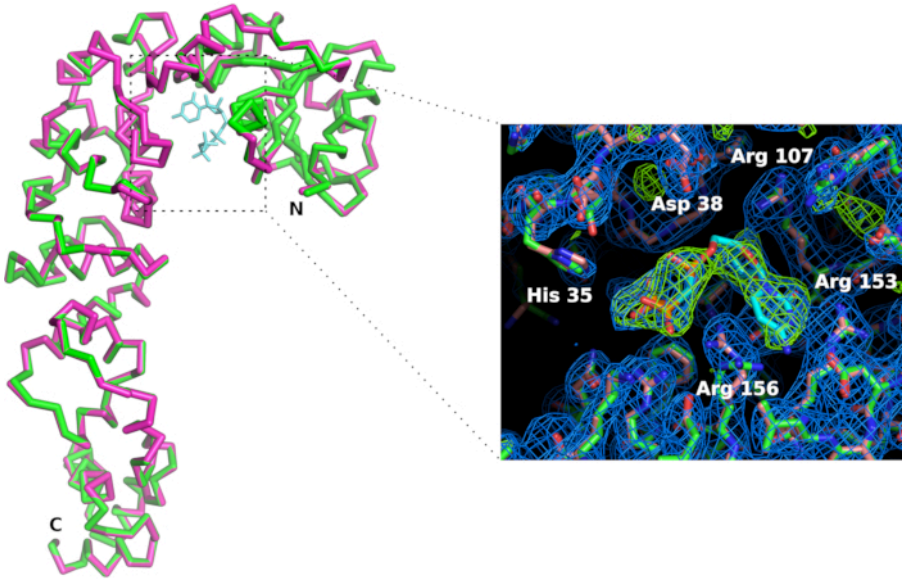


Figure 7: Comparison of CCA-adding enzyme active site in the apo form (in pink) and in the complex with a CTP analog (in green). Although the overall conformation of the enzyme is not affected, the binding of the CMPcPP ligand is accompanied by a slight reorganization of side chains in the active site. The 2*Fo-Fc* electron density map (in blue) is contoured at 1.2 sigma. The difference electron density map contoured at 4 sigma (in green) confirms the presence of the ligand in the active site. Please click here to view a larger version of this figure.

Copyright © 2021 JoVE Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported License

Crystallized sample	CCA-adding enzyme	CCA-adding enzyme + CMPcPP
Crystal analysis		
X-ray beamline	SLS – X06DA	SLS – X10SA
Wavelength (Å)	1.000	1.000
Temperature (K)	293	293
Detector	Pilatus 2M-F	Pilatus 6M
Crystal-detector distance (mm)	300	400
Crystals collected	6	14
Crystals selected	5	5
Rotation range per image (°)	0.1	0.2
Exposure time per image (s)	0.1	0.1
No. of images selected	1000	540
Total rotation range (°)	100	108
Space group	$P4_32_12$	$P4_32_12$
a, c (Å)	71.5, 293.8	71.4, 293.6
Mean mosaicity (°)	0.04	0.04
Resolution range (Å)	46 – 2.54 (2.6 – 2.54)	48 – 2.3 (2.4 – 2.3)
Total No. of reflections	176105 (9374)	232642 (32937)
No. of unique reflections	23922 (1598)	34862 (4066)
Completeness (%)	90.6 (84.6)	99.5 (100.0)
Redundancy	7.5 (6.0)	6.7 (8.1)
$\langle I/\sigma(I)\rangle$	8.1 (1.3)	6.9 (0.7)
Rmeas (%) §	18.6 (126.0)	18.0 (231.2)
CC1/2 (%) £	98.7 (55.0)	98.7 (46.9)
Overall B factor from Wilson plot ($Å^2$)	57.4	60.6
Crystallographic refinement		
No. of reflections, working set / test set	23583 / 1180	34840 / 3405

Copyright © 2021 JoVE Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported License

Final Rcryst (%) / R_{free} (%)	18.8 / 21.4	20.0 / 22.9
No. of non-H atoms: overall / protein / ligand / solvent	2998 / 2989 / 0 / 9	3057 / 2989 / 29 / 10
R.m.s. deviations for bonds (Å) / angles (°)	0.009 / 1.23	0.010 / 1.22
Average B factors ($Å^2$): overall / protein / ligand / solvent	60.1 / 60.1 / 0 / 52.7	62.5 / 62.6 / 60.1 / 55.5
Ramachandran plot: most favored (%) / allowed (%)	98.1 / 1.9	97.2 / 2.8
PDB id	6IBP	6Q52

Table 1: Data collection and refinement statistics

§ Redundancy-independent $Rmeas = \Sigma_{hkl}(N/N-1)^{1/2}\Sigma_i \mid I_i(hkl) - \langle I(hkl)\rangle \mid / \Sigma_{hkl}\Sigma_i I_i(hkl)$, where N is the data multiplicity [17].

£ Data with low <I/σ(I)> in outer shell (<2.0) were included based on CC1/2 criterion (correlation between two random halves of the dataset > 50%) as proposed by Karplus & Diederichs [18].

Discussion

Current protocols in biocrystallography involve the preparation of crystals using methods such as vapor diffusion or batch[13, 14], and their transfer into a microloop for cryo-cooling[15, 16] before performing the diffraction analysis in a nitrogen jet at cryogenic conditions. In contrast, direct crystal cryo-cooling is not possible in ChipX[3] and crystals cannot be extracted from their microfluidic channel, which can be seen as limitations of this setup. However, the protocol described in the article provides a fully integrated pipeline for the determination of crystal structures at room temperature (i.e., in more physiological conditions). Even though data collection at room-temperature causes an increase of radiation damage[19], this effect is counterbalanced by fast data acquisition time (a maximum of 60° rotation is collected on each crystal) and by merging of several partial datasets. Both ChipX design and material were optimized to reduce background scattering and diffraction signal attenuation[3], and data collection can be performed on crystals with dimensions equivalent to half the size of the channels (40 µm)[4].

To summarize, the main advantages of the protocol are the following. The crystals are produced in a convection-free environment (microfluidic channels), which is very favorable to the growth of high-quality crystals. The counter-diffusion method implemented in ChipX is very efficient at screening the supersaturation landscape; the diffusion of crystallants into the chip channel creates a concentration and supersaturation wave that helps determine appropriate nucleation and growth conditions[5]. Crystals are never directly handled, but are analyzed in situ, inside the chip, which preserves their genuine diffraction properties (i.e.,

Copyright © 2021 JoVE Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported License

does not alter crystal mosaicity by physical interaction or cryocooling)[20]. The diffraction analysis is performed on a series of crystals distributed along the chip channels with low dose exposure to minimize radiation damage, and a full dataset is assembled by merging partial data from the series. The standard footprint and simple design of ChipX will allow in the future a complete automation of *in situ* data collection using synchrotron or XFEL facilities. All steps of the protocol are carried out in ChipX. From the experimenter point-of-view, chip setup is simple and easy to perform with standard pipets and does not require any extra equipment. The tree-like channel connection at the sample inlet minimizes dead volumes in the system, which is important when working with samples that are difficult to purify or that are only available in limited quantity.

In conclusion, the lab-on-a-chip approach implemented in ChipX simplifies and efficiently miniaturizes the process of crystallization by counter-diffusion and crystal structure determination, allowing to go from the sample to its 3D structure in a single device. It is widely applicable and offers a user-friendly, cost-effective solution for routine serial biocrystallography investigations at room temperature.

Disclosures

The authors have nothing to disclose.

Acknowledgments

The authors acknowledge the Swiss Light Source (Villigen, Switzerland) for beamtime allocation to the project on beamlines X10SA (PXII) and X06DA (PXIII), Alexandra Bluhm for her contribution to structure refinement, Clarissa Worsdale for the recording of the voiceover and François Schnell (Université de Strasbourg) for his assistance in video editing and SFX. This work was supported by the French Centre National de la Recherche Scientifique (CNRS), the University of Strasbourg, the LabEx consortium "NetRNA" (ANR-10-LABX-0036_NETRNA), a PhD funding to R.dW from the Excellence initiative (IdEx) of the University of Strasbourg in the frame of the French National Program "Investissements d'Avenir", a PhD funding to K.R. from the French-German University (UFA-DFH, grant no. CT-30-19), the Deutsche Forschungsgemeinschaft (grant no. Mo 634/10-1). The authors benefitted from the PROCOPE Hubert Curien cooperation program (French Ministry of Foreign Affair and Deutscher Akademischer Austauschdienst).

References

1. Giegé, R., Sauter, C. Biocrystallography: past, present, future. *HFSP Journal.* **4** (3-4), 109-121 (2010).
2. Dhouib, K. et al. Microfluidic chips for the crystallization of biomacromolecules by counter-diffusion and on-chip crystal X-ray analysis. *Lab on a Chip.* **9** (10), 1412-1421 (2009).
3. Pinker, F. et al. ChipX: A Novel Microfluidic Chip for Counter-Diffusion Crystallization of Biomolecules and in Situ Crystal Analysis at Room Temperature. *Crystal Growth & Design.* **13** (8), 3333-3340 (2013).
4. de Wijn, R. et al. A simple and versatile microfluidic device for efficient biomacromolecule crystallization and structural analysis by serial crystallography. *IUCrJ.* **6** (3), 454-464S2052252519003622 (2019).
5. García-Ruiz, J.Ma. et al. A supersaturation wave of protein crystallization. *Journal of Crystal Growth.* **232** (1-4), 149-155 (2001).
6. Otálora, F., Gavira, J.A., Ng, J.D., García-Ruiz, J.M. Counterdiffusion methods applied to protein

Copyright © 2021 JoVE Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported License

crystallization. *Progress in Biophysics and Molecular Biology.* **101** (1-3), 26-37 (2009).

7. Pusey, M., Barcena, J., Morris, M., Singhal, A., Yuan, Q., Ng, J. Trace fluorescent labeling for protein crystallization. *Acta Crystallographica Section F Structural Biology Communications.* **71** (7), 806-814 (2015).
8. Meyer, A., Betzel, C., Pusey, M. Latest methods of fluorescence-based protein crystal identification. *Acta Crystallographica Section F Structural Biology Communications.* **71** (2), 121-131 (2015).
9. Betat, H., Rammelt, C., Mörl, M. tRNA nucleotidyltransferases: ancient catalysts with an unusual mechanism of polymerization. *Cellular and Molecular Life Sciences.* **67** (9), 1447-1463 (2010).
10. Ernst, F.G.M., Erber, L., Sammler, J., Jühling, F., Betat, H., Mörl, M. Cold adaptation of tRNA nucleotidyltransferases: A tradeoff in activity, stability and fidelity. *RNA Biology.* **15** (1), 144-155 (2018).
11. Kabsch, W. XDS. *Acta Crystallographica. Section D, Biological Crystallography.* **66** (2), 125-132 (2010).
12. Adams, P.D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica. Section D, Biological Crystallography.* **66** (2), 213-221 (2010).
13. Dessau, M.A., Modis, Y. Protein Crystallization for X-ray Crystallography. *Journal of Visualized Experiments.* **47** (2011).
14. Sauter, C., Lorber, B., McPherson, A., Giegé, Richard Crystallization - General Methods. *In International Tables of Crystallography, Vol. F, Crystallography of Biological Macromolecules (2nd edition), E. Arnold, D.M. Himmel & M.G. Rossmann (eds).* 99-120 (2012).
15. Garman, E. "Cool" crystals: macromolecular cryocrystallography and radiation damage. *Current Opinion in Structural Biology.* **13** (5), 545-551 (2003).
16. Li, D., Boland, C., Aragao, D., Walsh, K., Caffrey, M. Harvesting and Cryo-cooling Crystals of Membrane Proteins Grown in Lipidic Mesophases for Structure Determination by Macromolecular Crystallography. *Journal of Visualized Experiments.* **67** (2012).
17. Diederichs, K., Karplus, P.A. Improved R-factors for diffraction data analysis in macromolecular crystallography. *Nature Structural Biology.* **4** (4), 269-275 (1997).
18. Karplus, P.A., Diederichs, K. Linking Crystallographic Model and Data Quality. *Science.* **336** (6084), 1030-1033, (2012).
19. de la Mora, E., Coquelle, N., Bury, C.S., Rosenthal, M., Holton, J. M., Carmichael, I., Garman, E. F., Burghammer, M., Colletier, J-P., and Weik, M. Radiation damage and dose limits in serial synchrotron crystallography at cryo- and room temperatures. *Proceedings of the National Academy of Sciences of the United States of America.* **117** (8), 4142-4151 (2020).
20. Nave, C. A., Description of Imperfections in Protein Crystals. *Acta Crystallographica. Section D, Biological Crystallography.* **54** (5), 848-853 (1998).

Copyright © 2021 JoVE Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported License

VII. CCA-addition in the cold: Structural characterization of the psychrophilic CCA-adding enzyme from the permafrost bacterium *Planococcus halocryophilus*

Since the discovery and first studies on extremophile (living in hot, acidic, cold or rich heavy metal environments) organisms in the 1970's, many questions were raised regarding the adaptation of molecular mechanisms allowing life in such conditions. To date, the main trend for cold adaptation stipulates that an increased flexibility should confer enzymes the capacity to fulfill their role at low temperature (De Maayer *et al.*, 2014; Feller & Gerday, 2003). Within all living organisms, DNA and RNA polymerization constitute an essential pillar for cell survival and requires elements acting with both a defined flexibility and high fidelity. Thus, CCA-adding enzymes are perfect examples of tRNA nucleotidyltransferases operating via structural rearrangement of functional domains during the catalytic process (Tomita *et al.*, 2004).

In the following chapter, we explored the functional and structural properties of the CCA-adding enzyme from *Planococcus halocryophilus* (PhaCCA), a psychrophilic bacterium isolated from Canadian Arctic permafrost (Mykytczuk *et al.*, 2012). Previous biochemical investigations (e.g., affinity for substrates, activity and efficiency a different temperatures) published by the Leipzig team in 2017 showed a higher error-rate of PhaCCA compared to meso- and thermophilic homologs (Ernst *et al.*, 2018). Shortly after, high resolution crystal structures (1.8 Å) in its apo form as well as in complex with CTP were solved by the same collaborators (De Wijn *et al.*, 2018). Here, we provide a detailed comparative study of PhaCCA structure versus the first structure of this family ever solved, that of the closely related enzyme from thermophilic *Geobacillus stearothermophilus* (GstCCA) (Li *et al.*, 2002).

As expected our results showed an overall increased flexibility of the cold-adapted enzyme, mostly affecting the non-conserved C-terminal side where a reduced amount (14%) of secondary structures was observed. In addition, the N-terminal motif C, acting as a spring element for head and neck reorientation during nucleotide switching, looses a helix capping element generally present on helix α5. In GstCCA this capping element is composed of two arginines (R137, R138) forming ionic bridges with E134, hence stabilizing the whole area. In PhaCCA these arginines are replaced by a lysine (K133) and an asparagine (N134) which are not able to interact with other side chains. Taken together our high-resolution structures could pinpoint adaptation strategies used by RNA polymerases in extreme cold conditions. Our results have recently been published in Computational and Structural Biotechnology Journal.

Depuis la découverte et les premières études sur les organismes extrêmophiles (vivant dans des environnements chauds, acides, froids ou riches en métaux lourds) dans les années 1970, de nombreuses questions ont été soulevées concernant l'adaptation des mécanismes moléculaires permettant la vie dans de telles conditions. A ce jour, la principale tendance pour l'adaptation au froid stipule qu'une flexibilité accrue devrait conférer aux enzymes la capacité de remplir leur rôle à basse température (De Maayer *et al.* , 2014 ; Feller & Gerday, 2003). Au sein de tous les organismes vivants, la polymérisation de l'ADN et de l'ARN constitue un pilier essentiel à la survie cellulaire et nécessite des éléments agissant à la fois avec une flexibilité définie et une haute fidélité. Ainsi, les enzymes d'addition de CCA sont de parfaits exemples de nucléotidyltransférases d'ARNt opérant par réarrangement structurel de domaines fonctionnels au cours du processus catalytique (Tomita *et al.* , 2004).

Dans le chapitre suivant, nous avons exploré les propriétés fonctionnelles et structurelles de l'enzyme d'addition du CCA de *Planococcus halocryophilus* (PhaCCA), une bactérie psychrophile isolée du pergélisol de l'Arctique canadien (Mykytczuk *et al.* , 2012). Des études biochimiques antérieures (par exemple, affinité pour les substrats, activité et efficacité à différentes températures) publiées par l'équipe de Leipzig en 2017 ont montré un taux d'erreur plus élevé de PhaCCA par rapport aux homologues méso- et thermophiles (Ernst *et al.* , 2018). Peu après, des structures cristallines à haute résolution (1,8 Å) sous sa forme apo ainsi qu'en complexe avec le CTP ont été résolues par les mêmes collaborateurs (De Wijn *et al* , 2018). Nous présentons ici une étude comparative détaillée de la structure de PhaCCA par rapport à la première structure de cette famille jamais résolue, celle de l'enzyme étroitement apparentée de *Geobacillus stearothermophilus* thermophile (GstCCA) (Li *et al.*, 2002).

Comme prévu, nos résultats ont montré une flexibilité globale accrue de l'enzyme adaptée au froid, affectant principalement le côté C-terminal non conservé où une quantité réduite (14%) de structures secondaires a été observée. En outre, le motif C N-terminal, agissant comme un élément de ressort pour la réorientation de la tête et du cou pendant le changement de nucléotide, perd un élément de coiffage de l'hélice (“*helix capping*”) généralement présent sur l'hélice α5. Dans GstCCA, cet élément de coiffage est composé de deux arginines (R137, R138) formant des ponts ioniques avec E134, stabilisant ainsi toute la zone. Dans PhaCCA ces arginines sont remplacées par une lysine (K133) et une asparagine (N134) qui ne sont pas capables d'interagir avec d'autres chaînes latérales. Prises ensemble,

nos structures à haute résolution pourraient mettre en évidence les stratégies d'adaptation utilisées par les ARN polymérases dans des conditions de froid extrême. Nos résultats ont récemment été publiés dans la revue *Computational and Structural Biotechnology Journal*.

Seit der Entdeckung und den ersten Studien über extremophile (in heißen, sauren, kalten oder schwermetallreichen Umgebungen lebende) Organismen in den 1970er Jahren wurden viele Fragen zur Anpassung der molekularen Mechanismen aufgeworfen, die ein Leben unter solchen Bedingungen ermöglichen. Bis heute besagt der Haupttrend zur Kälteanpassung, dass eine erhöhte Flexibilität Enzymen die Fähigkeit verleihen sollte, ihre Aufgabe bei niedrigen Temperaturen zu erfüllen (De Maayer *et al.*, 2014; Feller & Gerday, 2003). In allen lebenden Organismen stellt die DNA- und RNA-Polymerisation eine wesentliche Säule für das Überleben der Zelle dar und erfordert Elemente, die sowohl mit einer bestimmten Flexibilität als auch mit einer hohen Zuverlässigkeit arbeiten. So sind CCA-addierende Enzyme perfekte Beispiele für tRNA-Nukleotidyltransferasen, die während des katalytischen Prozesses durch strukturelle Umstrukturierung funktioneller Domänen arbeiten (Tomita *et al.*, 2004).

Im folgenden Kapitel untersuchten wir die funktionellen und strukturellen Eigenschaften des CCA-addierenden Enzyms aus *Planococcus halocryophilus* (PhaCCA), einem psychrophilen Bakterium, das aus dem arktischen Permafrost Kanadas isoliert wurde (Mykytczuk *et al.* , 2012). Frühere biochemische Untersuchungen (z.B. Affinität für Substrate, Aktivität und Effizienz bei verschiedenen Temperaturen), die das Leipziger Team 2017 veröffentlichte, zeigten eine höhere Fehlerrate von PhaCCA im Vergleich zu meso- und thermophilen Homologen (Ernst *et al.* , 2018). Kurz darauf wurden von denselben Mitarbeitern hochauflösende Kristallstrukturen (1,8 Å) in seiner apo-Form sowie im Komplex mit CTP gelöst (De Wijn *et al.* , 2018). Hier bieten wir eine detaillierte vergleichende Studie der PhaCCA-Struktur mit der ersten jemals gelösten Struktur dieser Familie, der des eng verwandten Enzyms aus dem thermophilen *Geobacillus stearothermophilus* (GstCCA) (Li *et al.*, 2002).

Wie erwartet zeigten unsere Ergebnisse eine insgesamt erhöhte Flexibilität des kälteangepassten Enzyms, die sich hauptsächlich auf die nicht konservierte C-terminale Seite auswirkt, wo eine reduzierte Anzahl (14 %) von Sekundärstrukturen beobachtet wurde. Darüber hinaus verliert das N-terminale Motiv C, das als Federelement für die Neuausrichtung von Kopf und Hals während des Nukleotid-Spezifitätswechsels fungiert, ein Helix-Capping-Element, das normalerweise auf Helix α5 vorhanden ist. In GstCCA besteht dieses Kappen Element aus zwei Argininen (R137, R138), die mit E134 eine Ionenbrücke bilden und somit den gesamten Bereich stabilisieren. In PhaCCA sind diese Arginine durch

ein Lysin (K133) und ein Asparagin (N134) ersetzt, die nicht in der Lage sind, mit anderen Seitenketten zu interagieren. Zusammengenommen könnten unsere hochauflösenden Strukturen Anpassungsstrategien aufzeigen, die von RNA-Polymerasen unter extremen Kältebedingungen genutzt werden. Unsere Ergebnisse wurden kürzlich im *Computational and Structural Biotechnology Journal* veröffentlicht.

Computational and Structural Biotechnology Journal 19 (2021) 5845–5855

COMPUTATIONAL AND STRUCTURAL BIOTECHNOLOGY JOURNAL

journal homepage: www.elsevier.com/locate/csbj

CCA-addition in the cold: Structural characterization of the psychrophilic CCA-adding enzyme from the permafrost bacterium *Planococcus halocryophilus*

Raphaël de Wijn [a,1], Kévin Rollet [a,b,1], Felix G.M. Ernst [b], Karolin Wellner [b], Heike Betat [b], Mario Mörl [b,*], Claude Sauter [a,*]

[a] Architecture et Réactivité de l'ARN, Université de Strasbourg, CNRS, IBMC, 67084 Strasbourg, France
[b] Institute for Biochemistry, Leipzig University, Brüderstr. 34, 04103 Leipzig, Germany

ARTICLE INFO

Article history:
Received 14 June 2021
Received in revised form 12 October 2021
Accepted 12 October 2021
Available online 21 October 2021

Keywords:
CCA-adding enzyme
tRNA
Cold adaptation
Psychrophilic protein
X-ray crystallography
SAXS
Psychrophilic RNA polymerase

ABSTRACT

CCA-adding enzymes are highly specific RNA polymerases that add and maintain the sequence C-C-A at tRNA 3'-ends. Recently, we could reveal that cold adaptation of such a polymerase is not only achieved at the expense of enzyme stability, but also at the cost of polymerization fidelity. Enzymes from psychrophilic organisms usually show an increased structural flexibility to enable catalysis at low temperatures. Here, polymerases face a dilemma, as there is a discrepancy between the need for a tightly controlled flexibility during polymerization and an increased flexibility as strategy for cold adaptation. Based on structural and biochemical analyses, we contribute to clarify the cold adaptation strategy of the psychrophilic CCA-adding enzyme from *Planococcus halocryophilus*, a gram-positive bacterium thriving in the arctic permafrost at low temperatures down to −15 °C. A comparison with the closely related enzyme from the thermophilic bacterium *Geobacillus stearothermophilus* reveals several features of cold adaptation - a significantly reduced amount of alpha-helical elements in the C-terminal tRNA-binding region and a structural adaptation in one of the highly conserved catalytic core motifs located in the N-terminal catalytic core of the enzyme.

© 2021 The Author(s). Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

Psychrophilic bacteria are defined as cold-adapted microorganisms thriving at temperatures below 15 °C and tolerating 0 °C [1]. To be functional at temperatures at the freeze point or below, enzymes of such psychrophiles require specific evolutionary adaptations [2]. A widespread strategy is the destabilization of the enzyme/substrate complex to reduce the required activation energy [3–5]. However, such a destabilization is based on a higher structural flexibility of the enzyme, resulting in an increased temperature sensitivity of cold-adapted enzymes [4]. In general, enzymes do not rely on a uniform strategy for cold adaptation. Rather, structural flexibility can be achieved by a number of factors ranging from a decreased content of secondary structure elements to extended loop regions. In addition, the amount of disulfide bridges, electrostatic interactions and hydrogen bonds can be reduced, while amino acids promoting flexibility (glycine) or interfering with helix formation (proline) are enriched [3,4,6–10].

In DNA and RNA polymerases, a defined conformational flexibility is of vital importance for proper function and correct nucleotide addition [11–14]. Accordingly, an increased flexibility as a strategy for cold adaptation might interfere with polymerization speed and fidelity. A highly specialized RNA polymerase with an outstanding accuracy is the CCA–adding enzyme (tRNA nucleotidyltransferase) that adds and maintains the triplet sequence C-C-A at the 3′–end of tRNAs [15–17]. As members of the polymerase β group, CCA-adding enzymes are classified into class I (archaeal enzymes) and class II (bacterial and eukaryotic enzymes) [15,18,19]. During synthesis of the CCA sequence, the latter group exhibits defined structural rearrangements in the catalytic core that are required for correct NTP recognition and restricted addition of three positions [12,14]. Both CTP and ATP are bound in a single nucleotide binding pocket, where highly conserved amino acid residues form Watson-Crick-like hydrogen bonds to the base moiety of the

* Corresponding authors.
E-mail addresses: mario.moerl@uni-leipzig.de (M. Mörl), c.sauter@ibmc-cnrs.unistra.fr (C. Sauter).
[1] These authors contributed equally to the work.

https://doi.org/10.1016/j.csbj.2021.10.018
2001-0370/© 2021 The Author(s). Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology.
This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

incoming NTP [20]. To switch from CTP towards ATP incorporation, a defined reorientation of these residues is required in order to accommodate the ATP in a highly selective way. As these coordinated movements of the enzyme are essential for a correct and efficient CCA-addition, the increased enzyme flexibility required for activity at low temperatures leads to a conflict between cold adaptation and proper polymerization. Recently, we could show that the CCA-adding enzyme from *Planococcus halocryophilus* (*Pha*CCA), a bacterium from the arctic permafrost growing at −15 °C [21], is such a cold-adapted enzyme and catalyzes CCA-addition *in vitro* down to 0 °C [22]. Interestingly, both the N–terminal catalytic core as well as the C-terminal tRNA–binding region are adapted to the cold, indicating that the psychrophilic features are not restricted to a single protein domain, but are distributed over the whole enzyme. In addition, and consistent with the features of many cold-adapted enzymes, *Pha*CCA shows a decreased thermal stability. However, the price to be paid for this cold adaptation is not only a structural destabilization, but also a significant reduction in polymerization fidelity [22]. Here, we present biochemical and structural data contributing how the psychrophilic CCA–adding enzyme from *P. halocryophilus* is adapted to such low temperatures. Our results reveal that a combination of individual strategies – a global flexibility in the tRNA-binding C-terminus as well as a thermal adaptation in the highly conserved catalytic core motif C controlling interdomain movements during CCA-addition – contribute to the remarkable cold adaptation of this unusual RNA polymerase.

2. Materials and methods

2.1. Expression of recombinant enzymes

For biochemical analysis, the coding regions of interest were cloned into pET30Ek/Lic and recombinant enzymes were expressed as previously described with N-terminal 6xHis–tag in *E. coli* BL21 (DE3) *cca::cam*, lacking the endogenous CCA-adding enzyme in order to avoid any contaminating background CCA-adding activity [22]. Point mutations were introduced into the coding sequence of *P. halocryophilus* CCA-adding enzyme by site-directed mutagenesis using mutagenic primer pairs. Generation of the loop chimera was achieved via amplification of the respective loop coding region from the *G. stearothermophilus* CCA–adding enzyme gene to obtain a megaprimer for QuikChange site-directed mutagenesis in the *Pha*CCA open reading frame [12,22].

The pETDuet constructs provided both T7 promoter-based expression of *E. coli* RNase T as well as an autonomous expression platform including the native *cca* promotor for comparable expression of the recombinant tRNA nucleotidyltransferase and were established for *in vivo* complementation in *E. coli* strain JM109 (DE3) *cca::cam* as described [23]. Transformants were streaked in sectors onto LB agar plates supplemented with ampicillin (100 μg ml^{-1}) and were incubated for 2 days at the indicated temperatures.

2.2. Preparation and biochemical analysis of P. halocryophilus CCA-adding enzyme

Purification of the recombinant enzyme was done according to Ernst *et al.* [22] and stored at 4 °C in 50 mM TRIS/HCl, pH 7.5, 200 mM NaCl, 5 mM $MgCl_2$. Protein concentrations were determined using a NanoDrop ND–1000 spectrophotometer with extinction coefficients deduced from their amino acid composition (https://web.expasy.org/protparam/). Samples were ultracentrifuged for 1 h at 100,000 g prior to any further analysis or experiment, and their homogeneity at 20 °C was systematically verified by dynamic light scattering (DLS) using a Zetasizer Nano ZS instrument (Malvern) and a 20 μl quartz cuvette. The reported particle diameter corresponds to the Z–average for at least 5 measurements while fitting a unimodal distribution, as implemented in the instrument. *In vitro* CCA incorporation and CD spectroscopy were performed as described [22].

2.3. Electrophoretic mobility shift assay (EMSA)

Electrophoretic mobility shift assays were performed as recently described with minor deviations [24,25]. 0.3 pmol α–^{32}P–ATP-labeled yeast $tRNA^{Phe}$ lacking the CCA-terminus was heated for 2 min at 90 °C, cooled down to room temperature and incubated with 0 to 5 μM enzyme in 50 mM glycine/NaOH (pH 8.5), 10 mM $MgCl_2$ and 2 mM DTT for 10 min at temperatures where the enzymes show significant activity (4 °C and 30 °C for *Pha*CCA and 37 °C for *Gst*CCA) [22]. After addition of 80% m/v glycerol to a final concentration of 18.5% m/v, tRNAs were separated on a 5% m/v native polyacrylamide gel for 60 min at 3 W and 4 °C. To visualize complexed and free tRNA substrates, a Typhoon 9410 phosphorimager was used (Cytiva, Freiburg, Germany). Determination of the approximate dissociation constant of *Gst*CCA was performed by nonlinear regression using GraphPad Prism 7.

2.4. Crystallization of PhaCCA

*Pha*CCA crystals were obtained as described [26]. First diffracting crystals were grown by counter diffusion (CD) in capillaries plunged in 3 M ammonium sulfate, 0.1 M sodium acetate, pH 5.0, from a commercial CD kit (Triana Science & Technologies). Their diffraction patterns extended to a maximum of 2.65 Å and presented an ice-ring at about 3.7 Å. To improve data quality, another condition (1 M diammonium hydrogen phosphate, 100 mM sodium acetate pH 4.5) was used in combination with microseeding [26].

The seed stock was prepared as follows: three droplets of 1 μl containing large number of small crystals (<50 μm) were pooled, resuspended and crushed in 10 μl of the former solution with a standard P10 pipette. Large crystals (150–300 μm) were grown at 20 °C in 24–well plates (NeXtal Biotechnologies) within 1–3 weeks in 1–2 μl droplets made of 1/10th of seed stock, 4/10th of *Pha*CCA and 5/10th of reservoir solutions (in volume ratio).

Prior to data collection, crystals were soaked for 30 s in the reservoir solution supplemented with 20% m/v glycerol, mounted in MicroLoops (MiTeGen) and cryocooled in liquid nitrogen, except crystals grown by CD in ammonium sulfate which were directly cryocooled in their capillary. Crystals of enzyme:ligand complexes were obtained by a preliminary 30 s soaking in the mother liquor containing either 5 mM of CTP or 10 mM of CMPcPP (non-hydrolyzable CTP analog; Jena Bioscience, catalog number NU-438S).

2.5. Crystallographic analysis of PhaCCA

Crystal characterization was performed at X06DA beamline, Swiss Light Source (Villigen, Switzerland), equipped with a PILATUS 2 M detector or at Proxima 1 beamline, Soleil synchrotron (Saint-Aubin, France), equipped with an Eiger 16 M detector. All data sets were processed and scaled with the XDS Package (Kabsch, 2010), see statistics in Table 1. A selection of 5% of reflections was put apart and used throughout refinement to calculate the R–free value.

The anomalous signal of sulfur atoms present in the native protein (which contains 11 methionine residues, no cysteine) was exploited to determine a set of initial phases using Autosol from PHENIX package [27,28]. A complete model was built and refined with Autobuild and Phenix_refine.

Table 1
Data collection and refinement statistics.

	CCA-adding enzyme SSAD	CCA-adding enzyme	CCA-adding enzyme + CTP	CCA-adding enzyme + CMPcPP	CCA-adding enzyme + AS
X-ray beamline	SLS - PXIII	PROXIMA1	PROXIMA1	PROXIMA1	SLS - PXIII
Wavelength (Å)	2.075	0.9786	0.9786	0.9786	1.000
Temperature (K)	100	100	100	100	100
Detector	Pilatus 2 M−F	Pilatus 6 M	Pilatus 6 M	Eiger 16 M	Pilatus 2 M−F
Space group	P4(3)2(1)2	P4(3)2(1)2	P4(3)2(1)2	P4(3)2(1)2	P4(3)2(1)2
a, *c* (Å)	69.79, 291.21	70.03, 291.59	70.38, 291.49	69.711, 290.342	70.0, 292.2
Solvent content (%)	66.6	66.8	67.2	67.2	66.9
Mean mosaicity (°)	0.08	0.04	0.05	0.11	0.18
Resolution range (Å)	50 – 2.25 (2.31 – 2.25)	49 – 1.80 (1.91 – 1.80)	50 – 1.85 (1.96 – 1.85)	50 – 2.2 (2.34 – 2.2)	50 – 2.80 (2.9 – 2.80)
Total No. of reflections	5,020,925 (116071)	890,444 (139573)	776857(123819)	965,891 (160098)	1,036,294 (104514)
No. of unique reflections	64,470 (4607)	68,555 (10787)	63965(10146)	37,600 (5903)	18,615 (1819)
Completeness (%)	99.0 (96.2)	99.9 (99.3)	99.8 (99.2)	99.8 (99.3)	98.6 (100.0)
Redundancy	78 (25)	13 (13)	12 (12)	25.7 (27.1)	55.7 (57.5)
$\langle I/\sigma(I)\rangle$	30.7 (0.9)	19.5 (0.7)	22.1 (1.0)	16.84 (1.98)	21.3 (3.1)
*R*meas (%)	17.0 (311.3)	7.4 (264.8)	6.9 (239.4)	15 (161.6)	25.8 (204)
CC1/2 (%)	99.9 (42.5)	100 (59.0)	100 (45.1)	99.9 (78.6)	99.9 (100.0)
B factor from Wilson plot (Å 2)	47.1	44.8	45.0	50.9	59.4
Reflections in working/test sets	34,968/1686	68,408/3420	63,868/3195	35,845/1886	17,701/901
Final R_{work}/R_{free} (%)	18.9/21.1	20.7/23.2	20.2/21.9	20.6/24.2	22.0/25.6
No. of non-H atoms	3251	3310	3398	3214	3030
Protein/solvent/ligand/ion	3062/147/42/0	3043/238/29	3072/268/58/0	3005/99/145/0	2089/21/20/0
R.m.s.d. bonds (Å)/angles (°)	0.008/0.96	0.007/0.794	0.008/0.979	0.011/1.11	0.009/1.05
Average *B* factors (Å^2)	48.2	44.0	44.0	51.9	53.3
biomolecule/solvent (Å^2)	48.0/49.9	/42.5	/45.9	51.3/52.7	53.2/46.9
Ramachandran plot regions: most favored, allowed (%)	98.4/1.4	99.0/1.0	99.0/1.0	98.4/1.4	98.1/1.7
PDB id	7OTR	6QY6	6QXN	7OQX	7OTL

Other structures were obtained by molecular replacement using the sulfur-SAD model in Phaser [28] and Coot [29] for manual inspection and rebuilding. They were all refined using isotropic Atomic Displacement Parameters (ADPs) and TLS (translation/libration/screw) constraints. Water molecules were added where appropriate, as well as small molecules present in the mother liquor (glycerol, TRIS, acetate, phosphate, sulfate ions). While CTP appeared in a single conformation, two conformations were necessary to model the electron density present in the active site of the enzyme in the case of CMPcPP.

In all models, the residue Val215 was systematically slightly off the Ramachandran and flagged as an outlier due to a local loop conformation. Also, no density could be observed for the sidechain of His35 in the neighborhood of the CTP binding pocket, as well as for the last four residues (374–377) and residues from the N–terminal expression tag (43 residues including the 6xHis-tag). Model building started between position −4 and −2 (corresponding to aspartic acid residues) depending on the quality of the electron density map. Structure analysis and molecular representations were performed with PyMOL Molecular Graphics System, Version 2.4 (Schrödinger, LLC). Structure-based sequence alignments and 2D structure analysis were performed using ENDscript webservice [30].

2.6. Small-angle X-ray scattering (SAXS)

SAXS experiments were performed on the SWING beamline at synchrotron SOLEIL (Saint-Aubin, France). The beam wavelength was λ = 1.033 Å. The EigerX4M detector (Dectris) was positioned at a distance of 2,000 mm from the sample with the direct beam off-centered. The resulting exploitable q-range was 0.005 – 0.5 Å^{-1}, where the wave vector $q = 4\pi \sin\theta/\lambda$ and 2θ is the scattering angle.

A *Pha*CCA solution (60 μl at a concentration of 2.9 mg/ml) was loaded onto a size exclusion column (Bio SEC-3 with 300 Å pore size, Agilent Technologies) and SAXS measurements were performed throughout elution operated at 15 °C and a flow rate of 0.3 ml/min with a mobile phase containing 100 mM HEPES/NaOH, pH 7.5, 150 mM NaCl and 10 mM $MgCl_2$. The eluate was analyzed by SAXS in a continuous flow capillary cell with a frame duration of 990 ms at intervals of 10 ms. Data processing, interpretation, Rg evaluation over elution profiles were performed using Foxtrot [31] and data analysis (determination of Rg, Dmax) with the ATSAS package [32]. A complete model including the N–terminal extension was generated using Modeller [33] based on the crystal structure. It was refined under SAXS constraints using DADIMODO [34], a genetic algorithm-based refinement program, to better fit the SAXS profile owing to a goodness–of–fit calculated with CRYSOL [35]. Graphics were prepared with LibreOffice Draw (The Document Foundation) and BioXTAS-RAW [36].

2.7. Nanoscale differential scanning fluorimetry (nanoDSF)

Comparative thermal stability assays were carried out by nanoDSF in a Prometheus NT.48 instrument (NanoTemper Technologies). CCA-adding enzymes were incubated on ice at a concentration of 10 nM in their storage buffer, alone or in the presence of 30 nM of either ATP, CTP or GTP 15 min prior to experiment. Glass capillaries were filled with 10 μl of each solution and deposited on the sample holder. Excitation wavelength was set to 280 nm and intrinsic tryptophan fluorescence was recorded at 330 and 350 nm at temperatures ranging from 20 to 95 °C (1 °C/min, 1 measurement per 0.044 °C). Corresponding first derivative plots were calculated using the manufacturer's software (PR.ThermControl, version 2.1.2) to determine melting temperature (Tm), along with those of the scattering signal highlighting the unfolding of the proteins.

3. Results

3.1. Structural characterization of PhaCCA

In order to investigate the structural properties of *Pha*CCA leading to cold adaptation, crystals of the corresponding recombinant enzyme were grown. The crystallized construct consisted of the

P. halocryophilus wild type sequence (377 amino acids) with an N–terminal extension of 43 residues including a 6xHis affinity tag. First crystals of *Pha*CCA were obtained in the presence of ammonium sulfate as crystallant and diffracted to a maximum resolution of 2.8 Å [26]. They were used as seeds to screen for new conditions and led to crystals with similar cell parameters and symmetry (see Table 1) but showing significantly improved diffraction properties (extending to a resolution of 1.8 Å) in the presence of PEG 3350 or diammonium hydrogen phosphate. The latter crystallant was selected in combination with microseeding to ensure a reproducible crystal production [26]. Sulfur-SAD data collection was performed to determine a set of phases and calculate an experimental electron density map, which allowed to automatically build > 90% of *Pha*CCA residues (Fig. S1A–B) and revealed the characteristic architecture of bacterial and eukaryotic CCA-adding enzymes (Fig. 1) and its seahorse shape popularized by Li *et al.* [20].

An ensemble of five structures was determined under cryoconditions, in apo state or in complex with a substrate in the presence of various salts (Fig. S1C–E). All crystals showed a high solvent content (67%) with one enzyme monomer in the asymmetric unit. The refined structures are very similar, showing a root mean-square deviation (rmsd) ranging from 0.16 to 0.22 Å (for 370 superimposed Cα positions) when compared to the reference apo structure solved at the highest resolution (Fig. S2). Their N–terminal extension is disordered and the first visible residue varied from position −4 to −2 at the end of the extension. The four C-terminal residues (HSHT) are also disordered, as well as the central part of the flexible loop (residues 82–93) in the head domain, except for the complex with CTP where it is clearly visible (Fig. S1C).

Fast crystal soaking (30 s) with CTP or its non-hydrolyzable analog prior to cryocooling readily led to the binding of the ligand in the enzyme's catalytic site (Fig. S3). In contrast, soaking or co–crystallization attempts with ATP were not successful.

Early crystal structures of CCA-adding enzymes from *G. stearothermophilus* and *H. sapiens* were reported as dimers [20,37], but dimer interfaces were differing, and the existence of a functional dimeric enzyme has never been confirmed. *Pha*CCA is unambiguously monomeric in its crystalline state as well as in solution, as shown by SEC and SAXS analyses (Fig. 1B). Gyration radius *Rg* derived for *Pha*CCA from the Guinier plot is 30.7 Å (Fig. S4), in agreement with the theoretical value calculated for the crystal structure of the monomer (28.2 Å) to which a flexible N-terminal extension is appended. Also, the Kratky plot is consistent with a molecule showing compact multidomain core and a flexible tail (Fig. S4).

3.2. Psychrophilic adaptation of PhaCCA

*Pha*CCA was compared to a close homolog from *Geobacillus stearothermophilus* (*Gst*CCA), a bacterium belonging to the order of *Bacillales* like *P. halocryophilus* [20]. The obtained thermal stability profiles determined by NanoDSF are typical for psychrophilic and thermophilic enzymes with Tms of ∼ 43 °C *vs* 73 °C, respectively (Fig. 2). These values are consistent with Tm values obtained by CD ([22] and below). The slight Tm variation observed between the two methods is likely due to the fact that they track different steps of unfolding (the loss of 2D structures in CD, the increase of solvent exposure and resulting fluorescence quenching of aromatic residues in NanoDSF) and was already described for other enzymes [38]. Because small ligands can potentially stabilize a protein fold upon binding and induce an increase in Tm, the analysis was reproduced in the presence of nucleotides. However, no

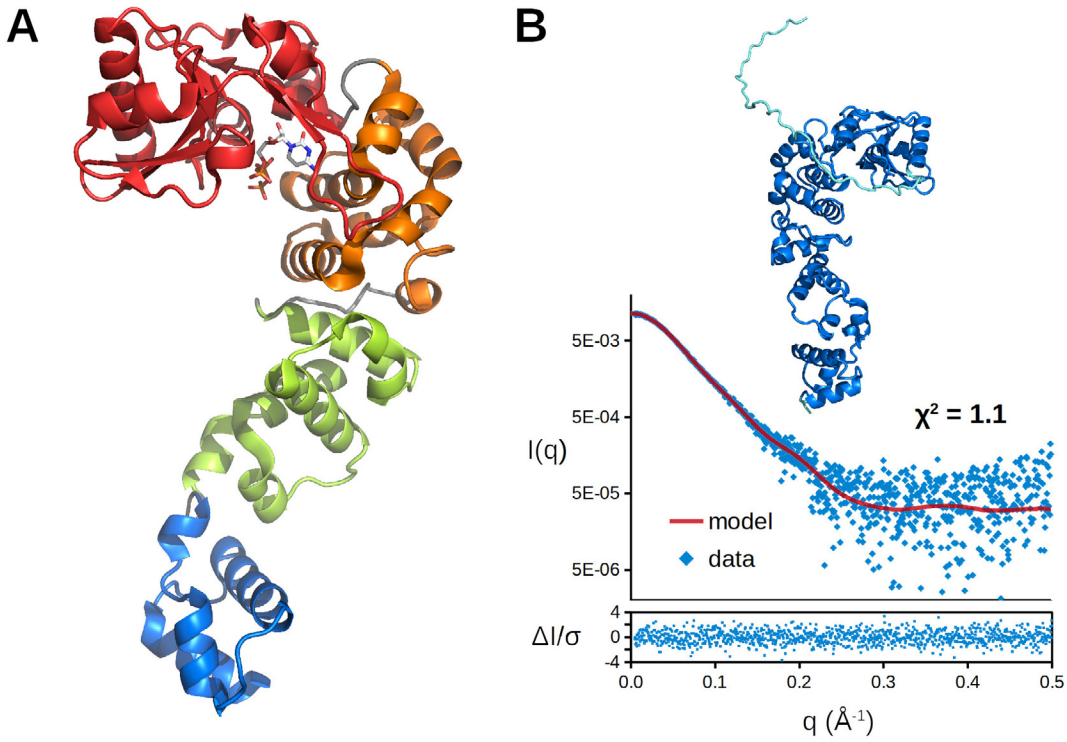
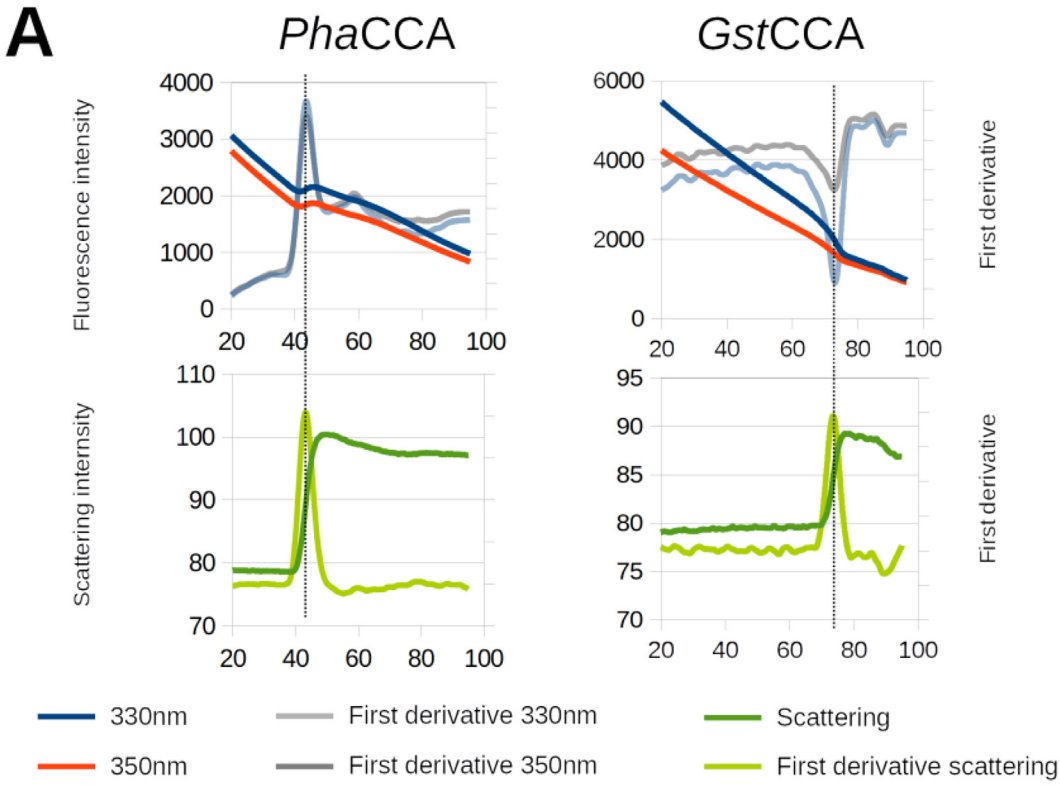


Fig. 1. Structural organization of psychrophilic *Pha*CCA. A) Crystal structure of *Pha*CCA in complex with CTP determined at 1.85 Å resolution. The enzyme monomer presents the characteristic seahorse shape similar to that described by Li et al. [20], with its head, neck, body and tail domains depicted in red, orange, green and blue, respectively. B) Solution structure of *Pha*CCA including the enzyme core (dark blue) and the flexible 43 residues long N–terminal expression tag (cyan). The theoretical SAXS profile (red line) corresponding to this monomeric model is in good agreement with the experimental data (blue dots), as indicated by the goodness-of-fit (χ^2). The error-weighted residual difference plot $\Delta I/\sigma = [I_{exp}(q) - I_{model}(q)]/\sigma(q)$ is given at the bottom. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)



B

Enzyme	Tm (°C)	+CTP	+ATP	+GTP
*Gst*CCA	73.4 ± 0.5	73.4 ± 0.8	73.1 ± 0.3	73.2 ± 0.3
*Pha*CCA	43.6 ± 0.3	45.1 ± 1.4	43.9 ± 0.4	43.4 ± 0.3

Fig. 2. Comparison of *Pha*CCA and *Gst*CCA thermal stability by NanoDSF. A) Representative curves are shown for the two apo enzymes with the variation of fluorescence at 330 and 360 nm, as well as the scattering signal, whose increase confirms protein unfolding. Melting temperatures (Tm) correspond to the inflection point of the first derivatives. B) Tm values in the absence and presence of nucleotides, and associated standard deviations calculated as the mean of three values (Tm from 330 nm, 350 nm fluorescence and from scattering) for at least 3 independent samples.

significant effect was observed with the natural substrates CTP and ATP or with GTP, which is not supposed to bind. Still, it can be an indication that CTP and ATP binding does not induce conformational changes large enough to alter the enzyme's stability (see Fig. S4).

The structures of *Pha*CCA and *Gst*CCA were compared to track signs of structural adaptation to cold. As shown in Fig. 3, they can be easily superimposed with an RMSD of 2.3 Å (taking into account 360 Cα positions). The sequence alignment in Fig. 3 illustrates the high sequence similarity in the N–terminal region, where the catalytic core is located in the head and neck domains, displaying the characteristic motifs A to E, as well as a flexible loop between motifs A and B. In this region, the two structures show comparable 2D structure contents with 58.8% and 59.0% of residues located in alpha-helical elements (Table 2). In contrast, the C-terminal region, encompassing the body and the tail domains and constituting the tRNA binding site, is much less conserved and its helix content (alpha and 3_{10} helices) shows a clear 14% drop in *Pha*CCA (71.7%) compared to *Gst*CCA (86.1%) (Table 2). An example of drastic helix reduction is illustrated in Fig. 3, with complete helices replaced by a loop in *Pha*CCA at the junction of the neck and the body domains, as well as in the body. To investigate whether this alpha-helical reduction affects the affinity of *Pha*CCA to a tRNA substrate, electrophoretic mobility shifts were performed (Fig. S5). For *Gst*CCA, an interaction with a tRNA substrate was clearly visible, resulting in a Kd value of approximately 1.3 μM, comparable to that of other class II CCA-adding enzymes [24,25]. In contrast, no gel shift was visible for *Pha*CCA, indicating a reduced tRNA interaction that does not allow a calculation of a binding constant. Taken together, while the reduced number of alpha-helices in the tRNA-binding C-terminus seems to decrease the substrate affinity of *Pha*CCA, it is likely to represent - in combination with elongated loop elements - a main feature of cold adaptation that provides the increased structural flexibility observed in this enzyme [22].

In contrast, the N–terminal catalytic core seems to retain all sequence and structure characteristics of bacterial CCA–adding enzymes. Yet, our previous investigations clearly indicated that this enzyme region also contributes to the psychrophilic adaptation of *Pha*CCA [22]. As structural flexibility is a prominent prerequisite for cold adaptation, we focused on regions in the N–terminal catalytic core that participate or contribute to structural rearrangements during CCA–addition, in particular the nucleotide binding pocket [20], the flexible loop [39] and the hinge element of motif C [12].

When the *Pha*CCA crystals were soaked with CTP or a non-hydrolyzable CTP analog, the diffraction patterns of the resulting

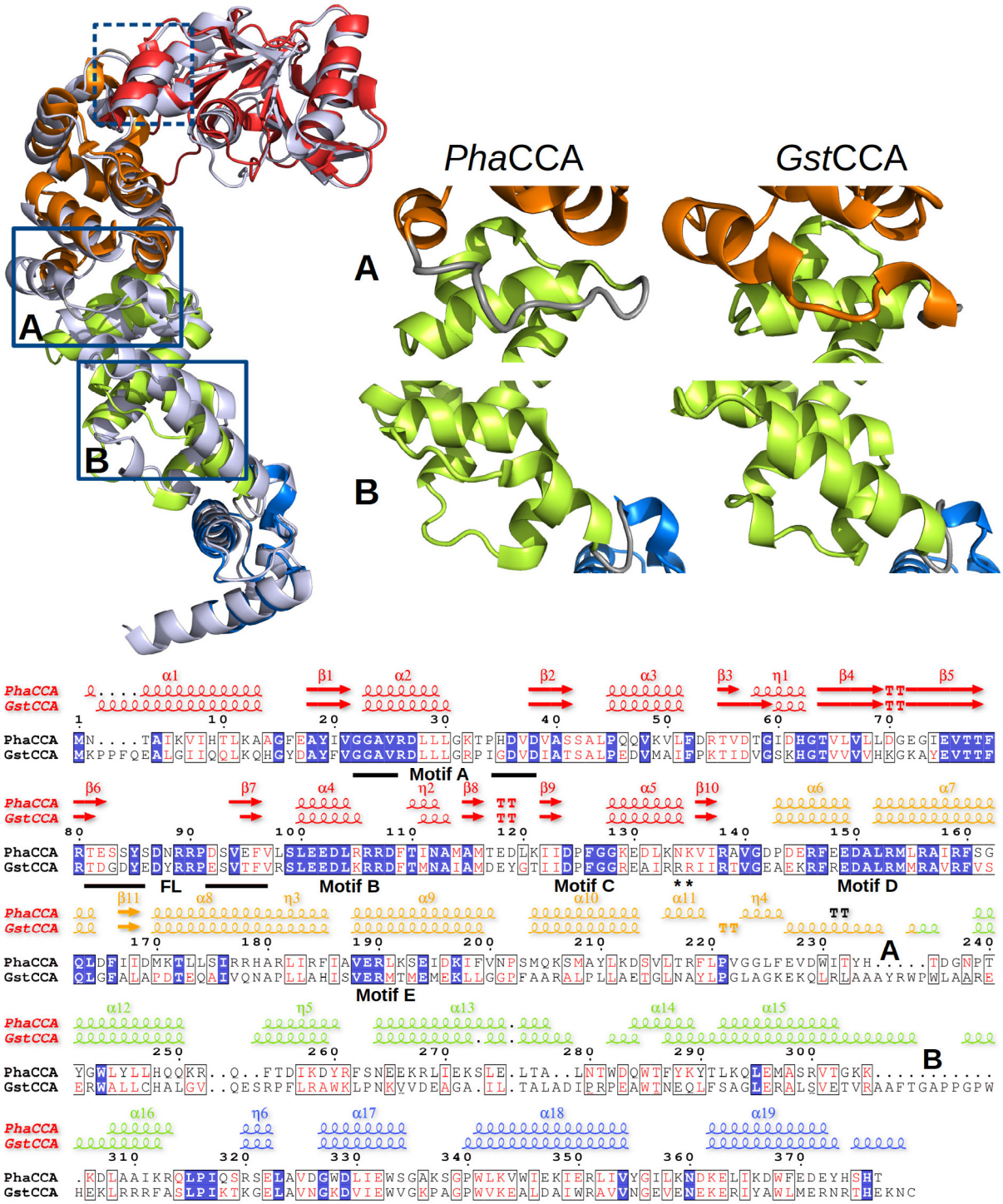


Fig. 3. Structural adaptations in the psychrophilic CCA-adding enzyme. On the left, the structural alignment of the *Pha*CCA enzyme (with its domains colored as in Fig. 1) and its thermophilic homolog *Gst*CCA [20] depicted in light blue, reveals typical cold adaptations in its structure. Examples of shorten alpha-helical elements in *Pha*CCA are boxed in blue (the dashed box highlights helix α5 in motif C discussed in Fig. 5). As illustrated by the close-up views on the right, the helix α11 in region A at the junction of the neck and body is replaced by a loop, while in region B in the body domain helices α15-16 are drastically reduced. These variations in helical regions can be easily spotted in the 2D structure elements presented above the sequence alignment for each structure (see labels A, B). Strictly conserved residues are highlighted in blue, similar residues in red. Catalytic core motifs A to E and the flexible loop (FL) are indicated. Asterisks indicate helix-capping residues R137 and R138 of *Gst*CCA. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

co–crystals revealed an orientation of the bound nucleotide that is identical to that found in the corresponding *Gst*CCA/CTP structure [20] (Fig. 4 and S4). Hence, the shape of the nucleotide binding pocket with the base-recognizing residues EDxxR of motif D does not differ between the thermophilic and the psychrophilic enzymes, excluding a prominent role in temperature adaptation.

A comparison of the apo enzyme with the two complexes shows that CTP binding does not involve any conformational change and the substrate is almost accommodated in a key-and-lock manner

Table 2
2D structure content of psychrophilic and thermophilic enzymes.

Enzyme	*Pha*CCA	*Gst*CCA
PDB id	6QY6	1MIW
Total number of residues	377	404
Number of residues in helices (α and 3_{10})	241	280
Number of residues in β-strands	37	39
Number of residues in 2D structures	278	319
% of residues in helices overall	**63.9**	**69.3**
% of residues in β-strands overall	9.8	9.7
% of residues in 2D structures overall	73.7	79.0
% of residues in helices in head + neck	58.8	59.0
% of residues in helices in body + tail	**71.7**	**86.1**

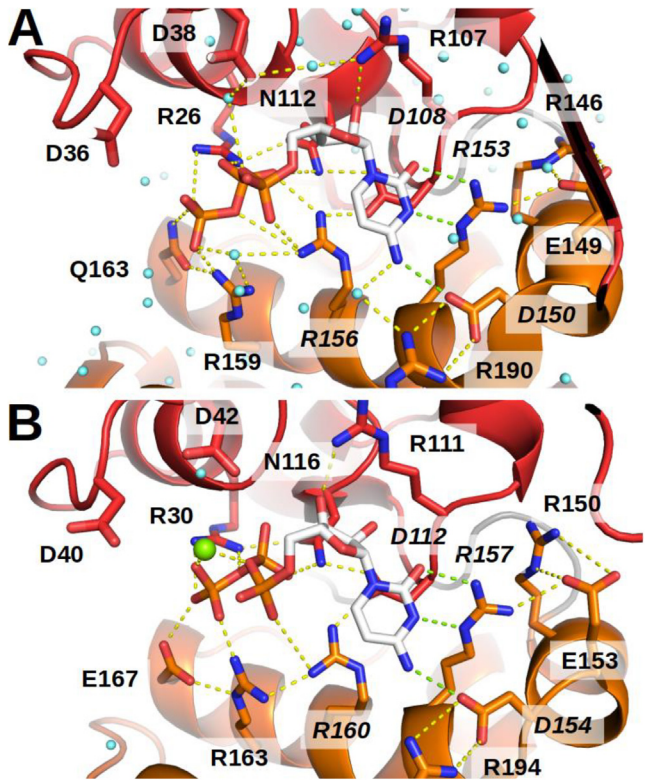


Fig. 4. A highly conserved CTP binding pocket. Close-up views of the catalytic core of *Pha*CCA (A) and *Gst*CCA (B) in the presence of CTP. Highly conserved residues with side chains contacting the nucleotide are indicated. Water molecules are represented as blue spheres and magnesium ion bound to *Gst*CCA as a green sphere. No magnesium site could be identified in *Pha*CCA structure. The dense network of interactions (with a distance < 4 Å) between protein side chains and the substrate is symbolized by dashed lines. Four side chains indicated in italics interact directly with the nucleobase and constitute the amino acid template in motif D: two couples of Arg/Asp residues, one forming a stacking platform below the base, the second making three hydrogen bonds indicated by green dashed lines with the Watson-Crick face of CTP. The latter residues are hold in place by additional connections to surrounding side chains. In both structures, the orientation of the CTP is almost identical and the amino acid template of motif D forms the same interactions to the bound base moiety, indicating that the templating function is not specifically cold adapted in *Pha*CCA. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

(Fig. S4). This is in agreement with the very low RMSDs between the different structures (Fig. S2).

As indicated above, our attempts to obtain *Pha*CCA crystals with ATP remained unsuccessful, and the ATP-bound state that could have been compared to the *Gst*CCA:ATP complex was not observed.

The reorganization of the amino acid template in motif D during the specificity switch from CTP to ATP binding is mediated by a flexible loop element [39,40]. It is discussed that this element, located between motifs A and B, acts as a lever to adjust the base-pairing amino acids for adenine interaction after the addition of two C residues to the tRNA substrate [41,42]. To investigate a possible contribution of the flexible loop to cold adaptation, we replaced this element in *Pha*CCA by the corresponding region of the thermophilic *Gst*CCA enzyme (Fig. S6A). While the loop sequence itself is not very conserved [41], it is flanked by constant sequence elements, allowing the identification of compatible fusion sites in both enzymes. As the N–terminal border, we selected the position separating the basic and the acidic residue in the conserved B/A motif (*Pha*CCA: T81, *Gst*CCA: T85), while the C-terminal border was the last variant position of the loop region (*Pha*CCA: L98, *Gst*CCA: R102), resulting in the loop chimera *PGP*CCA (PGP: first and last character describe the origin of the enzyme body, the second character indicates the origin of the loop) (Fig. 5B and S5B). CD spectral analysis of the recombinant protein revealed no differences in overall secondary structure compared to the parental enzymes. In addition, *PGP*CCA was fully active *in vitro* and added a complete CCA triplet on a yeast $tRNA^{Phe}$ transcript, a standard substrate for *in vitro* CCA–addition (Fig. S7). These data indicate that the loop replacement did not interfere with proper protein folding and catalytic activity. To determine a possible impact on thermal stability, unfolding of the *PGP*CCA chimera was investigated by CD spectroscopy at 222 nm over a temperature gradient ranging from 0 to 90 °C [22]. With a denaturation curve identical to that of *Pha*CCA, the chimera showed no differences to the parental enzyme (Fig. S6C). These data show that the flexible loop apparently does not contribute to the cold adaptation of the *Pha*CCA enzyme.

The third N-terminal element contributing to flexibility is motif C, a region acting as a spring element involved in the reorientation of head and neck domains during CCA-addition [12]. The sequence alignment exhibits a high degree of similarity between *Pha*CCA and *Gst*CCA (Figs. 3, 5B). The side chain of a strictly conserved aspartate – D124/D128 in *Pha*CCA/*Gst*CCA, respectively – makes a stabilizing interaction with the backbone of the two N–terminal residues of helix α5. This feature is described as helix-capping and is known to contribute to protein stability [43–45]. At the opposite extremity of this helix, the two arginine residues (R137, R138) interact with E134 in the crystal structure of *Gst*CCA. This corresponds to another feature of helix capping with side chains of amino acids at positions C1 and C2, stabilizing the alpha–helical structure through interaction with amino acids within the helix [46,47]. In the thermophilic *Gst*CCA, the C1 and C2 residues form ionic bridges with E134 (Fig. 5B). This feature is absent in *Pha*CCA, where the arginines are replaced by K133 and N134. To investigate a potential cold adaptation of motif C in *Pha*CCA in more detail, K133/N134 in *Pha*CCA were replaced by the corresponding *Gst*CCA positions R137 and R138, resulting in *Pha*CCA–RR (Fig. 5A). Similar to the loop chimera, *Pha*CCA–RR was catalytically active in CCA-addition, and CD spectral analysis indicated that its structure does not deviate from that of the corresponding wild type enzyme (Fig. S7). However, thermal denaturation analysis revealed that the two helix-capping arginines from *Gst*CCA lead to a stabilization of the *Pha*CCA structure, increasing the melting temperature by approximately 4 °C (Fig. 5C). It is interesting to note that the stabilizing effect of single amino acids in motif C has also been observed for the structurally related human CCA–adding enzyme. Here, replacement of aspartate at position 139 by alanine abolished the N–terminal helix capping and led to a strongly reduced A–adding activity [12]. Taken together, whereas the nucleotide binding pocket appears unchanged and the flexible loop is not involved in cold adaptation of *Pha*CCA, the absence of C-terminal helix-capping residues in motif C seems to contribute to psychrophilic adaptation at the price of a simultaneously reduced tRNA substrate binding.

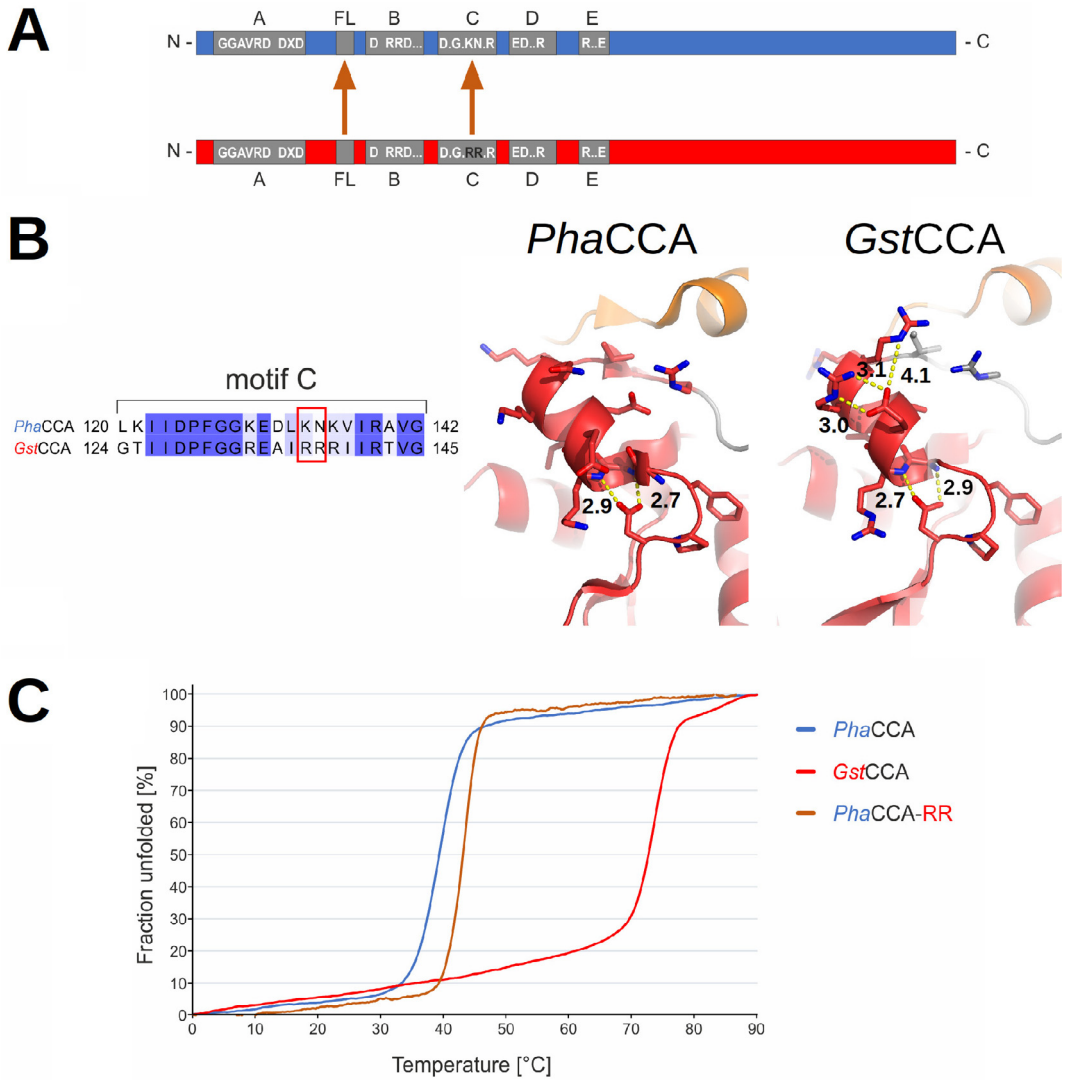


Fig. 5. Replacement of flexible elements in *Pha*CCA by their thermophilic counterparts. A) Bar diagrams of *Pha*CCA (blue) and *Gst*CCA (red). Elements of the catalytic core are indicated in grey. In *Pha*CCA, the flexible loop (FL) and motif C were replaced by the corresponding regions of *Gst*CCA. B) In *Gst*CCA, helix α5 in motif C (see Fig. 3) carries two R residues that are absent in *Pha*CCA (left panel). This arginine pair stabilizes an alpha-helical element by a C-terminal helix-capping (right panel). Note the presence of a conserved N-terminal helix-capping involving a strictly conserved aspartic acid of motif C. C) Thermal unfolding of CCA-adding enzymes and *Pha*CCA with introduced arginine positions involved in C-terminal helix-capping. With a melting temperature of approximately 39 °C, *Pha*CCA is a true psychrophilic protein (blue). Introducing the helix-capping RR pair increases the melting temperature of *Pha*CCA–RR to 43 °C (brown). *Gst*CCA shows a typical thermophilic profile and melts at 73 °C (red). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3.3. Thermostabilizing effect of PhaCCA-RR is not sufficient in vivo

To investigate their temperature-dependent catalytic activity in a natural environment, both *Pha*CCA wild type and *Pha*CCA–RR enzyme were subjected to *in vivo* complementation in a *cca*–deficient *E. coli* strain (Fig. 6). To this end, the corresponding genes were cloned into a previously described pETDuet-1-based dual expression system [23]. This system allows to study an enzyme's ability to repair truncated tRNA CCA–ends. A catalytically inactive version of the *E. coli* enzyme (*Eco*CCA AxA) does not complement and served as negative control when evaluating growth fitness behavior. Introduction of the *Pha*CCA constructs revealed that both *Pha*CCA wt and *Pha*CCA–RR promote growth at 30 °C. To examine whether the increased thermostability of *Pha*CCA–RR is reflected by a CCA–repair activity at elevated temperatures, transformants were incubated at 34 and 37 °C. Interestingly, we observed no differences in growth behavior between wt and RR variant at the investigated temperatures. While both enzyme versions enabled the cells to thrive at 34 °C, no growth was observed at 37 °C. Obviously, the structure-stabilizing effect of the helix-capping R residues is not sufficient for proper activity at this temperature, and it is likely that other catalytically important regions of *Pha*CCA suffer from subtle unfolding that abolish catalytic activity.

4. Discussion

CCA-adding enzymes constitute a specific class of RNA polymerases that function without nucleic acid template. They can accommodate all tRNAs present in the cell to ensure that they carry an intact 3′–CCA tail necessary for aminoacylation. Synthesis of this CCA sequence is accompanied by conformational changes in the enzyme to correctly position the tRNA for each polymerization step and select the correct nucleotide [12–14]. This work provides the first detailed insight into the structure of a psychrophilic member of this polymerase family. Cold-living organisms are known to produce more flexible enzymes in order to lower their activation energy [2] and the central question for CCA-adding enzymes is how this adaptation can be achieved without compromising the

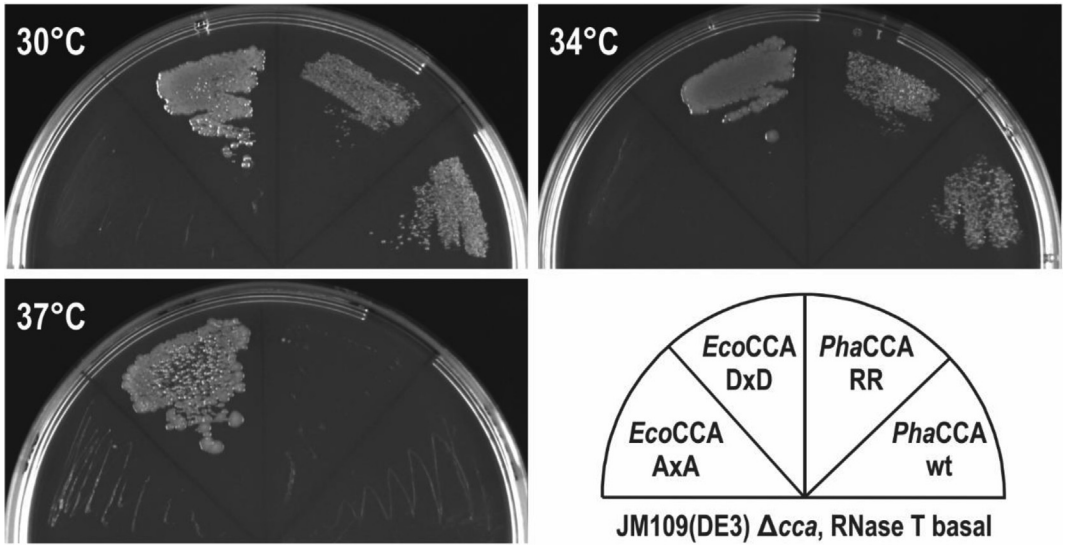


Fig. 6. Temperature-dependent activity of *Pha*CCA–RR *in vivo*. *Pha*CCA–RR complements for the endogenous CCA-adding activity in *E. coli* at temperatures up to 34 °C. At 37 °C, however, the enzyme is not sufficiently active, indicating that the helix-capping effect does not provide enough stability to render the enzyme active at higher temperatures. *Eco*CCA, *E. coli* CCA–adding enzyme; DxD, catalytically active wild type enzyme; AxA, inactive *Eco*CCA variant; *Pha*CCA–RR, *P. halocryophilus* CCA–adding enzyme with K133R and N134R replacements; wt, wild type.

specificity of the reaction, which requires an accurate selection of incorporated nucleotides. To address this question, we compared the structural properties of the psychrophilic *Pha*CCA and the thermophilic *Gst*CCA, both originating from bacteria belonging to the order of *Bacillales*. With a Tm difference of 30–34 °C (Figs. 2, 5), their thermal stability profile is characteristic of their respective ecosystem. Their evolutive proximity (39% sequence identity overall, 50% in the catalytic head and neck domains) allows to easily spot 1D-2D-3D structural differences as illustrated in Fig. 3. While *Pha*CCA shows the characteristic shape of bacterial CCA-adding enzymes, it also presents clear signs of cold adaptation. Over the past two decades, a variety of psychrophilic enzymes has been compared to their mesophilic and thermophilic homologs, showing that the global reduction of the number of non-covalent interactions (H-bonds and salt bridges) is a general hallmark of their adaptation. This is commonly correlated with a loss of 2D structure elements, a reduction of arginine residues involved in ion pairings on the protein surface and of the size of non-polar side chains in hydrophobic clusters as well as clusters of glycines instead of prolines in loops in order to increase the flexibility of enzymes and their activity at low temperatures [2,7–9,48–53]. Variations in amino acid composition, such as shifts from arginine to lysine or proline to glycine, are not seen in *Pha*CCA, but the ratio of negatively to positively charged residues is slightly increased, causing a drop of pI to 6.1 (8.4 for *Gst*CCA), which is another characteristic of psychrophilic proteins. The most obvious feature is the reduction of helical structures (Fig. 3) at the hinge between N- and C-terminal regions of the enzyme, as well as in the C-terminal part involved in tRNA interaction and positioning. The local weakening of hydrogen bond network resulting from replacing helices by loops certainly provides additional flexibility to the body domain and, considering its central position, to the whole enzyme, which functions as a clamp or vise to position and check the tRNA substrate status [54,55].

The N-terminal region is much more conserved and secondary structure contents are very similar in *Pha*CCA and *Gst*CCA. Hence, this region confers substrate specificity and is responsible for the catalytic activity. It is interesting to note that CTP binding occurs without structural change (Fig. S4) and the strict conservation of the amino acids forming Watson-Crick-like hydrogen bonds to the incoming nucleotide [20] excludes any cold specific adaptation of the nucleotide binding pocket. However, the specificity switch from CTP to ATP is clearly affected. At 4 °C, the Km value for ATP incorporation is 64-fold higher than that of *Gst*CCA, suggesting a strongly reduced binding affinity for ATP [22], in accordance with the impossibility to soak this nucleotide in the crystalline enzyme. This - as well as the observed reduced tRNA affinity - is also in line with the common adaptation strategy of cold-adapted enzymes: the destabilization of the enzyme/substrate complex leads to a reduction of the required activation energy [3–5].

In bacterial and eukaryotic CCA-adding enzymes, different flexible regions are important for highly coordinated movements during nucleotide addition as well as to monitor the structural intactness of the tRNA substrate [12–14,55]. To switch the enzymes' specificity from CTP towards ATP, the templating amino acids of the nucleotide binding pocket need to be reoriented, which demands for a rearrangement of the N–terminal region [20,23,56–58]. In this transition, the flexible loop acts as a lever that forces the amino acid template into its base-specific specific position [39,41]. While similar surface-exposed flexible loops are often targets for the thermal adaptation of proteins [8], our data indicate that the flexible loop of *Pha*CCA does not contribute to its activity in the cold.

In contrast, the second flexible element motif C, which coordinates interdomain rearrangements during catalysis, is an interesting spot of adaptation [12]. It presents a C-terminal helix capping in *Gst*CCA, while this stabilizing interaction is absent in *Pha*CCA. When C1 and C2 residues are restored, *Pha*CCA gains 4 °C of Tm, reinforcing the idea that the identity of these side chains contributes to the cold adaptation of this enzyme. These results are in line with data from other psychrophilic enzymes, where N- or C-caps have been found to weaken alpha-helices in order to adapt to low temperatures [6]. In addition, single amino acid substitutions in the near active site of a psychrophilic adenylate kinase lead to an increase in T_m by ~ 8 °C [12,52]. In conclusion, our results indicate that the highly conserved motif C is not only essential to ensure proper CCA-addition, but is equally important to modulate the temperature stability in psychro- and thermophilic enzymes.

Although the introduction of two helix-capping amino acids from *Gst*CCA into *Pha*CCA (*Pha*CCA–RR) led to thermostabilization of the *Pha*CCA variant *in vitro*, no such effect could be observed when complementing under natural conditions in *E. coli* (Fig. 6).

Possibly, the thermostabilizing effect observed *in vitro* is not sufficient for the rather stringent *in vivo* system demanding for an efficient CCA-repair activity [23]. Further, motif C destabilization is probably not the only cold adaptation in *Pha*CCA, and additional cumulative adaptations distributed over the whole protein render the enzyme truly psychrophilic. As a result, such regions are obviously destabilized at increased temperature, and their unfolding inactivates *Pha*CCA-RR, despite an increased stability of motif C due to the introduced helix-capping arginine residues. This is supported by the fact that chimeric enzymes consisting of N- and C-terminal halves from *Pha*CCA and *Gst*CCA show temperature adaptations in both regions [22]. Such cumulative structural changes involved in adaptation to temperature have also been reported elsewhere, such as in isocitrate dehydrogenase, which - depending on the respective bacterium - has up to three different regions contributing to the thermal properties [59].

Taken together, the crystallographic and biochemical analysis of the psychrophilic *Pha*CCA enzyme shows that cold adaptation is not only achieved by the local uncapping of an alpha-helix in the catalytic core but also by a general reduction of alpha-helical parts in the C-terminus of this enzyme. While it is a point of debate whether proteins are cold-adapted due to an increase in global or local flexibility [60,61], the CCA-adding enzyme of *P. halocryophilus* seems to make use of both strategies.

5. Conclusion

Our structural data contribute to clarify the mechanism of cold adaptation in the tRNA nucleotidyltransferase of *P. halocryophilus*. We confirm recent observations indicating that this enzyme acquired increased flexibility as a strategy for cold adaptation [22]. The crystal structures reveal a globally conserved scaffold with local adaptations to provide the structural plasticity compatible with catalysis at low temperature. The clearly visible reduction of helical regions appears in the less conserved C-terminal part of the enzyme that shows promiscuous binding of all tRNAs as substrates. The N-terminal catalytic core, however, which is catalytically active and shows high substrate specificity for CTP and ATP, is only marginally affected, though lacking a helix-capping feature that potentially brings 4 °C of thermal stability. Further investigations will clarify whether other RNA polymerases follow a similar adaptive strategy for efficient synthesis in the cold.

6. Funding information

This work used the Integrated Structural Biology platform of the Strasbourg Instruct-ERIC (PID 12797) center IGBMC-CBI supported by FRISBI (ANR-10-INBS-05–01). This work was supported by the Deutsche Forschungsgemeinschaft (grant no. Mo 634/10–1 to M. M.), the French Centre National de la Recherche Scientifique (CNRS), the University of Strasbourg, the LabEx consortia "NetRNA" (ANR-10-LABX-0036_NETRNA to C.S.), a PhD funding to R.dW. from the Excellence initiative (IdEx) of the University of Strasbourg in the frame of the French National Program "Investissements d'Avenir", a PhD funding to K.R. from the French-German University (DFH-UFA grant no. CT-30–19). The authors benefitted from the PROCOPE Hubert Curien cooperation program (French Ministry of Foreign Affair and DAAD to C.S. and M.M.).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors thank the following synchrotron facilities and associated scientists for beamtime allocation to the project and assistance during data collection: beamlines X06DA (PXIII) at the Swiss Light Source (Villigen, Switzerland), PROXIMA 1 and SWING at SOLEIL synchrotron (Saint-Aubin, France). They also acknowledge O. Hennig for protein preparation, B. Masquida and V. Oliéric for collecting *Pha*CCA S–SAD dataset, C. Birck for assistance and advice during NanoDSF experiments and the FRISBI–Instruct-ERIC from IGBMC (Illkirch, France) for access to the biophysics platform. We also thank S. Bonin for expert technical assistance.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.csbj.2021.10.018.

References

[1] Morita RY. Psychrophilic bacteria. Bacteriol Rev 1975;39(2):144–67.
[2] Feller G. Psychrophilic enzymes: from folding to function and biotechnology. Scientifica (Cairo) 2013;2013:1–28. https://doi.org/10.1155/2013/512840.
[3] Feller G, Gerday C. Psychrophilic enzymes: Hot topics in cold adaptation. Nat Rev Microbiol 2003;1(3):200–8. https://doi.org/10.1038/nrmicro773.
[4] Siddiqui KS, Williams TJ, Wilkins D, Yau S, Allen MA, Brown MV, et al. Psychrophiles. Annu Rev Earth Planet Sci 2013;41(1):87–115. https://doi.org/10.1146/annurev-earth-040610-133514.
[5] D'Amico S, Collins T, Marx J-C, Feller G, Gerday C, Gerday C. Psychrophilic microorganisms: challenges for life. EMBO Rep 2006;7(4):385–9. https://doi.org/10.1038/sj.embor.7400662.
[6] Struvay C, Feller G. Optimization to low temperature activity in psychrophilic enzymes. Int J Mol Sci 2012;13(9):11643–65. https://doi.org/10.3390/ijms130911643.
[7] Yang L-L, Tang S-K, Huang Y, Zhi X-Y. Low Temperature Adaptation Is Not the Opposite Process of High Temperature Adaptation in Terms of Changes in Amino Acid Composition. Genome Biol Evol 2015;7(12):3426–33. https://doi.org/10.1093/gbe/evv232.
[8] Kovacic F, Mandrysch A, Poojari C, Strodel B, Jaeger K-E. Structural features determining thermal adaptation of esterases. Protein Eng Des Sel 2016;29(2):65–76. https://doi.org/10.1093/protein/gzv061.
[9] Parvizpour S, Hussin N, Shamsir MS, Razmara J. Psychrophilic enzymes: structural adaptation, pharmaceutical and industrial applications. Appl Microbiol Biotechnol 2021;105(3):899–907. https://doi.org/10.1007/s00253-020-11074-0.
[10] Cavicchioli R, Charlton T, Ertan H, Mohd Omar S, Siddiqui KS, Williams TJ. Biotechnological uses of enzymes from psychrophiles. Microb Biotechnol 2011;4(4):449–60. https://doi.org/10.1111/j.1751-7915.2011.00258.x.
[11] Darst SA, Opalka N, Chacon P, Polyakov A, Richter C, Zhang G, et al. Conformational flexibility of bacterial RNA polymerase. Proc Natl Acad Sci USA 2002;99(7):4296–301. https://doi.org/10.1073/pnas.052054099.
[12] Ernst FGM, Rickert C, Bluschke A, Betat H, Steinhoff H-J, Mörl M. Domain movements during CCA-addition: a new function for motif C in the catalytic core of the human tRNA nucleotidyltransferases. RNA Biol 2015;12(4):435–46. https://doi.org/10.1080/15476286.2015.1018502.
[13] Kim S, Liu C, Halkidis K, Gamper HB, Hou Y-M. Distinct kinetic determinants for the stepwise CCA addition to tRNA. RNA 2009;15(10):1827–36. https://doi.org/10.1261/rna.1669109.
[14] Tomita K, Fukai S, Ishitani R, Ueda T, Takeuchi N, Vassylyev DG, et al. Structural basis for template-independent RNA polymerization. Nature 2004;430(7000):700–4. https://doi.org/10.1038/nature02712.
[15] Betat H, Rammelt C, Mörl M. tRNA nucleotidyltransferases: ancient catalysts with an unusual mechanism of polymerization. Cell Mol Life Sci 2010;67(9):1447–63. https://doi.org/10.1007/s00018-010-0271-4.
[16] Weiner AM. tRNA maturation: RNA polymerization without a nucleic acid template. Curr Biol 2004;14(20):R883–5. https://doi.org/10.1016/j.cub.2004.09.069.
[17] Xiong Y, Steitz TA. A story with a good ending: tRNA 3'-end maturation by CCA-adding enzymes. Curr Opin Struct Biol 2006;16(1):12–7. https://doi.org/10.1016/j.sbi.2005.12.001.
[18] Aravind L, Koonin EV. DNA polymerase beta-like nucleotidyltransferase superfamily: identification of three new families, classification and evolutionary history. Nucleic Acids Res 1999;27(7):1609–18. https://doi.org/10.1093/nar/27.7.1609.
[19] Yue D, Maizels N, Weiner AM. CCA-adding enzymes and poly(A) polymerases are all members of the same nucleotidyltransferase superfamily: characterization of the CCA-adding enzyme from the archaeal hyperthermophile Sulfolobus shibatae. RNA 1996;2(9):895–908.
[20] Li F, Xiong Y, Wang J, Cho HD, Tomita K, Weiner AM, et al. Crystal Structures of the Bacillus stearothermophilus CCA-Adding Enzyme and Its Complexes with

ATP or CTP. Cell 2002;111(6):815–24. https://doi.org/10.1016/S0092-8674(02)01115-7.
[21] Mykytczuk N, Foote S, Omelon C, Southam G, Greer C, Whyte L. Bacterial growth at-15 degrees C; molecular insights from the permafrost bacterium Planococcus halocryophilus Or1. ISME J 2013;7(6):1211–26. https://doi.org/10.1038/ismej.2013.8.
[22] Ernst FGM, Erber L, Sammler J, Jühling F, Betat H, Mörl M. Cold adaptation of tRNA nucleotidyltransferases: A tradeoff in activity, stability and fidelity. RNA Biol 2018;15(1):144–55. https://doi.org/10.1080/15476286.2017.1391445.
[23] Wellner K, Pöhler M-T, Betat H, Mörl M. Dual expression of CCA-adding enzyme and RNase T in Escherichia coli generates a distinct cca growth phenotype with diverse applications. Nucleic Acids Res 2019;47(7):3631–9. https://doi.org/10.1093/nar/gkz133.
[24] Erber L, Hoffmann A, Fallmann J, Hagedorn M, Hammann C, Stadler PF, et al. Unusual Occurrence of Two Bona-Fide CCA-Adding Enzymes in Dictyostelium discoideum. Int J Mol Sci 2020;21(15):5210. https://doi.org/10.3390/ijms21155210.
[25] Hennig O, Philipp S, Bonin S, Rollet K, Kolberg T, Jühling T, et al. Adaptation of the Romanomermis culicivorax CCA-Adding Enzyme to Miniaturized Armless tRNA Substrates. Int J Mol Sci 2020;21(23):9047. https://doi.org/10.3390/ijms21239047.
[26] de Wijn R, Hennig O, Ernst FGM, Lorber B, Betat H, Mörl M, et al. Combining crystallogenesis methods to produce diffraction-quality crystals of a psychrophilic tRNA-maturation enzyme. Acta Crystallogr F Struct Biol Commun 2018;74(11):747–53. https://doi.org/10.1107/S2053230X18014590.
[27] Adams PD, Afonine PV, Bunkóczi G, Chen VB, Davis IW, Echols N, et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr 2010;66(2):213–21. https://doi.org/10.1107/S0907444909052925.
[28] McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser crystallographic software. J Appl Crystallogr 2007;40(Pt 4):658–74. https://doi.org/10.1107/S0021889807021206.
[29] Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr 2004;60(Pt 12 Pt 1):2126–32. https://doi.org/10.1107/S0907444904019158.
[30] Robert X, Gouet P. Deciphering key features in protein structures with the new ENDscript server. Nucleic Acids Res 2014;42. https://doi.org/10.1093/nar/gku316.
[31] David G, Pérez J. Combined sampler robot and high-performance liquid chromatography: a fully automated system for biological small-angle X-ray scattering experiments at the Synchrotron SOLEIL SWING beamline. J Appl Crystallogr 2009;42(5):892–900. https://doi.org/10.1107/S0021889809029288.
[32] Manalastas-Cantos K, Konarev PV, Hajizadeh NR, Kikhney AG, Petoukhov MV, Molodenskiy DS, et al. ATSAS 3.0: expanded functionality and new tools for small-angle scattering data analysis. J Appl Crystallogr 2021;54(1):343–55. https://doi.org/10.1107/S1600576720013412.
[33] Webb B, Sali A. Comparative Protein Structure Modeling Using MODELLER. Curr Protoc Bioinformatics 2014;47(1). https://doi.org/10.1002/0471250953.2014.47.issue-110.1002/0471250953.bi0506s47.
[34] Rudenko O, Thureau A, Perez J. In: Evolutionary refinement of the 3D structure of multi-domain protein complexes from small angle X-ray scattering data. New York, NY, United States: Association for Computing Machinery; 2019. p. 401–2.
[35] Svergun D, Barberato C, Koch MHJ. CRYSOL – a Program to Evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates. J Appl Crystallogr 1995;28(6):768–73. https://doi.org/10.1107/S0021889895007047.
[36] Hopkins JB, Gillilan RE, Skou S. BioXTAS RAW: improvements to a free open-source program for small-angle X-ray scattering data reduction and analysis. J Appl Crystallogr 2017;50(Pt 5):1545–53. https://doi.org/10.1107/S1600576717011438.
[37] Augustin MA, Reichert AS, Betat H, Huber R, Mörl M, Steegborn C. Crystal Structure of the Human CCA-adding Enzyme: Insights into Template-independent Polymerization. J Mol Biol 2003;328(5):985–94. https://doi.org/10.1016/S0022-2836(03)00381-4.
[38] Sauter C, Lorber B, Gaudry A, Karim L, Schwenzer H, Wien F, et al. Neurodegenerative disease-associated mutants of a human mitochondrial aminoacyl-tRNA synthetase present individual molecular signatures. Sci Rep 2015;5(1). https://doi.org/10.1038/srep17332.
[39] Neuenfeldt A, Just A, Betat H, Mörl M. Evolution of tRNA nucleotidyltransferases: a small deletion generated CC-adding enzymes. Proc Natl Acad Sci USA 2008;105(23):7953–8. https://doi.org/10.1073/pnas.0801971105.
[40] Erber L, Franz P, Betat H, Prohaska S, Mörl M. Divergent Evolution of Eukaryotic CC- and A-Adding Enzymes. Int J Mol Sci 2020;21(2):462. https://doi.org/10.3390/ijms21020462.
[41] Hoffmeier A, Betat H, Bluschke A, Günther R, Junghanns S, Hofmann H-J, et al. Unusual evolution of a catalytic core element in CCA-adding enzymes. Nucleic Acids Res 2010;38(13):4436–47. https://doi.org/10.1093/nar/gkq176.
[42] Toh Y, Takeshita D, Numata T, Fukai S, Nureki O, Tomita K. Mechanism for the definition of elongation and termination by the class II CCA-adding enzyme. EMBO J 2009;28(21):3353–65. https://doi.org/10.1038/emboj.2009.260.
[43] Presta LG, Rose GD. Helix signals in proteins. Science 1988;240(4859):1632–41. https://doi.org/10.1126/science.2837824.
[44] Richardson JS, Richardson DC. Amino acid preferences for specific locations at the ends of alpha helices. Science 1988;240(4859):1648–52. https://doi.org/10.1126/science.3381086.
[45] Segura J, Oliva B, Fernandez-Fuentes N. CAPS-DB: a structural classification of helix-capping motifs. Nucleic Acids Res 2012;40:D479–85. https://doi.org/10.1093/nar/gkr879.
[46] Aurora R, Rose GD. Helix capping. Protein Sci 1998;7(1):21–38. https://doi.org/10.1002/pro.5560070103.
[47] Rose GD. Lifting the lid on helix-capping. Nat Chem Biol 2006;2(3):123–4. https://doi.org/10.1038/nchembio0306-123.
[48] Yang G, Yang G, Aprile L, Turturo V, Pucciarelli S, Pucciarelli S, et al. Characterization and comparative analysis of psychrophilic and mesophilic alpha-amylases from Euplotes species: A contribution to the understanding of enzyme thermal adaptation. Biochem Biophys Res Commun 2013;438(4):715–20. https://doi.org/10.1016/j.bbrc.2013.07.113.
[49] Tiberti M, Papaleo E. Dynamic properties of extremophilic subtilisin-like serine-proteases. J Struct Biol 2011;174(1):69–83. https://doi.org/10.1016/j.jsb.2011.01.006.
[50] Hiteshi K, Gupta R. Thermal adaptation of α-amylases: a review. Extremophiles 2014;18(6):937–44. https://doi.org/10.1007/s00792-014-0674-5.
[51] Bae E, Phillips GN. Structures and analysis of highly homologous psychrophilic, mesophilic, and thermophilic adenylate kinases. J Biol Chem 2004;279(27):28202–8. https://doi.org/10.1074/jbc.M401865200.
[52] Moon S, Kim J, Koo J, Bae E. Structural and mutational analyses of psychrophilic and mesophilic adenylate kinases highlight the role of hydrophobic interactions in protein thermal stability. Struct Dyn 2019;6(2):24702. https://doi.org/10.1063/1.5089707.
[53] Mandelman D, Ballut L, Wolff DA, Feller G, Gerday C, Haser R, et al. Structural determinants increasing flexibility confer cold adaptation in psychrophilic phosphoglycerate kinase. Extremophiles 2019;23(5):495–506. https://doi.org/10.1007/s00792-019-01102-x.
[54] Betat H, Mörl M. The CCA-adding enzyme: A central scrutinizer in tRNA quality control. BioEssays 2015;37(9):975–82. https://doi.org/10.1002/bies.201500043.
[55] Kuhn C-D, Wilusz JE, Zheng Y, Beal PA, Joshua-Tor L. On-enzyme refolding permits small RNA and tRNA surveillance by the CCA-adding enzyme. Cell 2015;160(4):644–58. https://doi.org/10.1016/j.cell.2015.01.005.
[56] Cho HD, Verlinde CLMJ, Weiner AM. Reengineering CCA-adding enzymes to function as (U, G)- or dCdCdA-adding enzymes or poly(C, A) and poly(U, G) polymerases. Proc Natl Acad Sci USA 2006;104(1):54–9. https://doi.org/10.1073/pnas.0606961104.
[57] Lizano E, Scheibe M, Rammelt C, Betat H, Mörl M. A comparative analysis of CCA-adding enzymes from human and E. coli: differences in CCA addition and tRNA 3'-end repair. Biochimie 2008;90(5):762–72. https://doi.org/10.1016/j.biochi.2007.12.007.
[58] Betat H, Mede T, Tretbar S, Steiner L, Stadler PF, Mörl M, et al. The ancestor of modern Holozoa acquired the CCA-adding enzyme from Alphaproteobacteria by horizontal gene transfer. Nucleic Acids Res 2015;43(14):6739–46. https://doi.org/10.1093/nar/gkv631.
[59] Mouri Y, Takada Y. Contribution of Three Different Regions of Isocitrate Dehydrogenases from Psychrophilic and Psychrotolerant Bacteria to Their Thermal Properties. Curr Microbiol 2018;75(11):1523–9. https://doi.org/10.1007/s00284-018-1554-5.
[60] Papaleo E, Tiberti M, Invernizzi G, Pasi M, Ranzani V. Molecular determinants of enzyme cold adaptation: comparative structural and computational studies of cold- and warm-adapted enzymes. Curr Protein Pept Sci 2011;12(7):657–83. https://doi.org/10.2174/1389203711109070657.
[61] Siddiqui KS, Cavicchioli R. Cold-adapted enzymes. Annu Rev Biochem 2006;75(1):403–33. https://doi.org/10.1146/biochem.2006.75.issue-110.1146/annurev.biochem.75.103004.142723.

Supplementary material

CCA addition in the cold: Structural characterization of the psychrophilic CCA-adding enzyme from the permafrost bacterium *Planococcus halocryophilus*

Raphaël de Wijn[a,#], Kévin Rollet[a,b,#], Felix G.M. Ernst[b], Karolin Wellner[b], Heike Betat[b], Mario Mörl[b*], Claude Sauter[a*]

[a]Architecture et Réactivité de l'ARN, Université de Strasbourg, CNRS, IBMC, 67084 Strasbourg, France

[b]Institute for Biochemistry, Leipzig University, Brüderstr. 34, 04103 Leipzig, Germany

[#]these authors contributed equally to the work

*corresponding authors

List of figures:

Figure S1. Experimental maps and structural details of *Pha*CCA

Figure S2. Comparison of *Pha*CCA structures determined in various conditions

Figure S3. SEC-SAXS analysis of *Pha*CCA

Figure S4. Nucleotide binding site with CTP / CMPcPP bound or in apo form

Figure S5. Binding behavior of *Gst*CCA and *Pha*CCA to a tRNA substrate

Figure S6. Replacement of flexible loop region does not lead to thermostabilization in *Pha*CCA

Figure S7. Analysis of both secondary structure and *in vitro* activity for respective chimeric variants

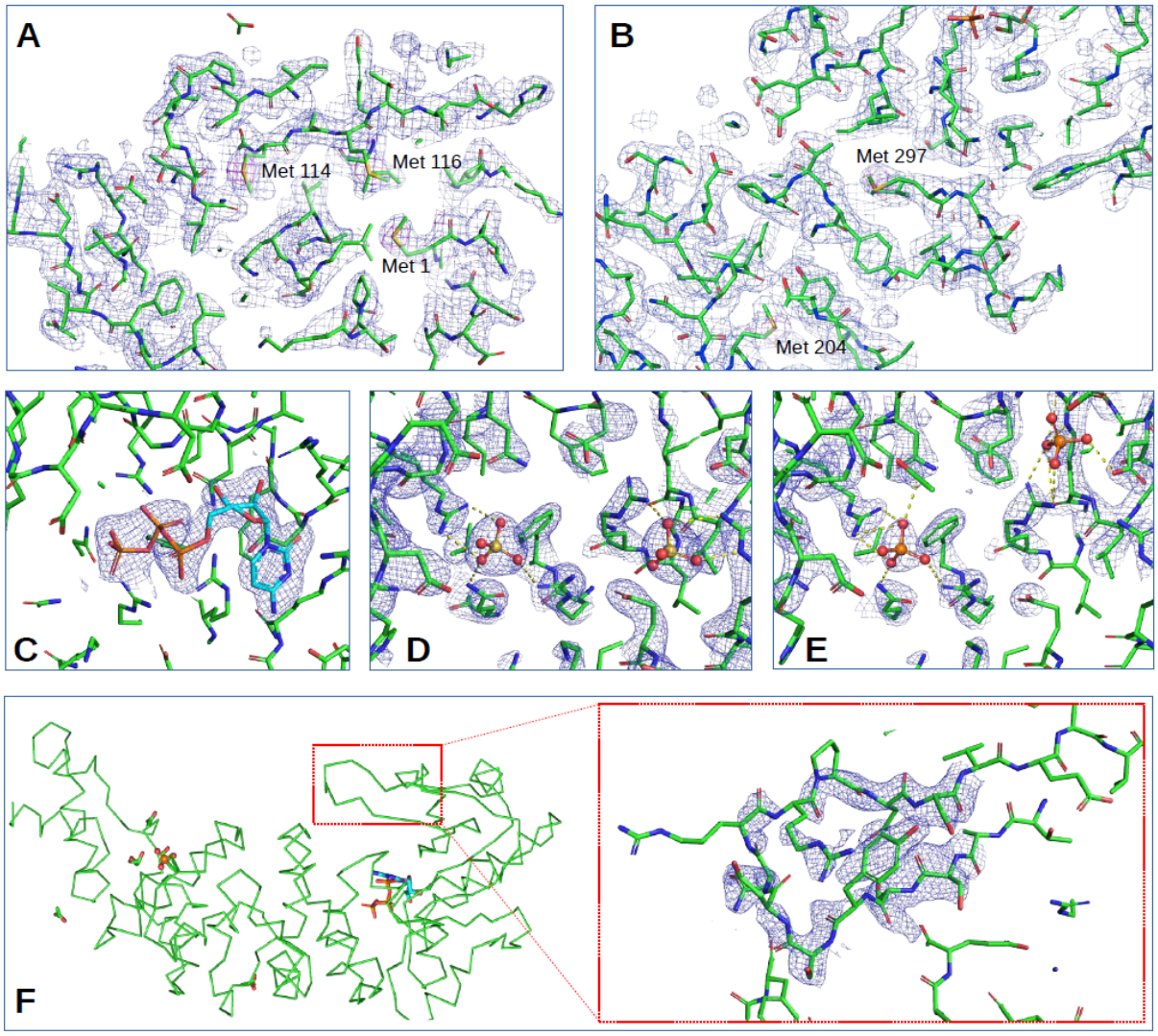


Figure S1. Experimental maps and structural details of *Pha*CCA. A) N-terminal region of sulfur SAD-phasing structure with associated *2Fo-Fc* experimental map contoured at 1.2 σ in blue. The sulfur atoms of Met1, Met104 and Met106 are clearly visible in the anomalous difference map contoured at 2.5 σ in orange. B) C-terminal region of sulfur SAD-phasing structure with associated *2Fo-Fc* experimental map contoured at 1.2 σ in blue. Sulfur atoms of Met204 and Met297 are visible in the anomalous difference map contoured at 2.5 σ in orange. C) D) E) Catalytic site of structure containing CTP, structure obtained in ammonium sulfate and sulfur SAD-phasing structure, respectively, with associated *2Fo-Fc* maps contoured at 1.2 σ. Polar contacts between SO_4^{2-} or PO_4^{2-} and side chains are represented by yellow dashed lines. F) Flexible loop (residues Ser83 to Ser93) in the structure of the complex with CTP and associated *2Fo-Fc* experimental map contoured at 1.0 σ.

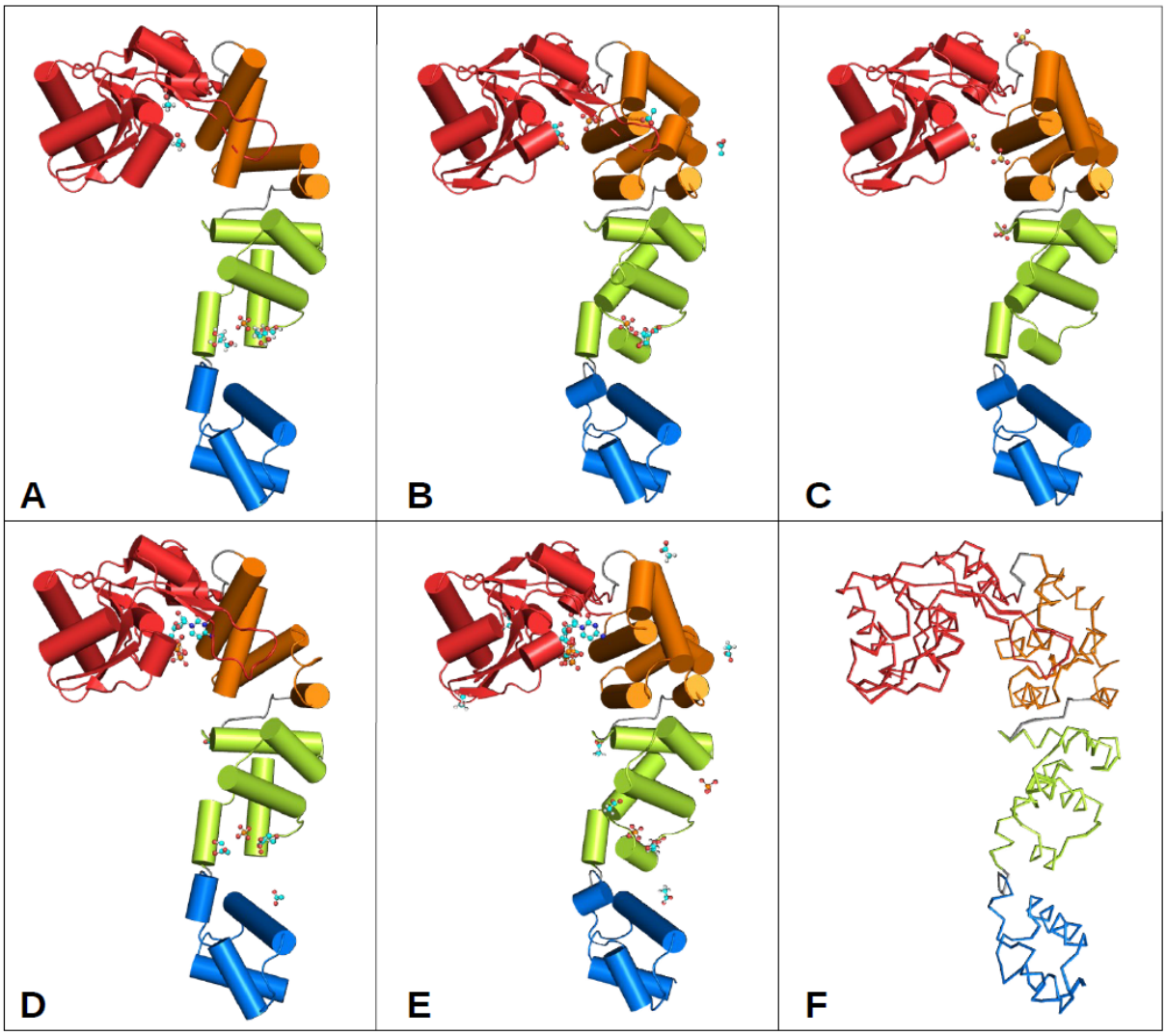


Figure S2. **Comparison of *Pha*CCA structures determined in various conditions.** Head, neck, body and tail domains are colored in red, orange, green and blue, respectively. Small molecules visible at the surface of the structures in their solvent shell include *Pha*CCA ligands (CTP, CMPcPP), buffer compounds (acetate, phosphate or sulfate ions) and the cryoprotectant (glycerol). They are represented in ball and stick mode. RMSD values were calculated with Phenix_Compare, using the native structure (A) determined at the highest resolution (1.8 Å, see Table 1 for details) as reference. A) Native structure. B) Native sulfur SAD structure (RMSD = 0.20 Å). C) Structure obtained in the presence of ammonium sulfate (RMSD = 0.22 Å). D) Structure obtained from a crystal soaked in 5 mM CTP (RMSD = 0.16 Å). E) Structure obtained from a crystal soaked in 10 mM CMPcPP (RMSD = 0.19 Å). Structures A, B, D, E were obtained from crystals grown in 100 mM sodium acetate pH 4.5, 1 M di-ammonium hydrogen phosphate, structure C from a crystal grown in 100 mM sodium acetate pH 5, 3 M ammonium sulfate. F) Superposition of all structures represented in ribbon mode, illustrating the absence of conformational transition between the apo and CTP-bound state.

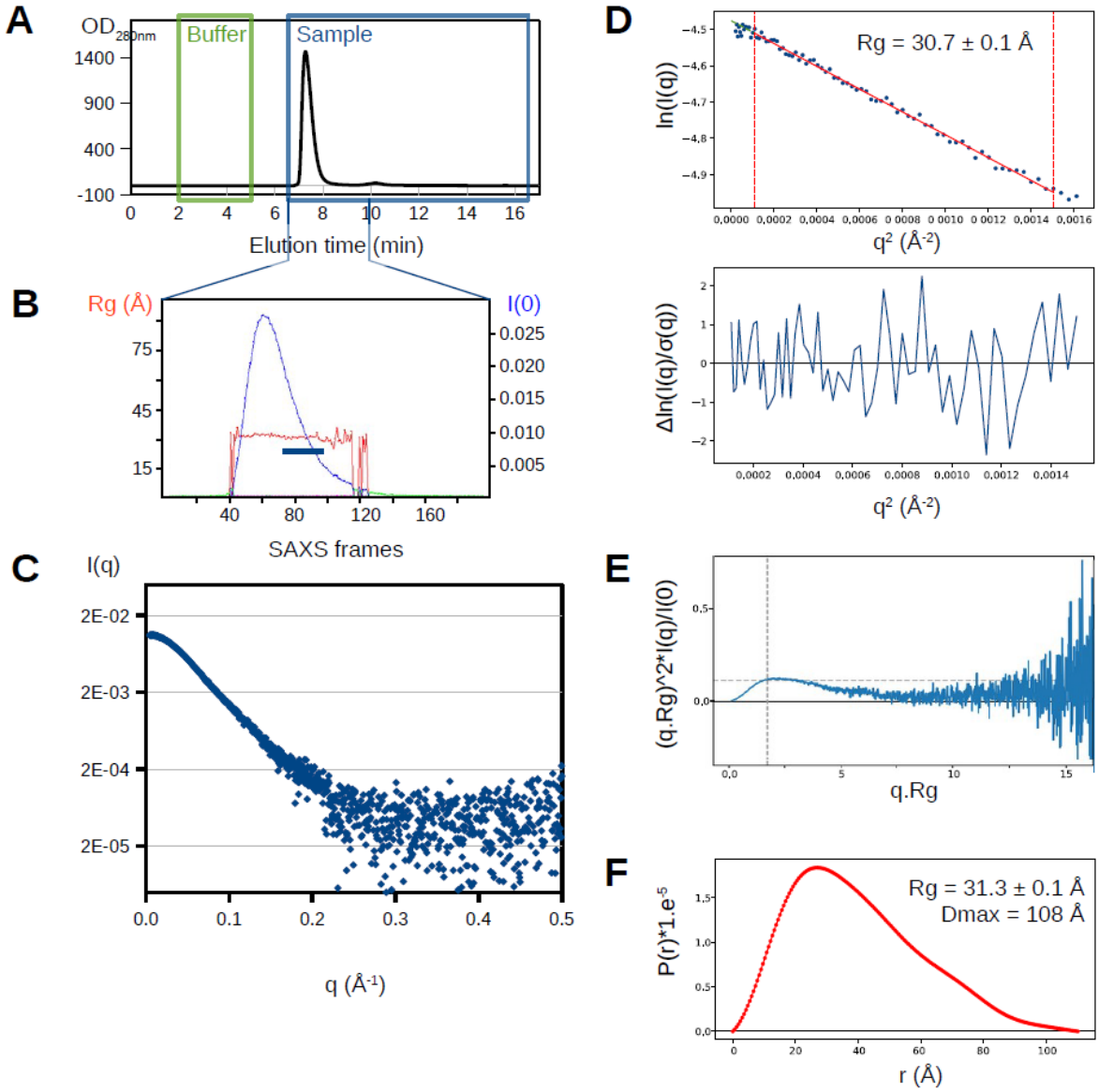


Figure S3: SEC-SAXS analysis of *Pha*CCA. A) SEC chromatogram showing a single population. The green and blue windows correspond to SAXS analyses to collect the buffer (180 images) and sample (630 images) signals, respectively. B) Estimate of Rg and I(0) for the 200 first sample images. C) SAXS profile obtained after solvent subtraction by merging individual profiles in the region indicated by the blue bar in B) (images 75-95) where Rg values are stable. D) Guinier plot, associated weighted residual and derived gyration radius (Rg). E) Dimensionless Kratky plot consistent with a compact multidomain protein with a flexible extension [1]. F) Pair atom distance distribution P(r) with derived Rg and Dmax.

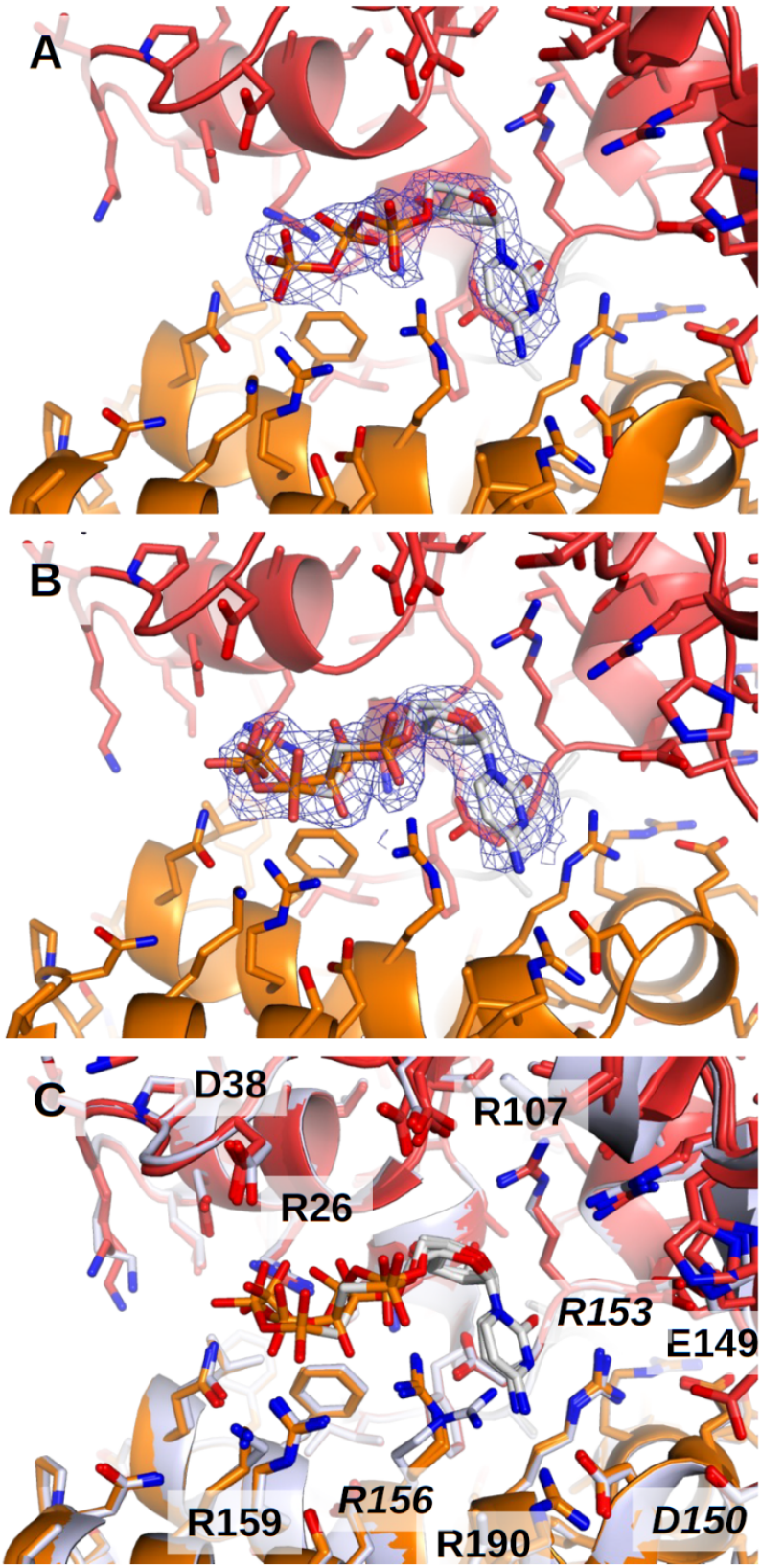


Figure S4. Nucleotide binding site with CTP / CMPcPP bound or in apo form.
A) Complex with CTP, B) with CMPcPP (in two conformations) with the electron density map *2fo-fc* map contoured at 1.3 σ. Residues of the head and neck domains are colored in red and orange, respectively. C) Overlay of *Pha*CCA in apo-form (in grey) and the two complexes

highlighting the absence of conformational change and suggesting a key-and-lock binding process. The only side chain slightly affected is Arg156. Residues in italic are in direct contact with the ligand (see also **Figure 4**).

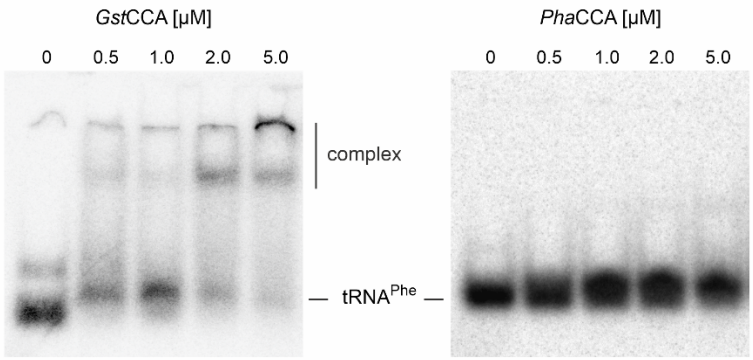


Figure S5. Binding behavior of *Gst*CCA and *Pha*CCA to a tRNA substrate. Images of representative electrophoretic mobility shift assays with *Gst*CCA (left) and *Pha*CCA (right). *Gst*CCA exhibits efficient binding to the substrate $tRNA^{Phe}$. After densitometrical quantitation of the radioactive signals, an approximate binding constant of 1.3 µM was determined. In contrast, *Pha*CCA shows no substrate interaction over the whole tRNA concentration range (right). This reduced substrate affinity represents a typical strategy of cold-adapted enzymes to low temperatures.

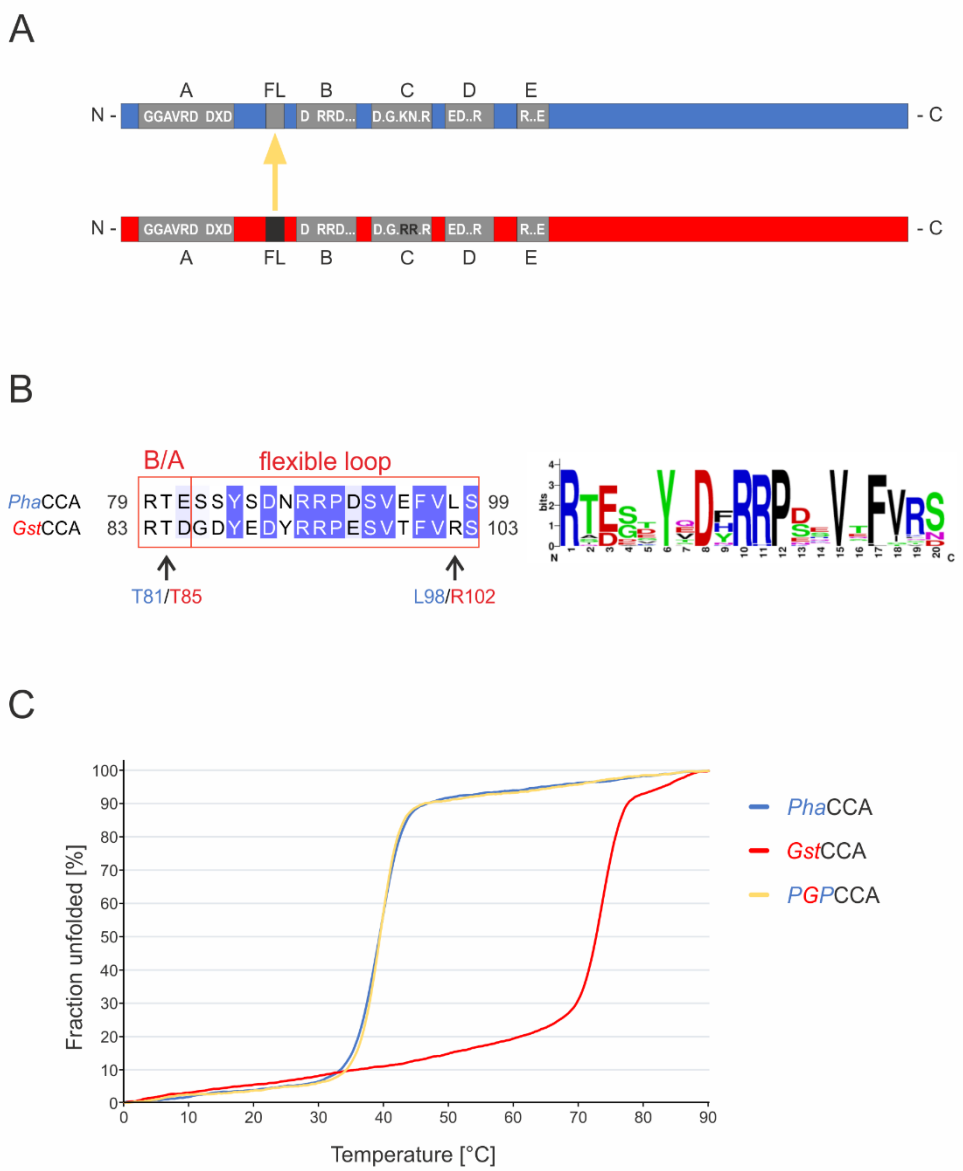


Figure S6: Replacement of flexible loop region does not lead to thermostabilization in *Pha*CCA. A) Bar diagrams of *Pha*CCA (blue) and *Gst*CCA (red). Elements of the catalytic core are indicated in grey. Yellow arrow depicts the flexible loop (FL) replacement in *Pha*CCA by the corresponding regions of *Gst*CCA. B) Sequence alignment presenting the fusion positions (T81/T85; L98/R102) of the loop transplantation. On the right, a *Weblogo* [2] serves to illustrate the loop consensus sequence in *Bacillales* [3]. C) Thermal unfolding of CCA-adding enzymes and respective loop chimera; the replacement of the flexible loop by the corresponding *Gst*CCA sequence has no stabilizing effect (yellow).

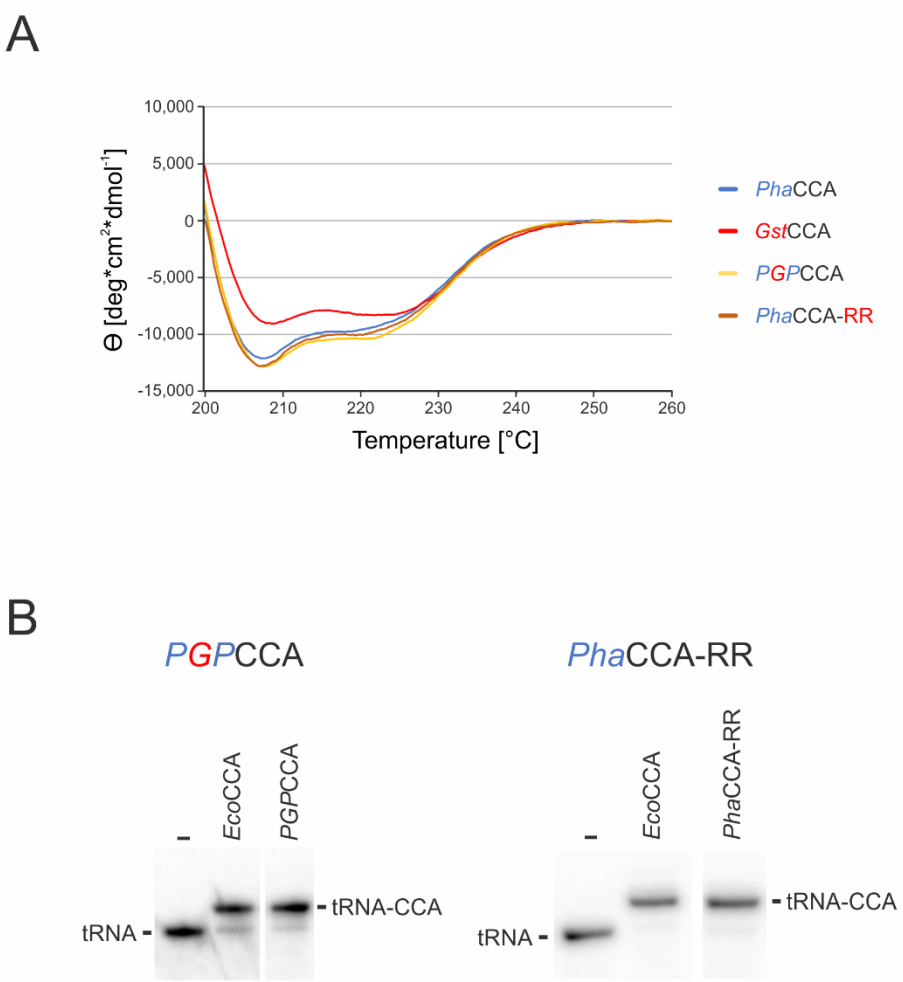


Figure S7: Analysis of both secondary structure and *in vitro* activity for respective chimeric variants. A) CD spectral analysis of chimeric variants shows no differing overall secondary structure compared to the parental enzymes. B) As determined by CCA-incorporation on a radiolabeled yeast $tRNA^{Phe}$ transcript, the investigated chimeric variants are active *in vitro* despite the manipulation within the core motifs; *E. coli* CCA-adding enzyme (*Eco*CCA) served as positive control.

VIII. Conclusion & perspectives

My thesis work was carried out in the frame of a French-German collaboration between the University of Strasbourg, the University of Leipzig and lately the University of Hamburg. The first aspect of my project was focused on the functional characterization of the particular class of nucleotidyltransferases called CCA-adding enzymes. In this purpose, I could benefit from the widely recognized expertise on these catalysts from my host laboratory in Leipzig, led by Prof. Dr. Mario Mörl. As described in this thesis, three members of the enzyme family were chosen for their unique adaptation (PhaCCA and RcuCCA) or to explore early evolutionary events of this group of enzymes (AncCCA). Then, the second part of my project took place in Strasbourg under the direction of Dr. Claude Sauter at the Institute of Molecular and Cell Biology. There, we aimed to determine crystal structures of targeted enzymes in their apo form or in complex with ligands. On the one hand, using the Xtal controller apparatus and the ChipX3 device, we sought to obtain macrocrystals suitable for X-ray diffraction experiments on a standard synchrotron beamline. On the other hand, we were looking for reproducible conditions to grow well characterized microcrystals suitable for time-resolved experiments either at synchrotron (e.g., P14/T-REXX beamline in Hamburg) or XFEL sources. Overall, the ultimate goal of my thesis work was to combine results from biochemical and structural investigations in order to lay the foundations for time-resolved x-ray crystallography experiments and the production of a real-time “molecular movie” of the 3’CCA addition on tRNAs.

1. Exploring RcuCCA biochemical properties to improve its crystallization

During my stay in Leipzig, I had the opportunity to strengthen my skills in protein purification, but also in biochemical methods such as EMSA to measure enzyme affinity for different substrates or steady-state Michaelis–Menten kinetics to determine kinetics parameters. Hence, I took part in the study of *R. culicivorax* CCA-adding enzyme which is adapted to process miniaturized tRNAs found in mitochondria from this nematode. Thanks to the work of Dr. Oliver Hennig, Susanne Philipp and Sonja Bonin, we could establish that only RcuCCA is able to process such “bizarre” hairpin-like substrates. Then, with the design of RcuCCA/HsaCCA chimeric enzymes our team successfully determined which part of the

enzyme, and more precisely, which residues from the N-terminal part are responsible for the binding and full maturation of 3' end of small tRNAs. I could further establish and compare the affinity of selected chimeras and wild type (wt) enzymes for the particular substrates. This is of great importance for the next part of the project which is the co-crystallization of the enzyme with its substrates, i.e., mini tRNAs and/or nucleotides (i.e., ATP and CTP). Getting insights into *R. culicivorax* CCA-adding enzyme structure would be key to understanding miniaturized mitochondrial tRNA processing mechanisms. As previously mentionned, crystallization trials for the apo from or in complex with miniaturized tRNAs have not been successful yet. However, my host team from Leipzig demonstrated (see chapter IV) that the introduction of RcuCCA K74 residue at corresponding position in the β-turn of the human enzyme (HsaCCA) could confer the chimera a similar catalytic capacity as wt RcuCCA on these miniaturized tRNAs. The wt human CCA-adding enzyme has previously been crystallized and its structure was also solved by my German team at a resolution of 3.4 Å (Augustin et al., 2003) and more recently to 1.9 Å (Kuhn et al 2015). As a point mutation should not disturb crystal formation dramatically, our new strategy will focus on this single point mutant of the human enzyme having similar affinity for mini tRNAs as RcuCCA to get crystals of the apo enzyme and in the presence of various substrates. The latter include mini helices which should form more compact complexes less likely to disrupt the crystal packing upon RNA reorientation on the enzyme during CCA addition.

2. Combining advanced crystallization methods for a reproducible crystal production

The main part of my work was conducted under the direction of Dr. Claude Sauter in Strasbourg. It consisted to determine reproducible crystallization conditions for the three target enzymes with the two following objectives: (i) obtaining macrocrystals useful for conventional synchrotron analysis and (ii) producing microcrystalline material suitable for time-resolved experiments on specialized synchrotron beamlines or at an XFEL source. In order to maximize our success, we combined conventional and advanced crystallization methods developed in our laboratory. In particular, as presented in chapter VI, we demonstrated the advantages of the counter-diffusion technique used in our microfluidic ChipX3 device and applied to PhaCCA enzyme. ChipX3 is not only easy to handle but it also requires very low amount of material to screen a wide range of crystallization conditions. On

the top of that, we could show that the device is perfectly adapted to *in situ* X-ray diffraction experiments at room temperature on every standard synchrotron beamline. In this regard, we are convinced that ChipX3 will constitute an efficient and low-cost alternative to current microplates for future crystallization projects of macromolecules.

To continue with the idea of facilitating and improving the research for optimal crystallization conditions, we largely demonstrated the potential and strengths of the Xtal Controller 900 . Indeed, as explained in chapter V the XC900 can be compared to a "crystallization joystick" allowing users to navigate throughout the whole phase diagram of a macromolecule. Taking PhaCCA and hen white egg lysozyme combined with Tb-Xo4 nucleant as models, we hcould show that this instrument gives access in real time to early nucleation events followed by a growth step and finally the formation of macrocrystals suitable for X-ray diffraction experiments. This said, the XC900 appears as powerful add-on for a deeper exploration of the crystallization of tricky macromolecules. Moreover, in the case of PhaCCA and AncCCA, we could monitor the formation of microcrystalline precipitate that could be reversed and driven toward the formation of micro and/or macrocrystals within few hours to days. Consistent with the increasing need for well characterized samples suitable for time-resolved experiments (i.e., microcrystal suspensions), we are planning to implement and promote the usage of this tool in the preparation process of such material.

3. Toward a molecular movie of enzymatic catalysis

Enzymatic catalysis and, to a larger extent, biological processes are *per se* dynamic and complex mechanisms. Until recently, our understanding of how biocatalysts work was based on two types of approaches. On the one side, a series of biochemical investigations centered for instance on kinetics, affinity for substrates and on the determination of essential parts of an enzyme necessary to its action were explored. On the other side, crystal structures of its apo form or in complex with its substrate(s) providing structural snapshots could confirm and bring complementary information to those obtained with biochemical assays. While conventional crystallography provides robust data on the architecture of an enzyme at definite time points, it tells us little about intermediate states and conformational transistions occurring during the catalytic reaction. Nevertheless, fantastic developments in time-resolved serial X-ray crystallography technology have recently changed the game and allow now users

to visualize the whole catalytic process taking place at the very heart of macromolecule crystals.

To do so, the preparation of well characterized nano (for SFX at XFEL) or microcrystals (for SSX at synchrotron) constitutes an absolute prerequisite to perform such experiments. As a matter of fact, microcrystals, among other advantages, are also perfect material for fast soaking with a specific ligand (Makinen & Fink, 1977; Geremia *et al.*, 2005) during a time-resolved experiment at a dedicated beamline (Kupitz *et al.*, 2016). Such "fast mixing experiments" have been successfully conducted at XFEL (Stagno *et al.*, 2017; Dasgupta *et al.*, 2019) and are currently under active development on specialized synchrotron beamlines. To date, the majority of experiments have been performed on model systems, which are well known and characterized since decades. In this regard, my thesis project was a attempt to push CCA-adding enzymes studied in my German host laboratory toward time-resolved experiments either at an XFEL or synchrotron light source. My thesis demonstrated that we were able (i) to determine the biochemical properties of a target enzyme (RcuCCA) (Hennig *et al.*, 2020), (ii) to establish a reproducible pipeline to control the growth of microcrystalline samples (PhaCCA and AncCCA) (de Wijn *et al.*, 2020) and (iii) to solve the enzyme structure in its apo form or in complex with ligands (de Wijn *et al.*, 2019) and decipher the adaptation mechanism of one member of the family (PhaCCA) to extreme life conditions (de Wijn *et al.*, 2021). Taken together and in combination with previous work between my French and German laboratories, our results lay the foundation for the future time-resolved experiments which should soon lead to a real-time molecular movie of the 3'CCA addition on a tRNA.

Bibliography

Aravind, L., & Koonin, E. V. (1999). DNA polymerase β-like nucleotidyltransferase superfamily: Identification of three new families, classification and evolutionary history. *Nucleic Acids Research*, *27*(7), 1609–1618. https://doi.org/10.1093/NAR/27.7.1609

Augustin, M. A., Reichert, A. S., Betat, H., Huber, R., Mörl, M., & Steegborn, C. (2003). Crystal structure of the human CCA-adding enzyme: Insights into template-independent polymerization. *Journal of Molecular Biology*, *328*(5), 985–994. https://doi.org/10.1016/S0022-2836(03)00381-4

Ban, N., Nissen, P., Hansen, J., Moore, P. B., & Steitz, T. A. (2000). The complete atomic structure of the large ribosomal subunit at 2.4 A resolution. *Science (New York, N.Y.)*, *289*(5481), 905–920. https://doi.org/10.1126/SCIENCE.289.5481.905

Barciszewska, M. Z., Perrigue, P. M., & Barciszewski, J. (2015). tRNA – the golden standard in molecular biology. *Molecular BioSystems*, *12*(1), 12–17. https://doi.org/10.1039/C5MB00557D

Bergfors, T. (2007). *Succeeding with seeding: some practical advice*. 1–10. https://doi.org/10.1007/978-1-4020-6316-9_1

Bergfors, T. (2009). Protein Crystallization: Second Edition. *IUL Biotechnology Series*, 504. http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-143243

Betat, H., Rammelt, C., & Mörl, M. (2010). tRNA nucleotidyltransferases: ancient catalysts with an unusual mechanism of polymerization. *Cellular and Molecular Life Sciences 2010 67:9*, *67*(9), 1447–1463. https://doi.org/10.1007/S00018-010-0271-4

Boutet, S., Fromme, P., & Editors, M. S. H. (2018). X-ray Free Electron Lasers. In S. Boutet, P. Fromme, & M. S. Hunter (Eds.), *X-ray Free Electron Lasers*. Springer International Publishing. https://doi.org/10.1007/978-3-030-00551-1

Chapman, H. N., Caleman, C., & Timneanu, N. (2014). Diffraction before destruction. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *369*(1647). https://doi.org/10.1098/RSTB.2013.0313

Chapman, H. N., Fromme, P., Barty, A., White, T. A., Kirian, R. A., Aquila, A., Hunter, M. S., Schulz, J., Deponte, D. P., Weierstall, U., Doak, R. B., Maia, F. R. N. C., Martin, A. V., Schlichting, I., Lomb, L., Coppola, N., Shoeman, R. L., Epp, S. W., Hartmann, R., … Spence, J. C. H. (2011). Femtosecond X-ray protein nanocrystallography. *Nature*, *470*(7332), 73–78. https://doi.org/10.1038/nature09750

Chayen, N. E., Saridakis, E., El-Bahar, R., & Nemirovsky, Y. (2001). Porous silicon: an effective nucleation-inducing material for protein crystallization. *Journal of Molecular Biology*, *312*(4), 591–595. https://doi.org/10.1006/JMBI.2001.4995

D'Arcy, A., Bergfors, T., Cowan-Jacob, S. W., & Marsh, M. (2014). Microseed matrix screening for optimization in protein crystallization: What have we learned? *Acta Crystallographica Section:F Structural Biology Communications*, *70*(January 2016), 1117–1126. https://doi.org/10.1107/S2053230X14015507

D'Arcy, A., Sweeney, A. Mac, & Haber, A. (2003). Using natural seeding material to generate nucleation in protein crystallization experiments. *Acta Crystallographica Section D*, *59*(7), 1343–1346. https://doi.org/10.1107/S0907444903009430

Dauter, Z. (2017). Collection of X-ray diffraction data from macromolecular crystals. *Methods in Molecular Biology (Clifton, N.J.)*, *1607*, 165. https://doi.org/10.1007/978-1-4939-7000-1_7

de Wijn, R., Hennig, O., Roche, J., Engilberge, S., Rollet, K., Fernandez-Millan, P., Brillet, K., Betat, H., Mörl, M., Roussel, A., Girard, E., Mueller-Dieckmann, C., Fox, G. C., Olieric, V., Gavira, J. A., Lorber, B., & Sauter, C. (2019). A simple and versatile microfluidic device for efficient biomacromolecule crystallization and structural analysis by serial crystallography. *Urn:Issn:2052-2525*, *6*(3), 454–464. https://doi.org/10.1107/S2052252519003622

De Wijn, Raphaël, Hennig, O., Ernst, F. G. M. M., Lorber, B., Betat, H., Mörl, M., & Sauter, C. (2018). Combining crystallogenesis methods to produce diffraction-quality crystals of a psychrophilic tRNA-maturation enzyme. *Acta Crystallographica Section F: Structural Biology Communications*, *74*(Pt 11), 747–753. https://doi.org/10.1107/S2053230X18014590

de Wijn, Raphaël, Rollet, K., Engilberge, S., McEwen, A. G. A. G., Hennig, O., Betat, H., Mörl, M., Riobé, F., Maury, O., Girard, E., Bénas, P., Lorber, B., & Sauter, C. (2020). Monitoring the Production of High Diffraction-Quality Crystals of Two Enzymes in Real Time Using In Situ Dynamic Light Scattering. *Crystals*, *10*(2), 65. https://doi.org/10.3390/cryst10020065

de Wijn, Raphaël, Rollet, K., Ernst, F. G. M., Wellner, K., Betat, H., Mörl, M., & Sauter, C. (2021). CCA-addition in the cold: Structural characterization of the psychrophilic CCA-adding enzyme from the permafrost bacterium Planococcus halocryophilus. *Computational and Structural Biotechnology Journal*, *19*, 5845–5855. https://doi.org/10.1016/J.CSBJ.2021.10.018

Desbois, S., Seabrook, S. A., & Newman, J. (2013). Some practical guidelines for UV imaging in the protein crystallization laboratory. *Acta Crystallographica Section F*, *69*(Pt 2), 201. https://doi.org/10.1107/S1744309112048634

DeTitta, G. T., Luft, J. R., & IUCr. (1995). Rate of water equilibration in vapor-diffusion crystallization: dependence on the residual pressure of air in the vapor space. *Urn:Issn:0907-4449*, *51*(5), 786–791. https://doi.org/10.1107/S090744499500299X

Diederichs, K., & Wang, M. (2017). Serial Synchrotron X-Ray Crystallography (SSX). *Methods in Molecular Biology*, *1607*, 239–272. https://doi.org/10.1007/978-1-4939-7000-1_10

Doerr, A. (2011). Diffraction before destruction. *Nature Methods 2011 8:4*, *8*(4), 283–283. https://doi.org/10.1038/nmeth0411-283

Dong, A., Xu, X., & Edwards, A. M. (2007). In situ proteolysis for protein crystallization and structure determination. *Nature Methods 2007 4:12*, *4*(12), 1019–1021. https://doi.org/10.1038/nmeth1118

Dubochet, J. (1982). *The mounting of macromolecules for electron microscopy with particular reference to surface phenomena and the treatment of support films by glow discharge.*

Dupasquier, M., Kim, S., Halkidis, K., Gamper, H., & Hou, Y. M. (2008). tRNA Integrity Is a Prerequisite for Rapid CCA Addition: Implication for Quality Control. *Journal of Molecular Biology*, *379*(3), 579–588. https://doi.org/10.1016/j.jmb.2008.04.005

Elgert, C., Rühle, A., Sandner, P., & Behrends, S. (2020). Thermal shift assay: Strengths and weaknesses of the method to investigate the ligand-induced thermostabilization of soluble guanylyl cyclase. *Journal of Pharmaceutical and Biomedical Analysis*, *181*, 113065. https://doi.org/10.1016/J.JPBA.2019.113065

Engilberge, S., Riobé, F., Di Pietro, S., Lassalle, L., Coquelle, N., Arnaud, C. A., Pitrat, D., Mulatier, J.

C., Madern, D., Breyton, C., Maury, O., & Girard, E. (2017). Crystallophore: A versatile lanthanide complex for protein crystallography combining nucleating effects, phasing properties, and luminescence. *Chemical Science*, *8*(9), 5909–5917. https://doi.org/10.1039/c7sc00758b

Engilberge, S., Riobé, F., Wagner, T., Di Pietro, S., Breyton, C., Franzetti, B., Shima, S., Girard, E., Dumont, E., & Maury, O. (2018). Unveiling the Binding Modes of the Crystallophore, a Terbium-based Nucleating and Phasing Molecular Agent for Protein Crystallography. *Chemistry - A European Journal*, *24*(39), 9739–9746. https://doi.org/10.1002/chem.201802172

Ernst, F. G. M., Erber, L., Sammler, J., Jühling, F., Betat, H., & Mörl, M. (2018). Cold adaptation of tRNA nucleotidyltransferases: A tradeoff in activity, stability and fidelity. *RNA Biology*, *15*(1), 144–155. https://doi.org/10.1080/15476286.2017.1391445

Ernst, F. G. M., Rickert, C., Bluschke, A., Betat, H., Steinhoff, H. J., & Mör, M. (2015). Domain movements during CCA-addition: A new function for motif C in the catalytic core of the human tRNA nucleotidyltransferases. *RNA Biology*, *12*(4), 435–446. https://doi.org/10.1080/15476286.2015.1018502

Fernández-Millán, P., Schelcher, C., Chihade, J., Masquida, B., Giegé, P., & Sauter, C. (2016). Transfer RNA: From pioneering crystallographic studies to contemporary tRNA biology. *Archives of Biochemistry and Biophysics*, *602*, 95–105. https://doi.org/10.1016/J.ABB.2016.03.005

Friedberg, E. C. (2014). Discovering Transfer RNA. *A Biography of Paul Berg*, 77–87. https://doi.org/10.1142/9789814569057_0007

Furukawa, R., Toma, W., Yamazaki, K., & Akanuma, S. (2020). Ancestral sequence reconstruction produces thermally stable enzymes with mesophilic enzyme-like catalytic properties. *Scientific Reports 2020 10:1*, *10*(1), 1–13. https://doi.org/10.1038/s41598-020-72418-4

García-Ruiz, J. M. (2003). Nucleation of protein crystals. *Journal of Structural Biology*, *142*(1), 22–31. https://doi.org/10.1016/S1047-8477(03)00035-2

Garman, E. F., & Owen, R. L. (2005). Cryocooling and radiation damage in macromolecular crystallography. *Urn:Issn:0907-4449*, *62*(1), 32–47. https://doi.org/10.1107/S0907444905034207

Ghisaidoobe, A. B. T., & Chung, S. J. (2014). Intrinsic Tryptophan Fluorescence in the Detection and Analysis of Proteins: A Focus on Förster Resonance Energy Transfer Techniques. *International Journal of Molecular Sciences 2014, Vol. 15, Pages 22518-22538*, *15*(12), 22518–22538. https://doi.org/10.3390/IJMS151222518

Giegé, R., & McPherson, A. (2006). General methods. *International Tables for Crystallography*, 81–93. https://doi.org/10.1107/97809553602060000660

Giege, R., Moras, D., & Thierry, J. C. (1977). Yeast transfer RNAAsp: A new high-resolution X-ray diffracting crystal form of a transfer RNA. *Journal of Molecular Biology*, *115*(1), 91–96. https://doi.org/10.1016/0022-2836(77)90248-0

Gowthaman, R., Guest, J. D., Yin, R., Adolf-Bryfogle, J., Schief, W. R., & Pierce, B. G. (2021). CoV3D: a database of high resolution coronavirus protein structures. *Nucleic Acids Research*, *49*(D1), D282–D287. https://doi.org/10.1093/NAR/GKAA731

Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N., & Altman, S. (1983). The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell*, *35*(3 Pt 2), 849–857. https://doi.org/10.1016/0092-8674(83)90117-4

Henderson, R. (1995). The potential and limitations of neutrons, electrons and X-rays for atomic

resolution microscopy of unstained biological molecules. *Quarterly Reviews of Biophysics*, *28*(2), 171–193. https://doi.org/10.1017/S003358350000305X

Holm, L., & Sander, C. (1995). DNA polymerase β belongs to an ancient nucleotidyltransferase superfamily. *Trends in Biochemical Sciences*, *20*(9), 345–347. https://doi.org/10.1016/S0968-0004(00)89071-4

Hopper, A. K., & Nostramo, R. T. (2019). TRNA processing and subcellular trafficking proteins multitask in pathways for other RNAs. *Frontiers in Genetics*, *10*(FEB), 1–14. https://doi.org/10.3389/fgene.2019.00096

Houben, L., Weissman, H., Wolf, S. G., & Rybtchinski, B. (2020). A mechanism of ferritin crystallization revealed by cryo-STEM tomography. *Nature 2020 579:7800*, *579*(7800), 540–543. https://doi.org/10.1038/s41586-020-2104-4

James Sumner, B. B. (1926). *THE ISOLATION AND CRYSTALLIZATION OF THE ENZYME UREASE*.

Jaskolski, M., Dauter, Z., & Wlodawer, A. (2014). A brief history of macromolecular crystallography, illustrated by a family tree and its Nobel fruits. *FEBS Journal*, *281*(18), 3985–4009. https://doi.org/10.1111/FEBS.12796

Jeangerard, D., Savko, M., Ciccone, L., Desjardins, K., Le Jollec, A., Haouz, A., & Shepard, W. (2018). *From Plate Screening to Artificial Intelligence: Innovative developments on PROXIMA 2A at Synchrotron SOLEIL; From Plate Screening to Artificial Intelligence: Innovative developments on PROXIMA 2A at Synchrotron SOLEIL*. https://doi.org/10.18429/JACoW-MEDSI2018-WEPH36

Jeruzalmi, D., & Steitz, T. A. (1997). Use of organic cosmotropic solutes to crystallize flexible proteins: application to T7 RNA polymerase and its complex with the inhibitor T7 lysozyme. *Journal of Molecular Biology*, *274*(5), 748–756. https://doi.org/10.1006/JMBI.1997.1366

Jühling, T., Duchardt-Ferner, E., Bonin, S., Wöhnert, J., Pütz, J., Florentz, C., Betat, H., Sauter, C., & Mörl, M. (2018). Small but large enough: structural properties of armless mitochondrial tRNAs from the nematode Romanomermis culicivorax. *Nucleic Acids Research*, *46*(17), 9170–9180. https://doi.org/10.1093/NAR/GKY593

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., … Hassabis, D. . (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, *596*, 583. https://doi.org/10.1038/s41586-021-03819-2

Kabsch, W. (2010). *XDS*. *Acta Crystallographica Section D Biological Crystallography*, *66*(2), 125–132. https://doi.org/10.1107/S0907444909047337

Kaischew, R., & Stranski, I. N. (1934). Zur Theorie der linearen Kristallisationsgeschwindigkeit. *Zeitschrift Für Physikalische Chemie*, *170A*(1), 295–299. https://doi.org/10.1515/ZPCH-1934-17027

Kedersha, N. L., Heuser, J. E., Chugani, D. C., & Rome, L. H. (1991). Vaults. III. Vault ribonucleoprotein particles open into flower-like structures with octagonal symmetry. *The Journal of Cell Biology*, *112*(2), 225–235. https://doi.org/10.1083/JCB.112.2.225

Kendrew, J. C., Bodo, G., Dintzis, H. M., Parrish, R. G., Wyckoff, H., & Phillips, D. C. (1958). A Three-Dimensional Model of the Myoglobin Molecule Obtained by X-Ray Analysis. *Nature 1958*

181:4610, *181*(4610), 662–666. https://doi.org/10.1038/181662a0

Kern, J., Loll, B., Lüneberg, C., DiFiore, D., Biesiadka, J., Irrgang, K. D., & Zouni, A. (2005). Purification, characterisation and crystallisation of photosystem II from Thermosynechococcus elongatus cultivated in a new type of photobioreactor. *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, *1706*(1–2), 147–157. https://doi.org/10.1016/J.BBABIO.2004.10.007

Kim, S. H., Yoo, H. J., Park, E. J., & Na, D. H. (2021). Nano Differential Scanning Fluorimetry-Based Thermal Stability Screening and Optimal Buffer Selection for Immunoglobulin G. *Pharmaceuticals*, *15*(1), 29. https://doi.org/10.3390/PH15010029

Kim, S., Liu, C., Halkidis, K., Gamper, H. B., & Hou, Y.-M. (2009). Distinct kinetic determinants for the stepwise CCA addition to tRNA. *RNA*, *15*(10), 1827–1836. https://doi.org/10.1261/RNA.1669109

Kissick, D. J., Wanapun, D., & Simpson, G. J. (2011). Second-order nonlinear optical imaging of chiral crystals. *Annual Review of Analytical Chemistry (Palo Alto, Calif.)*, *4*, 419–437. https://doi.org/10.1146/ANNUREV.ANCHEM.111808.073722

Koclega, K. D., Chruszcz, M., Zimmerman, M. D., Bujacz, G., & Minor, W. (2009). 'Hot' macromolecular crystals. *Crystal Growth & Design*, *10*(2), 580. https://doi.org/10.1021/CG900971H

Kovermann, M., Rogne, P., & Wolf-Watz, M. (2016). Protein dynamics and function from solution state NMR spectroscopy. *Quarterly Reviews of Biophysics*, *49*, e6. https://doi.org/10.1017/S0033583516000019

Kresge, N., Simoni, R. D., & Hill, R. L. (2005). The Discovery of tRNA by Paul C. Zamecnik. *Journal of Biological Chemistry*, *280*(40), e37–e39. https://doi.org/10.1016/S0021-9258(20)79029-0

Kühlbrandt, W., Wang, D. N., & Fujiyoshi, Y. (1994). Atomic model of plant light-harvesting complex by electron crystallography. *Nature 1994 367:6464*, *367*(6464), 614–621. https://doi.org/10.1038/367614a0

Kuhn, C. D., Wilusz, J. E., Zheng, Y., Beal, P. A., & Joshua-Tor, L. (2015). On-enzyme refolding permits small RNA and tRNA surveillance by the CCA-adding enzyme. *Cell*, *160*(4), 644–658. https://doi.org/10.1016/j.cell.2015.01.005

Kupitz, C., Basu, S., Grotjohann, I., Fromme, R., Zatsepin, N. A., Rendek, K. N., Hunter, M. S., Shoeman, R. L., White, T. A., Wang, D., James, D., Yang, J. H., Cobb, D. E., Reeder, B., Sierra, R. G., Liu, H., Barty, A., Aquila, A. L., Deponte, D., … Fromme, P. (2014). Serial time-resolved crystallography of photosystem II using a femtosecond X-ray laser. *Nature 2014 513:7517*, *513*(7517), 261–265. https://doi.org/10.1038/nature13453

Kupitz, C., Grotjohann, I., Conrad, C. E., Roy-Chowdhury, S., Fromme, R., & Fromme, P. (2014). Microcrystallization techniques for serial femtosecond crystallography using photosystem II from Thermosynechococcus elongatus as a model system. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1647). https://doi.org/10.1098/rstb.2013.0316

Kupitz, C., Olmos, J. L., Holl, M., Tremblay, L., Pande, K., Pandey, S., Oberthür, D., Hunter, M., Liang, M., Aquila, A., Tenboer, J., Calvey, G., Katz, A., Chen, Y., Wiedorn, M. O., Knoska, J., Meents, A., Majriani, V., Norwood, T., … Schmidt, M. (2017). Structural enzymology using X-ray free electron lasers. *Structural Dynamics*, *4*(4), 044003. https://doi.org/10.1063/1.4972069

Leslie, A. G. W. (2005). The integration of macromolecular diffraction data. *Urn:Issn:0907-4449*, *62*(1), 48–57. https://doi.org/10.1107/S0907444905039107

Li, F., Xiong, Y., Wang, J., Cho, H. D. D., Tomita, K., Weiner, A. M., & Steitz, T. A. (2002a). Crystal structures of the Bacillus stearothermophilus CCA-adding enzyme and its complexes with ATP or CTP. *Cell*, *111*(6), 815–824. https://doi.org/10.1016/S0092-8674(02)01115-7

Li, F., Xiong, Y., Wang, J., Cho, H. D. D., Tomita, K., Weiner, A. M., & Steitz, T. A. (2002b). Crystal Structures of the Bacillus stearothermophilus CCA-Adding Enzyme and Its Complexes with ATP or CTP. *Cell*, *111*(6), 815–824. https://doi.org/10.1016/S0092-8674(02)01115-7

Liebschner, D., Afonine, P. V., Baker, M. L., Bunkoczi, G., Chen, V. B., Croll, T. I., Hintze, B., Hung, L. W., Jain, S., McCoy, A. J., Moriarty, N. W., Oeffner, R. D., Poon, B. K., Prisant, M. G., Read, R. J., Richardson, J. S., Richardson, D. C., Sammito, M. D., Sobolev, O. V., … Adams, P. D. (2019). Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Urn:Issn:2059-7983*, *75*(10), 861–877. https://doi.org/10.1107/S2059798319011471

Littlechild, J. A. (2015). Enzymes from extreme environments and their industrial applications. *Frontiers in Bioengineering and Biotechnology*, *3*(OCT), 161. https://doi.org/10.3389/FBIOE.2015.00161/BIBTEX

LORBER, B., KERN, D., MEJDOUB, H., BOULANGER, Y., REINBOLT, J., & GIEGÉ, R. (1987). The microheterogeneity of the crystallizable yeast cytoplasmic aspartyl-tRNA synthetase. *European Journal of Biochemistry*, *165*(2), 409–417. https://doi.org/10.1111/J.1432-1033.1987.TB11454.X

Lorber, B., Sauter, C., Théobald-Dietrich, A., Moreno, A., Schellenberger, P., Robert, M. C., Capelle, B., Sanglier, S., Potier, N., & Giegé, R. (2009). Crystal growth of proteins, nucleic acids, and viruses in gels. *Progress in Biophysics and Molecular Biology*, *101*(1–3), 13–25. https://doi.org/10.1016/j.pbiomolbio.2009.12.002

Luft, J. R., Albright, D. T., Baird, J. K., & DeTitta, G. T. (1996). The rate of water equilibration in vapor-diffusion crystallizations: Dependence on the distance from the droplet to the reservoir. *Acta Crystallographica Section D: Biological Crystallography*, *52*(6), 1098–1106. https://doi.org/10.1107/S0907444996006932/FULL

Luft, J. R., Wolfley, J. R., & Snell, E. H. (2011). What's in a drop? Correlating observations and outcomes to guide macromolecular crystallization experiments. *Crystal Growth & Design*, *11*(3), 651. https://doi.org/10.1021/CG1013945

Maier, R., Zocher, G., Sauter, A., Da Vela, S., Matsarskaia, O., Schweins, R., Sztucki, M., Zhang, F., Stehle, T., & Schreiber, F. (2020). Protein Crystallization in the Presence of a Metastable Liquid-Liquid Phase Separation. *Crystal Growth and Design*, *20*(12), 7951–7962. https://doi.org/10.1021/ACS.CGD.0C01219/SUPPL_FILE/CG0C01219_SI_002.PDF

McPherson, A. (1976). The growth and preliminary investigation of protein and nucleic acid crystals for X-ray diffraction analysis. *Methods of Biochemical Analysis*, *23*(0), 249–345. https://doi.org/10.1002/9780470110430.CH4

McPherson, A., Malkin, A. J., & Kuznetsov, Y. G. (2000). Atomic force microscopy in the study of macromolecular crystal growth. *Annual Review of Biophysics and Biomolecular Structure*, *29*, 361–410. https://doi.org/10.1146/ANNUREV.BIOPHYS.29.1.361

McPherson, Alexander, & Kuznetsov, Y. G. (2014). Mechanisms, kinetics, impurities and defects: consequences in macromolecular crystallization. *Acta Crystallographica. Section F, Structural Biology Communications*, *70*(Pt 4), 384. https://doi.org/10.1107/S2053230X14004816

Mehrabi, P., Bücker, R., Bourenkov, G., Ginn, H. M., von Stetten, D., Müller-Werkmeister, H. M., Kuo, A., Morizumi, T., Eger, B. T., Ou, W. L., Oghbaey, S., Sarracini, A., Besaw, J. E., Pare-Labrosse, O., Meier, S., Schikora, H., Tellkamp, F., Marx, A., Sherrell, D. A., … Miller, R. J. D. (2021). Serial femtosecond and serial synchrotron crystallography can yield data of equivalent quality: A systematic comparison. *Science Advances*, *7*(12). https://doi.org/10.1126/SCIADV.ABF1380/ASSET/9DDDB450-BA99-48CF-B814-1EA8FEBABB1A/ASSETS/GRAPHIC/ABF1380-F3.JPEG

Mehrabi, Pedram, Müller-Werkmeister, H. M., Leimkohl, J.-P., Schikora, H., Ninkovic, J., Krivokuca, S., Andriček, L., Epp, S. W., Sherrell, D., Owen, R. L., Pearson, A. R., Tellkamp, F., Schulz, E. C., & Miller, R. J. D. (2020). The HARE chip for efficient time-resolved serial synchrotron crystallography. *Journal of Synchrotron Radiation*, *27*(2), 360–370. https://doi.org/10.1107/S1600577520000685

Mehrabi, Pedram, Muller-Werkmeister, H. M., Leimkohl, J. P., Schikora, H., Ninkovic, J., Krivokuca, S., Andricek, L., Epp, S. W., Sherrell, D., Owen, R. L., Pearson, A. R., Tellkamp, F., Schulz, E. C., & Dwayne Miller, R. J. (2020). The HARE chip for efficient time-resolved serial synchrotron crystallography. *Journal of Synchrotron Radiation*, *27*(2), 360–370. https://doi.org/10.1107/S1600577520000685/WZ5004SUP2.ZIP

Mikol, V, Hirsch, E., & Giegc, R. (1989). *Monitoring protein crystallization by dynamic light scattering*. *258*(1), 63–66.

Mikol, Vincent, Rodeau, J. L., & Giegé, R. (1990). Experimental determination of water equilibration rates in the hanging drop method of protein crystallization. *Analytical Biochemistry*, *186*(2), 332–339. https://doi.org/10.1016/0003-2697(90)90091-M

Mykytczuk, N. C. S., Foote, S. J., Omelon, C. R., Southam, G., Greer, C. W., & Whyte, L. G. (2013). Bacterial growth at -15 °C; molecular insights from the permafrost bacterium Planococcus halocryophilus Or1. *ISME Journal*, *7*(6), 1211–1226. https://doi.org/10.1038/ismej.2013.8

Mykytczuk, N. C. S., Wilhelm, R. C., & Whyte, L. G. (2012). Planococcus halocryophilus sp. nov., an extreme sub-zero species from high Arctic permafrost. *International Journal of Systematic and Evolutionary Microbiology*, *62*(Pt_8), 1937–1944. https://doi.org/10.1099/IJS.0.035782-0

Nakane, T., Kotecha, A., Sente, A., McMullan, G., Masiulis, S., Brown, P. M. G. E., Grigoras, I. T., Malinauskaite, L., Malinauskas, T., Miehling, J., Uchański, T., Yu, L., Karia, D., Pechnikova, E. V., de Jong, E., Keizer, J., Bischoff, M., McCormack, J., Tiemeijer, P., … Scheres, S. H. W. (2020). Single-particle cryo-EM at atomic resolution. *Nature*, *587*(7832), 152–156. https://doi.org/10.1038/S41586-020-2829-0

Neuenfeldt, A., Just, A., Betat, H., Mörl, M., A, N., A, J., H, B., & M, M. (2008). Evolution of tRNA nucleotidyltransferases: A small deletion generated CC-adding enzymes. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(23), 7953–7958. https://doi.org/10.1073/pnas.0801971105

Newman, J., Fazio, V. J., Lawson, B., & Peat, T. S. (2010). The C6 Web Tool: A Resource for the Rational Selection of Crystallization Conditions. *Crystal Growth and Design*, *10*(6), 2785–2792. https://doi.org/10.1021/CG1004209

Ng, J. D., Lorber, B., Witz, J., Théobald-Dietrich, A., Kern, D., & Giegé, R. (1996). The crystallization of biological macromolecules from precipitates: evidence for Ostwald ripening. *Journal of Crystal Growth*, *168*(1–4), 50–62. https://doi.org/10.1016/0022-0248(96)00362-4

Ng, J. D., Sauter, C., Lorber, B., Kirkland, N., Arnez, J., & Giegé, R. (2002). Comparative analysis of

space-grown and earth-grown crystals of an aminoacyl-tRNA synthetase: Space-grown crystals are more useful for structural determination. *Acta Crystallographica Section D: Biological Crystallography*. https://doi.org/10.1107/S0907444902003177

Okimoto, R., & Wolstenholme, D. R. (1990). A set of tRNAs that lack either the TψC arm or the dihydrouridine arm: Towards a minimal tRNA adaptor. *EMBO Journal*, *9*(10), 3405–3411. https://doi.org/10.1002/j.1460-2075.1990.tb07542.x

Otálora, F., Gavira, J. A., Ng, J. D., & García-Ruiz, J. M. (2009). Counterdiffusion methods applied to protein crystallization. *Progress in Biophysics and Molecular Biology*, *101*(1–3), 26–37. https://doi.org/10.1016/J.PBIOMOLBIO.2009.12.004

Pan, B., Xiong, Y., & Steitz, T. A. (2010). How the CCA-adding enzyme selects adenine over cytosine at position 76 of tRNA. *Science*, *330*(6006), 937–940. https://doi.org/10.1126/SCIENCE.1194985/SUPPL_FILE/PAN.SOM.PDF

Pearson, A. R., & Mehrabi, P. (2020). Serial synchrotron crystallography for time-resolved structural biology. *Current Opinion in Structural Biology*, *65*, 168–174. https://doi.org/10.1016/J.SBI.2020.06.019

Perutz, M. F., & Lehmann, H. (1968). *Molecular Pathology of Human Haemoglobin*.

Perutz, M. F., Rossmann, M. G., Cullis, A. F., Muirhead, H., Will, G., & North, A. C. T. (1960). Structure of Hæmoglobin: A Three-Dimensional Fourier Synthesis at 5.5-Å. Resolution, Obtained by X-Ray Analysis. *Nature 1960 185:4711*, *185*(4711), 416–422. https://doi.org/10.1038/185416a0

Perutzt, M. F. (1976). Fundamental research in molecular biology: relevance to medicine*. *Nature*, *262*.

Phizicky, E. M., & Hopper, A. K. (2010). tRNA biology charges to the front. *Genes & Development*, *24*(17), 1832–1860. https://doi.org/10.1101/GAD.1956510

Purslow, J. A., Khatiwada, B., Bayro, M. J., & Venditti, V. (2020). NMR Methods for Structural Characterization of Protein-Protein Complexes. *Frontiers in Molecular Biosciences*, *7*(9). https://doi.org/10.3389/FMOLB.2020.00009

Pusey, M., Barcena, J., Morris, M., Singhal, A., Yuan, Q., & Ng, J. (2015). Trace fluorescent labeling for protein crystallization. *Acta Crystallographica Section:F Structural Biology Communications*, *71*(i), 806–814. https://doi.org/10.1107/S2053230X15008626

Ramakrishnan, V. (2002). Ribosome Structure and the Mechanism of Translation. *Cell*, *108*(4), 557–572. https://doi.org/10.1016/S0092-8674(02)00619-0

Richard, S. B., Madern, D., Garcin, E., & Zaccai, G. (2000). *Halophilic Adaptation: Novel Solvent Protein Interactions Observed in the 2.9 and 2.6 Å Resolution Structures of the Wild Type and a Mutant of Malate Dehydrogenase from Haloarcula marismortui ‡*. https://doi.org/10.1021/bi991001a

Robert, X., & Gouet, P. (2014). Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Research*, *42*(W1), W320–W324. https://doi.org/10.1093/NAR/GKU316

Robertus, J. D., Ladner, J. E., Finch, J. T., Rhodes, D., Brown, R. S., Clark, B. F. C., & Klug, A. (1974). Structure of yeast phenylalanine tRNA at 3 Å resolution. *Nature 1974 250:5467*, *250*(5467), 546–551. https://doi.org/10.1038/250546a0

Ross, C. M., Foley, G., Boden, M., & Gillam, E. M. J. (2022). Using the Evolutionary History of Proteins to Engineer Insertion-Deletion Mutants from Robust, Ancestral Templates Using Graphical Representation of Ancestral Sequence Predictions (GRASP). *Methods in Molecular Biology*, *2397*, 85–110. https://doi.org/10.1007/978-1-0716-1826-4_6

Salemme, F. R. (1972). A free interface diffusion technique for the crystallization of proteins for X-ray crystallography. *Archives of Biochemistry and Biophysics*, *151*(2), 533–539. https://doi.org/10.1016/0003-9861(72)90530-9

Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989). *Molecular cloning : a laboratory manual*. https://books.google.com/books/about/Molecular_Cloning.html?hl=fr&id=8WViPwAACAAJ

Sanger, F., Nicklen, S., & Coulson, A. R. (1977). *DNA sequencing with chain-terminating inhibitors (DNA polymerase/nucleotide sequences/bacteriophage 4X174)*. *74*(12), 5463–5467.

Saridakis, E., & Chayen, N. E. (2009). Towards a 'universal' nucleant for protein crystallization. *Trends in Biotechnology*, *27*(2), 99–106. https://doi.org/10.1016/J.TIBTECH.2008.10.008

Savino, R., & Monti, R. (1996). Buoyancy and surface-tension-driven convection in hanging-drop protein crystallizer. *Journal of Crystal Growth*, *165*(3), 308–318. https://doi.org/10.1016/0022-0248(96)00151-0

Sawaya, M. R., Pelletier, H., Kumar, A., Wilson, S. H., & Kraut, J. (1994). Crystal Structure of Rat DNA Polymerase β: Evidence for a Common Polymerase Mechanism. *Science*, *264*(5167), 1930–1935. https://doi.org/10.1126/SCIENCE.7516581

Schmitz, K. S. (1990). *An introduction to dynamic light scattering by macromolecules*. 449.

Schrödinger, LLC. (2015). *The {PyMOL} Molecular Graphics System, Version~1.8*.

Schubert, R., Meyer, A., Dierks, K., Kapis, S., Reimer, R., Einspahr, H., Perbandt, M., & Betzel, C. (2015). Reliably distinguishing protein nanocrystals from amorphous precipitate by means of depolarized dynamic light scattering. *Urn:Issn:1600-5767*, *48*(5), 1476–1484. https://doi.org/10.1107/S1600576715014740

Schürer, H., Lang, K., Schuster, J., & Mörl, M. (2002). A universal method to produce in vitro transcripts with homogeneous 3′ ends. *Nucleic Acids Research*, *30*(12), e56. https://doi.org/10.1093/NAR/GNF055

Sharma, M., & Luthra-Guptasarma, M. (2009). Degradation of proteins upon storage at near-neutral pH: Indications of a proteolytic/gelatinolytic activity associated with aggregates. *Biochimica et Biophysica Acta (BBA) - General Subjects*, *1790*(10), 1282–1294. https://doi.org/10.1016/J.BBAGEN.2009.06.010

Shi, D., Lewis, M. R., Young, H. S., & Stokes, D. L. (1998). Three-dimensional crystals of Ca2+-ATPase from sarcoplasmic reticulum: merging electron diffraction tilt series and imaging the (h, k, 0) projection. *Journal of Molecular Biology*, *284*(5), 1547–1564. https://doi.org/10.1006/JMBI.1998.2283

Shi, P. Y., Maizels, N., & Weiner, A. M. (1998). CCA addition by tRNA nucleotidyltransferase: Polymerization without translocation? *EMBO Journal*, *17*(11), 3197–3206. https://doi.org/10.1093/EMBOJ/17.11.3197

Shtukenberg, A. G., García-Ruiz, J. M., & Kahr, B. (2021). Punin Ripening and the Classification of Solution-Mediated Recrystallization Mechanisms. *Crystal Growth and Design*, *21*(2), 1267–1277. https://doi.org/10.1021/acs.cgd.0c01545

Skarzynski, T. (2013). Collecting data in the home laboratory: evolution of X-ray sources, detectors and working practices. *Acta Crystallographica Section D: Biological Crystallography*, *69*(Pt 7), 1283. https://doi.org/10.1107/S0907444913013619

Slade, A., Kattini, R., Campbell, C., & Holcik, M. (2020). *Diseases Associated with Defects in tRNA CCA Addition*. 1–14. https://doi.org/10.3390/ijms21113780

Sleutel, M., & Van Driessche, A. E. S. (2018). Nucleation of protein crystals – a nanoscopic perspective. *Nanoscale*, *10*, 12256. https://doi.org/10.1039/c8nr02867b

Snell, E. H., & Helliwell, J. R. (2005). Macromolecular crystallization in microgravity. *Reports on Progress in Physics*, *68*(4), 799–853. https://doi.org/10.1088/0034-4885/68/4/R02

Some, D., Amartely, H., Tsadok, A., & Lebendiker, M. (2019). Characterization of proteins by size-exclusion chromatography coupled to multi-angle light scattering (Sec-mals). *Journal of Visualized Experiments*, *2019*(148). https://doi.org/10.3791/59615

Stetefeld, J., McKenna, S. A., & Patel, T. R. (2016). Dynamic light scattering: a practical guide and applications in biomedical sciences. *Biophysical Reviews*, *8*(4), 409. https://doi.org/10.1007/S12551-016-0218-6

Stokes, G. G. (2009). On the Theories of the Internal Friction of Fluids in Motion, and of the Equilibrium and Motion of Elastic Solids. *Mathematical and Physical Papers Vol.1*, 75–129. https://doi.org/10.1017/CBO9780511702242.005

Stranski, I. N., & Kaischew, R. (1934). Über den Mechanismus des Gleichgewichtes kleiner Kriställchen. *Zeitschrift Für Physikalische Chemie*, *26B*(1), 100–113. https://doi.org/10.1515/ZPCH-1934-2610

Strutt, J. W. (1871). XXXVI. On the light from the sky, its polarization and colour. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, *41*(273), 274–279. https://doi.org/10.1080/14786447108640479

Suhre, K., & Sanejouand, Y. H. (2004). ElNémo: A normal mode web server for protein movement analysis and the generation of templates for molecular replacement. *Nucleic Acids Research*, *32*(WEB SERVER ISS.), 610–614. https://doi.org/10.1093/nar/gkh368

Suzuki, T. (2021). The expanding world of tRNA modifications and their disease relevance. *Nature Reviews Molecular Cell Biology 2021 22:6*, *22*(6), 375–392. https://doi.org/10.1038/s41580-021-00342-0

Suzuki, T., Nagao, A., & Suzuki, T. (2011). Human mitochondrial trnas: Biogenesis, function, structural aspects, and diseases. *Annual Review of Genetics*, *45*, 299–329. https://doi.org/10.1146/ANNUREV-GENET-110410-132531

Théobald-Dietrich, A., de Wijn, R., Rollet, K., Bluhm, A., Rudinger-Thirion, J., Paulus, C., Lorber, B., Thureau, A., Frugier, M., & Sauter, C. (2020). Structural Analysis of RNA by Small-Angle X-ray Scattering. In *Methods in Molecular Biology* (Vol. 2113, pp. 189–215). Humana Press Inc. https://doi.org/10.1007/978-1-0716-0278-2_14

Toh, Y., Takeshita, D., Nagaike, T., Numata, T., & Tomita, K. (2011a). Article Mechanism for the Alteration of the Substrate Specificities of Template-Independent RNA Polymerases. *Structure/Folding and Design*, *19*, 232–243. https://doi.org/10.1016/j.str.2010.12.006

Toh, Y., Takeshita, D., Nagaike, T., Numata, T., & Tomita, K. (2011b). Mechanism for the alteration of the substrate specificities of template-independent RNA polymerases. *Structure*, *19*(2), 232–243. https://doi.org/10.1016/j.str.2010.12.006

Toh, Y., Takeshita, D., Numata, T., Fukai, S., Nureki, O., & Tomita, K. (2009). Mechanism for the definition of elongation and termination by the class II CCA-adding enzyme. *The EMBO Journal*, *28*(21), 3353–3365. https://doi.org/10.1038/EMBOJ.2009.260

Tomari, Y., Suzuki, T., & Ueda, T. (2002). tRNA Recognition by CCA-adding enzyme. *Nucleic Acids Research Supplement*, *2*, 77–78. http://www.mips.biochem.mpg.de/cgi-bin/proj/medgen/mit

Tomita, K., Fukai, S., Ishitani, R., Ueda, T., Takeuchi, H., Vassylyev, D. G., & Nureki, O. (2004). Structural basis for template-independent RNA polymerization. *Nature*, *430*(7000), 700–704. https://doi.org/10.1038/nature02712

Tomita, K., & Weiner, A. M. (2002). Closely Related CC- and A-adding Enzymes Collaborate to Construct and Repair the 3′-Terminal CCA of tRNA in Synechocystis sp. and Deinococcus radiodurans. *Journal of Biological Chemistry*, *277*(50), 48192–48198. https://doi.org/10.1074/JBC.M207527200

Tournier, V., Topham, C. M., Gilles, A., David, B., Folgoas, C., Moya-Leclair, E., Kamionka, E., Desrousseaux, M., Texier, H., Gavalda, S., Cot, M., Guémard, E., Dalibey, M., Nomme, J., Cioci, G., Barbe, S., Chateau, M., André, I., Duquesne, S., & Marty, A. (2020). An engineered PET depolymerase to break down and recycle plastic bottles. *Nature*, *580*. https://doi.org/10.1038/s41586-020-2149-4

Vekilov, P. G., Thomas, B. R., & Rosenberger, F. (1998). *Effects of Convective Solute and Impurity Transport in Protein Crystal Growth*. https://pubs.acs.org/sharingguidelines

Vonrhein, C., Flensburg, C., Keller, P., Sharff, A., Smart, O., Paciorek, W., Womack, T., & Bricogne, G. (2011). Data processing and analysis with the autoPROC toolbox. *Acta Crystallographica. Section D, Biological Crystallography*, *67*(Pt 4), 293–302. https://doi.org/10.1107/S0907444911007773

Watanabe, Y., & Inoko, Y. (2009). Size-exclusion chromatography combined with small-angle X-ray scattering optics. *Journal of Chromatography A*, *1216*(44), 7461–7465. https://doi.org/10.1016/J.CHROMA.2009.02.053

Webb, B., & Sali, A. (2016). Comparative Protein Structure Modeling Using MODELLER. *Current Protocols in Bioinformatics*, *54*(1), 5.6.1-5.6.37. https://doi.org/10.1002/CPBI.3

Węgrzyn, G., & Węgrzyn, A. (2008). Is tRNA only a translation factor or also a regulator of other processes? *Journal of Applied Genetics 2008 49:1*, *49*(1), 115–122. https://doi.org/10.1007/BF03195257

Wellner, K., Pöhler, M. T., Betat, H., & Mörl, M. (2019). Dual expression of CCA-adding enzyme and RNase T in Escherichia coli generates a distinct cca growth phenotype with diverse applications. *Nucleic Acids Research*, *47*(7), 3631–3639. https://doi.org/10.1093/nar/gkz133

Wende, S., Bonin, S., Götze, O., Betat, H., & Mörl, M. (2015). The identity of the discriminator base has an impact on CCA addition. *Nucleic Acids Research*, *43*(11), 5617–5629. https://doi.org/10.1093/NAR/GKV471

Wende, S., Platzer, E. G., Jühling, F., Pütz, J., Florentz, C., Stadler, P. F., & Mörl, M. (2014). Biological evidence for the world's smallest tRNAs. *Biochimie*, *100*(1), 151–158. https://doi.org/10.1016/J.BIOCHI.2013.07.034

Wernimont, A., & Edwards, A. (2009). In Situ Proteolysis to Generate Crystals for Structure Determination: An Update. *PLOS ONE*, *4*(4), e5094. https://doi.org/10.1371/JOURNAL.PONE.0005094

White, T. A., Kirian, R. A., Martin, A. V., Aquila, A., Nass, K., Barty, A., & Chapman, H. N. (2012). CrystFEL: a software suite for snapshot serial crystallography. *Urn:Issn:0021-8898*, *45*(2), 335–341. https://doi.org/10.1107/S0021889812002312

Williamson, A. R. (2000). Creating a structural genomics consortium. *Nature Structural Biology 2000 7:11*, *7*(11), 953–953. https://doi.org/10.1038/80726

Wilson, W. W. (1990). Monitoring crystallization experiments using dynamic light scattering: Assaying and monitoring protein crystallization in solution. *Methods*, *1*(1), 110–117. https://doi.org/10.1016/S1046-2023(05)80154-9

Wilusz, J. E., Whipple, J. M., Phizicky, E. M., & Sharp, P. A. (2011). tRNAs marked with CCACCA are targeted for degradation. *Science*, *334*(6057), 817–821. https://doi.org/10.1126/SCIENCE.1213671

Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G. W., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A., & Wilson, K. S. (2011). Overview of the CCP4 suite and current developments. *Urn:Issn:0907-4449*, *67*(4), 235–242. https://doi.org/10.1107/S0907444910045749

Winter, G. (2010). xia2: an expert system for macromolecular crystallography data reduction. *Journal of Applied Crystallography*, *43*(1), 186–190. https://doi.org/10.1107/S0021889809045701

Winter, Graeme, Beilsten-Edmands, J., Devenish, N., Gerstel, M., Gildea, R. J., McDonagh, D., Pascal, E., Waterman, D. G., Williams, B. H., & Evans, G. (2022). DIALS as a toolkit. *Protein Science : A Publication of the Protein Society*, *31*(1), 232–250. https://doi.org/10.1002/PRO.4224

Wolstenholme, D. R., Macfarlane, J. L., Okimoto, R., Clary, D. O., & Wahleithner, J. A. (1987). Bizarre tRNAs inferred from DNA sequences of mitochondrial genomes of nematode worms. *Proceedings of the National Academy of Sciences of the United States of America*, *84*(5), 1324–1328. https://doi.org/10.1073/pnas.84.5.1324

Yamashita, S., & Tomita, K. (2016). Mechanism of 3′-Matured tRNA Discrimination from 3′-Immature tRNA by Class-II CCA-Adding Enzyme. *Structure*, *24*(6), 918–925. https://doi.org/10.1016/j.str.2016.03.022

Yue, D., Maizels, N., & Weiner, A. M. (1996). CCA-adding enzymes and poly(A) polymerases are all members of the same nucleotidyltransferase superfamily: characterization of the CCA-adding enzyme from the archaeal hyperthermophile Sulfolobus shibatae. *RNA*, *2*(9), 895. /pmc/articles/PMC1369424/?report=abstract

Zhang, F., Gavira, J. A., Lee, G. W., & Zahn, D. (2021). Nonclassical Nucleation—Role of Metastable Intermediate Phase in Crystal Nucleation: An Editorial Prefix. *Crystals 2021, Vol. 11, Page 174*, *11*(2), 174. https://doi.org/10.3390/CRYST11020174

Zhu, D. W., Lorber, B., Sauter, C., Ng, J. D., Bénas, P., Le Grimellec, C., & Giegé, R. (2001). Growth kinetics, diffraction properties and effect of agarose on the stability of a novel crystal form of Thermus thermophilus aspartyl-tRNA synthetase-1. Acta Crystallographica Section D: Biological

Crystallography. https://doi.org/10.1107/S0907444901001743

Annexes

Annexes

Annexes.....192

1. Authors contribution.....194

1.1. Adaptation of the Romanomermis culicivorax CCA-Adding Enzyme to Miniaturized Armless tRNA Substrates.....194

1.2. Monitoring the Production of High Diffraction-Quality Crystals of Two Enzymes in Real Time Using In Situ Dynamic Light Scattering.....194

1.3. Crystallization and Structural Determination of an Enzyme:Substrate Complex by Serial Crystallography in a Versatile Microfluidic Chip.....194

1.4. CCA-addition in the cold: Structural characterization of the psychrophilic CCA-adding enzyme from the permafrost bacterium Planococcus halocryophilus.....195

2. Supplementary figures.....196

3. A simple and versatile microfluidic device for efficient biomacromolecule crystallization and structural analysis by serial crystallography.....198

4. Structural Analysis of RNA by Small-Angle X-ray Scattering.....211

5. Posters and presentations.....239

5.1. New microfluidic design for crystallization of biomolecules and their in situ analysis by serial crystallography.....239

5.2. Monitoring the Production of High Diffraction-Quality Crystals in Real Time Using In Situ Dynamic Light Scattering.....239

5.3. Discovering the world of RNA: meeting and sharing science with the general public 239

1. Authors contribution

1.1. Adaptation of the Romanomermis culicivorax CCA-Adding Enzyme to Miniaturized Armless tRNA Substrates.

H.B., C.S. and M.M. designed the experiments; S.P., O.H., T.K., S.B., K.R., and T.J. performed the experiments; H.B., C.S., M.M., S.P., K.R., and O.H. analyzed the data; O.H., S.P., H.B., and M.M. wrote the manuscript. All authors have read and agreed to the published version of the manuscript.

My contribution to the published work consisted in the purification of the chimeric enzyme and the *in vitro* transcription of radioactively labeled tRNAs. Furthermore, I used this material to determine the binding affinity of wild type and chimeric CCA-adding enzymes for selected tRNA substrates.

1.2. Monitoring the Production of High Diffraction-Quality Crystals of Two Enzymes in Real Time Using In Situ Dynamic Light Scattering

Formal analysis, R.D.W., K.R., E.G., P.B., B.L. and C.S.; funding acquisition, M.M., O.M., E.G. and C.S.; investigation, R.D.W., K.R., S.E., A.G.M., O.H., O.M., E.G., P.B., B.L. and C.S.; resources, F.R. and O.M.; supervision, M.M., O.M., E.G. and C.S.; writing—original draft, C.S.; writing—review and editing, R.D.W., K.R., S.E., A.G.M., O.H., H.B., M.M., F.R., O.M., E.G., P.B., B.L. and C.S. All authors have read and agreed to the published version of the manuscript.

My contribution to this work was focused on the crystallization of PhaCCA enzyme using the Xtal Controller device. I could establish two reproducible protocols to grow either high quality macro-crystals or micro-crystalline samples of the enzyme.

1.3. Crystallization and Structural Determination of an Enzyme:Substrate Complex by Serial Crystallography in a Versatile Microfluidic Chip

Formal analysis: R.D.W., K.R.,V.O., O.H., N.T., C.N., C.P., B.L., H.B., M.M., C.S.; funding acquisition: M.M. and C.S.; investigation, R.D.W., K.R., V.O., N.T., C.S.; video shooting: R.D.W., C.P., K.R., C.S.; supervision: M.M., and C.S.; writing—original draft: R.D.W., C.S.;

writing—review and editing: R.D.W., K.R., C.S.. All authors have read and agreed to the published version of the manuscript.

My participation to this work was centered on the ChipX preparation for video shooting. I took part also in the redaction and reviewing process of the manuscript.

1.4. CCA-addition in the cold: Structural characterization of the psychrophilic CCA-adding enzyme from the permafrost bacterium Planococcus halocryophilus

Formal analysis: R.D.W., K.R., F.G.M.E, K.W., H.B., M.M., C.S.; funding acquisition: M.M. and C.S.; investigation, R.D.W., K.R., F.G.M.E, K.W., H.B., M.M., C.S.; supervision, M.M., and C.S.; writing—original draft, M.M., H.B., C.S.; writing—review and editing, K.R., K.W., H.B., M.M., C.S. All authors have read and agreed to the published version of the manuscript.

My contribution to this work concerned the crystallization, data acquisition (PDB ID: 7OQX) and refinement of PhaCCA models (PDB ID: 7OQX, 7OTR, 7OTL). In addition, I performed nanoDSF experiments to explore the thermal stability of enzyme:nucleotide complexes compared to the apo form of the enzyme. In the end, I compared both secondary and tertiary structures of the psychrophilic PhaCCA and the thermophilic GstCCA enzymes.

Prof. Dr. Mario Mörl
Institute for Biochemistry
Leipzig University
Brüderstr. 34
04103 Leipzig
Germany

Dr. Claude Sauter
ARN – IBMC – CNRS
Université de Strasbourg
2, Allée Conrad Roentgen
67084 Strasbourg
France

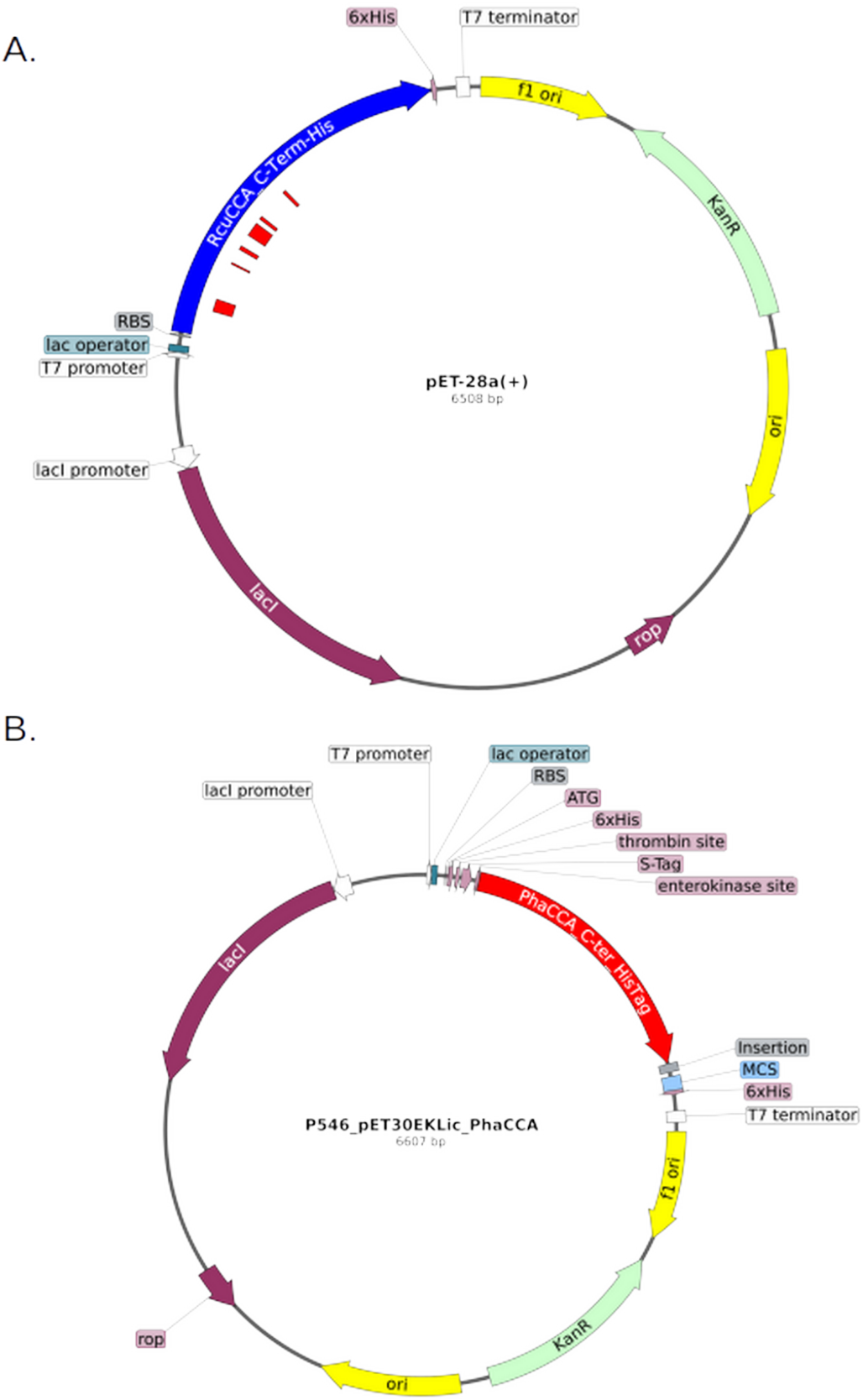
2. Supplementary figures

>Romanomermis culicivorax CCA adding enzyme with N-terminal His-Tag
MGSSHHHHHHSSGLVPRGSHMASMTGGQQMGRGSMKIDSPQFRSIFTPELRIVSDLFEKNGFELRIAG
GAVRDILLGQVPHDIDFATTATPTQMINLFNKEGIRMLNKRGEKHGTITCRINEKVNFEITTLRIDKL
CDGRRAEVEFTTDWYKDANRRDLTVNSIFLGLDGTVYDYFDGIEHLKARKILFVGDAKTRIQEDYLRI
LRYFRFFGRLTITPDDHDPMTLLAIKDNVHGLKNVSGERLWTEFKRIVTGRFAGSVVKIMLECNIGSY
LGLPETCNIDEFIRLCDVGCLAHNPMSMTMVSALLNFEDDIYNLDGRLKLSNKERFLGNFIIEHRSKI
SRKKSKDDSVESYYKDLLIGFQSPSIKLENSKEYVLELAKYNGDFDISRILEDYQAPKFPVDGVMLLQ
WGVKKGPAMKAVLSKLFEFWKIKQYDVTNDEFFQQIDDCYKRDRSRSPIHSKQ-

>Planococcus halocryophilus CCA adding enzyme with N-terminal His-Tag
MHHHHHHSSGLVPRGSGMKETAAAKFERQHMDSPDLGTDDDDKMNTAIKVIHTLKAAGFEAYIVGGAV
RDLLLGKTPHDVDVASSALPQQVKVLFDRTVDTGIDHGTVLVLLDGEGIEVTTFRTESSYSDNRRPDS
VEFVLSLEEDLRRRDFTINAMAMTEDLKIIDPFGGKEDLKNKVIRAVGDPDERFEEDALRMLRAIRFS
GQLDFIIDMKTLLSIRRHARLIRFIAVERLKSEIDKIFVNPSMQKSMAYLKDSVLTRFLPVGGLFEVD
WITYHTDGNPTYGWLYLLHQQKRQFTDIKDYRFSNEEKRLIEKSLELTALNTWDQWTFYKYTLKQLEM
ASRVTGKKKDLAAIKRQLPIQSRSELAVDGWDLIEWSGAKSGPWLKVWIEKI
ERLIVYGILKNDKELIKDWFEDEYHSHT

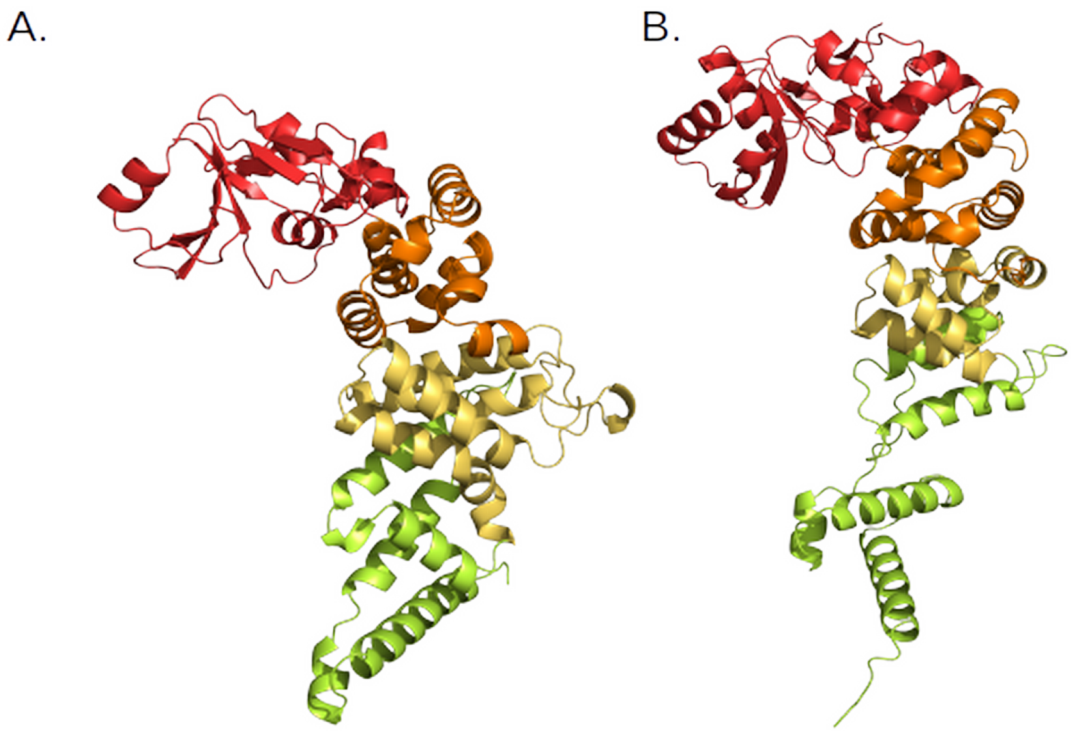
> Gammaproteobacteria Ancestral CCA-adding enzyme with C-terminal Hi-tag
MEIYLVGGAVRDRLLGRPVKERDWVVVGATPEQMLAQGFRPVGKDFPVFLHPETGEEYALARTERKTA
PGYHGFTFHASPDVTLEEDLARRDLTINAMAQDEDGTLIDPYNGQADLQARVLRHVSPAFAEDPVRIL
RVARFAARFAPLGFTVAPETLALMREMVASGEVDALVPERVWQEMEKALAEPTPSAFLRVLRECGALA
VLFPEVDRLFGVPQPPKWHPEIDTGVHTLMVLDQAARLSADLAVRFAALVHDLGKGLTPAEQLPRHHG
HEQRGVPLVRALCERLRVPSEYRDLAVLVARYHLHCHRAFELRPATILKLLEGLDAFRRPERLEQFLL
ACEADYRGRLGFEDRPYPQADYLRAAFAAAAAVSARPFVAQGLKGPAIGEALRQARIQAIAELKASNA
PALEHHHHHH

Supplementary figure 1: Target CCA-adding enzymes sequences



Supplementary figure 2: maps of plasmids used for CCA-adding enzymes overexepression

A. pET-28a(+)-RcuCCA. B. pET30EKLic-PhaCCA



Supplementary figure 3: AncCCA and RcuCCA computational models obtained with AlphaFold2

A. AncCCA model. B. RcuCCA model. The head, neck, body and tail are colored in red, orange, yellow and green, respectively.

3. A simple and versatile microfluidic device for efficient biomacromolecule crystallization and structural analysis by serial crystallography

My contribution to this work was initiated during my master internship under the direction of Dr. Claude Sauter. During this period, I could take part in crystallization experiments using the microfluidic ChipX, in particular centered on the aspartyl-tRNA synthetase from *Thermus thermophilus*.

IUCrJ
ISSN 2052-2525
BIOLOGY | MEDICINE

A simple and versatile microfluidic device for efficient biomacromolecule crystallization and structural analysis by serial crystallography

Raphaël de Wijn,[a] Oliver Hennig,[b] Jennifer Roche,[c] Sylvain Engilberge,[d] Kevin Rollet,[a] Pablo Fernandez-Millan,[a] Karl Brillet,[a] Heike Betat,[b] Mario Mörl,[b] Alain Roussel,[c] Eric Girard,[d] Christoph Mueller-Dieckmann,[e] Gavin C. Fox,[f] Vincent Olieric,[g] José A. Gavira,[h] Bernard Lorber[a] and Claude Sauter[a]*

[a]Architecture et Réactivité de l'ARN, UPR 9002, CNRS, Institut de Biologie Moléculaire et Cellulaire (IBMC), Université de Strasbourg, 15 Rue René Descartes, 67084 Strasbourg, France, [b]Institute for Biochemistry, Leipzig University, Bruederstrasse 34, 04103 Leipzig, Germany, [c]Architecture et Fonction des Macromolécules Biologiques, UMR 7257 CNRS–Aix Marseille University, 163 Avenue de Luminy, 13288 Marseille, France, [d]Université Grenoble Alpes, CEA, CNRS, IBS, 38000 Grenoble, France, [e]Structural Biology, European Synchrotron Radiation Facility, 38043 Grenoble, France, [f]PROXIMA 2A beamline, Synchrotron SOLEIL, L'Orme des Merisiers, Saint-Aubin, 91192 Gif-sur-Yvette, France, [g]Paul Scherrer Institute, Swiss Light Source, Forschungsstrasse 111, 5232 Villigen PSI, Switzerland, and [h]Laboratorio de Estudios Cristalográficos, IACT, CSIC–Universidad de Granada, Avenida Las Palmeras 4, 18100 Armilla, Granada, Spain. *Correspondence e-mail: c.sauter@ibmc-cnrs.unistra.fr

Received 13 December 2018
Accepted 14 March 2019

Edited by Z.-J. Liu, Chinese Academy of Sciences, China

Keywords: macromolecule; crystallization; counter-diffusion; microfluidics; seeding; ligand soaking; trace fluorescent labeling; serial crystallography; room temperature; protein structure; ChipX3.

PDB references: CCA-adding enzyme, 6ibp; CCA-adding enzyme + CMPcPP, 6q52; nanobody 02, 6gzp; protease 1, 6q3t; lipase, 6hw1; RNA duplex, 6ibq

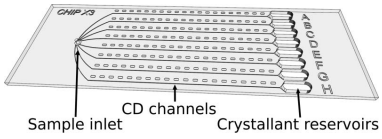
Supporting information: this article has supporting information at www.iucrj.org

Determining optimal conditions for the production of well diffracting crystals is a key step in every biocrystallography project. Here, a microfluidic device is described that enables the production of crystals by counter-diffusion and their direct on-chip analysis by serial crystallography at room temperature. Nine 'non-model' and diverse biomacromolecules, including seven soluble proteins, a membrane protein and an RNA duplex, were crystallized and treated on-chip with a variety of standard techniques including micro-seeding, crystal soaking with ligands and crystal detection by fluorescence. Furthermore, the crystal structures of four proteins and an RNA were determined based on serial data collected on four synchrotron beamlines, demonstrating the general applicability of this multipurpose chip concept.

1. Introduction

Crystallography plays a central role in contemporary biology because it enables the visualization of the 3D architecture of biological macromolecules, which provides insights into their cellular functions and partnerships on the atomic scale (Giegé & Sauter, 2010; Jaskolski *et al.*, 2014). Over the past two decades, the advent of structural genomics and associated high-throughput (HTP) technologies (Vincentelli *et al.*, 2003; Pusey *et al.*, 2005), together with dramatic improvements in experimental setups and the computational environment at synchrotron facilities (Terwilliger *et al.*, 2009; Owen *et al.*, 2016), have revolutionized the field and led to a torrent of new crystal structures. This productivity boost is clear from the number of structures deposited in the Protein Data Bank (PDB), which recently exceeded 150 000 entries.

In spite of such advances, the time-consuming and costly mapping of reagents and phase space to identify conditions that yield diffraction-quality crystals from a limited amount of the macromolecule remains a bottleneck in crystallographic studies (McPherson & Gavira, 2014; Luft *et al.*, 2014; Giegé, 2017). This process generally involves a trial-and-error sampling of chemical and physical space by screening hundreds of different cocktails composed of buffers at different pH



OPEN ACCESS

values, various crystallants (salts, alcohols and polymers) and temperature to find at least one appropriate solvent and the right supersaturation conditions. The miniaturization of crystallization assays in microplates with drop volumes of 0.1–1 µl (typically containing 1–10 µg of the macromolecule) and automation of the screening procedure have made this task considerably more efficient, making it possible to successfully conduct a project with only a few milligrams of pure sample (Sauter *et al.*, 2012).

With the introduction of the first microfluidic systems dedicated to HTP screening 15 years ago, the sample volume required for a single experiment was reduced by another order of magnitude, down to a few nanolitres (Hansen *et al.*, 2002; Zheng *et al.*, 2003). Indeed, microfluidics was immediately regarded as a major breakthrough, especially for biochemists dealing with samples that are difficult to purify in large quantities, such as macromolecules from higher eukaryotes, large biological assemblies and membrane proteins (Hansen & Quake, 2003; van der Woerd *et al.*, 2003). However, despite their potential, microfluidic technologies have not yet been massively adopted by the global community for crystal growth, as illustrated by the limited number of PDB entries (only about 30 as of March 2019) that specifically cite the use of microfluidic systems. This can be partly explained by the cost of these microsystems and their associated equipment, but also by the difficulty in successfully extracting fragile crystals from the chips or the requirement to reproduce them using conventional crystallization methods before they can be subjected to crystallographic analysis.

To expand the functionality and attractiveness of microchips beyond crystallization and HTP screening, several teams have explored the possibility of analyzing crystals directly in their microfluidic environment (Yadav *et al.*, 2005; Ng *et al.*, 2008; Sauter *et al.*, 2007; Dhouib *et al.*, 2009; Emamzadah *et al.*, 2009; Hansen *et al.*, 2006; Stojanoff *et al.*, 2011). Various geometries and materials have been tested and have led to promising results in terms of data collection, anomalous phasing or time-resolved applications (Pinker *et al.*, 2013; Khvostichenko *et al.*, 2014; Perry *et al.*, 2013, 2014). The difficulty of cryopreserving crystals to protect them from radiation damage inside chips, owing to the wide flat surfaces of the device causing vapor condensation and ice formation in the cryojet, was first perceived as an obstacle. However, the recent revival of multi-crystal data-collection techniques at room temperature by the X-ray free-electron laser (XFEL) community has changed the paradigm and popularized serial crystallography (Chapman *et al.*, 2011; Stellato *et al.*, 2014; Ayyer *et al.*, 2015). In this context, microfluidic systems provide promising solutions to prepare, handle and analyze crystals at both synchrotron beamlines and XFELs (Heymann *et al.*, 2014; Sui *et al.*, 2016).

In this report, we describe a versatile and low-cost microfluidic chip for crystal production and characterization. This chip was initially designed to miniaturize and facilitate the identification of crystal-growth conditions using the counter-diffusion method and its efficient self-optimizing process (Dhouib *et al.*, 2009; Pinker *et al.*, 2013). The latest version of the chip design, called ChipX3, incorporates several improvements in terms of sample injection, reservoir loading and design to allow low-cost manufacturing by injection molding. With ChipX3, we demonstrate that crystals can (i) be easily produced by seeding, (ii) be soaked *in situ* with ligands or (iii) be visualized by fluorescence imaging. In addition, the chip provides a stable platform for crystal storage, handling, shipment and *in situ* analysis by serial crystallography. We illustrate a range of applications for ChipX3 by the crystallization of seven soluble proteins, a membrane protein and an RNA duplex, as well as the structure determination of five 'non-model' macromolecules at room temperature using data collected on four beamlines at three different synchrotron sites. This lab-on-a-chip approach simplifies and efficiently miniaturizes the crystallographic structure-determination process, from the sample to its 3D structure, in a single device. It offers a user-friendly, cost-effective solution for routine biocrystallographic investigations at room temperature.

2. Materials and methods

2.1. Biomacromolecules, biochemicals and chemicals

The recombinant proteins used in this work include protease 1 from *Pyrococcus horikoshii* (PhP1), the llama nanobody PorM_02 (Nb02), a lipase from *Thermomyces lanuginosus* (Lip; provided by Macrocrystal Oy, Finland), the CCA-adding enzyme from the psychrophilic bacterium *Planococcus halocryophilus* (CCA), the TonB-dependent heme/hemoglobin outer membrane transporter (OMT) ShuA from the pathogen *Shigella dysenteriae* (OMT ShuA), the human mitochondrial aspartyl-tRNA synthetase (hmDRS) and aspartyl-tRNA synthetase 1 from the bacterium *Thermus thermophilus* (ttDRS), which were purified as described previously (Engilberge *et al.*, 2018; Duhoo *et al.*, 2017; Ernst *et al.*, 2018; Brillet *et al.*, 2009; Sauter *et al.*, 2015; Zhu *et al.*, 2001). Horse hemoglobin was purchased from Sigma. The nine-base-pair RNA duplex $[r(CGUGAUCG)dC]_2$ was prepared as described by Masquida *et al.* (1999). Stock concentrations and storage buffers are indicated in Table 1.

To facilitate the detection of CCA crystals by trace fluorescent labeling (TFL; Pusey *et al.*, 2015), the protein was fluorescently labeled with carboxyrhodamine-succinimidyl ester (Invitrogen, catalog No. C6157) as described by de Wijn *et al.* (2018). The labeled protein solution was stored at 277 K and mixed with the protein stock solution just before preparing crystallization assays as a fraction corresponding to less than 1% of the total protein stock. This solution will be referred to as 'CCA-TFL'.

The nonhydrolyzable analog of cytidyl triphosphate (CTP) that was soaked into the CCA crystals, cytidine-5′-[(α,β)-methyleno]triphosphate (CMPcPP), was purchased from Jena Bioscience (catalog No. NU-438). The lanthanide complex Tb-Xo4 (commercial name Crystallophore) used to crystallize PhP1 was synthesized as described by Engilberge *et al.* (2017).

2.2. ChipX3 manufacturing

ChipX3 devices were designed at IBMC, Strasbourg, France in collaboration with Synchrotron SOLEIL, Saint-Aubin,

Table 1
Biomolecules and crystallization conditions.

	Biological source	No. of residues/ molecular mass (kDa)	Biomolecule concentration (mg ml^{-1})	Biomolecule buffer solution	Crystallant solution
CCA-adding enzyme	*Planococcus halocryophilus*	420/48.5	5.5	50 m*M* Tris–HCl pH 7.5, 200 m*M* NaCl, 5 m*M* $MgCl_2$	30%(*m*/*v*) PEG 3350, 200 m*M* sodium formate pH 6.6
Nanobody 02	Llama	129/14.5	13.8	10 m*M* HEPES–NaOH pH 7.25, 150 m*M* NaCl	20%(*m*/*v*) PEG 3000, 0.1 *M* trisodium citrate pH 5.5
Protease 1	*Pyrococcus horikoshii*	6 × 166/111.6	7.4	20 m*M* Tris–HCl pH 7.5, 10 m*M* Xo4	3.4 *M* malonate pH 7.5
Lipase	*Thermomyces lanuginosus*	269/29.3	30	25 m*M* HEPES–NaOH pH 7.5	0.3 *M* sodium/potasssium phosphate, 50 m*M* sodium acetate pH 4.5
Aspartyl-tRNA synthetase 1	*Thermus thermophilus*	2 × 580/132	19	50 m*M* Tris–HCl pH 7.2, 1 m*M* EDTA, 1 m*M* DTT	10%(*m*/*v*) PEG 8000
Mitochondrial aspartyl-tRNA synthetase	*Homo sapiens*	2 × 630/140	30	50 m*M* HEPES–NaOH pH 7.5, 150 m*M* NaCl, 10% glycerol, 1 m*M* DTT	100 m*M* Tris–HCl pH 7.0, 40%(*m*/*v*) PEG 3350, 0.2 *M* sodium thiocyanate
OMT ShuA	*Shigella dysenteriae*	632/69.5	20	10 m*M* Tris–HCl pH 8.0, 1.4% *β*-D-octyl-glucoside	0.1 *M* sodium acetate, 20%(*m*/*v*) PEG 400, 15%(*m*/*v*) PEG 4000, 10%(*m*/*v*) PEG 8000 pH 5.0
RNA duplex	Synthetic	2 × 9/5.8	10	10 m*M* sodium cacodylate pH 6.0, 5 m*M* $MgCl_2$	2.6 *M* ammonium sulfate, 50 m*M* sodium cacodylate pH 6.0, 5 m*M* $MgSO_4$, 1 m*M* spermine
Hemoglobin	*Equus caballus*	574/62	20	50 m*M* potassium phosphate pH 7.5	3.3 *M* ammonium sulfate, 50 m*M* potassium phosphate pH 7.5

France and were manufactured by MicroLIQUID, Arraste, Spain. The fluidic layer (thickness 1 mm) was produced in cyclic olefin copolymer (COC; TOPAS 5013F-04) by injection molding. Channels and reservoirs were sealed with a second layer of COC (thickness 100 µm). The bonding process was carried out at 398 K and a pressure of 500 kPa. The straight section of the microfluidic channels is 4 cm long with a cross-section of 80 × 80 µm, to give a volume of 260 nl. The reservoir at their extremity has a volume of 10 µl (Fig. 1).

2.3. Sample loading and crystallization

Crystallization experiments in the ChipX3 were set up in three steps with a conventional 10 µl micropipet (Gilson) and regular tips (StarLab). Firstly, 4–6 µl of macromolecule solution was injected into the sample inlet connecting all channels to fill the entire arborescence up to the reservoirs. Secondly, 1 µl of paraffin oil (Fluka) was injected into the sample inlet to isolate the channels from each other and the inlet was sealed with CrystalClear tape (Hampton Research) to prevent evaporation and solution movements. The third and last step consisted of filling the reservoirs with 5 µl crystallization solution before sealing them with CrystalClear tape. The solutions used to set up the chips are listed in Table 1. All experiments were incubated at 293 K, except for the RNA duplex, which was crystallized at 310 K.

2.4. Crystallization by seeding

The condition producing the best CCA crystals (de Wijn *et al.*, 2018) was found using the microseed matrix screening (MMS) method described by D'Arcy *et al.* (2007, 2014). Small crystals grown by the hanging-drop method using a reservoir consisting of 1 *M* diammonium hydrogen phosphate, 100 m*M* sodium acetate pH 4.5 (condition E8 from the commercial screen JCSG++ from Jena Biosciences) were recovered, vigorously resuspended, vortexed and diluted in 50 µl of the same crystallant solution. This suspension was stored at 277 K and is referred to as the 'seed stock'. Protein crystallization solutions were prepared by mixing 6 µl enzyme solution (5.5 mg ml^{-1}), 1.5 µl seed stock (either the original or diluted solution) and 1 µl CCA-TFL and were immediately injected into the chip channels. Crystallization of the ttDRS enzyme in the ChipX3 was also performed using seeds. The 'seed stock' suspension was prepared as described for the CCA enzyme by crushing crystals grown by vapor diffusion in hanging drops with a reservoir consisting of 7%(*m*/*v*) PEG 8000, 10 m*M* $MgCl_2$. ttDRS crystallization solutions were prepared as a mixture consisting of 6.5 µl enzyme solution (19 mg ml^{-1}) and 1.5 µl seed stock, either the original or diluted solution, and were immediately injected into the chip channels.

2.5. Crystal soaking with substrate

To soak CCA crystals grown in the ChipX3, the tape covering the reservoirs was removed and 3 µl of 10 m*M* CMPcPP solution was added to selected reservoirs (final concentration of 3.75 m*M*) before sealing them again. This step was performed a week before data collection to ensure good diffusion along the microfluidic channels and in an attempt to maximize site occupancy in the crystals.

2.6. X-ray data collection and analysis

Diffraction data were collected either (i) on beamline PXII (Fuchs *et al.*, 2014) equipped with a PILATUS 6M detector or beamline PXIII (Bingel-Erlenmeyer *et al.*, 2011) equipped with a MAR CCD or a PILATUS 2M-F detector at the Swiss Light Source (SLS), Villigen, Switzerland, (ii) on the PROXIMA-2A (PX2A) beamline (Duran *et al.*, 2013) equipped with an EIGER X 9M detector at SOLEIL, Saint-Aubin, France or (iii) on beamline ID30B (McCarthy *et al.*, 2018) equipped with a PILATUS3 6M detector at the ESRF, Grenoble, France.

All serial data collections were performed at room temperature (T = 293–298 K) on crystals inside ChipX3, owing to the reduced scattering background of the chip (Pinker *et al.*, 2013). In most cases a dedicated 3D-printed holder mounted on a standard goniometer (see Fig. 5 and Supplementary Fig. S2) was used for data collections. To collect the widest possible rotation angle for each crystal in ChipX3, the channel containing the crystal was aligned with the rotation axis of the goniometer. Crystal alignment was performed either by standard low-dose grid screening at SLS and SOLEIL, or by a one-click procedure at ESRF as described by McCarthy *et al.* (2018). To avoid collisions with the surrounding equipment (beamstop and collimator), we typically collected 30° rotations per crystal or crystal sector between goniometer positions −30° and +30° (where 0° corresponding to the channels being perpendicular to the X-ray beam). Two data-collection strategies were used to obtain complete data: either merging several partial data sets (sweeps) from the same crystal (one orientation and a wide rotation range) or merging several data sets from different crystals (several orientations and a smaller rotation range per crystal). Table 2 provides details of data collection and processing.

Partial data sets were individually processed with *XDS* (Kabsch, 2010). When their number did not exceed ten, they were manually merged with *XSCALE* to find the best combination and determine the appropriate resolution range. In the case of the PhP1 enzyme, *ccCluster* (Santoni *et al.*, 2017) was used to determine the best partial data sets to merge among the 35 available. For all remaining steps, the *PHENIX* package was used (Adams *et al.*, 2010). Phases were determined by molecular replacement using the following structures: PDB entries 1miv (Li *et al.*, 2002) for CCA, 5lmw (Duhoo *et al.*, 2017) for Nb02, 1g2i (Du *et al.*, 2000) for PhP1, 4gwl (P. K. Shukla, M. Sinha, J. Mukherjee, M. N. Gupta, P. Kaur, S. Sharma & T. P. Singh, unpublished work) for Lip and 485d (Masquida *et al.*, 1999) for the RNA. The latter crystals (space group $H3$) presented translational pseudo-symmetry owing to the intrinsic symmetry of the duplex and merohedral twinning (twin fraction 0.21–0.39). Hence, the structure was refined using the twin law $h, -k-h, -l$. All structures were built and refined with *Coot* and *phenix.refine* (Emsley & Cowtan, 2004; Adams *et al.*, 2010).

3. Results and discussion

3.1. ChipX3 design and setup

ChipX3 was designed to perform counter-diffusion (CD) experiments and take advantage of convection-free conditions (a prerequisite of CD) in channels of small cross-section (width 80 µm, depth 80 µm) to enable the creation of crystallant concentration gradients by pure diffusion [Fig. 1(*a*)]. The channels, with a length of 4 cm, allow a broad screening of supersaturation states, as does conventional CD in microcapillaries (García-Ruiz *et al.*, 2001; Otálora *et al.*, 2009).

The geometry of the sample inlet was adapted to fit standard P2/P10 micropipet tips for chip loading using standard laboratory materials. No extra equipment (such as a pump) is required. The standard micropipet is used to inject the solution into the fluidic system. The branching channel configuration [Fig. 1(*b*)] allows the simultaneous loading of the eight channels in a single manipulation, thus limiting the loading time and solution dead volumes. Note that in the case of membrane-protein samples containing a detergent (such as ShuA in this work), solutions enter and fill the channels spontaneously owing to capillary action and the native wettability of the COC material. Labels embossed along the channels facilitate crystal location and grid mapping on synchrotron beamlines [Fig. 1(*c*)].

Once the channels have been filled with the macromolecule solution and the inlet closed with tape, crystallant solutions are deposited in the reservoirs [Fig. 1(*d*)]. The setup is fully compatible with viscous solutions such as the PEG mixtures used in CD screens (González-Ramírez *et al.*, 2017). The funnel-like channel shape has been optimized to facilitate the contact between the crystallization and macromolecule solutions and to avoid trapping air bubbles, which could prevent the diffusion process. 1 µl low-gelling temperature agarose solution at 1%(*w*/*v*) can optionally be deposited in the funnel prior to the crystallization cocktail to constitute a physical buffer at the entry to the channels that stabilizes the diffusion interface.

The loading procedure of ChipX3 is fast and straightforward. Setting up a

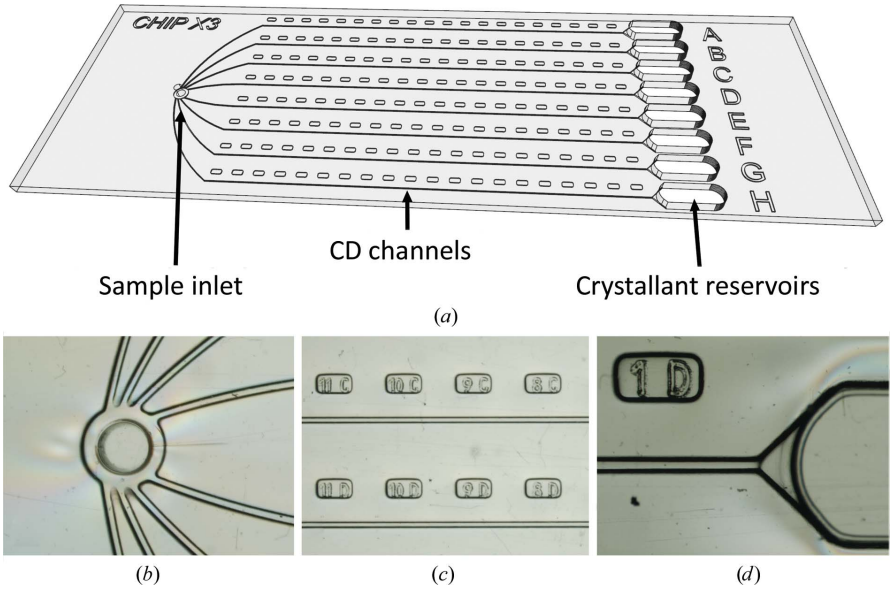


Figure 1
ChipX3 setup. (*a*) Schematic view of the chip, which has the dimensions of a microscope slide (75 × 25 mm) and eight channels with a straight segment of 4 cm and a cross-section of 80 × 80 µm. Close-up views are shown of (*b*) the inlet for the biomacromolecule solution, (*c*) the channels and labels, and (*d*) the end of the channel and the crystallant reservoir.

Table 2
Data-collection and refinement statistics.

Values in parentheses correspond to the high resolution range.

	CCA-adding enzyme	CCA-adding enzyme + CMPcPP	Nanobody 02	Protease 1	Lipase	RNA duplex
X-ray beamline	PXIII, SLS	PXII, SLS	PX2A, SOLEIL	PXIII, SLS	ID30B, ESRF	PXIII, SLS
Wavelength (Å)	1.000	1.000	0.826	1.240	0.976	0.826
Temperature (K)	293	293	293	293	293	293
Detector	PILATUS 2M-F	PILATUS 6M	EIGER	PILATUS 2M-F	PILATUS3 6M	MAR CCD
Crystal-to-detector distance (mm)	300	400	154	150/200	502	200
Crystals collected	6	14	9	1/11	14	3
Crystals selected	5	5	1	8	2	3
Rotation range per image (°)	0.1	0.2	0.1	0.2	0.1	2–3
No. of images selected	1000	540	500	1300	600	80
Total rotation range (°)	100	108	50	260	60	155
Exposure time per image (s)	0.1	0.1	0.1	0.1	0.02	1–2
Space group	$P4_32_12$	$P4_32_12$	$P4_32_12$	$P4_12_12$	$P6_1$	$R3$
a, c (Å)	71.5, 293.8	71.4, 293.6	66.7, 91.8	125.6, 133.9	142.6, 80.7	40.0, 69.1
Solvent content (%)	68.3	67.8	65.0	74.0	68.6	54.7
Mean mosaicity (°)	0.04	0.04	0.07	0.04	0.03	0.15
Resolution range (Å)	46–2.54 (2.60–2.54)	48–2.30 (2.40–2.30)	50–2.10 (2.18–2.10)	50–2.15 (2.21–2.15)	49.06–2.50 (2.60–2.50)	23–1.55 (1.59–1.55)
Total No. of reflections	176105 (9374)	232642 (32937)	45307 (4574)	1095436 (85346)	102820 (11312)	21681 (605)
No. of unique reflections	23922 (1598)	34862 (4066)	12281 (1196)	57690 (4522)	31982 (3668)	5485 (304)
Completeness (%)	90.6 (84.6)	99.5 (100.0)	97.2 (98.3)	98.5 (99.6)	98.5 (98.9)	91.5 (69.7)
Multiplicity	7.5 (6.0)	6.7 (8.1)	3.7 (3.8)	19.0 (18.9)	3.2 (3.1)	3.9 (2.0)
$\langle I/\sigma(I)\rangle$	8.1 (1.3)	6.9 (0.7)	11.3 (1.8)	12.0 (1.4)	6.3 (0.8)	6.1 (1.8)
R_{meas} (%)	18.9 (126.0)	18.0 (231.2)	7.5 (84.7)	17.4 (206.4)	8.6 (86.8)	17.9 (45.6)
$CC_{1/2}$ (%)	98.7 (55.0)	98.7 (46.9)	99.7 (73.5)	99.7 (69.4)	99.4 (49.4)	98.8 (75.5)
B factor from Wilson plot (Å^2)	57.4	60.6	45.2	50.8	63.3	23.6
Reflections in working/test sets	23583/1180	34840/3405	11053/1228	57659/5758	31516/1573	5484/382
Final R_{work}/R_{free} (%)	18.8/21.4	20.0/22.9	16.9/21.1	16.2/18.4	17.2/19.9	19.2/22.3
No. of non-H atoms						
Total	2998	3028	970	4017	4446	390
Protein	2989	2989	947	3921	4404	342
Solvent	9	10	23	96	47	43
Ligand	0	29	0	0	33	0
Ion	0	0	0	0	2	5
R.m.s.d., bonds (Å)	0.009	0.010	0.008	0.012	0.004	0.004
R.m.s.d., angles (°)	1.23	1.22	0.897	1.43	1.08	0.680
Average B factors (Å^2)						
Overall	60.1	62.6	53.8	57.1	83.8	17.8
Biomolecule	60.1	60.1	53.8	57.1	82.9	17.6
Solvent	52.7	55.5	50.9	56.1	62.6	15.1
Ramachandran plot regions (%)						
Most favored	98.1	97.2	95.8	98.4	96.6	
Allowed	1.9	2.8	4.2	1.6	3.2	
PDB code	6ibp	6q52	6gzp	6q3t	6hw1	6ibq
PDB code at 100 K†	6qy6	6qxn	5lmw	6hf6	4zgb	485d
R.m.s. distance (Å^2)/ΔV_c (%)†	0.79/5.0	0.83/3.8	0.79/6.6	0.47/4.3	1.0/3.9	0.40/2.7

† Structures solved at room temperature (this work) are compared with equivalent structures determined at cryogenic temperatures (100 K). R.m.s. distances are calculated taking into account all biomolecule atoms and ΔV_c quantifies the increase in the unit-cell volume (V_c) at room temperature.

chip with eight different conditions takes less than 5 min even for untrained experimenters, as attested by numerous assays performed in the five laboratories involved in this work and by the many participants of crystallization workshops [FEBS courses in 2014–2018 in Nové Hrady, Czech Republic; International School of Biological Crystallization (ISBC) 2015–2017 in Granada, Spain].

3.2. Crystallization in ChipX3

After a prototyping phase of small batches made by hot embossing (Pinker *et al.*, 2013), a 3D mold was machined with the new ChipX3 specifications to produce a larger batch by injection molding. This enabled validation of the concept using real cases beyond classical model proteins such as lysozyme or thaumatin. We report here on eight proteins of different sizes and sources (from bacteria to human) and an RNA oligomer (Table 1) crystallized in ChipX3. Crystallization conditions were adapted from those initially used in vapor diffusion or batch crystallization: while the biomacromolecule concentration was kept unchanged, the crystallant concentration was increased by a factor of 1.5–2, as recommended by Otálora *et al.* (2009). Representative examples ranging from small microcrystals to large crystals filling a portion of the channel are shown in Fig. 2. Typical counter-diffusion patterns can be observed along the concentration gradient, with microcrystalline material close to the reservoirs where supersaturation is maximal and larger crystals towards

the other extremity of the channels (see Supplementary Fig. S1).

Crystals appeared after a few hours or days and could be visualized under polarized light. To facilitate the detection of small crystals, we exploited different fluorescence approaches such as classical UV excitation (Meyer *et al.*, 2015), the fluorescent lanthanide compound Tb-Xo4 developed by Engilberge *et al.* (2017) and trace fluorescent labeling (TFL) as developed by Pusey *et al.* (2015). All three approaches were compatible with ChipX3, but the Tb-Xo4 molecule and TFL gave a much brighter signal (Fig. 3). Fluorescence has the advantage of rapidly localizing samples in the channels and may be used in the future to automate and speed up serial analysis.

3.3. Advanced crystallogenesis strategies

In addition to providing an efficient screening of supersaturation conditions, the CD process has other practical benefits, including the possibility of diffusing anomalous scatterers into pregrown crystals for phasing, or cryoprotecting with compounds such as glycerol (Gavira *et al.*, 2002; Ng *et al.*, 2003). In a previous study, we demonstrated the feasibility of on-chip SAD phasing at room temperature using crystals soaked by CD with a lanthanide complex (Pinker *et al.*, 2013). Along the same lines, CMPcPP, a nonhydrolyzable analog of CTP, which is a substrate of CCA-adding enzymes, was added to the reservoirs once CCA crystals had grown and one week before the synchrotron session. The resulting X-ray structures confirmed that the crystals were derivatized by smooth diffusion without any sign of damage (Figs. 4 and 6).

Microseeding can also be used together with CD crystallization (Bergfors, 2003; Gavira *et al.*, 2011) to bypass the nucleation step and promote rapid crystal growth. Hence, CCA and ttDRS crystals were grown by a combination of CD and seeding. Microseeds were added to the protein solution just before it was injected into the chips and the first crystals appeared in the channels after a few days. Seeding proved to be an effective way to trigger rapid and abundant crystal production, which is of particular interest for serial analysis.

We also used a new nucleant called crystallophore or Xo4 (Engilberge *et al.*, 2018) in the case of the protein PhP1, for which the crystallization conditions (Table 1) were determined only in the presence of this terbium complex. Tb-Xo4 was added to the protein solution before filling the channels. It triggered the nucleation and the growth of large PhP1 crystals, which completely filled the available volume. An added value for macromolecules crystallized in the presence of Xo4 is the strong luminescence when illuminated by UV light (see Fig. 3).

3.4. Serial crystal analysis inside ChipX3

The ChipX3 was designed for *in situ* characterization. Its overall thickness was optimized to give a good compromise between material rigidity and X-ray absorption/scattering (Pinker *et al.*, 2013). The COC material produces a characteristic diffuse scattering ring [Fig. 5(*c*)] in the resolution range 4–6 Å (see also Fig. 4 in Dhouib *et al.*, 2009 and Fig. 10.4 in Martiel *et al.*, 2018), which hardly affects data processing and quality. During data collection, the chip is oriented with its thickest layer facing the direct beam and the thinnest face

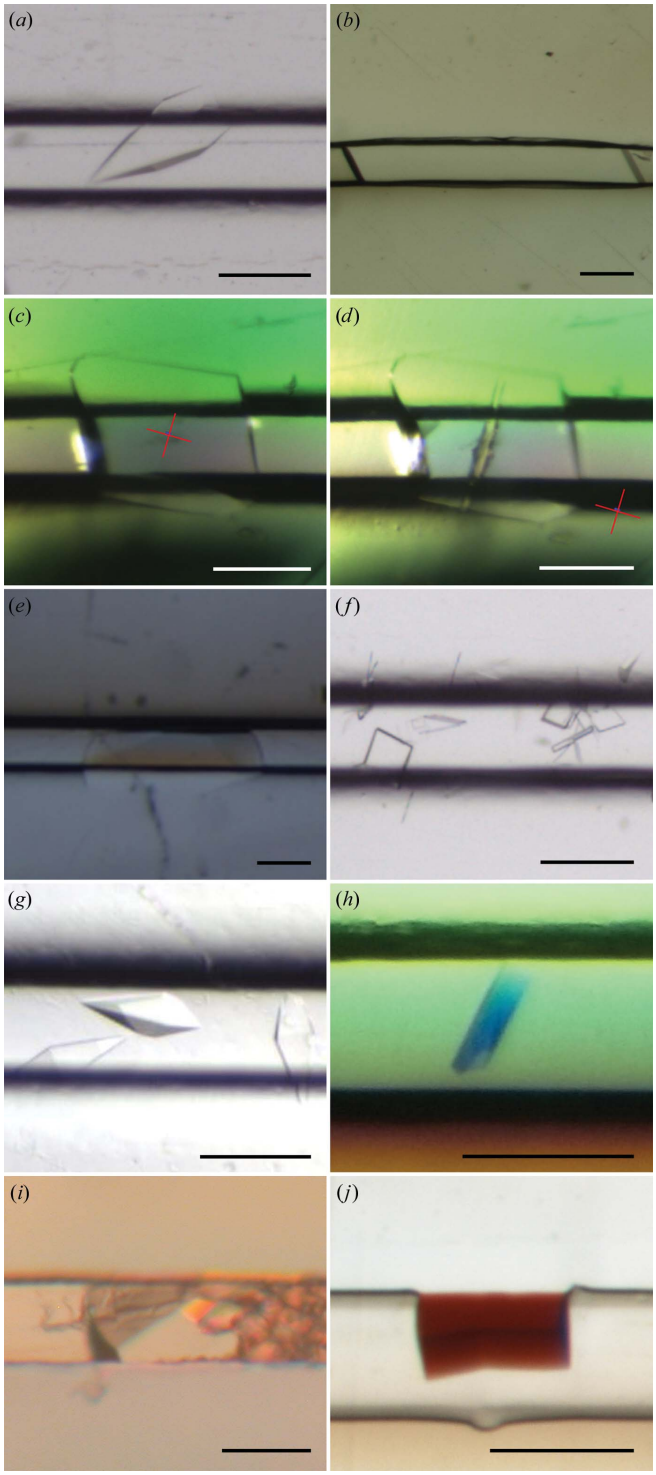


Figure 2
Examples of crystals obtained in ChipX3. Crystals were grown as described in Table 1. (*a*) CCA, (*b*) PhP1, (*c*, *d*) Nb02 before (*c*) and after (*d*) data collection, with the X-ray beam footprint, (*e*) lipase, (*f*) ttDRS, (*g*) hmDRS, (*h*) OMT ShuA, (*i*) oligo RNA duplex and (*j*) hemoglobin. The scale bar is 0.1 mm in length.

behind the crystal to minimize the attenuation of the diffraction signal (Fig. 5). Labels embossed along the channels enable the easy localization of crystals before analysis, with a view to future automation of the procedure on synchrotron beamlines. The chip can also be positioned in the beam using a plate gripper, as illustrated in Supplementary Fig. S3. To avoid intervention from beamline staff to mount/unmount the gripper, we developed a light chip holder that can be directly attached to a standard goniometer. This chip holder is manufactured by 3D printing (Fig. 5, Supplementary Fig. S2) and integrates a standard metal base (B5, SPINE-style; MiTeGen) that is in contact with the goniometer magnet. The holder can be used with any flat device of microscope-slide dimensions on synchrotron beamlines and laboratory-based instruments. The 3D description file for printing this device is provided as supporting information.

To illustrate the general applicability of on-chip serial crystallography at room temperature, we present the results of structure determination in the 1.5–2.5 Å resolution range of four proteins (CCA, Lip, Nb02 and PhP1) and an RNA (Table 2, Figs. 5 and 6). Hemoglobin crystals also yielded complete data to 2.8 Å resolution (data not shown), whereas the microcrystals of the aspartyl-tRNA synthetases and OMT ShuA only diffracted to low resolution and could not be used for structure determination without further optimization.

The data collections were carried out on series of crystals and their parameters were adapted for crystal size and sensitivity to radiation damage. When collecting several paths from the same crystal signs of radiation damage could clearly be seen [see Fig. 2(*d*)], accompanied by the formation of gas bubbles as described by Meents *et al.* (2010) and by the deterioration of data-collection statistics (data not shown).

Note that performing *in situ* analysis, *i.e.* without direct handling of the crystals, is a guarantee that their genuine diffraction properties have been preserved. Comparative tests on thaumatin or lipase crystals in ChipX3 sent by regular postal mail or carried to the synchrotron by experimenters did not show significant differences (results not shown), indicating that the chip is a stable and robust container for crystal storage and transport.

Final crystal structures were obtained either from a single large crystal and two wedges (Nb02) or from combining partial data sets from several individual crystals (RNA, CCA, Lip and PhP1). In the latter case, the use of *ccCluster* considerably facilitated the choice of partial data sets to be merged. The comparison of these structures with equivalent structures solved at cryogenic temperatures only showed small differences (see the r.m.s. distances in Table 2), although the unit-cell volumes were significantly larger (2.7–6.6%) at room temperature than at 100 K because of crystal shrinkage occurring during cryocooling.

The high sensitivity and low background of the latest hybrid pixel detectors (HPDs) compared with CCD detectors (Pinker *et al.*, 2013), and the very short analysis time (seconds) of the largest wedge of reciprocal space from single crystals are crucial to outrun radiation damage for room-temperature data collection. The analysis in shutterless mode also limits systematic errors in crystal orientation and thus improves the data quality. For example, the highest apparent mosaicity of the RNA crystals (see Table 2), which were analyzed at an early stage of this work with a MAR CCD detector, is a direct symptom of the data-collection strategies used before the advent of HPDs. In the future, the widespread integration of HPD technology at synchrotron sites

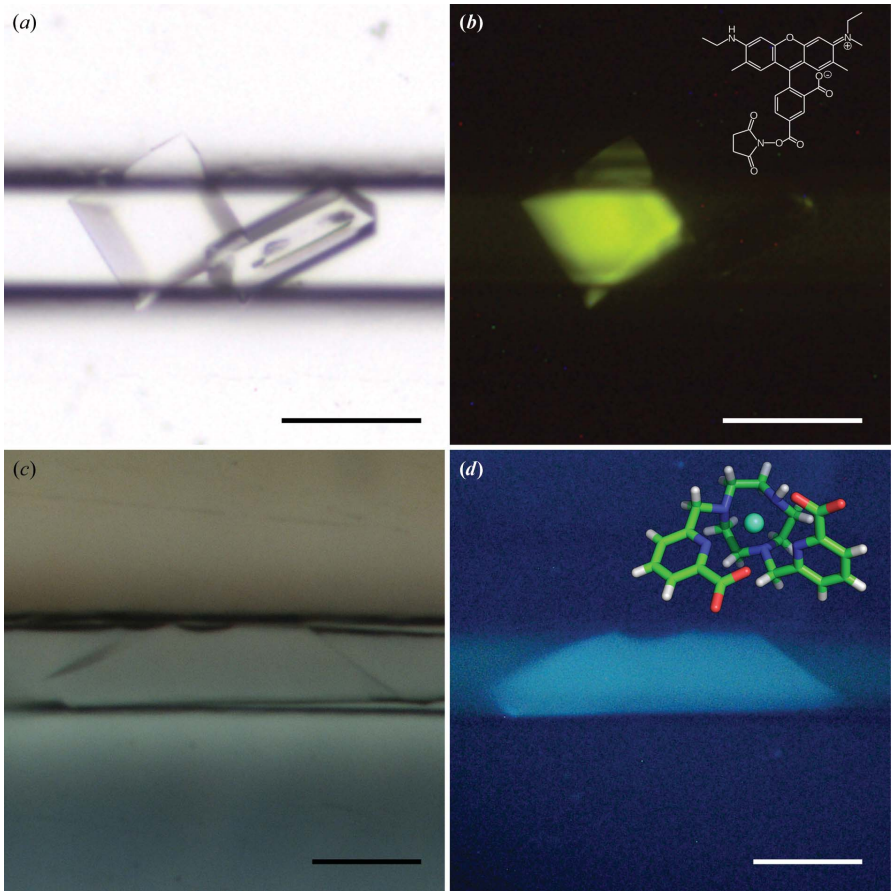


Figure 3
Crystal detection in ChipX3 by fluorescence. (*a*, *b*) CCA crystals grown as described in Table 1 with 0.6% CCA-TFL; (*c*, *d*) PhP1 crystals grown as described in Table 1 with 10 m*M* Tb-Xo4. (*a*, *c*) Crystals illuminated with white light. (*b*) Crystal illuminated with a 520 nm light source and image taken with a low-pass filter at 550 nm (LP550); inset, structure of carboxyrhodamine-succinimidyl ester. (*d*) Crystal illuminated with a 280–380 nm UV source; inset, structure of Tb-Xo4. The scale bar is 0.1 mm in length.

and on laboratory-based X-ray sources will undoubtedly facilitate the development of serial crystallography.

The concept of serial crystallography was introduced with XFEL sources and their extremely intense X-ray pulses that destroy the sample upon signal emission (a process called 'diffraction before destruction'). As a consequence, large numbers (thousands) of micro/nanocrystals are necessary to obtain a complete data set from series of individual still images. The serial approach has been extended to room-temperature data collection using synchrotron radiation. However, with a lower beam intensity (compared with XFELs) crystals can be used to collect more than a single image and up to several degrees of rotation. With very stable crystals (see Nb02 in Table 2), a single crystal may even be sufficient to collect complete data with the help of high symmetry and rapid analysis using HPDs. More generally, the number of crystals that are required for structure determination will depend on their size, their symmetry and their sensitivity to radiation damage. Most of our structures were derived from rather small series of 6–14 crystals and the combination of best data sets (Table 2). With highly sensitive samples such as membrane proteins, *in situ* room-temperature serial crystallography can still be carried out successfully using several hundred crystals (Huang *et al.*, 2015). In this context, ChipX3 provides a convenient means to produce batches of crystals distributed along chip channels and, in the future, automatic crystal detection and characterization should contribute to speeding up data collection and popularizing this kind of serial RT analysis.

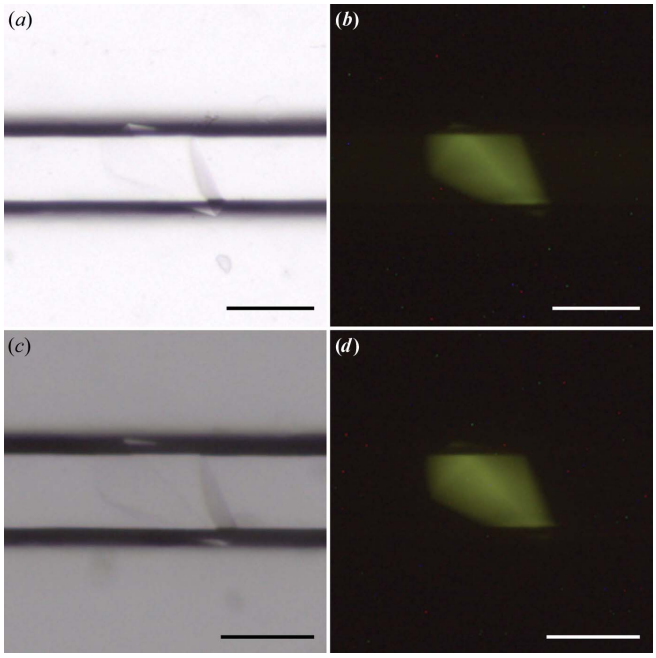


Figure 4
Crystals before and after soaking in ChipX3. Images of CCA crystals grown as described in Table 1 with 0.6% CCA-TFL. (*a*, *b*) Before soaking. (*c*, *d*) Images taken six days after soaking with CMPcPP at a final concentration of 3.75 m*M*. (*a*, *c*) White-light illumination. (*b*, *d*) Images taken with a 520 nm light source and a low-pass filter at 550 nm (LP550). The scale bar is 0.1 mm in length.

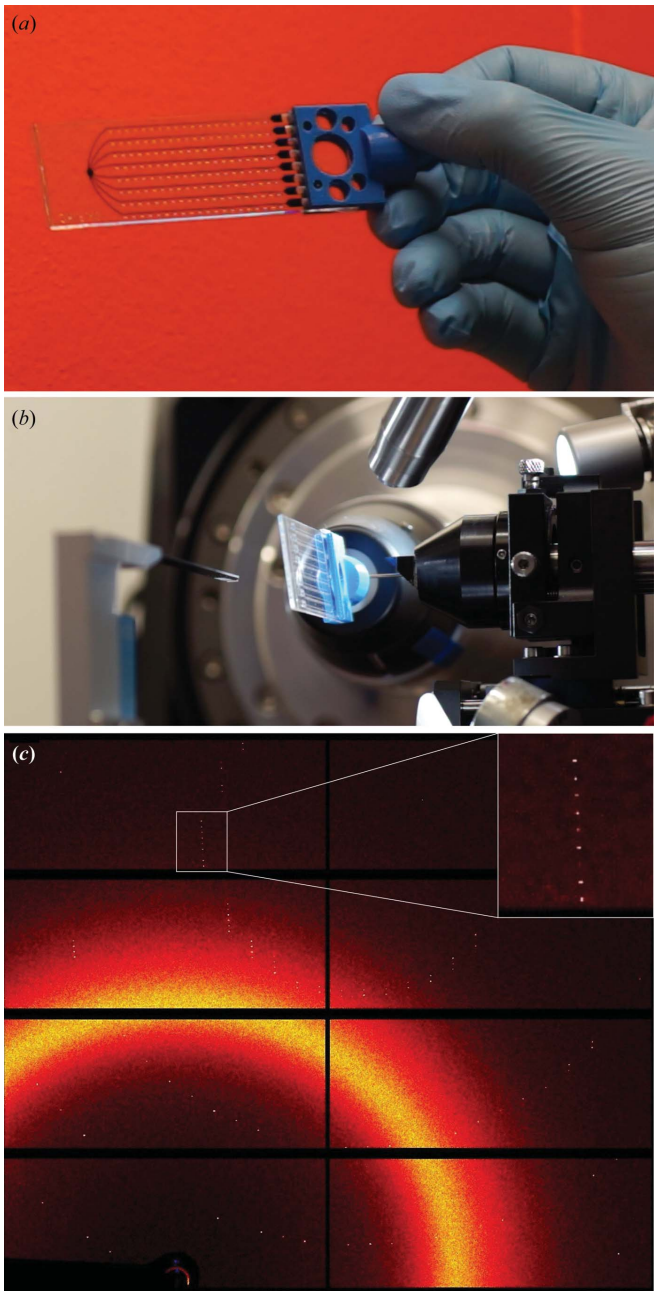


Figure 5
Diffraction analysis in ChipX3. (*a*) ChipX3 on its holder. (*b*) ChipX3 on beamline PXIII at the SLS synchrotron. (*c*) Example of a diffraction pattern of the CCA adding-enzyme in ChipX3 at room temperature (exposure 0.1 s, rotation 0.2°).

4. Conclusion

Microfluidics has demonstrated its value in terms of miniaturization for macromolecular crystallization experiments and HTP screening. With ChipX3, we propose a versatile tool that integrates all of the steps of a crystallographic study on a

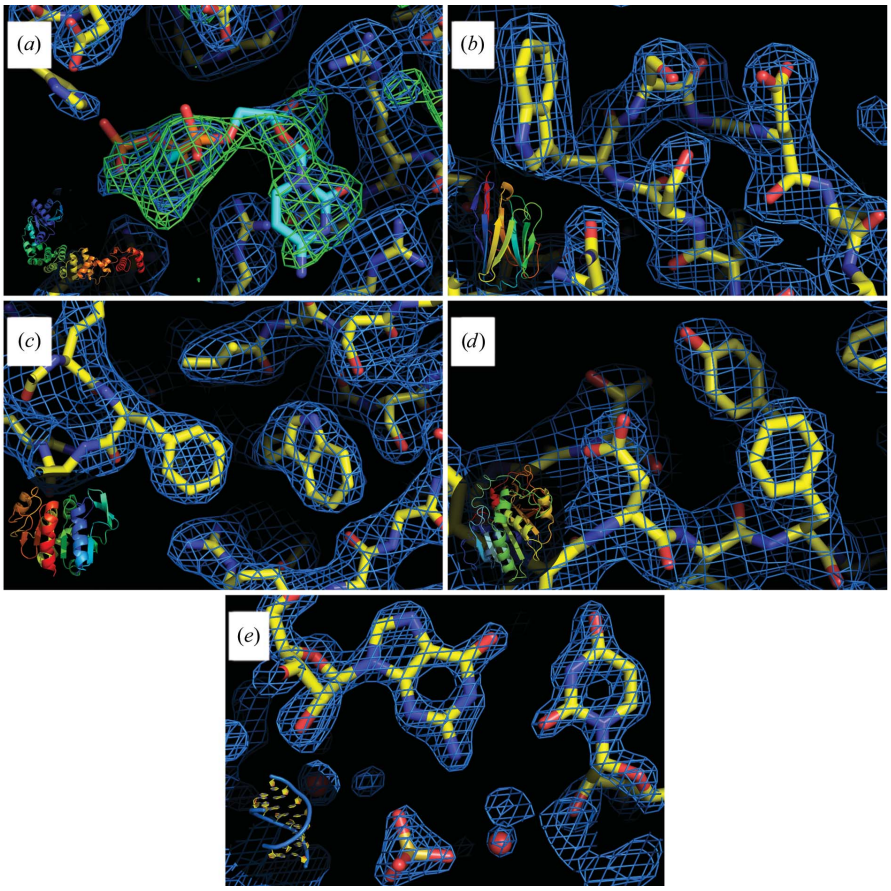


Figure 6
Electron-density maps and structures of target macromolecules. (*a*) CCA-adding enzyme with the positive density from the ligand, (*b*) nanobody, (*c*) protease 1, (*d*) lipase, (*e*) RNA duplex. Insets: schematic representations of the whole macromolecules. This figure was prepared using *PyMOL* (v1.8.6; Schrödinger) with $2F_o - F_c$ electron-density maps (in blue) contoured at 1.2σ and the difference map (in green) contoured at 4σ.

single device with the size of a microscope slide. The same chip serves to produce crystals by counter-diffusion (including seeding techniques), to soak them with ligands (for substrate catalysis, ligand screening in fragment-based drug design or phasing purposes) and to perform their diffraction analysis by *in situ* serial crystallography. The latter step, which is carried out on-chip at room temperature, no longer requires any crystal handling: neither fishing, nor mounting nor cryo-cooling. This guarantees the preservation of the intrinsic crystal quality, with the chip being a safe means of sample storage and transportation. ChipX3 is easy to use with standard laboratory equipment for sample loading and crystal observation, making it cost-effective, with minimal training or expertise required. We show the general applicability of this lab-on-chip concept with several case studies. Sample fluorescent labeling, as exemplified in this work, may be exploited to detect and center individual crystals in the X-ray beam and to perform their characterization fully automatically. Such microfluidic devices show great promise in the future in the combination of serial analysis pipelines developed at advanced X-ray sources (XFELs and synchrotrons) for routine structure determination at temperatures close to physiological conditions (Martin-Garcia *et al.*, 2016; Johansson *et al.*, 2017).

Acknowledgements

The authors thank the following synchrotron facilities and associated scientists for beamtime allocation to the project and for assistance during data collection: beamlines X06DA (PXIII) and X10SA (PXII) at the Swiss Light Source, Villigen, Switzerland, PROXIMA 2A at the SOLEIL synchrotron, Saint-Aubin, France and A. McCarthy for support on ID30B at the European Synchrotron Radiation Facility, Grenoble, France. They also acknowledge F. Riobé, O. Maury from Laboratoire de Chimie at ENS-Lyon and Polyvalan (Lyon, France) for providing Tb-Xo4, and V. Vanel and the team of the Shadok fablab (Strasbourg, France) for assistance with 3D printers, as well as the participants and organizers of crystallization schools at Nové Hrady in the Czech Republic (FEBS courses 2014–2018) and at Granada in Spain (ISBC 2013–2015) for their active participation in operational tests of ChipX3 and their technical feedback.

Funding information

The following funding is acknowledged: Agence Nationale de la Recherche (contract No. ANR-11-LABX-0057_MITOCROSS to Claude Sauter, Bernard Lorber; contract No. ANR-10-LABX-0036_NETRN to Claude Sauter, Bernard Lorber; contract No. ANR-13-BS07-0007-01 to Eric Girard, Sylvain Engilberge); Ministère des Affaires Etrangères (contract No. PROCOPE Hubert Curien to Claude Sauter, Mario Mörl); Deutsche Forschungsgemeinschaft (contract No. Mo 634/10-1 to Mario Mörl, Heike Betat); Université de Strasbourg [grant No. Initiative d'excellence (IDEX) to Claude Sauter, Raphaël de Wijn]; Centre National de la Recherche Scientifique (grant No. MRCT-2012_PTI_UPR9002 to Claude Sauter).

References

Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L.-W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C. & Zwart, P. H. (2010). *Acta Cryst.* D**66**, 213–221.
Ayyer, K., Geloni, G., Kocharyan, V., Saldin, E., Serkez, S., Yefanov, O. & Zagorodnov, I. (2015). *Struct. Dyn.* **2**, 041702.
Bergfors, T. (2003). *J. Struct. Biol.* **142**, 66–76.
Bingel-Erlenmeyer, R., Olieric, V., Grimshaw, J. P. A., Gabadinho, J., Wang, X., Ebner, S. G., Isenegger, A., Schneider, R., Schneider, J.,

Glettig, W., Pradervand, C., Panepucci, E. H., Tomizaki, T., Wang, M. & Schulze-Briese, C. (2011). *Cryst. Growth Des.* **11**, 916–923.
Brillet, K., Meksem, A., Thompson, A. & Cobessi, D. (2009). *Acta Cryst.* F**65**, 402–405.
Chapman, H. N., Fromme, P., Barty, A., White, T. A., Kirian, R. A., Aquila, A., Hunter, M. S., Schulz, J., DePonte, D. P., Weierstall, U., Doak, R. B., Maia, F. R. N. C., Martin, A. V., Schlichting, I., Lomb, L., Coppola, N., Shoeman, R. L., Epp, S. W., Hartmann, R., Rolles, D., Rudenko, A., Foucar, L., Kimmel, N., Weidenspointner, G., Holl, P., Liang, M., Barthelmess, M., Caleman, C., Boutet, S., Bogan, M. J., Krzywinski, J., Bostedt, C., Bajt, S., Gumprecht, L., Rudek, B., Erk, B., Schmidt, C., Hömke, A., Reich, C., Pietschner, D., Strüder, L., Hauser, G., Gorke, H., Ullrich, J., Herrmann, S., Schaller, G., Schopper, F., Soltau, H., Kühnel, K.-U., Messerschmidt, M., Bozek, J. D., Hau-Riege, S. P., Frank, M., Hampton, C. Y., Sierra, R. G., Starodub, D., Williams, G. J., Hajdu, J., Timneanu, N., Seibert, M. M., Andreasson, J., Rocker, A., Jönsson, O., Svenda, M., Stern, S., Nass, K., Andritschke, R., Schröter, C.-D., Krasniqi, F., Bott, M., Schmidt, K. E., Wang, X., Grotjohann, I., Holton, J. M., Barends, T. R. M., Neutze, R., Marchesini, S., Fromme, R., Schorb, S., Rupp, D., Adolph, M., Gorkhover, T., Andersson, I., Hirsemann, H., Potdevin, G., Graafsma, H., Nilsson, B. & Spence, J. C. H. (2011). *Nature (London)*, **470**, 73–77.
D'Arcy, A., Bergfors, T., Cowan-Jacob, S. W. & Marsh, M. (2014). *Acta Cryst.* F**70**, 1117–1126.
D'Arcy, A., Villard, F. & Marsh, M. (2007). *Acta Cryst.* D**63**, 550–554.
Dhouib, K., Khan Malek, C., Pfleging, W., Gauthier-Manuel, B., Duffait, R., Thuillier, G., Ferrigno, R., Jacquamet, L., Ohana, J., Ferrer, J.-L., Théobald-Dietrich, A., Giegé, R., Lorber, B. & Sauter, C. (2009). *Lab Chip*, **9**, 1412–1421.
Du, X., Choi, I. G., Kim, R., Wang, W., Jancarik, J., Yokota, H. & Kim, S.-H. (2000). *Proc. Natl Acad. Sci. USA*, **97**, 14079–14084.
Duhoo, Y., Roche, J., Trinh, T. T. N., Desmyter, A., Gaubert, A., Kellenberger, C., Cambillau, C., Roussel, A. & Leone, P. (2017). *Acta Cryst.* F**73**, 286–293.
Duran, D., Couster, S. L., Desjardins, K., Delmotte, A., Fox, G., Meijers, R., Moreno, T., Savko, M. & Shepard, W. (2013). *J. Phys. Conf. Ser.* **425**, 012005.
Emamzadah, S., Petty, T. J., De Almeida, V., Nishimura, T., Joly, J., Ferrer, J.-L. & Halazonetis, T. D. (2009). *Acta Cryst.* D**65**, 913–920.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Engilberge, S., Riobé, F., Di Pietro, S., Lassalle, L., Coquelle, N., Arnaud, C.-A., Pitrat, D., Mulatier, J.-C., Madern, D., Breyton, C., Maury, O. & Girard, E. (2017). *Chem. Sci.* **8**, 5909–5917.
Engilberge, S., Riobé, F., Wagner, T., Di Pietro, S., Breyton, C., Franzetti, B., Shima, S., Girard, E., Dumont, E. & Maury, O. (2018). *Chem. Eur. J.* **24**, 9739–9746.
Ernst, F. G. M., Erber, L., Sammler, J., Jühling, F., Betat, H. & Mörl, M. (2018). *RNA Biol.* **15**, 144–155.
Fuchs, M. R., Pradervand, C., Thominet, V., Schneider, R., Panepucci, E., Grunder, M., Gabadinho, J., Dworkowski, F. S. N., Tomizaki, T., Schneider, J., Mayer, A., Curtin, A., Olieric, V., Frommherz, U., Kotrle, G., Welte, J., Wang, X., Maag, S., Schulze-Briese, C. & Wang, M. (2014). *J. Synchrotron Rad.* **21**, 340–351.
García-Ruiz, J. M., Otálora, F., Novella, M. L., Gavira, J. A., Sauter, C. & Vidal, O. (2001). *J. Cryst. Growth*, **232**, 149–155.
Gavira, J. A., Hernandez-Hernandez, M. A., Gonzalez-Ramirez, L. A., Briggs, R. A., Kolek, S. A. & Shaw Stewart, P. D. (2011). *Cryst. Growth Des.* **11**, 2122–2126.
Gavira, J. A., Toh, D., Lopéz-Jaramillo, J., García-Ruíz, J. M. & Ng, J. D. (2002). *Acta Cryst.* D**58**, 1147–1154.
Giegé, R. (2017). *IUCrJ*, **4**, 340–349.
Giegé, R. & Sauter, C. (2010). *HFSP J.* **4**, 109–121.
González-Ramírez, L. A., Ruiz-Martínez, C. R., Estremera-Andújar, R. A., Nieves-Marrero, C. A., García-Caballero, A., Gavira, J. A., López-Garriga, J. & García-Ruiz, J. M. (2017). *Cryst. Growth Des.* **17**, 6780–6786.
Hansen, C. & Quake, S. R. (2003). *Curr. Opin. Struct. Biol.* **13**, 538–544.
Hansen, C. L., Classen, S., Berger, J. M. & Quake, S. R. (2006). *J. Am. Chem. Soc.* **128**, 3142–3143.
Hansen, C. L., Skordalakes, E., Berger, J. M. & Quake, S. R. (2002). *Proc. Natl Acad. Sci. USA*, **99**, 16531–16536.
Heymann, M., Opthalage, A., Wierman, J. L., Akella, S., Szebenyi, D. M. E., Gruner, S. M. & Fraden, S. (2014). *IUCrJ*, **1**, 349–360.
Huang, C.-Y., Olieric, V., Ma, P., Panepucci, E., Diederichs, K., Wang, M. & Caffrey, M. (2015). *Acta Cryst.* D**71**, 1238–1256.
Jaskolski, M., Dauter, Z. & Wlodawer, A. (2014). *FEBS J.* **281**, 3985–4009.
Johansson, L. C., Stauch, B., Ishchenko, A. & Cherezov, V. (2017). *Trends Biochem. Sci.* **42**, 749–762.
Kabsch, W. (2010). *Acta Cryst.* D**66**, 125–132.
Khvostichenko, D. S., Schieferstein, J. M., Pawate, A. S., Laible, P. D. & Kenis, P. J. A. (2014). *Cryst. Growth Des.* **14**, 4886–4890.
Li, F., Xiong, Y., Wang, J., Cho, H. D., Tomita, K., Weiner, A. M. & Steitz, T. A. (2002). *Cell*, **111**, 815–824.
Luft, J. R., Newman, J. & Snell, E. H. (2014). *Acta Cryst.* F**70**, 835–853.
Martiel, I., Olieric, V., Caffrey, M. & Wang, M. (2018). *Protein Crystallography: Challenges and Practical Solutions*, edited by K. Beis & G. Evans, pp. 1–27. Cambridge: Royal Society of Chemistry.
Martin-Garcia, J. M., Conrad, C. E., Coe, J., Roy-Chowdhury, S. & Fromme, P. (2016). *Arch. Biochem. Biophys.* **602**, 32–47.
Masquida, B., Sauter, C. & Westhof, E. (1999). *RNA*, **5**, 1384–1395.
McCarthy, A. A., Barrett, R., Beteva, A., Caserotto, H., Dobias, F., Felisaz, F., Giraud, T., Guijarro, M., Janocha, R., Khadrouche, A., Lentini, M., Leonard, G. A., Lopez Marrero, M., Malbet-Monaco, S., McSweeney, S., Nurizzo, D., Papp, G., Rossi, C., Sinoir, J., Sorez, C., Surr, J., Svensson, O., Zander, U., Cipriani, F., Theveneau, P. & Mueller-Dieckmann, C. (2018). *J. Synchrotron Rad.* **25**, 1249–1260.
McPherson, A. & Gavira, J. A. (2014). *Acta Cryst.* F**70**, 2–20.
Meents, A., Gutmann, S., Wagner, A. & Schulze-Briese, C. (2010). *Proc. Natl Acad. Sci. USA*, **107**, 1094–1099.
Meyer, A., Betzel, C. & Pusey, M. (2015). *Acta Cryst.* F**71**, 121–131.
Ng, J. D., Clark, P. J., Stevens, R. C. & Kuhn, P. (2008). *Acta Cryst.* D**64**, 189–197.
Ng, J. D., Gavira, J. A. & García-Ruiz, J. M. (2003). *J. Struct. Biol.* **142**, 218–231.
Otálora, F., Gavira, J. A., Ng, J. D. & García-Ruiz, J. M. (2009). *Prog. Biophys. Mol. Biol.* **101**, 26–37.
Owen, R. L., Juanhuix, J. & Fuchs, M. (2016). *Arch. Biochem. Biophys.* **602**, 21–31.
Perry, S. L., Guha, S., Pawate, A. S., Bhaskarla, A., Agarwal, V., Nair, S. K. & Kenis, P. J. A. (2013). *Lab Chip*, **13**, 3183.
Perry, S. L., Guha, S., Pawate, A. S., Henning, R., Kosheleva, I., Srajer, V., Kenis, P. J. A. & Ren, Z. (2014). *J. Appl. Cryst.* **47**, 1975–1982.
Pinker, F., Brun, M., Morin, P., Deman, A.-L., Chateaux, J.-F., Oliéric, V., Stirnimann, C., Lorber, B., Terrier, N., Ferrigno, R. & Sauter, C. (2013). *Cryst. Growth Des.* **13**, 3333–3340.
Pusey, M., Barcena, J., Morris, M., Singhal, A., Yuan, Q. & Ng, J. (2015). *Acta Cryst.* F**71**, 806–814.
Pusey, M. L., Liu, Z.-J., Tempel, W., Praissman, J., Lin, D., Wang, B.-C., Gavira, J. A. & Ng, J. D. (2005). *Prog. Biophys. Mol. Biol.* **88**, 359–386.
Santoni, G., Zander, U., Mueller-Dieckmann, C., Leonard, G. & Popov, A. (2017). *J. Appl. Cryst.* **50**, 1844–1851.
Sauter, C., Dhouib, K. & Lorber, B. (2007). *Cryst. Growth Des.* **7**, 2247–2250.
Sauter, C., Lorber, B., Gaudry, A., Karim, L., Schwenzer, H., Wien, F., Roblin, P., Florentz, C. & Sissler, M. (2015). *Sci. Rep.* **5**, 17332.
Sauter, C., Lorber, B., McPherson, A. & Giegé, R. (2012). *International Tables for Crystallography*, Vol. *F*, edited by E. Arnold, D. M. Himmel & M. G. Rossmann, pp. 99–121. Chester: International Union of Crystallography.

Stellato, F., Oberthür, D., Liang, M., Bean, R., Gati, C., Yefanov, O., Barty, A., Burkhardt, A., Fischer, P., Galli, L., Kirian, R. A., Meyer, J., Panneerselvam, S., Yoon, C. H., Chervinskii, F., Speller, E., White, T. A., Betzel, C., Meents, A. & Chapman, H. N. (2014). *IUCrJ*, **1**, 204–212.

Stojanoff, V., Jakoncic, J., Oren, D. A., Nagarajan, V., Navarro Poulsen, J.-C., Adams-Cioaba, M. A., Bergfors, T. & Sommer, M. O. A. (2011). *Acta Cryst.* F**67**, 971–975.

Sui, S., Wang, Y., Kolewe, K. W., Srajer, V., Henning, R., Schiffman, J. D., Dimitrakopoulos, C. & Perry, S. L. (2016). *Lab Chip*, **16**, 3082–3096.

Terwilliger, T. C., Stuart, D. & Yokoyama, S. (2009). *Annu. Rev. Biophys.* **38**, 371–383.

Vincentelli, R., Bignon, C., Gruez, A., Canaan, S., Sulzenbacher, G., Tegoni, M., Campanacci, V. & Cambillau, C. (2003). *Acc. Chem. Res.* **36**, 165–172.

Wijn, R. de, Hennig, O., Ernst, F. G. M., Lorber, B., Betat, H., Mörl, M. & Sauter, C. (2018). *Acta Cryst.* F**74**, 747–753.

Woerd, M. van der, Ferree, D. & Pusey, M. (2003). *J. Struct. Biol.* **142**, 180–187.

Yadav, M. K., Gerdts, C. J., Sanishvili, R., Smith, W. W., Roach, L. S., Ismagilov, R. F., Kuhn, P. & Stevens, R. C. (2005). *J. Appl. Cryst.* **38**, 900–905.

Zheng, B., Roach, L. S. & Ismagilov, R. F. (2003). *J. Am. Chem. Soc.* **125**, 11170–11171.

Zhu, D.-W., Lorber, B., Sauter, C., Ng, J. D., Bénas, P., Le Grimellec, C. & Giegé, R. (2001). *Acta Cryst.* D**57**, 552–558.

4. Structural Analysis of RNA by Small-Angle X-ray Scattering

At the beginning of my PhD contract, I could take part in structural analysis of RNA using Small-Angle X-ray Scattering (SAXS). My contribution consisted in the synthesis by *in vitro* transcription of selected tRNAs. Then I could take part in the data acquisition at Soleil synhcrotron (Paris) and have an introduction to SAXS data processing using ATSAS package.

Chapter 14

Structural Analysis of RNA by Small-Angle X-ray Scattering

Anne Théobald-Dietrich, Raphaël de Wijn, Kévin Rollet, Alexandra Bluhm, Joëlle Rudinger-Thirion, Caroline Paulus, Bernard Lorber, Aurélien Thureau, Magali Frugier, and Claude Sauter

Abstract

Over the past two decades small-angle X-ray scattering (SAXS) has become a popular method to characterize solutions of biomolecules including ribonucleic acid (RNA). In an integrative structural approach, SAXS is complementary to crystallography, NMR, and electron microscopy and provides information about RNA architecture and dynamics. This chapter highlights the practical advantages of combining size-exclusion chromatography and SAXS at synchrotron facilities. It is illustrated by practical case studies of samples ranging from single hairpins and tRNA to a large IRES. The emphasis is also put on sample preparation which is a critical step of SAXS analysis and on optimized protocols for in vitro RNA synthesis ensuring the production of mg amount of pure and homogeneous molecules.

Key words SEC-SAXS, RNA, tRNA, IRES, Structure, Integrative structural biology

1 Introduction

Ribonucleic acid (RNA) molecules constitute a very diverse biopolymer family in terms of sequence, size, and shape, and getting insights into their 3D structure is essential to understand their cellular functions and partnerships. Small-angle X-ray scattering (SAXS) is a powerful means to study structural and dynamic properties of RNA in solution, complementary to methods such as crystallography, nuclear magnetic resonance (NMR), or cryoelectron microscopy (cryoEM) [1]. Although SAXS does not provide atomic details, it requires no crystal or isotopic labeling and is performed in solution. The method is compatible with a wide range of solvents and gives a rapid access to structural parameters such as size, oligomeric state, shape of isolated molecules, or complexes [2, 3]. SAXS was already employed in the 60s in early structural investigations on tRNAs [4–6] and has since been applied to samples ranging from compact hairpins to particles as large as the ribosome [7] to study RNA architecture, folding, dynamics, and

Veronique Arluison and Frank Wien (eds.), *RNA Spectroscopy: Methods and Protocols*, Methods in Molecular Biology, vol. 2113, https://doi.org/10.1007/978-1-0716-0278-2_14, © Springer Science+Business Media, LLC, part of Springer Nature 2020

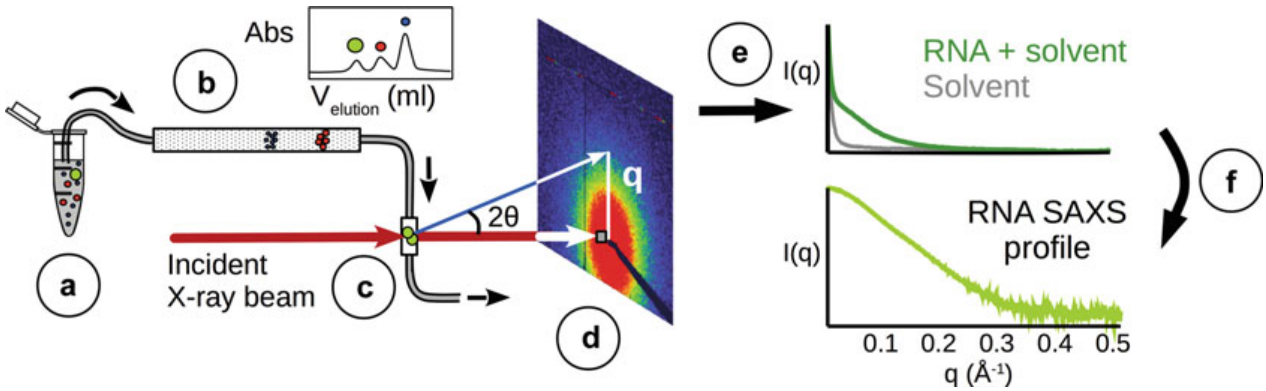


Fig. 1 Principle of a SEC-SAXS setup. The sample solution (**a**) is loaded onto a HPLC-SEC column (**b**) to separate the RNA from higher molecular mass populations (aggregates, multimers). The eluate (monitored by UV absorbance) is directly injected in the SAXS capillary cell (**c**) and scattering patterns are collected on a 2D detector (**d**). 2D patterns are radially averaged (**e**) to produce unidimensional scattering profiles for the solvent (grey curve) and for the solution containing the RNA of interest in the same solvent (green curve). The solvent signal is subtracted from that of the sample (**f**) to obtain a profile corresponding to the RNA. This profile or SAXS curve describes the variation of the scattered intensity as a function of the amplitude of the scattering vector or momentum transfer $q = 4\pi\sin\theta/\lambda$ expressed in Å^{-1} or nm^{-1}, where θ is half the angle between the scattered and incident radiations and λ the wavelength. Structural information on the size and the shape of the scattering particle in solution can then be extracted from the curve $I(q)$ vs. q (*see* Fig. 7)

complexes with proteins [8–15]. Over the last two decades it has grown in popularity among the biologists thanks to the development of software packages dedicated to the characterization of biomolecules and an easy access to performant synchrotron facilities [16].

In practice, a SAXS analysis consists of shining X-rays on a solution of particles and collecting the resulting scattered patterns (Fig. 1). The signal of the solvent is recorded separately and subtracted from the signal of the RNA sample in the solvent to obtain a SAXS profile corresponding to the shape of the RNA. Solvent subtraction is a critical step in data processing. To ensure that the solvent taken as a reference is the same as the solvent in which the RNA is dissolved, it is generally recommended to carry out dialysis or size-exclusion chromatography (SEC). This chapter will indeed focus on the SEC-SAXS approach [17–20], which is available at most synchrotron facilities and implements a FPLC/HPLC-SEC directly upstream the SAXS chamber or capillary (Fig. 1) to separate molecules according to their hydrodynamic properties. The advantage of such a setup is twofold: it guarantees (1) good solvent matching and (2) monodispersity of the RNA entering the measurement cell. Because all molecular entities present in solution contribute to the scattered signal as the square of their volume, it is crucial to separate the population of interest (e.g., monomeric RNA) from aggregates or multimers (likely to form at high concentrations) prior to collecting the SAXS signal [20, 21].

Despite the progress of recent pixel detectors in terms of sensitivity and low background, SAXS, alike other structural methods, is still sample consuming. A single SEC-SAXS run typically requires 20–80 μL of pure RNA solution at 2–5 mg/mL (i.e., 40–400 μg), taking into account that SEC will dilute the sample 5–10 times. Sample preparation is another critical step and we recommend to start with at least a milligram of pure, homogeneous, and monodisperse RNA sample to work comfortably. Among RNA preparation methods detailed in [22], in vitro transcription allows the synthesis of almost any unmodified RNA sequence and has found applications in all areas of RNA research. These transcriptional systems are based on the cloning of synthetic DNA encoding the RNA sequence of interest downstream of the promoter of a phage DNA-dependent RNA polymerase. Among the different phage RNA polymerases available (SP6, T7, T5, T3), the T7 RNA polymerase has found the widest applications. Detailed protocols to produce RNA using the T7 RNA polymerase can be found in [23].

The first part of this chapter presents experimental conditions for efficient T7 transcription and practical tricks to obtain high amount of homogenous transcripts for SAXS experiments (Fig. 2). When designing an RNA molecule, several constraints have to be considered such as the nature of the 5′ sequence of the construct, its total length, and the homogeneity of the 3′ end. Three protocols adapted to the production of RNAs of different sizes by in vitro transcription are described, as well as RNA purification techniques, and the usefulness of a strict quality control of the sample to guarantee the reproducibility and the success of SAXS measurements.

The second part of this chapter is an introduction to standard SEC-SAXS analysis of RNA using synchrotron radiation. The description is based on the experimental setup of the SAXS beamline SWING [19] at SOLEIL synchrotron (Saint-Aubin, France), but the principle of data collection and analysis is applicable to any experimental setup. Besides the determination of structural parameters such as the radius of gyration (*Rg*) and maximal dimension (*Dmax*) of the particle, its Porod volume (*Vp*), and characteristic *I* (0) directly related to its molecular mass (MM), SAXS offers a wide range of possibilities, including the study of conformational changes, flexibility, disorder, or aggregation. The experimental section illustrates data processing and analysis performed using the widespread *ATSAS* suite [24] on different RNA samples including single RNA hairpins, tRNAs and variants, and internal ribosomal entry site (IRES) present in the 5′ UTR of the mRNA of human glycyl-tRNA synthetase (GlyRS; M. Frugier, to be published). The usefulness of SAXS in an integrative structural approach is discussed in the notes and exemplified with the characterization of atypical tRNAs and tRNA/protein complexes.

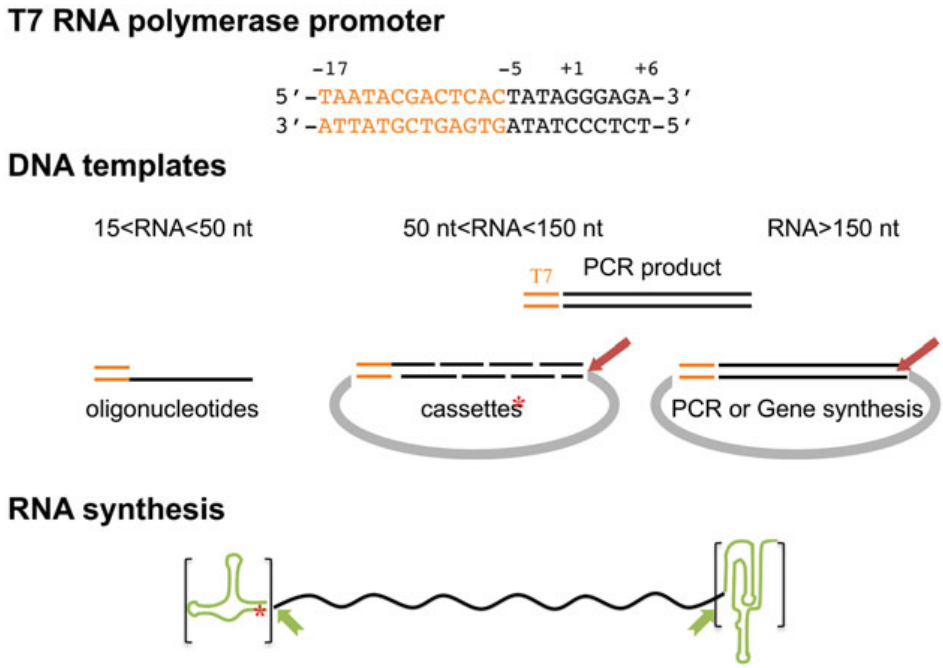


Fig. 2 RNA sample preparation by in vitro T7 transcription. The T7 RNA polymerase promoter encompasses 23 nucleotides (nt) and contains the recognition domain (orange, −17 through −5) and the initiation domain (black, −5 through +6); transcription initiates at position +1 with 3 Gs giving rise to an RNA transcript complementary to the template strand. The nature of the template may differ depending on the size of the desired product: hybridized oligonucleotides, PCR products, or linearized (red arrows) plasmids. In order to produce transcripts with homogeneous 5′ and 3′ extremities, ribozymes (green) may be inserted. Red asterisks indicate protocols detailed in this chapter

Readers eager to learn more about theoretical aspects of the method will find reviews on the topic [25–27], some of which specifically dedicated to SAXS and RNA [8, 20, 28, 29].

2 Materials

2.1 Cloning of Cassettes

1. Desalted DNA oligonucleotides.
2. Restriction enzymes: *Bam*H1, *Hind*III, *Bst*NI, or *Nsi*I.
3. Other enzymes: T4 polynucleotide kinase, T4 DNA ligase, FastAP.
4. Dry heat block.
5. LB/ampicillin agar plates.

2.2 Purification of RNA on Polyacrylamide Gels and Band Visualization

1. 33 × 40 cm^2 glass plates.
2. 2-mm-thick spacers.
3. 2-mm-thick combs.
4. 10× TBE: 0.89 M Tris–borate pH 8.0, 20 mM EDTA.

5. 300-mL denaturing polyacrylamide gel (acrylamide/bisacrylamide 19:1 ratio, 8 M urea, 1× TBE) or native polyacrylamide gel (acrylamide/bisacrylamide 37.5:1 ratio, 1× TBE).
6. 2× RNA denaturing gel loading dye solution: 95% (v/v) formamide, 10 mM EDTA, 25 mM Tris–HCl pH 8.0, 0.025% (w/v) xylene cyanol, and 0.025% (w/v) bromophenol blue.
7. 2× RNA native gel loading dye solution: 50% (w/v) glycerol, 25 mM Tris–HCl pH 8.0, 0.025% (w/v) xylene cyanol and 0.025% (w/v) bromophenol blue.
8. Plastic wrap.
9. Scalpel blade.
10. TLC plate with fluorescent dye (Polygram Gel 400 UV 254, Macherey-Nagel).
11. UV light source.

2.3 Autocatalytic Cleavage of RNA with Ribozyme

1. Dry heat block.
2. Transcription buffer (5×): 200 mM Tris–HCl pH 8.1, 150 mM $MgCl_2$, 5 mM spermidine, 25 mM DTE, 0.05% (v/v) Triton X-100, 20 mM of ATP, CTP, GTP, and UTP.
3. Dilution buffer: 40 mM Tris–HCl pH 8.1 and 30 mM $MgCl_2$.
4. Incubate 1 hour at 60°C.

2.4 RNA Electroelution

1. BioRad Sub-Cell GT or Mini-Sub Cell GT apparatus (BioRad).
2. Sterile 1× TBE.
3. Power supply.
4. Elutrap system and Biotrap membranes BT1 and BT2 (Schleicher & Schuell).
5. Alternatively to 4: Dialysis tube Spectra/Por of appropriate cutoff (Spectrum Laboratories, Inc.) and clips to seal dialysis tubes.

2.5 RNA Purification by Chromatography

1. Bio-Scale Q5 anion exchange column (BioRad).
2. Superdex 75 increase 10/300 column (GE Healthcare Life Sciences).
3. Superdex 200 increase 10/300 column (GE Healthcare Life Sciences).
4. Superose 6 increase 3.2/300 column for longer transcripts (GE Healthcare Life Sciences).
5. Anion exchange buffer QA: 50 mM Tris–HCl pH 7.5, 50 mM KCl, 5 mM $MgCl_2$.

Table 1
HPLC-SEC columns

Column	Manufacturer	Exclusion limits/volume/flow rate/matrix
Bio SEC3-100	Agilent	0.1–100 kDa/4.6 mL/0.3 mL/min/silica
Bio SEC3-150	Agilent	0.5–150 kDa/4.6 mL/0.3 mL/min/silica
Bio SEC3-300	Agilent	5–1250 kDa/4.6 mL/0.3 mL/min/silica
Superdex 200 increase	GE Healthcare	10–600 kDa/2.4 mL/0.15 mL/min/agarose

Table 2
SAXS structural parameters of selected RNA molecules

Name	Hairpin tetraloop	Hairpin T-loop	$tRNA^{Phe}$	gly-IRES
Nucleotides	32	35	76	237
Rg Guinier (Å)	15.59 ± 0.02	18.25 ± 0.04	22.04 ± 0.03	50.58 ± 0.06
Rg p(*r*) (Å)	15.74 ± 0.01	18.51 ± 0.01	22.86 ± 0.02	50.51 ± 0.02
D_{max} (Å)	51	62	74	157
Vp ($Å^3$)	11,600	14,600	28,100	169,000
MM_{theor} (kDa)	10.2	11.4	23.5	71.6
MM_{Vc} (kDa)[a]	10.1	11.0	21.8	106

[a]Molecular mass estimate according to [33]

6. Anion exchange buffer QB: 50 mM Tris–HCl pH 7.5, 2 M KCl, 5 mM $MgCl_2$.
7. Size-exclusion buffer: 50 mM Tris–HCl pH 7.5, 150 mM KCl, 5 mM $MgCl_2$.
8. Amicon® Ultra Centrifugal Filters (Merck) with adapted volume and molecular mass cut-off.

2.6 HPLC-SEC Chromatography

1. Size-exclusion columns: *see* examples in Table 1.
2. HPLC system (Agilent LC System 1200–1260 series).
3. 2-mL Vials and conical glass inserts with polymer feet for the HPLC injectors (Agilent).
4. Size-exclusion buffer: 50 mM HEPES-Na pH 7.5, 150 mM KCl, 5 mM $MgCl_2$, filtered and degassed on filtration unit with 0.22-μm filter.

2.7 Data Processing and Analysis

1. Data processing is performed with *Foxtrot* [19] as implemented on SWING beamline.

2. Data analysis is performed with the *ATSAS* suite [24], a widely used package in the bioSAXS community and freely available to academics at https://www.embl-hamburg.de/biosaxs/software.html. Other packages are available and may be used as well, such as *BIOXTAS-RAW* and *SCATTER* [30, 31].
3. Model building and inspection are carried out using *PyMOL* (Schrödinger LTD).

3 Methods

3.1 RNA Production by In Vitro Transcription (See Fig. 2 and Note 1)

3.1.1 RNA Cloning Using the Cassette Technique

1. Reconstitute the gene with overlapping oligonucleotides as shown for a synthetic tRNA gene in Fig. 3.
2. Design 18- to 21-mers that overlap by 7–10 nucleotides.
3. Include in the construct (1) the upstream T7 RNA polymerase promoter sequence (−17 to −1), (2) *Hind*III and *Bam*HI restriction sites for cloning in pUC118 vector, and (3) the downstream *Bst*NI restriction site coincidental with the 3′-end of the sequence to ensure the correct CCA end of the tRNA.
4. If *Bst*NI is present in the tRNA sequence, then it may be replaced by *Nsi*I.
5. By substituting two oligonucleotides (one on each strand), point variants of the RNA gene can be easily cloned in parallel.

3.1.2 Phosphorylation of Oligonucleotides

1. Phosphorylate individually each oligonucleotide (400 pmol) with 10 U (1 μL) T4 polynucleotide kinase in 50 μL.
2. Incubate for 15 min at 37 °C.
3. Inactivate the kinase by rising the temperature to 65 °C for 10 min.

3.1.3 Hybridization of Phosphorylated Oligonucleotides

1. Mix 2 μL (16 pmol) of each phosphorylated oligonucleotide ($2 \times 10 = 20$ μL total volume).
2. Add 20 μL of buffer 150 mM Tris-HCl pH 7.5, 30 mM $MgCl_2$, 15 mM dithioerythritol.
3. Incubate the mixture at 85 °C for 10 min.
4. Switch off the heater block and leave the tube to slowly cool down to 30 °C (2–3 h).

3.1.4 Ligation of the RNA Gene into pUC 118 Vector

1. Digest successively pUC118 with *Bam*HI and *Hind*III.
2. Perform a dephosphorylation step of the digested vector with FastAP, stop the reaction with phenol extraction, precipitate and adjust the concentration at 1 μg/μL.

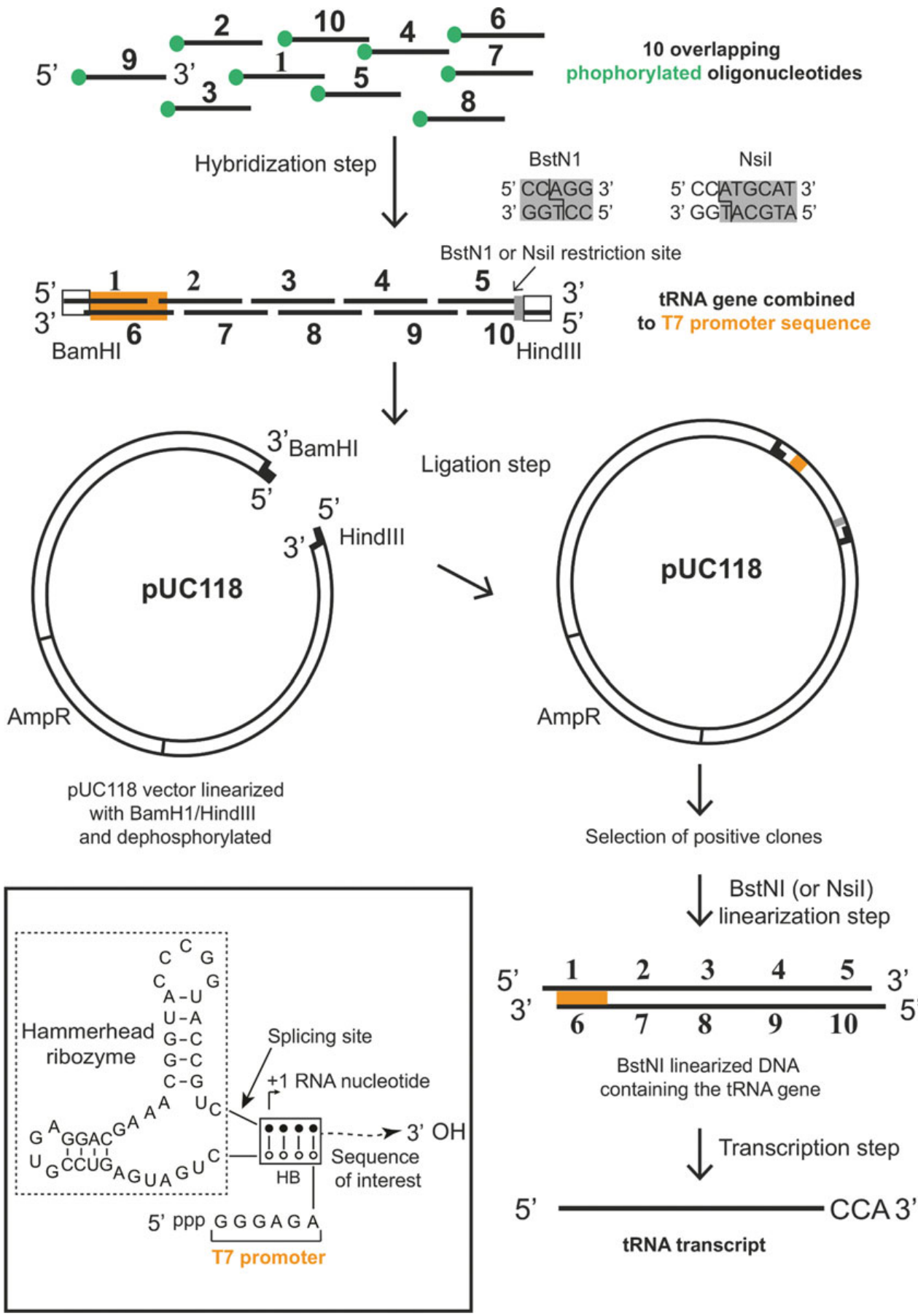


Fig. 3 Production of RNA transcripts using the cassette technique. The RNA synthetic gene (here a tRNA gene) in combination with the T7 RNA polymerase promoter and BstNI site is assembled before its ligation into the

3. Incubate 10 μL of the hybridization mixture (obtained in Subheading 3.1.3) with 1 μg pUC118 (digested and dephosphorylated) in the presence of 5 μL ATP 10 mM and 5 U (5 μL) T4 DNA ligase (Thermo Scientific) overnight at 16 °C.

3.1.5 Bacterial Transformation and Search for Positive Clones

1. Transform 200 μL of $CaCl_2$ chimio-competent *E. coli* TG1 cells with the ligation mixture.
2. Spread on LB/ampicillin plates and incubate overnight at 37 °C.
3. Positive clones are verified by sequencing.
4. Perform DNA maxi-preparation of the positive plasmid.

3.1.6 In Vitro T7 Transcription

1. Digest 50 μg of plasmid with *Bst*NI.
2. Prepare a 500 μL transcription mix containing 100 μL of transcription buffer (5×), 25 U T7 RNA polymerase, 50 μg of linearized plasmid and 4 mM of each nucleotide triphosphate (pH adjusted to 7.5).
3. Incubate at 37 °C for 3 h in a dry heat block.
4. Stop the reaction by ethanol precipitation (2.5 volumes) in the presence of 0.3 M acetate-Na pH 5.2.
5. Centrifuge the solution for 15 min at 15,000 *g* and store the dry RNA pellet at −20 °C.

3.1.7 RNA Cloning Using the Transzyme Strategy

1. Assemble the sequence of the synthetic transzyme gene with cassettes (*see* Subheading 3.1.1) as illustrated in Fig. 3.
2. Oligonucleotide#2, coding for the hammerhead ribozyme, remains identical for all sequences (boxed in dashed lines in Fig. 3).
3. The four nucleotides on the 5′ side of the hybridization box (HB), downstream of the strong T7 internal promoter, must be complementary to the first four nucleotides of the sequence of interest that form the 3′ side of HB.

Fig. 3 (continued) pUC118 plasmid. The tRNA gene-containing plasmid is then linearized with *Bst*NI and T7 RNA transcription is performed. Note that *Bst*NI enables the production of RNA molecules ending with the 3′-CCA sequence universally conserved in tRNA. In the transzyme strategy (insert), the oligonucleotide#2 incorporates the hammerhead ribozyme. The transzyme construct, i.e., a transcript integrating the self-cleavable ribozyme, contains, from 5′ to 3′, the strong internal promoter sequence for the T7 RNA polymerase, the hammerhead ribozyme sequence (boxed in dashed lines), and the domain called hybridization box (HB). The splicing of the RNA occurs between the hammerhead ribozyme and the HB. The open circles correspond to the four variable nucleotides immediately downstream of the T7 promoter. They are chosen to be complementary to the starting 5′ sequence of the RNA of interest (filled circle) allowing the formation of the base-paired structure (HB) that directs the specific cleavage of the ribozyme. The autocatalytic cleavage releases the desired RNA sequence with a 5′OH terminus

3.1.8 Transcription and Autocatalytic Cleavage

1. Prepare a 500 μL transcription mix containing 100 μL of transcription buffer (5×), 25 U T7 RNA polymerase, 50 μg of linearized plasmid in a 1.5-mL Eppendorf tube and 4 mM of each nucleotide triphosphate (pH adjusted to 7.5).
2. Incubate at 37 °C for 3 h in a dry heat block.
3. Dilute the transcription mix with 2 mL of dilution buffer (*see* **Note 2**).
4. Incubate at 60 °C for 1 h in a dry heat block.
5. Stop the reaction by ethanol precipitation (*see* Subheading 3.1.6).

3.2 RNA Purification (Fig. 4)

3.2.1 RNA Purification on Gel (See **Note 3**)

1. Resolubilize the precipitated transcript in water and RNA denaturing loading buffer (1:1 mix).
2. Clean glass plates, spacers, and combs.
3. Assemble gel plates using tape to seal the bottom carefully and put the plates upright.
4. Prepare 25 mL of gel and pour the "bottom" layer (5 cm high) of the gel; let it polymerize.
5. Pour the rest of the gel slowly, remove air bubbles, insert combs, and let it polymerize horizontally for at least 15 min.
6. Remove both the tape and the comb and rinse thoroughly the well with 1× TBE.
7. Perform the electrophoresis overnight at 650 V/55 mA at room temperature.
8. To insure an optimal separation, run the electrophoresis as long as possible being careful not to let the RNA migrate outside the gel. For example, to purify a 76-nt-long tRNA transcript, the transcription reaction is loaded on a 12% gel and the electrophoresis is stopped 2 h after the xylene cyanol dye has left the bottom of the gel.
9. Remove the gel from the plates at the end of the electrophoresis process, and cover it with a clean plastic wrapping film.
10. Place the gel on the TLC plate and illuminate with the UV source (254 nm). DNA and RNA molecules will appear as dark purple bands (*see* **Note 3**).
11. Cut the gel piece containing the RNA of interest, store it at −20 °C until needed, or use it immediately for electroelution.

3.2.2 RNA Extraction by Electroelution Using Biotrap Device

1. Before use, sterilize the Biotrap device wrapped in aluminum foil.
2. Insert a BT2 membrane in the second trap insert and two BT1 membranes at each end of the device (Fig. 5a).

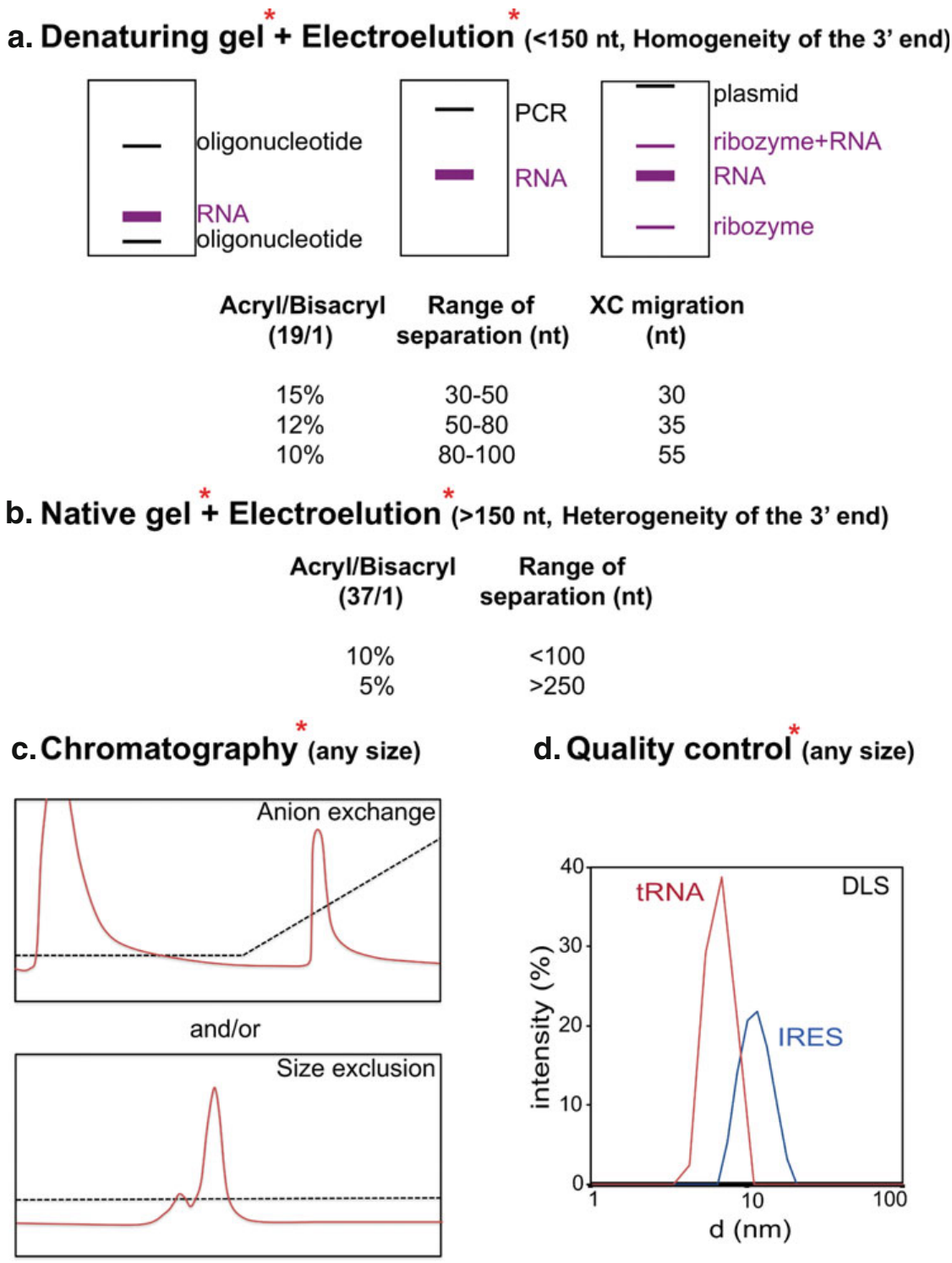


Fig. 4 Large-scale purification of RNA for SAXS. The purification procedure varies with the size of the RNA. (**a**) Molecules of less than 150 nt can be purified on denaturing polyacrylamide gels. Several profiles can be observed depending on the chosen DNA template (black) and the produced RNA molecules (purple). The percentage of the gel has to be adapted to the size (denaturing gels) or the shape (native) of the RNA. For longer RNAs (>150 nt), it is recommended to use (**b**) native gels and/or (**c**) chromatography (anion exchange or size exclusion) to remove unincorporated nucleotides and incomplete or unfolded transcripts. In all cases, native and denaturing gels are used to assess the purity and homogeneity of the sample, along with SEC and DLS, before SAXS analysis. (**d**) DLS is a simple, rapid, and noninvasive method ideal for quality control [44]. In the example two samples—yeast $tRNA^{Phe}$ and the IRES of human GlyRS mRNA—show a monomodal size distribution that confirms their homogeneity. Red asterisks indicate protocols detailed in this chapter

3. Put RNA-containing gel pieces in the device and immerse them in sterile 1× TBE buffer.
4. Perform a first electroelution during 1 h at 150 V in a cold room (4 °C).
5. Recover the eluate from the elution chamber (800 μL) and divide it in two times 400 μL.
6. Perform a second elution with fresh 1 × TBE and split again the recovered eluate.
7. Add 40 μL of sodium acetate 3 M pH 5.2 and 1 mL ethanol to each sample; let it precipitate overnight at −20 °C.
8. Centrifuge for 45 min at 15,000 *g*, wash, and dry the pellet.
9. The recovery yield of RNA from the gel with the Biotrap is about 80–90%.

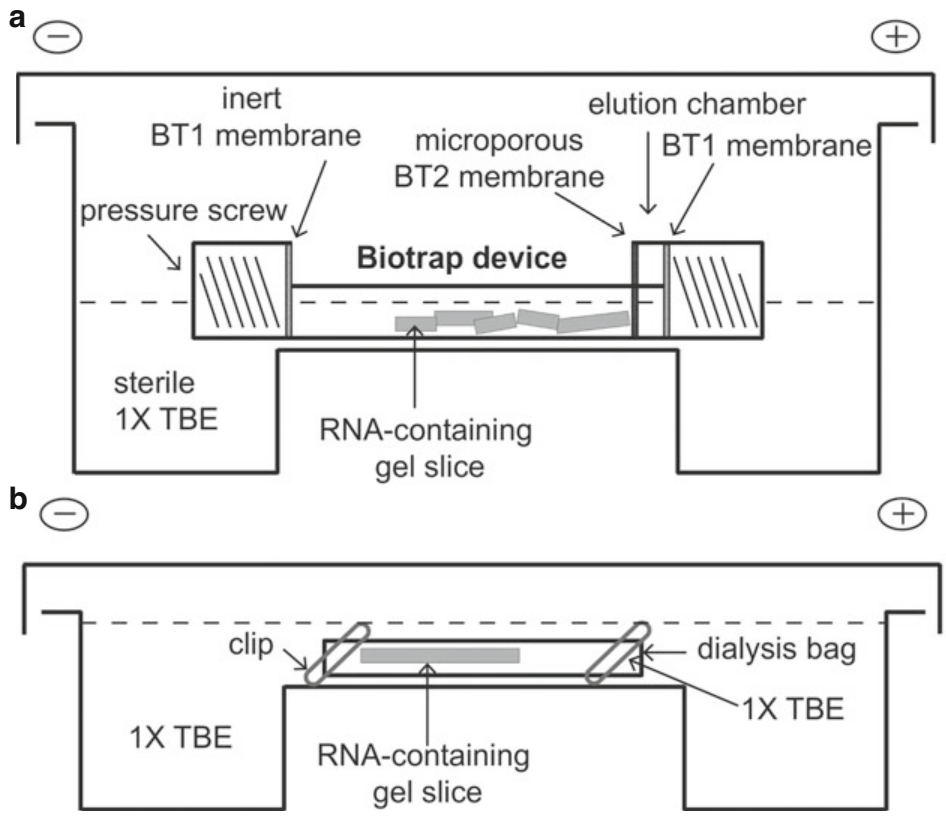


Fig. 5 Electroelution of RNA transcript from polyacrylamide gel. (**a**) Using the Biotrap apparatus: transcript-containing gel slices are introduced in a Biotrap device closed by BT1 membranes on each side and are submerged with 10-mL sterile 1× TBE. A BT2 membrane is inserted to form a chamber of adjustable volume in which the transcript will be trapped. The Biotrap is placed in the electrophoresis Sub-Cell GT system filled with 1× TBE buffer (buffer levels in and out of the Biotrap device should be identical). The electric field permits the migration of the transcript out of the gel pieces into the elution chamber formed between BT1 and BT2 membranes. (**b**) Using a dialysis bag: the transcript-containing gel slice is introduced in a dialysis tube filled with sterile 1× TBE and sealed on each end with a dialysis clip. The bag is placed on a Mini-Sub Cell GT or in a Sub-Cell GT apparatus (depending on the size of the bag) so that the gel piece is perpendicular to the electrodes. After immersion with 1× TBE buffer, the electric field is applied to permit the migration of the transcript out of the gel slice inside the bag

3.2.3 RNA Extraction by Electroelution in a Dialysis Bag

1. Use this alternative as a less sophisticated and less expensive option (Fig. 5b).
2. Insert RNA-containing gel pieces in the dialysis tube (its cutoff depends on the molecular weight of the RNA to be extracted, e.g., for tRNA transcripts, use a cutoff of 12–14 kDa to keep the eluted molecules inside the tube).
3. Place the tube in a migration device (Mini-Sub Cell GT or Sub-Cell GT) and immerse it in 1× TBE buffer.
4. Perform two successive electroelutions at 150 V in a cold room (4 °C) for 20 min.
5. Before recovering the transcript, invert the current during 30 s to unstick the eluted transcript from the inner bag membrane.
6. Pipet the buffer from the dialysis bag.
7. Reduce the volume (which is usually comprised between 2 and 10 mL) by a concentration step in appropriate centrifugal filter units.
8. If necessary, perform an additional precipitation step with sodium acetate and ethanol.
9. The recovery yield of the transcript from the gel with the dialysis tube is estimated at about 50–60%.

*3.2.4 RNA Purification by Chromatography (See **Note 4**)*

1. Centrifuge the transcription mix for 5 min at 15,000 *g* to remove the white precipitate that contains pyrophosphate (PPi).
2. Add 9 volumes of water to the supernatant.
3. Equilibrate the anion exchange column with five volumes of buffer QA.
4. Load the sample on the column using a reduced flow rate (no more than 1 mL/min).
5. Wash the column with 20 column volumes of buffer QA. The flow rate can be increased up to 2 mL/min.
6. Run a salt gradient from 50 mM KCl (buffer QA) to 1.5 M (buffer QB) over 15 column volumes.
7. Monitor the elution using the absorbance at 254 nm.
8. Analyze the fractions on a native gel.
9. If additional bands are present (corresponding to unfolded RNA or RNA aggregates), then a SEC step may be added to remove unwanted conformers. Perform an ethanol precipitation of the RNA sample or concentrate it in a Centricon centrifugal unit and run a SEC in the corresponding buffer. The main peak contains the RNA of interest, freed of aggregates.

10. Collect the corresponding fractions and ethanol precipitate or concentrate the RNA in a Centricon centrifugal unit with the adapted cutoff.

3.2.5 RNA Quantification and Quality Control (See **Note 5**)

1. Dissolve the RNA pellet in distilled water and determine the concentration of the solution with a spectrophotometer at 260 nm (1 absorbance unit corresponds to 40 μg/mL).
2. Denaturing conditions can alter the proper folding of the transcript. Therefore, before use, a denaturation/renaturation step of the transcript is recommended. It consists of incubating the water-solved transcript for 2 min at 65 °C followed by a 10-min cooling at 25 °C and subsequent addition of $MgCl_2$ (at the final concentration needed for functional assays).
3. Assess the purity and quality of the eluted transcript by performing an analytical polyacrylamide gel electrophoresis on an aliquot.
4. Run an analytical SEC with the column used for SEC-SAXS analysis (*see* Subheading 3.3.2) to assess the homogeneity of the sample in solution.
5. When available, use multiangle (MALS) or dynamic light scattering (DLS) to confirm sample homogeneity before running a SAXS experiment (*see* Fig. 4d).

3.3 SEC-SAXS Experiment

3.3.1 Column and Sample Preparation (See **Note 6**)

1. Rinse the SEC column with water (three column volumes).
2. Equilibrate with the elution buffer (five column volumes recommended) at the flow rate used for sample elution (0.3 mL/min with BioSEC columns f.ex.; *see* Table 1).
3. Centrifuge the sample (10 min, 15,000 *g*, 4 °C).
4. Transfer an aliquot of 20–80 μL in an adapted HPLC vial, and place the vial in the injector of the HPLC system.

3.3.2 Test Elution (Optional)

1. This step should be performed before the synchrotron session as part of sample quality control.
2. Inject 5 μL of sample and run a test elution.
3. Follow the elution based on UV absorbance signal (Fig. 6a).
4. Define the SAXS acquisition windows for the solvent (at the beginning of the run) and for the sample (starting at the end of the void volume V0).

3.3.3 Data Acquisition

1. Inject 10–70 μL of sample and run the SEC elution.
2. Follow the elution based on UV absorbance (*see* **Note 6** and Fig. 6a).
3. Collect scattering patterns for the solvent and the solvent plus sample in appropriate time windows (Fig. 6a).

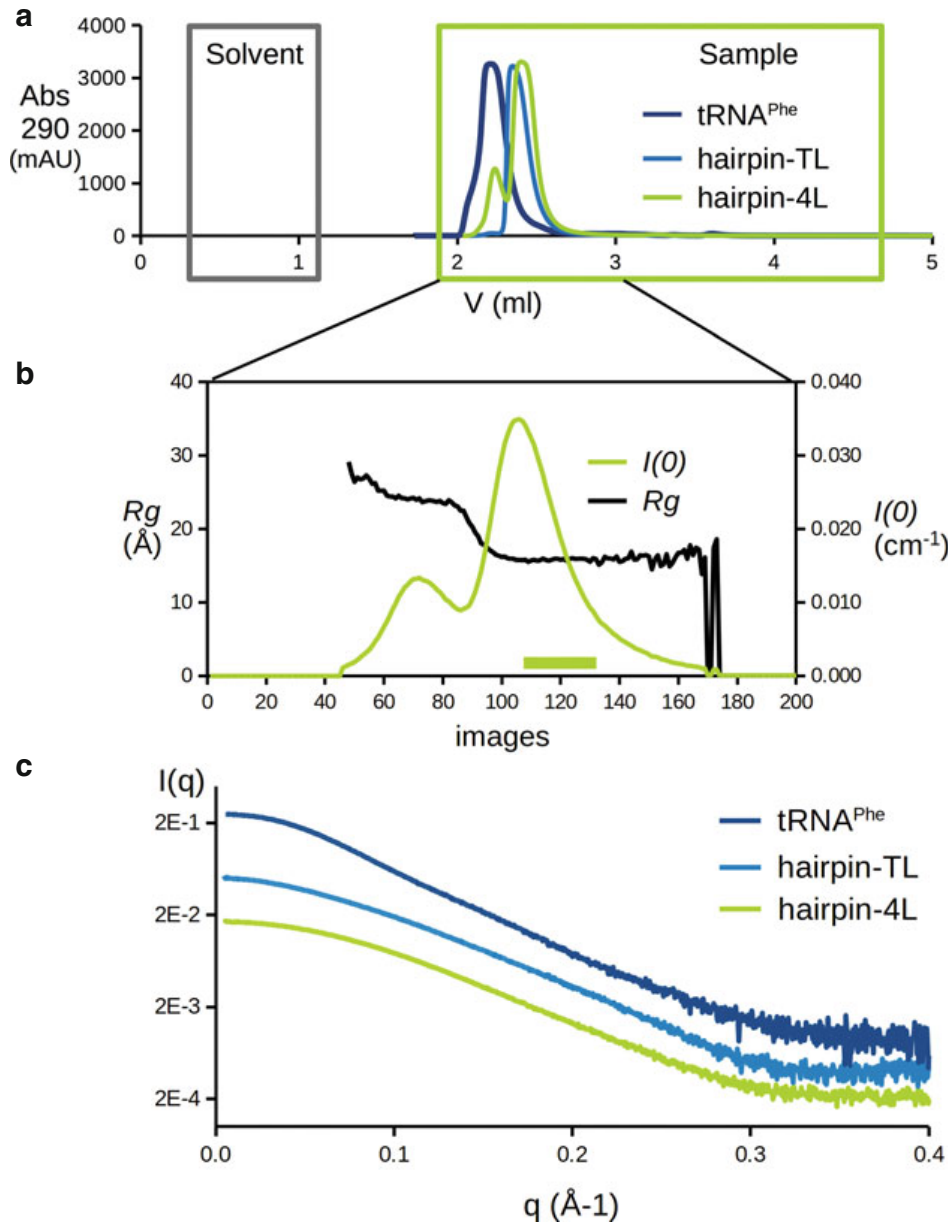


Fig. 6 Example of SEC-SAXS data collection and processing. (**a**) SEC chromatogram recorded at 290 nm to avoid signal saturation. The elution profiles performed on a BioSEC3-100 column at a flow of 0.3 mL/min are depicted for three samples: the native yeast $tRNA^{Phe}$ (in dark blue) and two hairpins corresponding to the acceptor stem of yeast $tRNA^{Phe}$ with its natural 7-nucleotide T-loop (hairpin TL in light blue) or with a tetraloop (hairpin 4L in green). The RNA quantity injected for these analyses was ~250 μg. The three RNAs elute in the expected order, from the largest (23.5 kDa) to the smallest (10.2 kDa). SAXS signal is recorded for the solvent in the grey window and then for the sample in the green window. The averaged solvent signal (*see* Fig. 1) is then subtracted from the curves obtained for each image in the green window to retrieve the sample signal from which X-ray scattering intensity *I*(0) and radius of gyration *Rg* can be estimated. (**b**) Evolution of *I*(0) and *Rg* values along the SEC elution of the hairpin 4L sample. A dataset of 640 SAXS images was collected in the green window, but the analysis is restricted to the region of the major elution peak (images 0–200). As in the chromatogram, two peaks are observed, the first characterized by *Rg* values of ~24 Å and likely containing a dimer of hairpin, the second with *Rg* values of ~16 Å corresponding to the monomeric RNA. SAXS profiles from images 112–130 (green bar) showing a stable *Rg* are averaged leading to the green profile for the hairpin 4L shown in (**c**). The same procedure was applied to the two other RNAs leading to the profiles displayed in (**c**). Resulting SAXS curves are represented on a relative scale with an offset to facilitate their comparison

3.3.4 Data Processing

1. Process the set of images collected with the solvent, i.e., create a unidimensional SAXS profile for each image by radial averaging of 2D patterns (Fig. 1e).
2. Superimpose and compare the solvent profiles to check that there is no drift in the signal.
3. Average these profiles to produce a low-noise solvent curve.
4. Process the set of images collected during the elution of the sample and create corresponding unidimensional SAXS profiles (*see* **Note 7**).
5. Subtract the averaged solvent profile to each profile to retrieve the characteristic SAXS signal of the solute(s) (Fig. 1f).
6. Run a global Guinier analysis to estimate the Rg and $I(0)$ of each subtracted frame (Fig. 6b).
7. Identify a region where the Rg is stable and average the profiles to obtain the SAXS curve of the sample (Fig. 6c).

3.4 Data Analysis

1. Display the SAXS curve of the sample in *PRIMUS* [24] (Fig. 7a).
2. In the analysis panel, select "radius of gyration" (*see* Fig. 7b and **Note 8**) and plot the data as $\ln(I)$ vs. q^2.
3. Adjust the high q value to match Guinier conditions ($qRg < 1.2$).
4. Adjust high and low q values to get a linear fit of the data and the low residual.
5. Record the Rg and associated standard deviation.
6. If the curve has been determined on an absolute scale (in cm^{-1}) with respect to a standard (water scattering, for example), then $I(0)$ should also be recorded for further molecular mass determination.
7. In the analysis panel, select "distance distribution" (*see* Fig. 7c and **Note 9**).
8. Starting from *AUTOGNOM* parameters, adjust the high and low q values to get a good fit of the SAXS curve (left panel).
9. Adjust the D_{max} to get a pair-distance distribution function (PDDF, right panel) or $p(r)$ that falls off smoothly at $p(D_{max}) = 0$.
10. For ab initio reconstructions, q_{max} should not exceed 0.25–0.3 ($qRg < 7$–8).
11. Record the Rg from $p(r)$ and D_{max}, and save the output (extension .out) file.
12. The panel "molecular weight" provides an estimate of the particle molecular mass based on various approaches [32, 33].

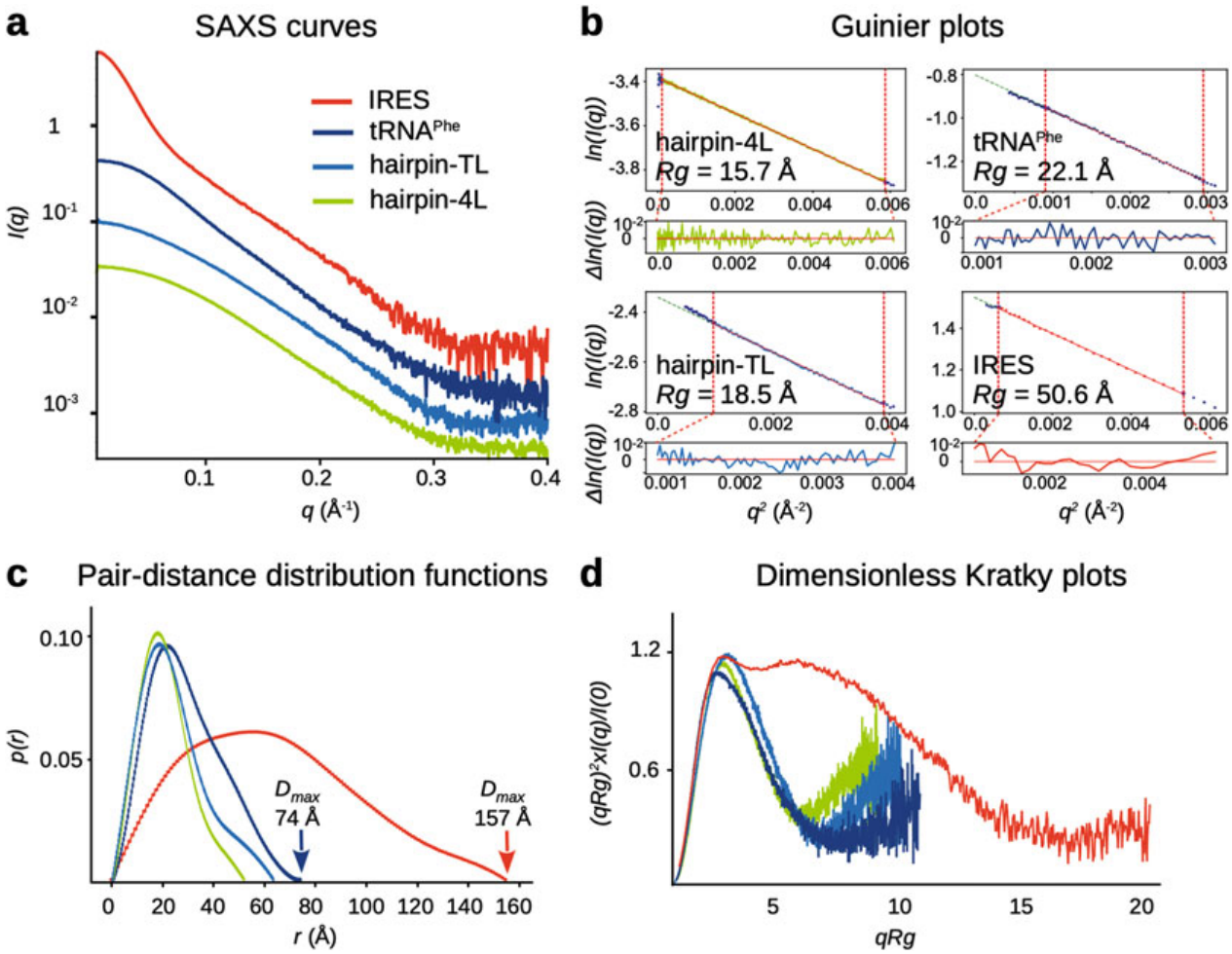


Fig. 7 SAXS data analysis on various RNAs. (**a**) SAXS curves of native yeast tRNA^{Phe} (in dark blue) and two derived acceptor arms, hairpin TL (light blue) and hairpin 4 L (green), and the IRES (red) present in the human GlyRS mRNA (*see* details in Table 2). Corresponding Guinier plots are shown in (**b**). In the Guinier approximation at low angles (typically $q_{max} < 1.2/Rg$ Å^{-1}), the data plotted as ln$I(q)$ vs. q^2 should be linear and the slope provides an estimation of *Rg. Rg values are indicated for the four RNAs.* The residual Δln($I(q)$) vs. q^2 shown below each Guinier plot indicates the quality of fit in the selected q interval (i.e., between red dotted lines). If the data are on absolute scale, the intercept at $q = 0$ gives the $I(0)$. (**c**) An indirect Fourier transform of the SAXS curves gives access to the pair-distance distribution function $p(r)$ of the sample, which allows to determine the maximal distance D_{max}, as well as a more accurate estimate of Rg and $I(0)$. The shape of the $p(r)$ is correlated to the shape of the particle. Hairpins and tRNA^{Phe} show similar $p(r)$ profiles typical of elongate objects, whereas the IRES looks more as a large multidomain particle. (**d**) The dimensionless Kratky representation $I(q)/I(0) * (qRg)^2$ vs. qRg (plotted using *BioXtas-RAW* [31]) is used to estimate the "globularity" of the particles and to detect flexible or unfolded regions. Unfolded particles produce a hyperbolic Kratky plot that plateaus at high q, whereas folded particles show a return to the baseline. Perfectly globular objects show a peak at about 1.1 for $qRg = \sqrt{3}$. It is the case with the four studied RNAs, the IRES displaying a second peak indicative of a multidomain particle

13. The panel "Porod volume" provides an estimate of the particle volume from the Porod invariant Q (*see* **Note 10**).
14. Plot the data as $I(q)q^4$ vs. q (Porod plot) with data up to $0.25 < q < 0.3$ to derive the Vp of the particle.

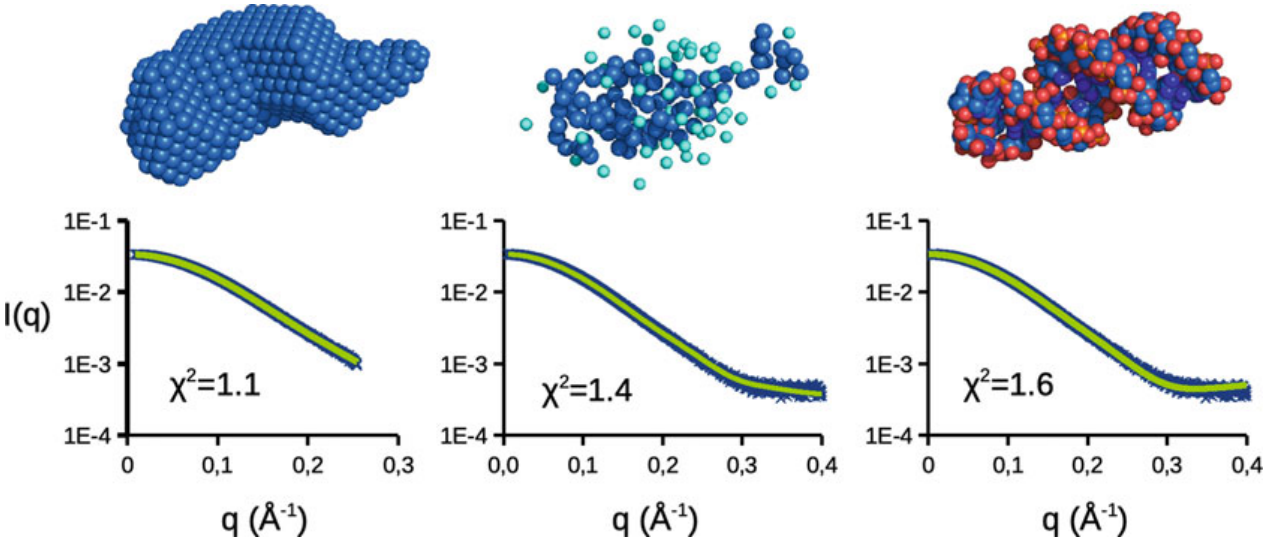


Fig. 8 SAXS-based RNA modeling. In this example dealing with the hairpin 4L (32 nt), the SAXS curve is used to test an atomic model or to generate ab initio models in combination with the *p*(*r*) (*see* Fig. 6). (Left) *DAMMIF* is run at low resolution to reconstruct molecular shapes made of dummy atoms (blue spheres) that reproduce (green curve) the SAXS curve (blue crosses). Twenty models were generated with a normalized spatial discrepancy (NSD) of 0.5 and the model showing the best goodness of fit (χ^2) is displayed. (Middle) Modeling against the full range of data can be performed with *GASBOR* creating chains made of dummy residues (blue spheres) folded by simulated annealing and surrounded by solvent molecules (light blue). The best model of an ensemble showing an NDS of 0.7 is displayed. (Right) When an atomic model is available, its theoretical SAXS curve (green curve) can be computed with *CRYSOL* and compared with experimental data (blue crosses). This model is derived from the acceptor arm of $tRNA^{Phe}$ (PDBid 1EHZ, ref. 56) in which the T-loop is replaced by a UUCG tetraloop. The minimized model is in excellent agreement with the SAXS data ($\chi^2 = 1.6$). Ab initio models—low-resolution envelopes and chains of dummy residues—provide a good idea of the overall shape of the hairpin. All models of hairpin 4L are aligned and displayed at the same scale for comparison

15. Plot the data as $I(q)q^2$ vs. q (Kratky in plot menu) to assess the compactness of the particle (*see* Fig. 7d and **Note 10**).

3.5 SAXS-Based RNA Modeling (See Note 11)

1. Use the output file of GNOM (**step 10** in Subheading 3.5) as input for *DAMMIF* [34] to generate a series of models made of dummy atoms (Fig. 8).
2. Run *DAMMIF* either from the *PRIMUS* interface or from the command line for a better control of the parameters.
3. Superimpose dummy atom models and compare them (*DAMAVER* procedure) to determine their normalized spatial discrepancy (NSD) using different tools of *ATSAS* [24]. The mean NSD should be lower than 0.7–0.9 for a unique solution.
4. Visualize a series of 10–30 ab initio models generated by *DAMMIF* in *PyMOL*.
5. Use the command "set grid_mode, on" in *PyMOL* to display the models on a grid.

6. Apply similar procedures with *GASBOR* [35] to data at higher angles ($0.4 < q < 0.6$ Å^{-1}) to create dummy residue models. The chain of dummy residues is folded by simulated annealing to fit the experimental data.

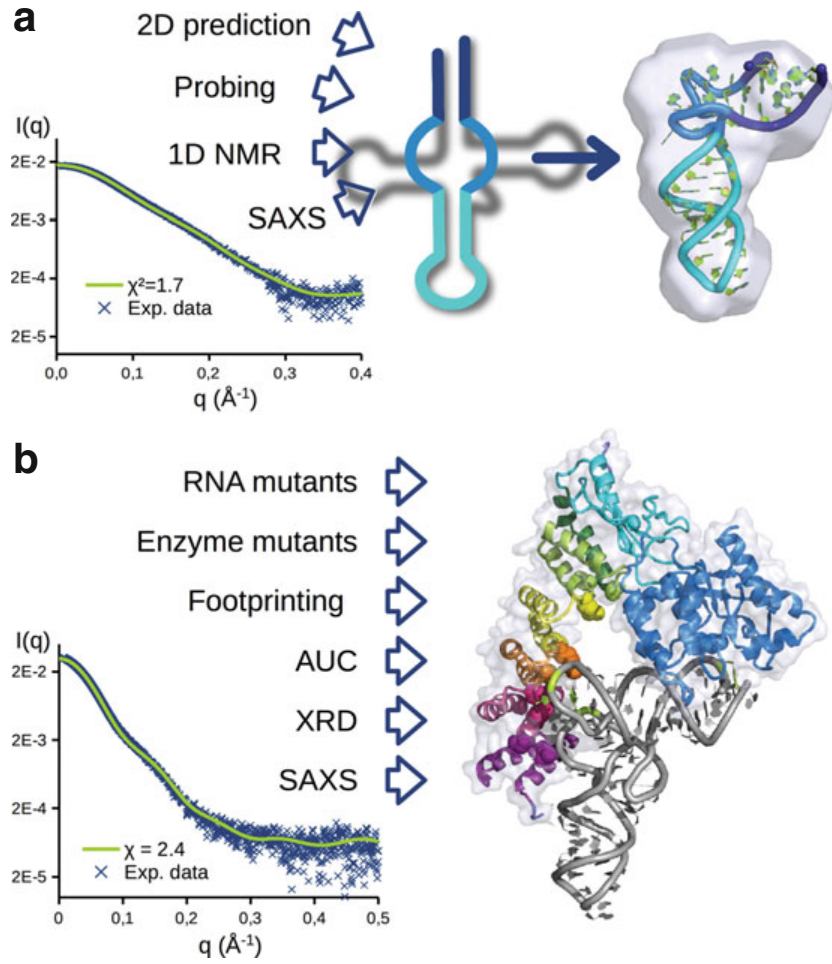


Fig. 9 SAXS, a powerful tool for integrative structural studies. (**a**) This example illustrates such an approach in the characterization of the world's smallest tRNAs, which are armless tRNAs found in the mitochondria of the nematode *Romanomermis culicivorax* [21]**.** A 2D structure prediction indicated that these tRNAs lacking both the D- and T-arms would fold as an elongated hairpin with a central bulge, as displayed at the center for $tRNA^{Leu}$ with a shadow suggesting a classical tRNA cloverleaf fold in background. The 2D fold was confirmed by in-line RNA probing; NMR only detected Watson-Crick interactions but no tertiary interactions. The SAXS analysis (the fit to the experimental data of the final model is displayed) helped fold the RNA in 3D and confirmed its boomerang shape, in spite of its severely shortened sequence (45 nt instead of 75–80). (**b**) The integrative approach is also very helpful to analyze RNA/protein complexes. For example, in the study of the protein-only RNase P (PRORP2) from *Arabidopsis thaliana* in association with a pre-tRNA substrate [57], the model of the tRNA 5′ maturation complex was built based on a set of distance constrains – interacting domains were determined by footprinting and by binding tests in ultracentrifugation (AUC) using RNA and enzyme mutants—and structural information on individual partners from X-ray diffraction (XRD), in combination with SAXS (the fit to the experimental data of the final model is displayed)

7. When an atomic model is available, calculate its theoretical SAXS curve with *CRYSOL* [36] and compare it to experimental data.
8. Inspect atomic models in *PyMOL*, modify them, and compare them to experimental data using the *SASPY* plugin [37] which provides a direct link to the *ATSAS* package and the possibility to iteratively adjust the model.
9. The *ATSAS* suite may be further exploited to model flexible regions of the molecule, to perform rigid-body or multicomponent fitting to improve the goodness of fit (i.e., minimize the χ^2) of the model to the data.
10. Use complementary biochemical and biophysical methods to derive additional structural constrains and help refine the model against SAXS data (Fig. 9).
11. Follow recommendations from [38] for the presentation of SAXS data in publications.

4 Notes

1. When designing an RNA molecule to be produced by T7 RNA polymerase, several constraints have to be considered, which are the overall length of the RNA, the nature of its 5′ sequence, and the desired homogeneity in 3′.

   For short RNAs (20–50 nt), the template can be reduced to a double-stranded promoter region followed by a single-stranded template sequence [39]. For medium RNA molecules (50–150 nt), synthetic genes can be assembled from oligonucleotides (cassettes) and cloned in a standard vector (Subheading 3.1.1). Finally, for longer sequences, transcription can be performed directly from the product of a PCR reaction. However, it is recommended to clone the DNA template into a plasmid in order to improve the transcription yield. Since the transcription termination occurs at the extremity of the DNA template, it is important to insert a restriction site, which (1) will be used to linearize the plasmid and (2) will generate the proper sequence at the end of the DNA template that matches the RNA sequence. For example, the *Bst*NI restriction enzyme is used to linearize the DNA template when producing tRNA transcripts with their characteristic 3′ CCA end. These considerations are detailed in [23].

   Despite the presence of a strong promoter, the T7 RNA polymerase has intrinsic limitations related to the nature of its promoter which starts at nucleotide −17 and extends to the nucleotide +6 of the sequence to be transcribed [40, 41]. RNA transcripts starting at position +1 with one or more G residues

are better transcribed by the T7 RNA polymerase and the presence of Cs and Ts in the +1 to +6 region of a transcript may be responsible for low transcription efficiency. However, the sequence can be modified to meet the T7 promoter wild-type sequence (GGGAGA or AGGGAG) if the 5′ extremity is not essential for RNA folding and function. Alternatively, a hammerhead ribozyme sequence can be introduced between the T7 RNA polymerase promoter and the sequence to be transcribed. Transcription of this construct gives rise to a "transzyme" molecule (Subheading 3.1.7). The autocatalytic activity of the ribozyme releases an RNA transcript starting with the proper 5′-OH nucleotide [42].

Importantly, the T7 RNA polymerase adds one or two nucleotides in a DNA-independent manner at the 3′ extremity of the transcript. The longer the transcript, the more difficult the elimination of this 3′ heterogeneity. For short RNA sequences (<100 nt), one can isolate the specific transcript from $n + 1$ and $n + 2$ contaminating transcripts by electrophoresis on thick denaturing gels (Subheading 3.2.1). However, for longer transcripts, the resolution of such gels does not allow an efficient purification. When the homogeneity of the 3′ end is essential, another autocatalytic ribozyme can be inserted in 3′ [43]. Longer transcripts can be purified either on native gels or by chromatography if the presence of one or two supplementary nucleotides at the 3′ extremity does not interfere with RNA folding or function (Subheadings 3.2.1 and 3.2.5).

2. Transcription conditions are adapted from a protocol described previously [23], with a $MgCl_2$ concentration of 30 mM instead of 22 mM and the final dilution and incubation time adapted to enhance the yield of autocatalytic cleavage. The latter may occur already during the transcription step. However, the yield of cleavage generally does not exceed 40% of the molecules. To increase the amount of cleavage, it is recommended to dilute the mix and incubate longer.

3. The percentage and concentration of acrylamide used in a purification gel has to be adapted to the size of the RNA of interest (Fig. 4) and the size of the wells should also be adapted to the amount of RNA to purify. For example, the product of an efficient 1 mL transcription reaction is phenol extracted, precipitated, and solubilized into 250 μL water and 250 μL loading dye and loaded in a $2 \times 10 \times 100$ mm^3 well (contains up to 2 mL sample). It is important to get a thin and intense band that can easily be visualized and separated from the contaminating products. When purifying RNA on a native gel, it is necessary to test the migration conditions on an aliquot of the transcription reaction before; indeed, the migration behavior of

an RNA in native conditions will depend essentially on the shape of the molecule and cannot be predicted.

When looking at the result of denaturing gel under UV, three scenarios are possible depending on the DNA template nature and the cloning strategy. Suppose one is purifying an *n*-nucleotide (nt)-long RNA: (1) DNA bands of 17 nt and *n* + 17 nt will be detected when using an oligonucleotide template and (2) a unique band at *n* + 17 nt when using a PCR template, while (3) a linearized plasmid should not enter efficiently the gel and appear as a smear at the top of the gel.

4. A chromatography on a strong anion exchanger can be used to purify the desired RNA transcript from unincorporated nucleotides, DNA template, T7 RNA polymerase, and abortive RNA fragments. The proposed protocol was designed for the Bio-Scale Q5 Anion Exchange Column (BioRad) and can be adapted to other anion exchange resins like Mono Q or DEAE sepharose [44]. At least two peaks should be observed at 254 nm along the elution gradient (Fig. 4). The first one contains free nucleotides that do not bind to the resin and the second one contains the RNA of interest. Additional peaks corresponding to short abortive transcripts may appear depending on the efficiency of the transcription reaction. Importantly, the main peak containing the full-length RNA transcript has to be analyzed by loading a sample of each collected fraction on an analytical denaturing polyacrylamide gel. The product should be identified as a strong band on the gel. Multiple additional bands indicate that the RNA transcript is not fully separated from the DNA template. Then a phenol: chloroform (pH 5.5) extraction step followed by ethanol precipitation at the end of the transcription reaction and before loading the column will remove both the DNA template and the T7 RNA polymerase from the sample.
5. As indicated in the introduction, it is essential to guarantee sample purity and monodispersity before performing a SAXS analysis because the scattered signal contains information about all entities present in the solution. If the solution contains even a low amount of large particles (oligomers, aggregates), structural interpretations may be compromised. A rigorous quality control should therefore include at least an analysis of the sample on native and denaturing polyacrylamide gels, a DLS [45], or MALS analysis and a SEC separation identical to that used during the SAXS experiment.
6. The choice of HPLC-SEC column depends on the size of the RNA and its protein partner(s) and should allow a good separation of the molecule of interest from aggregates or potential oligomers. Some proteins tend to adsorb on silica-based Bio SEC3 matrices and agarose-based columns may be a better

option. It is advisable to use the same column in the laboratory as on the beamline to optimize the elution conditions prior to synchrotron shifts. During the SEC-SAXS analysis, sample elution is followed by UV absorbance, but the signal at 260 nm is often saturated in the major peak due to the high amount of RNA injected to ensure a good SAXS signal. One can record the elution profile at a higher wavelength, such as 280 nm in the case of a complex with a protein, or 290 nm where RNA absorption is significantly lower to record the entire profile of the main peak (Fig. 6a).

7. Data processing is generally performed on the fly during data acquisition and softwares may vary from beamline to beamline but the principle remains the same. The most critical step is the subtraction of solvent and setup contribution (Fig. 1f). It is important to determine the so-called solvent curve (which also accounts for the background from the SAXS capillary) with a buffer solution that is as close as possible to the environment of the macromolecule. An advantage of the SEC-SAXS analysis over an analysis performed in batch mode is indeed to place the RNA of interest in a well-defined mobile phase (buffer solution), the signal of which is measured at the beginning of the elution after extensive equilibration of the column.
8. Examples of data analysis in Fig. 7 use the *ATSAS* suite (version 2.8.4) [24]. The first step is the Guinier analysis [2] at very small angles (Fig. 7b) to derive an estimate of the radius of gyration Rg, which characterizes the mass distribution in the particle, and the forward scattered intensity $I(0)$ if the curve is on absolute scale. $I(0)$ is proportional to particle concentration and to the square of their volume and can be used to determine the molecular mass of the particle. The Guinier plot should be as linear as possible with a low residual. Lack of linearity is an indication of either repulsive (downturn of the plot) or attractive (aggregation, upturn of the plot) interactions (*see* Fig. 3 in [26]). Repulsion may arise from the high negative charge of the RNA backbone at low counterion concentrations and can be fixed by increasing the salt content. Attractive interactions leading to aggregation may appear at high concentrations used for SAXS analysis or may be due to inappropriate buffer conditions causing particle instability. This can be solved by adjusting the ionic strength and the valence of the cations.
9. The PDDF or $p(r)$ is the indirect Fourier transform of the SAXS curve (Fig. 7c) and represents the distribution of distance between scattering elements in the particle. The $p(r)$ profile contains information on the shape and volume of the particle (*see* Fig. 3 in [27]) and provides an estimate of the maximum interatomic distance D_{max}. $I(0)$ and Rg can also be determined from the $p(r)$ taking into account the entire scattering profile.

This step is performed with *GNOM* [46] and the output file is used for ab initio reconstruction (*see* Subheading 3.5).

10. In the Porod representation ($I(q)q^2$ vs. q), the Porod invariant Q corresponds to the integrated area under the curve and is related to the particle (Porod) volume by $Vp = 2\pi \cdot I(0)/Q$ [3, 47, 48]. The Kratky representation ($I(q)q^4$ vs. q) or its dimensionless version ($I(q)/I(0) * (qRg)^2$ vs. qRg [49] detects particles with a compact core which display one peak and a rising tail at higher angles due to the presence of disordered and flexible regions attached to globular domains, while totally unfolded molecules display no peak but a continuously rising curve (Fig. 7d).

11. SAXS data provide an experimental basis for RNA modeling (Fig. 8). The first possibility is the ab initio reconstruction of the particle shape at low resolution ($q_{\max} < 8/Rg$ Å^{-1}) and does not require any a priori knowledge of the structure. It is carried out with *DAMMIF* [34] using dummy-atom ensembles which are adjusted to fit the $p(r)$ function and reproduce the SAXS curve with a low discrepancy measured by a χ^2. Another approach makes use of all available data (high resolution) and of *GASBOR* [35], which folds a chain of dummy residues to reproduce the SAXS profile. This procedure was developed for proteins made of amino acid chains, and owing to the difference in size and density with nucleotides, the number of dummy residues should be multiplied by three. In both approaches, one will seek models with χ^2 close to 1 and NSD of the series below 0.7–0.9.

    If a crystal, NMR, or EM structure is available, its theoretical SAXS profile can be readily computed with *CRYSOL* [36] and compared to the experimental SAXS profile, thus providing a means to adjust/improve 3D structure predictions using normal mode analysis, for instance [50]. When no structure is known, the 2D structure of an RNA may be inferred from sequence alignment and covariations and assembled into 3D models with programs like *ASSEMBLE* or *RNACOMPOSER* [51, 52]. Resulting models can then be evaluated with SAXS data. Other approaches based on coarse-grained representation or on RNA modular modeling have also been applied to RNA reconstruction filtered by SAXS information [53–55].

    Finally, while SAXS alone hardly provides an unambiguous and unique answer, it is a very powerful tool in integrative structural biology as illustrated in Fig. 9. The combination of complementary information from biochemical and biophysical methods is useful to sort a series of ab initio models and to support a final proposition.

Acknowledgments

The authors thank the team of the SWING beamline (SOLEIL synchrotron, Saint-Aubin, France), in particular Pierre Roblin and Javier Pérez, for beam time provision and assistance during data collection and processing, as well as José Teixeira (LLB, CEA Saclay, France) for his critical reading of the manuscript and fruitful discussion. The work described in this chapter was supported by the French Centre National de la Recherche Scientifique, the University of Strasbourg, the LabEx consortium "NetRNA" (ANR-10-LABX-0036_NETRNA).

References

1. Putnam CD, Hammel M, Hura GL et al (2007) X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution. Q Rev Biophys 40:191
2. Guinier A (1939) La diffraction des rayons X aux très petits angles: application à l'étude de phénomènes ultramicroscopiques. Ann Phys 11:161–237
3. Glatter O, Kratky O (1982) Small angle X-ray scattering. Academic Press Inc Ltd., London
4. Witz J (2003) 1964: The first model for the shape of a transfer RNA molecule. An account of an unpublished small-angle X-ray scattering study. Biochimie 85:1265–1268
5. Lake JA, Beeman WW (1967) Yeast transfer RNA: a small-angle X-ray study. Science 156:1371–1373
6. Ninio J, Luzzati V, Yaniv M (1972) Comparative small-angle X-ray scattering studies on unacylated, acylated and cross-linked Escherichia coli transfer RNAIVal. J Mol Biol 71:217–229
7. Svergun DI, Nierhaus KH (2000) A map of protein-rRNA distribution in the 70 S *Escherichia coli* ribosome. J Biol Chem 275:14432–14439
8. Chen Y, Pollack L (2016) SAXS studies of RNA: structures, dynamics, and interactions with partners. Wiley Interdiscip Rev RNA 7:512–526
9. Fang X-W, Yang X-J, Littrell K et al (2001) The Bacillus subtilis RNase P holoenzyme contains two RNase P RNA and two RNase P protein subunits. RNA 7:233–241
10. Russell R, Millett IS, Tate MW et al (2002) Rapid compaction during RNA folding. Proc Natl Acad Sci 99:4266–4271
11. Lipfert J, Ouellet J, Norman DG et al (2008) The complete vs ribozyme in solution studied by small-angle X-ray scattering. Structure 16:1357–1367
12. Hammond JA, Rambo RP, Filbin ME et al (2009) Comparison and functional implications of the 3D architectures of viral tRNA-like structures. RNA 15:294–307
13. Schmitt E, Panvert M, Lazennec-Schurdevin C et al (2012) Structure of the ternary initiation complex aIF2–GDPNP–methionylated initiator tRNA. Nat Struct Mol Biol 19:450–454
14. Jones CP, Cantara WA, Olson ED et al (2014) Small-angle X-ray scattering-derived structure of the HIV-1 5′ UTR reveals 3D tRNA mimicry. Proc Natl Acad Sci 111:3395–3400
15. Colussi TM, Costantino DA, Hammond JA et al (2014) The structural basis of transfer RNA mimicry and conformational plasticity by a viral RNA. Nature 511:366–369
16. Tuukkanen AT, Spilotros A, Svergun DI (2017) Progress in small-angle scattering from biological solutions at high-brilliance synchrotrons. IUCrJ 4:518–528
17. Mathew E, Mirza A, Menhart N (2004) Liquid-chromatography-coupled SAXS for accurate sizing of aggregating proteins. J Synchrotron Radiat 11:314–318
18. Watanabe Y, Inoko Y (2009) Size-exclusion chromatography combined with small-angle X-ray scattering optics. J Chromatogr A 1216:7461–7465
19. David G, Pérez J (2009) Combined sampler robot and high-performance liquid chroma-

tography: a fully automated system for biological small-angle X-ray scattering experiments at the Synchrotron SOLEIL SWING beamline. J Appl Crystallogr 42:892–900

20. Rambo RP, Tainer JA (2010) Improving small-angle X-ray scattering data for structural analyses of the RNA world. RNA 16:638–646
21. Jühling T, Duchardt-Ferner E, Bonin S et al (2018) Small but large enough: structural properties of armless mitochondrial tRNAs from the nematode Romanomermis culicivorax. Nucleic Acids Res 46:9170–9180
22. Baronti L, Karlsson H, Marušič M et al (2018) A guide to large-scale RNA sample preparation. Anal Bioanal Chem 410:3239–3252
23. Beckert B, Masquida B (2011) Synthesis of RNA by in vitro transcription. In: Nielsen H (ed) RNA. Humana Press, Totowa, pp 29–41
24. Franke D, Petoukhov MV, Konarev PV et al (2017) *ATSAS 2.8*: a comprehensive data analysis suite for small-angle scattering from macromolecular solutions. J Appl Crystallogr 50:1212–1225
25. Lipfert J, Doniach S (2007) Small-angle X-ray scattering from RNA, proteins, and protein complexes. Annu Rev Biophys Biomol Struct 36:307–327
26. Jacques DA, Trewhella J (2010) Small-angle scattering for structural biology—expanding the frontier while avoiding the pitfalls. Protein Sci 19:642–657
27. Mertens HDT, Svergun DI (2010) Structural characterization of proteins and complexes using small-angle X-ray solution scattering. J Struct Biol 172:128–141
28. Cantara WA, Olson ED, Musier-Forsyth K (2017) Analysis of RNA structure using small-angle X-ray scattering. Methods 113:46–55
29. Pollack L (2011) Time resolved SAXS and RNA folding. Biopolymers 95:543–549
30. Förster S, Apostol L, Bras W (2010) Scatter: software for the analysis of nano- and meso-scale small-angle scattering. J Appl Crystallogr 43:639–646
31. Hopkins JB, Gillilan RE, Skou S (2017) BioXTAS RAW: improvements to a free open-source program for small-angle X-ray scattering data reduction and analysis. J Appl Crystallogr 50:1545–1553
32. Fischer H, Neto MDO, Napolitano HB et al (2010) Determination of the molecular weight of proteins in solution from a single small-angle X-ray scattering measurement on a relative scale. J Appl Crystallogr 43:101–109
33. Rambo RP, Tainer JA (2013) Accurate assessment of mass, models and resolution by small-angle scattering. Nature 496:477–481
34. Franke D, Svergun DI (2009) *DAMMIF*, a program for rapid *ab-initio* shape determination in small-angle scattering. J Appl Crystallogr 42:342–346
35. Svergun DI, Petoukhov MV, Koch MH (2001) Determination of domain structure of proteins from X-ray solution scattering. Biophys J 80:2946–2953
36. Svergun D, Barberato C, Koch MHJ (1995) *CRYSOL*—a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. J Appl Crystallogr 28:768–773
37. Panjkovich A, Svergun DI (2016) SASpy: a PyMOL plugin for manipulation and refinement of hybrid models against small angle X-ray scattering data. Bioinformatics 32 (13):2062–2064
38. Trewhella J, Duff AP, Durand D et al (2017) 2017 publication guidelines for structural modelling of small-angle scattering data from biomolecules in solution: an update. Acta Crystallogr Sect Struct Biol 73:710–728
39. Francklyn C, Schimmel P (1989) Aminoacylation of RNA minihelices with alanine. Nature 337:478
40. Oakley JL, Strothkamp RE, Sarris AH et al (1979) T7 RNA polymerase: promoter structure and polymerase binding. Biochemistry 18:528–537
41. Rong M, He B, McAllister WT et al (1998) Promoter specificity determinants of T7 RNA polymerase. Proc Natl Acad Sci 95:515–519
42. Fechter P, Rudinger J, Giegé R, Théobald-Dietrich A (1998) Ribozyme processed tRNA transcripts with unfriendly internal promoter for T7 RNA polymerase: production and activity. FEBS Lett 436:99–103
43. Price SR, Ito N, Oubridge C et al (1995) Crystallization of RNA-protein complexes. I. Methods for the large-scale preparation of RNA suitable for crystallographic studies. J Mol Biol 249:398–408
44. Koubek J, Lin KF, Chen YR et al (2013) Strong anion-exchange fast performance liquid chromatography as a versatile tool for preparation and purification of RNA produced by in vitro transcription. RNA 19:1449–1459
45. Lorber B, Fischer F, Bailly M et al (2012) Protein analysis by dynamic light scattering: methods and techniques for students. Biochem Mol Biol Educ 40:372–382
46. Svergun DI (1992) Determination of the regularization parameter in indirect-transform methods using perceptual criteria. J Appl Crystallogr 25:495–503

47. Rambo RP, Tainer JA (2011) Characterizing flexible and intrinsically unstructured biological macromolecules by SAS using the Porod-Debye law. Biopolymers 95:559–571
48. Reyes FE, Schwartz CR, Tainer JA et al (2014) Methods for using new conceptual tools and parameters to assess RNA structure by small-angle X-ray scattering. Methods Enzymol 549:235–263
49. Durand D, Vivès C, Cannella D et al (2010) NADPH oxidase activator p67phox behaves in solution as a multidomain protein with semi-flexible linkers. J Struct Biol 169:45–53
50. Suhre K, Sanejouand Y-H (2004) ElNemo: a normal mode web server for protein movement analysis and the generation of templates for molecular replacement. Nucleic Acids Res 32: W610–W614
51. Jossinet F, Ludwig TE, Westhof E (2010) Assemble: an interactive graphical tool to analyze and build RNA architectures at the 2D and 3D levels. Bioinformatics 26:2057–2059
52. Biesiada M, Pachulska-Wieczorek K, Adamiak RW et al (2016) RNAComposer and RNA 3D structure prediction for nanotechnology. Methods 103:120–127
53. Yang S, Parisien M, Major F et al (2010) RNA structure determination using SAXS data. J Phys Chem B 114:10039–10048
54. Gajda MJ, Martinez Zapien D, Uchikawa E et al (2013) Modeling the structure of RNA molecules with small-angle X-ray scattering data. PLoS One 8:e78007
55. Bhandari YR, Fan L, Fang X et al (2017) Topological structure determination of RNA using small-angle X-ray scattering. J Mol Biol 429:3635–3649
56. Shi H, Moore PB (2000) The crystal structure of yeast phenylalanine tRNA at 1.93 A resolution: a classic structure revisited. RNA 6:1091–1105
57. Pinker F, Schelcher C, Fernandez-Millan P et al (2017) Biophysical analysis of *Arabidopsis* protein-only RNase P alone and in complex with tRNA provides a refined model of tRNA binding. J Biol Chem 292:13904–13913

5. Posters and presentations

5.1. New microfluidic design for crystallization of biomolecules and their in situ analysis by serial crystallography

This poster was presented at the following events:

- Doctoral School Days (ED days) 2019, 21/03/2019, Strasbourg
- UPR 9002 meeting 2019, 17.05.2019, Strasbourg
- FEBS advanced course:

5.2. Monitoring the Production of High Diffraction-Quality Crystals in Real Time Using In Situ Dynamic Light Scattering

This poster was presented at the following events:

- École de Biologie Structurale Intégrative, 18-25.06.2021, Oléron (France)
- FEBS advanced course: Advanced methods in macromolecular crystallization IX, 09-14.08.2021, České Budějovice (Czech Republic). It won the IUCrJ poster prize.
- International Union of Crystallography Congress, 14-22.08.2021, Prague (Czech Republic)

5.3. Discovering the world of RNA: meeting and sharing science with the general public

This poster was presented at the 2nd Integrative Structural Biology meeting (29.11-03.12.2021) in Paris-Saclay (France).

New microfluidic design for crystallization of biomolecules and their in situ analysis by serial crystallography

Raphaël de Wijn[1], Oliver Hennig[2], Jennifer Roche[3], Sylvain Engilberge[4], Kévin Rollet[1], Pablo Fernandez-Millan[1], Karl Brillet[1], Heike Betat[2], Mario Mörl[2], Alain Roussel[3], Eric Girard[4], Vincent Oliéric[5], Gavin C. Fox[6], Christoph Mueller-Dieckmann, Bernard Lorber[1], Claude Sauter[1]

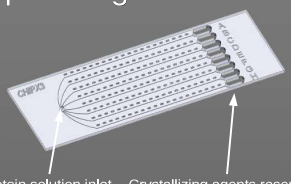
[1]Institut de Biologie Moléculaire et Cellulaire, UPR 9002 CNRS, Strasbourg, France
[2]Biochemistry and Molecular Biology, Inst. for Biochemistry, Leipzig University, Leipzig, Germany
[3]Architecture et Fonction des Macromolécules Biologiques, UMR 7257 CNRS, Marseille, France
[4]Institut de Biologie Structurale, UMR 5075 CEA-CNRS-UGA, Grenoble, France
[5]Paul Scherrer Institute, Swiss Light Source, ligne X06DA, Villigen, Switzerland
[6]Synchrotron SOLEIL, ligne PROXIMA 2A, Gif-sur-Yvette, France
[7]Structural Biology, European Synchrotron Radiation Facility, 38043 Grenoble France

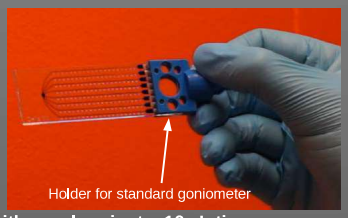
k.rollet@ibmc-cnrs.unistra.fr

Introduction

For about fifteen years, microfluidics has opened up new possibilities and brought many benefits for the crystallization of biomolecules. Indeed, microfluidic systems facilitate the manipulation of nanovolumes of sample solutions, as well as extreme miniaturization and parallelization of crystallization assays. In addition, they provide a convection-free environment that favors the growth of high-quality crystals [1]. We present a new multifunctional microchip design [2] combining: 1) search and optimization of crystallization conditions of biomolecules by the counter-diffusion method, 2) crystal identification by fluorescence microscopy, 3) microcrystalline seeding, 4) derivatization of crystals by substrate soaking, and 5) in situ crystal analysis at room temperature. This design was tested on a large panel of biomolecules including soluble or membrane proteins. Several crystal structures have been solved in situ by the serial crystallography approach under synchrotron radiation. Some are presented in this poster.

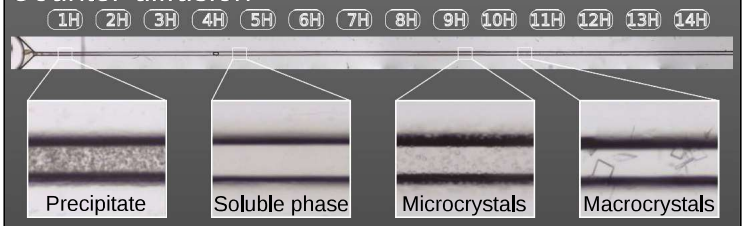
ChipX design





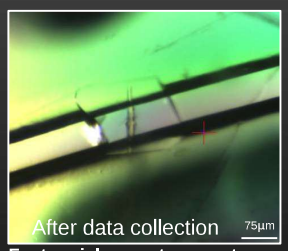
This design allows easy sample loading with regular pipets, 10µL tips and convenient use on conventional beamlines with a dedicated holder

Counter-diffusion

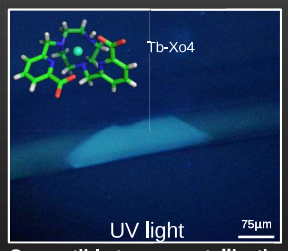


ChipX were designed to implement the counter-diffusion method which is well known for its efficiency based on a self-optimizing process: the crystallizing agent diffuses through an elongated crystallization chamber (the microfluidic channel), creates a gradient of concentration and each assay samples a broad range of supersaturation states. The longer the channel, the broader the screening.

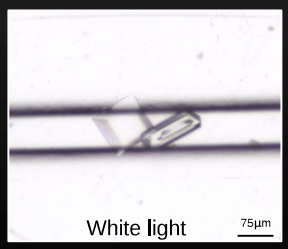
Advanced crystallogenesis strategies

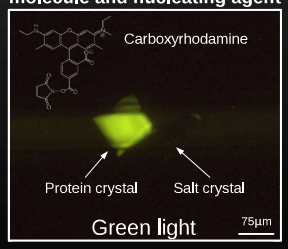


Fast serial room-temperature data collection



Compatible to co-crystallization with Tb-Xo4 as phasing agent, luminescent molecule and nucleating agent

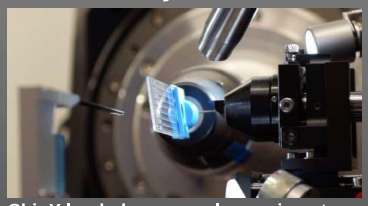




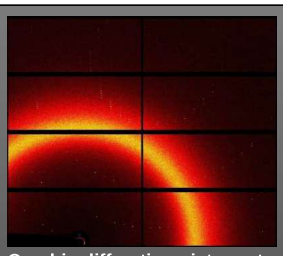
Compatible with microcrystalline seeding
- **→ seeds added to the protein solution before ChipX loading and soaking**
- **→ molecules can be added to the reservoir after crystal growth**

Compatible with fluo-labelling for visualization

In-situ analysis



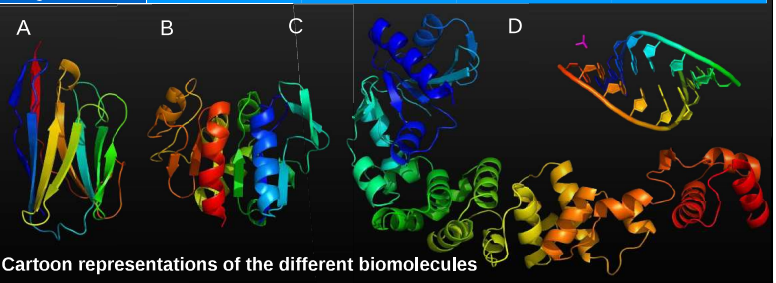
ChipX loaded on a regular goniometer head at the SLS – PXIII beamline



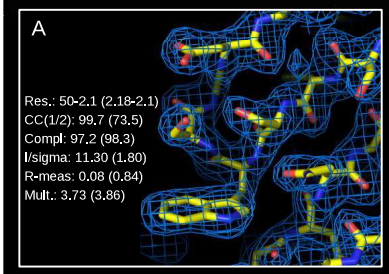
On-chip diffraction picture at room-temperature at the SLS – PXIII beamline

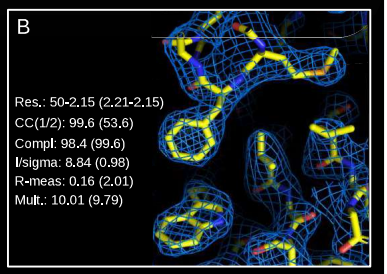
Table: On-chip room-temperature collections report

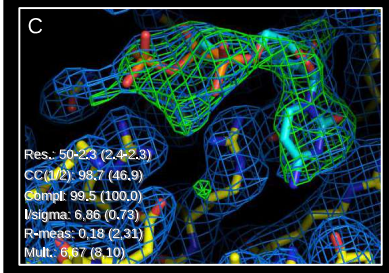
Protein and organism	A Camelid Nanobody	B Protease I (*Pyrococcus horikhoshii*)	C CCA-adding enzyme (*P. halocryophilus*) + ligand	D RNA duplex (Synthetic)
Beamline	SOLEIL – Proxima 2-A	SLS - PXIII	SLS - PXII	SLS - PXIII
Resolution	2.1 Å	2.15 Å	2.3 Å	1.55 Å
Space group	$P4_32_12$	$P4_12_12$	$P4_32_12$	R3
Crystals collected	9	12	14	3
Crystals selected	1	8	5	3
Oscillation range	0.1°	0.2 °	0.2°	2-3°
Images selected	500	1300	540	80

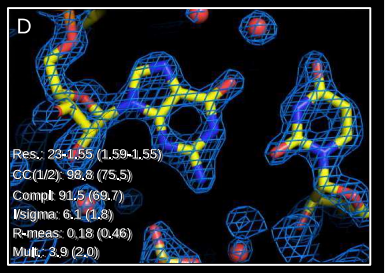


Cartoon representations of the different biomolecules









Electron density clouds snapshots of different molecules collected at room-temperature and after molecular replacement

In brief

A chip design dedicated to advanced crystallogenesis strategies:
- Crystallization by counter-diffusion
- Easy crystal visualization by fluo-labelling
- Easy crystal seeding and soaking
- Fast serial analysis at room-temperature

User-friendly, no extra equipment needed!

References

[1] Sauter C., Dhouib K., Lorber B. From macrofluidics to microfluidics in the crystallization of biological macromolecules. *Cryst. Growth Des.*, **2007**, 7, 2247-50.

[2] de Wijn R., Hennig O., Roche J., Engilberge S., Rollet K., Fernandez-Millan P., Brillet K., Betat H., Mörl M., Roussel A., Girard E., Mueller-Dieckmann C., Fox G.C., Olieric V., Gavira JA., Lorber B. and Sauter C. A simple and versatile microfluidic device for efficient biomacromolecule crystallization and structural analysis by serial crystallography, IUCrJ, **2019**, 6(3); 454-464.

Monitoring the Production of High Diffraction-Quality Crystals in Real Time Using In Situ Dynamic Light Scattering

Kévin Rollet[1,2], Raphaël de Wijn[1], Sylvain Engilberge[3], Alastair G. McEwen[4], Oliver Hennig[2], Heike Betat[2], Mario Mörl[2], François Riobé[5], Olivier Maury[5], Eric Girard[3], Philippe Bénas[1], Bernard Lorber[1], Claude Sauter[1]

[1]Institut de Biologie Moléculaire et Cellulaire, UPR 9002 CNRS, Strasbourg, France
[2]Biochemistry and Molecular Biology, Inst. for Biochemistry, Leipzig, Germany
[3]Institut de Biologie Structurale, UMR 5075 CEA-CNRS-UGA, Grenoble, France
[4]Plateforme de Biologie et Génomique Structurale.Institut de Génétique et de Biologie Moléculaire et Cellulaire, CNRS UMR 7104, INSERM U 1258, Illkirch, France
[5]Univ. Lyon, ENS Lyon, CNRS-UMR 5182, Lyon, France

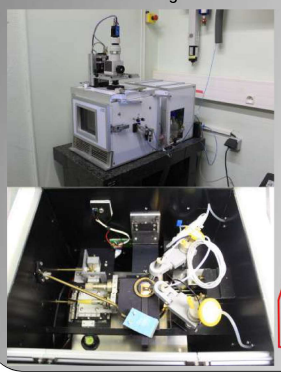
k.rollet@ibmc-cnrs.unistra.fr

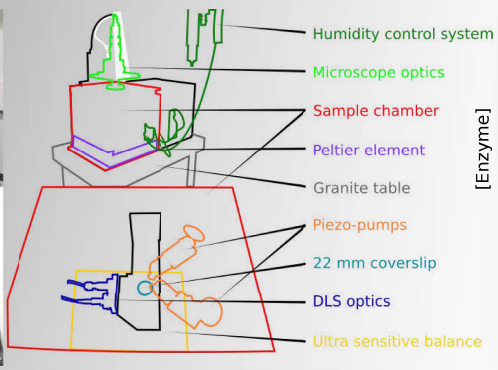
Introduction

The reproducible preparation of well-diffracting crystals is a prerequisite for every structural study based on crystallography. An instrument called XtalController [1] has recently been designed that allows the monitoring of crystallization assays using dynamic light scattering and microscopy, and integrates piezo pumps to alter the composition of the mother liquor during the experiment. We have applied this technology to study the crystallization of two enzymes, the CCA-adding enzyme of the psychrophilic bacterium *Planococcus halocryophilus*, and the lysozyme from hen egg white in the presence of a synthetic nucleation agent. We were able to (i) detect early nucleation events and (ii) drive the crystallization system (through cycles of dissolution/crystallization) toward growth conditions yielding crystals with excellent diffraction properties. This approach opens a way to the rational production of samples for crystallography, ranging from nanocrystals for electron diffraction, microcrystals for serial or conventional X-ray diffraction, to larger crystals for neutron diffraction.

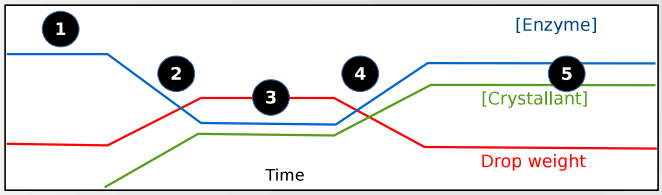
Xtal Controller XC900

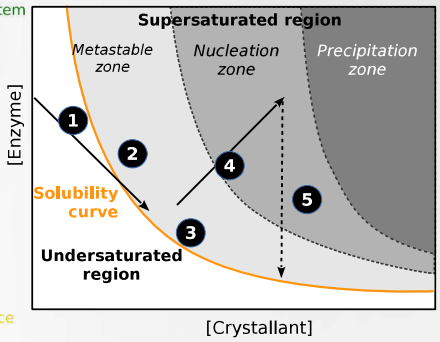
The XC900 consists of a crystallization chamber with precise temperature and humidity control. The composition of an initial drop (5–10 µl) of the target biomolecule solution can be modified by the injection of various solutions (e.g. water, buffer or crystallant) using two piezo injectors spraying 70 pL droplets. The sample can also be concentrated by evaporation of volatile chemicals and its composition is continuously calculated from its weight recorded to ±1 µg by an ultra-sensitive balance. The instrument provides diagnostic means to track the drop content along the experiment; the early occurrence of association events leading to nucleation, as well as nanocrystals can be detected in real time by dynamic light scattering (DLS) and the growth of crystals by video microscopy as soon as they reach a size exceeding a few microns.





Navigating in the phase diagram





1: incubation of the enzyme solution at constant concentration

2: addition of crystallant

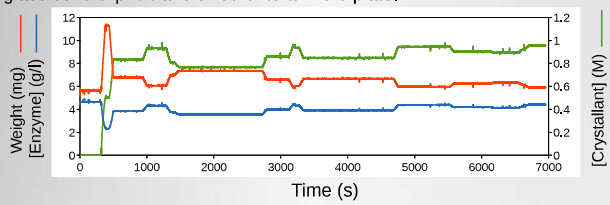
3: incubation, constant drop in weight to keep biomolecule and crystallant concentrations constant

4: controlled evaporation of the drop leading to increased concentrations of biomolecule and crystallant to reach the nucleation zone

5: incubation until crystals start to grow, consume part of the soluble enzyme stock and bring the system back to equilibrium on the solubility curve

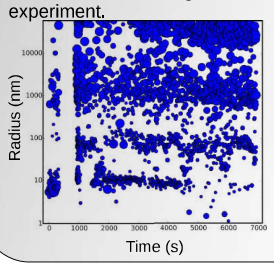
Case 1 From nano to macro crystals

A tRNA processing enzyme (CCA-adding enzyme) from psychrophilic bacteria *Planococcus halocryophilus* is crystallized after addition of crystallant (1.0M diammonium phosphate, 0.1M ammonium acetate pH4.5) then by playing on concentration/dilution steps. Too high concentration leads to a reversible precipitate. At the end of the experiment the droplet stays in-between precipitate/clear phase and the glass coverslip is transferred onto a linbro plate.

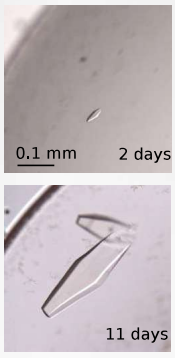


Radius plot

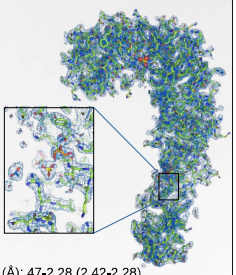
DLS signal is plotted against time and allows to track appearance or disappearance of particle populations with a size ranging from 10 to >10.000 nm during whole experiment.



Transferred droplet



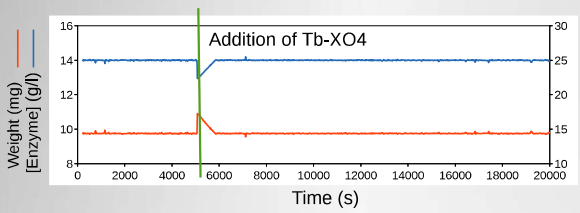
Final structure

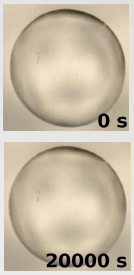


Res.(Å): 47-2.28 (2.42-2.28)
CC(1/2)(%): 99.9 (76.0)
Compl.(%): 99.5 (98.5)
I/sigma: 15.5 (1.5) ; Mult.: 6.9 (7.0)
R-meas: 0.09 (1.26) ; PDB ID : 6TVZ

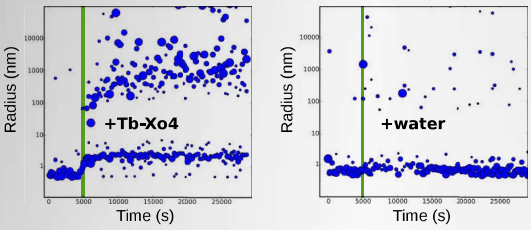
Case 2 Triggering nucleation with Tb-XO4

The lanthanide based crystallophore Tb-XO4 (Polyvalan SAS, Lyon, France) is known for its phasing and nucleating properties [2]. Here, we used hen egg-white lysozyme (HEWL) dissolved in pure water to demonstrate lysozyme self-association after addition of 1µL of Tb-XO4 at 100mM. In the control experiment, 1µL of pure water is injected instead of Tb-XO4.



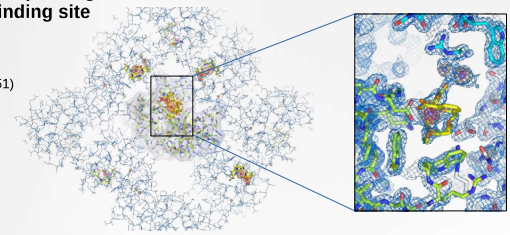


Radius plot showing immediate nucleation effect of Tb-XO4



Seikagaku HELW

View of tetragonal crystal packing of HEWL and zoom on Tb-XO4 binding site



Res.(Å): 35-1.51 (1.60-1.51)
CC(1/2)(%): 100.0 (78.1)
Compl.(%): 99.2 (96.5)
I/sigma: 29.7 (1.9)
R-meas: 0.04 (0.73)
Mult.: 8.4 (4.8)
PDB ID: 6TVY

In brief

The XC900 is an innovative system allowing:

- a full navigation into the phase diagram
- the monitoring and tracking of nucleation events in real time
- the crystallization of various biological systems
- the refinement of crystal growth conditions
- the design of growth protocols for nano/micro/macro-crystals

References

[1] de Wijn, R., Rollet, K., Engilberge, S., McEwen, A., Hennig, O., Betat, H., Mörl, M., Riobé, F., Maury, O., Girard, E., Bénas, P., Lorber, B., & Sauter, C. (2020). Monitoring the Production of High Diffraction-Quality Crystals of Two Enzymes in Real Time Using In Situ Dynamic Light Scattering. Crystals, 10(2), 65.
[2]Engilberge, S., Riobé, F., Di Pietro, S., Lassalle, L., Coquelle, N., Arnaud, C. A., Pitrat, D., Mulatier, J. C., Madern, D., Breyton, C., Maury, O., & Girard, E. (2017). Crystallophore: A versatile lanthanide complex for protein crystallography combining nucleating effects, phasing properties, and luminescence. Chemical Science, 8(9), 5909–5917.

IUCr Journals

The International Union of Crystallography (IUCr) is pleased to award an

IUCr Journals Poster Prize

at the

FEBS 2021 Advanced Course "Advanced Methods in Macromolecular Crystallization IX"

(Ceske Budejovice, Czech Republic; 9–14 August 2021)

to

Kévin Rollet

Université de Strasbourg, Leipzig University

for the presentation

"P13 MONITORING THE CRYSTALLIZATION OF TWO ENZYMES IN REAL TIME BY DYNAMIC LIGHT-SCATTERING."

S Lidin

Sven Lidin
IUCr President

Andrew Allen

Andrew Allen
Editor-in-Chief, IUCr Journals

Discovering the world of RNA: meeting and sharing science with the general public

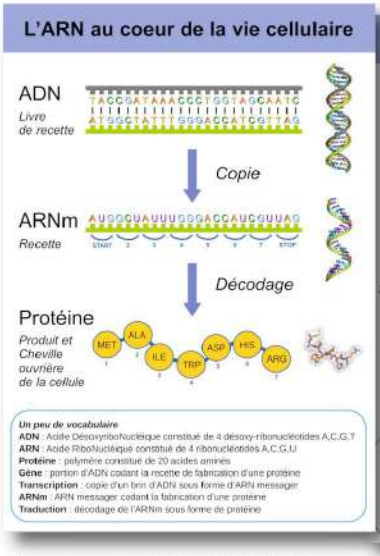
Morgan Baldaccini, Stéphanie Baudrey, Philippe Bénas, Léna Coudray, Alexis Duvergé, Agnès Gaudry, Lucas Herrgott, Franck Martin, Noémie Mercier, Mélanie Messmer, Julian Nommé, Charline Pasquier, Kévin Rollet, Claude Sauter & Monika Vilimova

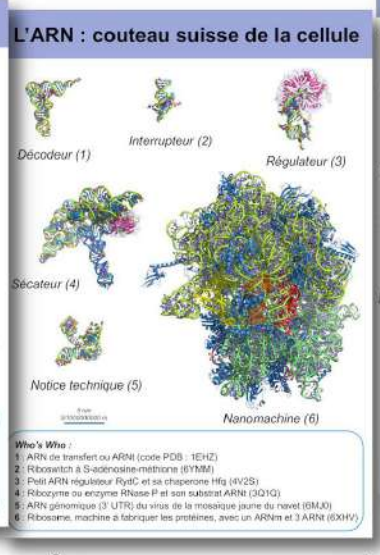
Architecture et Réactivité de l'ARN – IBMC – CNRS – Université de Strasbourg

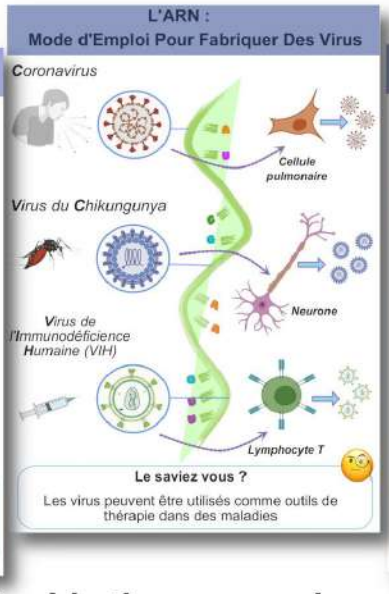
RNA, did you say RNA? Until two years ago and the occurrence of the coronavirus pandemic, few people had heard of RNA, whereas the DNA double helix image is well anchored in the collective unconscious. But in 2021 the situation has changed since everyone has heard about RNA viruses and RNA vaccines. This is why our laboratory, the well-named ARN for architecture and reactivity of RNA, decided to surf on the news and communicate on the world of RNA during the national science festival (Fête de la Science, October 2021). Various communication tools – posters, 3D models, movies, experiments – were deployed to answer questions of the greatest number of people.

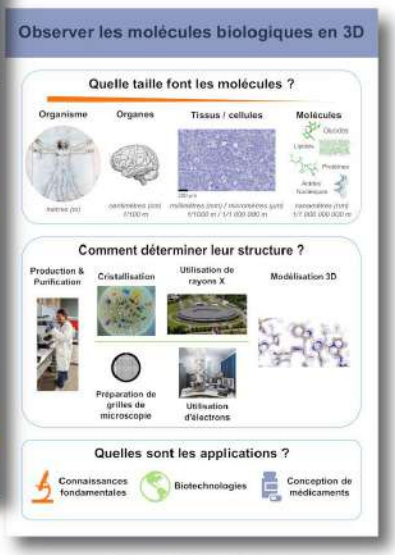
fête de la Science 2021

Setting the background

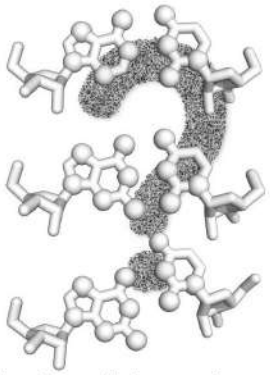






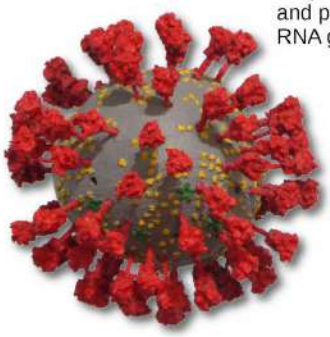


Playing with RNA basepairs



Building RNA Watson-Crick basepairs thanks to magnetic H-bonds of the RNA basepair puzzle [1]

Inside the coronavirus



3D printed model of SARS-CoV2 [2] and presentation of the virus and its RNA genome in virtual reality

Wow Amazing!!!!

Exploring biomolecules

3D screen and a series of PyMOL scripts [3]:
- to explore the architecture of proteins and nucleic acids,
- to recognize and build side chains in an electron density map at atomic resolution

Growing protein crystals

Recipe for rapid crystallisation of hen egg white lysozyme:

Low cost, fast, easy, and beautiful!

Useful links

[1] https://pymolwiki.org/index.php/Scene
[2] https://jkondo.wixsite.com/basepairpuzzle
[3] https://insidecorona.net/3d-print-corona/

In brief

- Structural biology is easy to share!
- Very motivating for young researchers
- Considerable interest from the public

Université de Strasbourg

École Doctorale des Sciences de la Vie et de la Santé STRASBOURG

Kévin ROLLET

De l'analyse fonctionnelle d'ARNt nucléotidyltransférases vers leur caractérisation par cristallographie sérielle sous rayonnement X

Résumé en français

La dernière décennie a été marquée par d'importants développements dans le domaine de la biologie structurale, en particulier dans la cristallographie sérielle aux rayons X résolue en temps (TR-SSX/SFX). Sous l'impulsion de la technologie des lasers à électrons libres (XFEL), les synchrotrons de nouvelle génération ont commencé à développer des lignes de lumière de biocristallographie dédiées à ces approches sérielles. L'objectif de cette thèse, menée dans le cadre d'une co-tutelle entre l'Université de Strasbourg (France) et l'Université de Leipzig (Allemagne), était d'appliquer ces approches à la caractérisation fonctionnelle et structurale de trois *CCA-adding enzymes*, qui polymérisent l'extrémité 3'CCA universellement conservée des ARN de transfert. En combinant biochimie et méthodes de cristallisation avancées, il s'agissait (I) de déterminer les paramètres cinétiques des réactions catalysées, (II) de trouver des conditions de reproductibles de production de cristaux pour les analyses standard et SSX et (III) de résoudre la structure de ces enzymes dans leur forme apo ou en complexe avec leurs ligands. La caractérisation biochimique réalisée à Leipzig a permis de déterminer l'affinité de liaison de divers substrats ainsi que les propriétés cinétiques de nos candidats. En mettant à profit la technologie Xtal Controller, il a été possible de suivre les événements de nucléation précoces et contrôler entièrement la cristallisation de certaines enzymes. En outre, la puce microfluidique ChipX, conçue au laboratoire à Strasbourg, a permis l'acquisition *in situ* de données sérielles sous rayonnement synchrotron à température ambiante. Dans leur ensemble, les résultats de ce travail constituent une base solide pour pousser le projet vers l'établissement d'un " film moléculaire " en temps réel de l'addition de la queue 3'CCA des ARNt par les nouvelles stratégies TR-SFX/SSX.

English abstract

The last decade has seen important developments in the field of structural biology, particularly in time-resolved serial X-ray crystallography (TR-SSX/SFX). Driven by free electron laser (XFEL) technology, new generation synchrotrons have started to develop biocrystallography beamlines dedicated to these serial approaches. The objective of this thesis, conducted in the framework of a co-supervision between the University of Strasbourg (France) and the University of Leipzig (Germany), was to apply these approaches to the functional and structural characterization of three CCA-adding enzymes, which polymerize the universally conserved 3'CCA end of transfer RNAs. By combining biochemistry and advanced crystallization methods, the aim was (I) to determine the kinetic parameters of the catalyzed reactions, (II) to find reproducible crystal production conditions for standard and SSX analyses and (III) to resolve the structure of these enzymes in their apo form or in complex with their ligands. The biochemical characterization performed in Leipzig allowed us to determine the binding affinity of various substrates as well as the kinetic properties of our candidates. By using the Xtal Controller technology, it was possible to monitor early nucleation events and fully control the crystallization of some enzymes. In addition, the ChipX microfluidic chip, designed in the laboratory in Strasbourg, allowed the *in situ* acquisition of serial data under synchrotron radiation at room temperature. Taken together, the results of this work provide a solid basis to push the project towards the establishment of a real-time "molecular movie" of tRNA 3'CCA tail addition by the new TR-SFX/SSX strategies.

UNIVERSITÄT LEIPZIG

Université franco-allemande Deutsch-Französische Hochschule