

**Improvement of Protein Crystal Diffraction Using Post-
Crystallization Methods: Infrared Laser Radiation
Controls Crystal Order**

Dissertation
zur Erlangung des Doktorgrades
an der Fakultät für Mathematik, Informatik und Naturwissenschaften
Fachbereich Physik
der
Universität Hamburg

vorgelegt von
Breyan Helmuth Ross Díaz

Hamburg

2018

Gutachter/innen der Dissertation:

Prof. Dr. Henry Chapman
Prof. Dr. Robert Huber

Zusammensetzung der Prüfungskommission:

Prof. Dr. Arwen Pearson
Prof. Dr. Henry Chapman
Prof. Dr. Robert Huber
Dr. Ruth Geiss-Friedlander

Vorsitzende/r der Prüfungskommission:

Prof. Dr. Daniela Pfannkuche

Datum der Disputation:

02.04.2019

Vorsitzender Fach-Promotionsausschusses PHSIK:

Prof. Dr. Wolfgang Hansen

Leiter des Fachbereichs PHYSIK:

Prof. Dr. Michael Potthoff

Dekan der Fakultät MIN:

Prof. Dr. Heinrich Graener

Acknowledgment

I would like to deeply thank Prof. Robert Huber, for always keeping a supportive and critical eye on my research. This contributed enormously to forge my scientific skills. In the same line, I am very grateful to Prof. Henry Chapman for his valuable advice and exciting collaboration.

I extend the acknowledgement to all people from whom I received support at Proteros Biostructures. In particular to Dr. Reiner Kiefersauer and Dr. Stephan Krapp, for those fruitful discussions about science and crystallography. Similarly, I thank Dr. Ruth Geiss-Friedlander for the sympathy and great scientific advice, which enabled me to develop an integrative thinking.

Finally, I thank all those people which in one way or the other helped in this journey. To my family, which although far, found always ways to keep me close, and to those who help me with manuscript corrections, Marcelo and Nena.

Munich, 26.11.18

Abstract

Protein crystal diffraction is intimately related to crystal order. Thus, the more ordered the crystal the better it diffracts. This feature is fundamental to solve the molecular structure of macromolecules, used by scientists to reveal functions of proteins and ultimately to treat diseases. However, protein crystals are hard to obtain, and most importantly, they often do not diffract sufficiently to provide relevant structural data. This challenging scenario lays the basis of this project, which aims at improving protein crystal order by means of post-crystallization techniques.

Protein crystals consist of two major components, protein and water. The equilibrium between these components is a hallmark of this study. To control crystal order, intra-crystalline water, which is mostly in an amorphous state, can be restructured without affecting protein conformation, by either dehydration, annealing or water exchange. The new crystal state, or phase, might present improved diffraction. The Free Mounting Laser System (FML) is an experimental setup which targets almost exclusively intra-crystalline water to induce higher crystal order. Therefore, it is proposed as a new non-contact method able to successfully improve diffraction quality of protein crystals.

In the present study infrared radiation was focused on protein crystals by FML to increase their temperature, thereby dehydrating them, inducing annealing, or exchanging water by stabilizer chemicals. The establishment of these methods allowed us to understand the responses of diverse model crystal systems to IR radiation. They offer significant advantages over previous systems, mostly related to speed of dehydration and mild non-contact crystal treatments. Moreover, the application of these techniques made it possible to solve the crystallographic structure of two new drug targets, the dipeptidyl peptidases 8 (DPP8) and 9 (DPP9). A detailed analysis of their molecular structure revealed a new binding mechanism, different from the related family member DPP4. DPP8 and DPP9 undergo a disorder-order transition of a 26 aa segment (R-segment). This segment has an arginine, which is responsible for substrate binding. The disordered “open” conformation is suggested to be the interaction target of SUMO1 protein. This non-covalent interaction could be disturbed by DPP8/9 inhibitors in biochemical and biophysical assays. We therefore propose this interaction to be a key checkpoint with physiological relevance in immune response and cancer.

Altogether, this study highlights the relevance of innovative protein crystallography methods. The development of new techniques is quintessential to contribute for future structural contributions to drug development and discovery of new drug targets.

Contents

List of Figures	viii
List of Tables.....	x
1. Introduction	1
1.1 Part I: Crystal order and diffraction improvement	1
1.1.1 Theory	1
1.1.1.1 Proteins and protein crystals.....	1
1.1.1.2 Phase transitions and crystals	1
1.1.1.3 Relative humidity and dew point.....	4
1.1.1.4 Infrared radiation and water absorption.....	4
1.1.1.5 Crystallography, phasing and refinement.....	5
1.1.2 Post-crystallization diffraction improvement	6
1.1.2.1 Protein crystal water content control	7
1.1.2.2 Protein crystal annealing	8
1.1.2.3 Protein crystal chemical modification	9
1.1.3 Aim and objectives of the study: infrared radiation modifies protein crystal order	11
1.2 Part II, study case: DPP8, DPP9 and SUMO1	12
1.2.1 Physiological context.....	12
1.2.1.1 Proteases	12
1.2.1.2 DPP4, DPP8 and DPP9: physiological context	12
1.2.1.3 DPP4, DPP8 and DPP9: structural relationship	13
1.2.1.4 SUMO1	14
1.2.1.5 DPP8 and DPP9 interact with SUMO1	15
1.2.2 Aim and objectives of the study: structural and functional characterization of DPP8 and DPP9 interaction with SUMO1.....	16
2. Materials and Methods	17
2.1 Materials.....	17
2.1.1 Chemicals: Protein constructs and crystallization conditions	17
2.1.2 DPP4, DPP8 and DPP9 inhibitors	18
2.2 Methods.....	18
2.2.1 Post-crystallization improvement methods	18
2.2.1.1 Free Mounting Laser System (FML) hardware setup	18
2.2.1.2 Humidity determination.....	19
2.2.1.3 Laser power and optics.....	20
2.2.1.4 FML dehydration experiments	22
2.2.1.5 Dropping experiments	23
2.2.1.6 FML annealing experiments.....	24

2.2.2	Molecular and structure determination methods	25
2.2.2.1	Molecular cloning	25
2.2.2.2	Protein expression and purification	25
2.2.2.3	Crystallization and x-ray measurements.....	26
2.2.2.4	Structure solution and data analysis	27
2.2.3	Functional and biophysical methods	27
2.2.3.1	SUMO1 oligomer formation	27
2.2.3.2	Surface plasmon resonance	27
2.2.3.3	DPP8 and DPP9 kinetics measurements	28
2.2.3.4	Pull-down assays.....	29
2.2.3.5	Molecular dynamics.....	29
3.	Results	30
3.1	Crystal order and post-crystallization treatments	30
3.1.1	Response of protein crystals to dehydration.....	30
3.1.2	Laser IR-radiation water absorption.....	31
3.1.3	Laser IR radiation increases temperature of protein crystals.....	32
3.1.4	Dehydration response of protein crystals to IR treatments.....	33
3.1.5	Crystal annealing.....	35
3.1.5.1	Crystal annealing under oil.....	35
3.1.5.2	Crystal annealing at 100 K: ice phase transition	36
3.1.5.3	DPP9 crystal annealing at 100 K using FML	37
3.1.6	Chemical modifications.....	38
3.1.6.1	DPP8 and DPP9 chemical treatment	38
3.1.6.2	Lysozyme crystal structures after FMS and FML treatments	40
3.2	Structure and mechanism of dipeptidyl peptidases 8 and 9.....	42
3.2.1	DPP8 and DPP9 structure solution as homodimer and monomer	42
3.2.2	DPP8 and DPP9 β -propeller domain.....	46
3.2.3	DPP8 and DPP9 α/β hydrolase domain	48
3.2.4	Inhibitor binding exposes active site architecture of DPP8 and DPP9.....	49
3.2.5	Structure of DPP8 interacting with Val-BoroPro	52
3.2.6	DPP8 and DPP9 β -propeller tunnel, active site cavity, and side opening.....	54
3.2.7	Inhibitor binding reveals allosteric and cooperative inhibition	54
3.3	DPP8 and DPP9 interact non-covalently with SUMO1	57
3.3.1	SUMO1 oligomers purification.....	57
3.3.2	DPP9 and SUMO1 complex formation.....	58
3.3.3	Binding of crosslinked SUMO1 to DPP9.....	59
3.3.4	SUMO1 does not affect the activity of DPP8 or DPP9.....	62

4.	Discussion	63
4.1	Post-crystallization diffraction improvement: dehydration.....	64
4.2	Post-crystallization diffraction improvement: annealing	65
4.3	Post-crystallization diffraction improvement: chemical modification	66
4.4	DPP8 and DPP9 as successful targets of post-crystallization treatments.....	66
4.4.1	DPP8 and DPP9 molecular structure.....	66
4.4.2	DPP8 and DPP9 display allosteric and cooperative binding.....	68
4.4.3	Molecular dynamics simulation	68
4.4.4	Bacterial DPP4 structural diversity	69
4.5	SUMO1 interacts with DPP8 and DPP9	70
5.	Conclusion.....	72
6.	Outlook.....	73
7.	References	74

List of Figures

Figure 1.1. Protein structural elements and crystals.....	1
Figure 1.2. 2D Escher phase transition.....	2
Figure 1.3. Phase diagrams.....	3
Figure 1.4. Water IR radiation absorption.....	5
Figure 1.5. Schematic representation of Bragg's law	5
Figure 1.6. Diffraction improvement of CODH after dehydration.....	8
Figure 1.7. Hanging drop method equilibrium.....	10
Figure 1.8. Protease cleavage site nomenclature.....	12
Figure 1.9. DPP4 molecular structure	14
Figure 1.10. SUMO1 / ubiquitin comparison and sumoylation pathway.....	15
Figure 2.1. DPP4, DPP8 and DPP9 inhibitors.....	18
Figure 2.2. Free mounting laser (FML) setup	19
Figure 2.3. FMS calibration using K_2SO_4	20
Figure 2.4. Laser energy rates	21
Figure 2.5. Laser optics and alignment	22
Figure 2.6. Pico-dropper devices.....	23
Figure 2.7. Dropping solutions using pico-dropper.....	24
Figure 2.8. Glass-based crystal holder design.....	24
Figure 3.1. Unit cell changes as a function of relative humidity.....	30
Figure 3.2. CODH anisotropic dehydration	31
Figure 3.3. Near-infrared absorption of molecules.....	32
Figure 3.4. CODH crystal temperature change during IR treatment.....	33
Figure 3.5. CODH dehydration and its dependence on laser power	33
Figure 3.6. Lysozyme dehydration speed limit	34
Figure 3.7. Unit cell dimensions of lysozyme crystals under oil	35
Figure 3.8. Lysozyme crystal IR irradiated under PFP AE oil cover.....	36
Figure 3.9. Cryogenic water phase transition induced by FML	37
Figure 3.10. FML annealing of DPP9 crystal at 100K.....	38
Figure 3.11. DPP8 crystal transformation induced by dropping DMSO.....	39
Figure 3.12. Nebulizer soaking method	39
Figure 3.13. Dropping of DPP9 crystal with TMAO as cryo-protectant.....	40
Figure 3.14. FMS and FML dropping effect comparison.....	41
Figure 3.15. Multiple sequence alignment of DPP4, 8 and 9 highlighting structural features.....	43
Figure 3.16. Homodimer alignment and comparison of DPP4/8 and DPP8/9	44
Figure 3.17. R-segment order/disorder transition in DPP8 and DPP9	47
Figure 3.18. Ordering of R-helix upon substrate binding.....	48

Figure 3.19. Comparison of peptides bound in the active site of DPP4 and DPP8.....	48
Figure 3.20. SLRFLYEG and 1G244 active site binding and induced fit.....	50
Figure 3.21. DPP8 bound SLRFLYEG peptide aligned with SUMO1	51
Figure 3.22. S2 subsite loop comparison between DPP4, DPP8, and DPP9.....	52
Figure 3.23. Val-BoroPro bound to DPP8.....	53
Figure 3.24. DPP4 and DPP8 pore size comparison	54
Figure 3.25. DPP8 and DPP9 interaction and enzyme kinetics with SLRFLYEG and 1G244.....	55
Figure 3.26. SUMO1 oligomers formation	58
Figure 3.27. DPP9 forms a stable complex with SUMO1 oligomers.....	59
Figure 3.28. Increased binding of DPP9 to oligomeric SUMO1.....	60
Figure 3.29. DPP9-SUMO1 complex is disrupted by small molecules binding	61
Figure 3.30. SUMO1-linker-SUMO1 oligomer does not increase binding to DPP9	61
Figure 3.31. DPP8 and DPP9 activity in presence of SUMO1	62
Figure 4.1. Protein data bank deposited structures.....	63
Figure 4.2. RMSD of DPP8 and DPP9 simulated systems	69
Figure 4.3. Structural comparative analysis of human DPP4/8 versus bacterial DPP4	70
Figure 4.4. Model of interaction between SUMO1 and DPP8/9.....	71

List of Tables

Table 2.1. Proteins used in this study.....	17
Table 2.2. Crystallization conditions of crystals used in this study.....	17
Table 2.3. Relative humidity equilibrium of standard salt solutions.....	19
Table 2.4. Beam size and optics position correlation.....	22
Table 3.1. Crystallographic parameters for lysozyme treated structures.....	41
Table 3.2. Crystallographic parameters for DPP8 and DPP9 structures.....	45
Table 3.3. Crystallographic parameters for DPP8-Val-BoroPro structure.....	53
Table 3.4. Inhibition data for DPP8 and DPP9.....	56

Abbreviations

AMC	7-Amino-4-methylcoumarin
BS3	Bis[sulfosuccinimidyl] suberate
CODH	Carbon monoxide dehydrogenase
DPP	Dipeptidyl peptidase
DTT	Dithiotreitol
FML	Free Mounting Laser System
FMS	Free Mounting System
GP-AMC	Gly-Pro-AMC
GST	Glutathione S-Transferase
HEPES	(4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid)
OD	Optical density
PFPAE	Perfluoropolyalkylether
PMSF	Phenylmethylsulfonyl fluoride
RH	Relative humidity
SUMO	Small ubiquitin-like modifier
TMAO	Trimethylamine N-oxide

1. Introduction

1.1 Part I: Crystal order and diffraction improvement

1.1.1 Theory

1.1.1.1 Proteins and protein crystals

Proteins are one of the four organic macromolecular building blocks of life. The unit-based monomers of proteins are 20 different natural amino acids, which combined are responsible for an enormous functional and structural protein diversity. Each amino acid bears a characteristic biochemical footprint. They differ amongst other things in charge, hydrophobicity and size. Their sequential disposition forming a chain is the basis on which the secondary structure is built. α -helices and β -strands are the core structural elements used to shape proteins. Additional loops and turns are less structured and flexible, usually found at the protein surface. At optimal conditions of concentration and precipitant agents, proteins tend to interact with each other, forming a regular grid growing three-dimensionally as a function of time ^[1]. As a result of this a crystal is formed (Fig. 1.1).

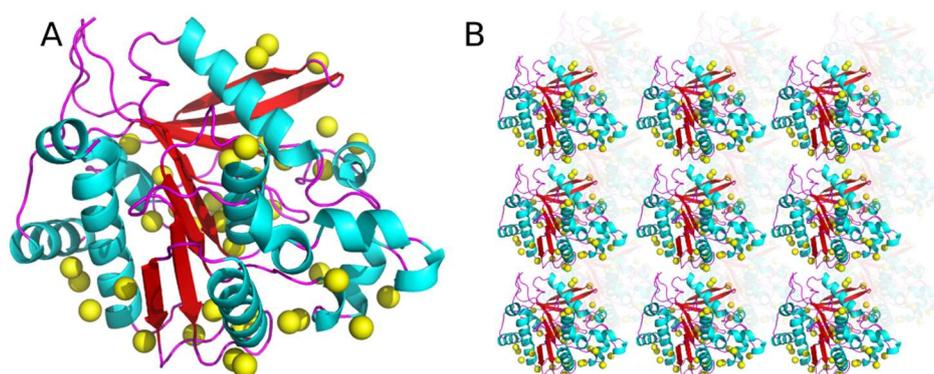


Figure 1.1. Protein structural elements and crystals. (A) α -helices (cyan), β -sheets (red), connecting loops (magenta) and water molecules (yellow). (B) Three-dimensional lattice formed by protein interactions. Hydrolase domain of DPP4 (PDB 1R9N) ^[2].

Protein crystals are composed mainly of protein, buffer components and water. The average amount of water per protein crystal is roughly 50%. A fraction of these water molecules interacts tightly with protein counterparts, known as structural water. The rest of water molecules are less fixed, located in deep water channels distributed all over the crystal, susceptible to be redistributed. These water molecules are known as bulk water ^[3]. Protein crystals are very fragile and sensitive to environmental conditions. Small changes in temperature, pressure or hydration might affect the crystal order.

1.1.1.2 Phase transitions and crystals

Crystals are materials or instruments in several fields of human endeavor. Their physicochemical properties are tightly related to their geometric arrangement, known as phases. Typically crystals are made of one phase. If a crystal transforms from one phase to a new one, it experiences a phase transition

[4]. Phase transitions are well characterized by inorganic crystals, but they occur in a broad range of materials, with a number of ways to induce them [5]. In most cases, temperature or pressure are the triggers of phase transition events, leading to quite remarkable modification of physical properties. Phase transitions are driven by a Gibbs free energy gradient, where both phases have the same free enthalpy at the transition point. Phase transitions are characterized by two phases, where the highest symmetry contains all symmetry operators of the lowest one, also referred to as high temperature (disordered) and low temperature phases (ordered), respectively [4]. A graphic example of a phase transition can be obtained by modifying a drawing by M. C. Escher: modifying the color of the birds renders a change in the symmetry, therefore generating a new phase (Fig. 1.2).

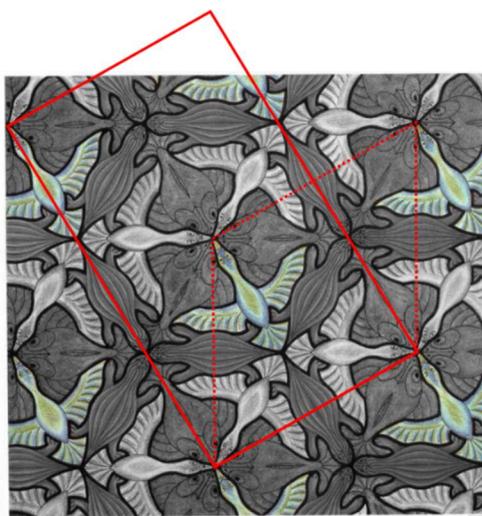


Figure 1.2. 2D Escher phase transition. Modifying the color of birds in north-west direction changes the symmetry of the image from a trigonal lattice (dotted-line) p3m1 (No. 14) to a centered rhombic (continuous-line) cm (No. 5), therefore corresponding to a phase transition in two dimensions. The 3-fold rotation axis is lost, while the mirror and glide planes are kept. The image was modified from the original M. C. Escher drawing Fishes, Birds and Turtles.

Phase transitions can be classified by several definitions. The Ehrenfest classification defines the existence of different orders, being the first order the most commonly observed with hysteresis as one characteristic feature (e.g. melting and boiling) and second order transitions, (e.g. crystallographic, ferro-, para-magnetic and electric) This classification is based on the order of the differential of the Gibbs free energy for which a jump is observed at the phase transition point. Thus, a phase transition is considered to be of the first order, if a measurable quantity like entropy or volume (first derivative of Gibbs free energy) presents a jump-like discontinuity at the phase transition (associated with a latent heat). Moreover, a second phase transition will present the jump-like discontinuity in the second derivative of the Gibbs free energy at the phase transition e.g. heat capacity. Conversely, the heat capacity of a first order phase transition goes to infinite values [6].

The Landau classification focuses on the second order phase transition, introducing the concept of “order parameter” which correspond to a measurable quantity. This major contribution helped to generalize the concept of phase transition by using thermodynamic equations of states, making them universal and independent of the subject of study (equation 1.1) [7].

$$U = TdS - pdV + \mu dN \quad (\text{Eq. 1.1})$$

Where U is the internal energy at constant pressure and variable temperature per molecule unit. The current model of phase transitions categorizes them into discontinues and continuous phase transitions. These two cases can be exemplified by using a volume-temperature and magnetization-temperature diagram for a particular element. Figure 1.3 describes the existence of two different states, separated by a transition curve. Each curve either meets a new phase transition or ends in a critical point, depending on symmetry constraints. The chart in figure 1.3A describes the volume change of a substance as a function of temperature at constant pressure. At temperatures below the critical point there are two possible phases, each of them with very different volumes. Transitions from one point to the other imply a discontinuous volume change. In contrast, above the critical point both phases coexist and a transition from one point to the next one does not produce a change in volume. Therefore a continuous phase transition occurs [8].

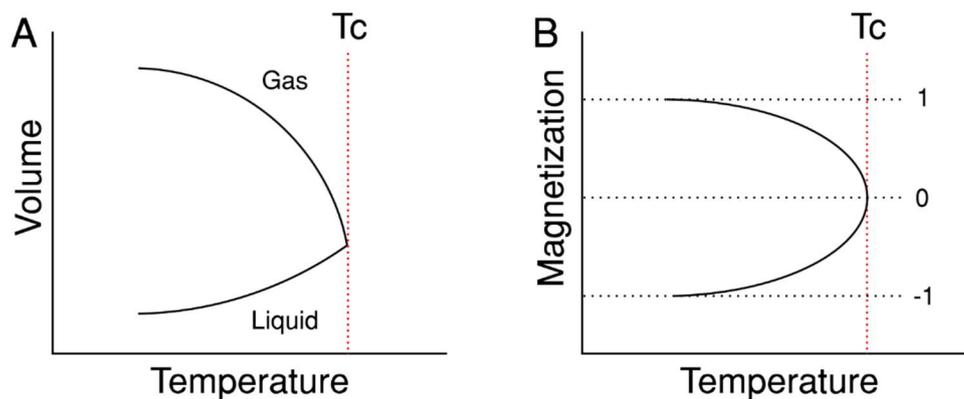


Figure 1.3. Phase diagrams. (A) Volume-temperature phase diagram. Below T_c the two possible phases, liquid and gas, are separated by a discontinuous first-order phase transition. Above T_c , both phases coexist simultaneously, undergoing a continuous second order phase transition. (B) Magnetization-temperature diagram. Magnetization of a material below T_c from 1 to -1 . Above the T_c the material is paramagnetic. In both cases pressure and the number of molecules are constant.

Phase transitions can occur from solid to solid (s-s) state as well. A crystal can change its properties after a certain treatment. S-s phase transitions follow the same fundamental theoretical concepts as liquid-gas, with a different state equation depending on the model of study. One of the best studied examples has to do with the phase transition of paramagnetic to ferromagnetic materials, where a magnetic field is induced after the critical point is overcome [9] (Fig. 1.3B). Iron crystals also have several phase transitions as temperature raises. They transition from a body centered (bcc) cubic (α -Fe) crystal structure to a face centered (γ -Fe) and a high-temperature bcc phase (δ -Fe). Finally, it reaches the melting point at 1809 K [10]. These materials and many others are relevant topics in metallurgy,

electro-induction and so forth. In macromolecular crystallography, protein crystals might also exhibit phase transitions, though are much less studied due to the complexity and fragility of these crystals [11]. Studies in the context of crystallographic protein structure determination show how using different post-crystallization methods can induce phase transitions, impacting directly the diffraction quality [12]. Diffraction resolution and mosaicity are important quantities as measurable parameters. Additionally, in some cases crystal volume can also be a parameter to determine the existence of a phase transition.

1.1.1.3 Relative humidity and dew point

Relative humidity (RH) is an important factor to keep in mind when controlling protein crystals water content. Relative humidity is defined as the ratio of the partial pressure of water vapor expressed as a percentage of the maximum amount of water at a certain temperature and pressure. Furthermore, the dew point is the temperature at which water contained in the vapor phase begins to condensate. A convenient expression to calculate relative humidity as a function of temperature and dew point is the Magnus equation 1.2 [13].

$$RH = \exp \left\{ a_w b_w \left[\frac{T_{dp} - T_{at}}{(b_w + T_{dp})(b_w + T_{at})} \right] \right\} \times 100\% \quad (\text{Eq. 1.2})$$

Where a_w and b_w are constants with 17.5043 K and 241.2 K, respectively, T_{dp} corresponds to the dew point temperature and T_{at} to the measured air temperature. This formula is considered valid from 0 ° to 60 °C of temperature, 1% to 100% relative humidity and dew point between 0 ° to 50 °C.

1.1.1.4 Infrared radiation and water absorption

Light is absorbed following the Beer-Lambert law [14]. This relation correlates the absorption of light to the amount of molecules of a sample in a wavelength dependent manner. Infrared (IR) electromagnetic radiation absorption by water is the major fraction of the electromagnetic spectra absorbed in the atmosphere, responsible for heat retention [15]. IR radiation can be separated into different regimes: near-infrared (0.75-1.4 μm), mid-infrared (1.4-15 μm) and far-infrared (15-1000 μm) wavelengths. The absorption spectrum of water reveals several absorption peaks maxima (Fig. 1.4) [16].

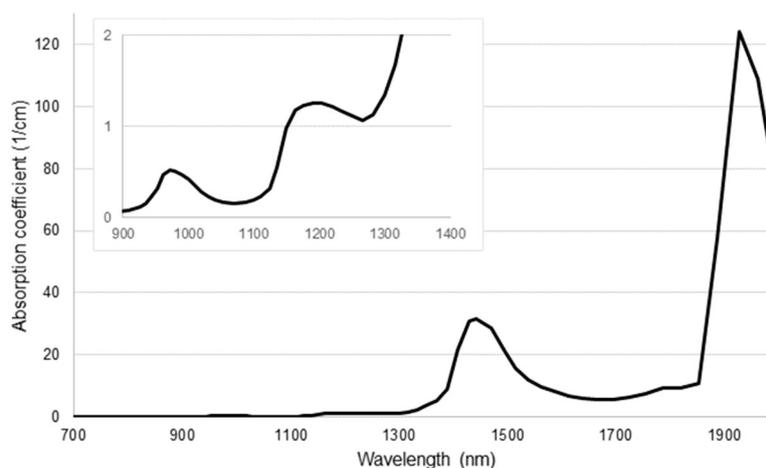


Figure 1.4. Water IR radiation absorption. Absorption spectra of water from 700 nm to 2000 nm. Two prominent peaks are visible at 1400 and 1900 nm. Both peaks correspond to water stretching motions. The inserted zoom-in chart shows small absorption peaks from 900 to 1400 nm^[17].

Water is a molecule with a non-zero dipolar momentum, conferred by a non-symmetrical charge distribution due to the molecular atomic electronegativity difference^[18]. This characteristic feature enables water to interact with the electric component field of an incoming electromagnetic radiation. Depending on the radiation frequency, different motions of water molecules, either rotations, bending or stretching might absorb energy. In liquid and solid water most rotational movements are decreased due to hydrogen bonds between water molecules. Each water motion has a particular resonance frequency. If that resonance frequency matches the frequency of the incoming electromagnetic radiation, a transition to a higher energetic state will be achieved. In figure 1.4, the peak at 1440 nm corresponds to the first overtone of water stretching motion, whereas the peak at 960 nm is the second overtone of water stretching^[19].

1.1.1.5 Crystallography, phasing and refinement

Protein crystallography can be regarded as a particular case of coherent diffracting imaging. Here, electron density is calculated as the content in the unit cell convolved with the crystal periodic lattice. Thus, the added scattering from each family of planes, as defined by Bragg's law in the early time of crystallography, enhances x-ray diffraction, as illustrated in figure 1.5 and the derived formula (Eq. 1.3)^[20].

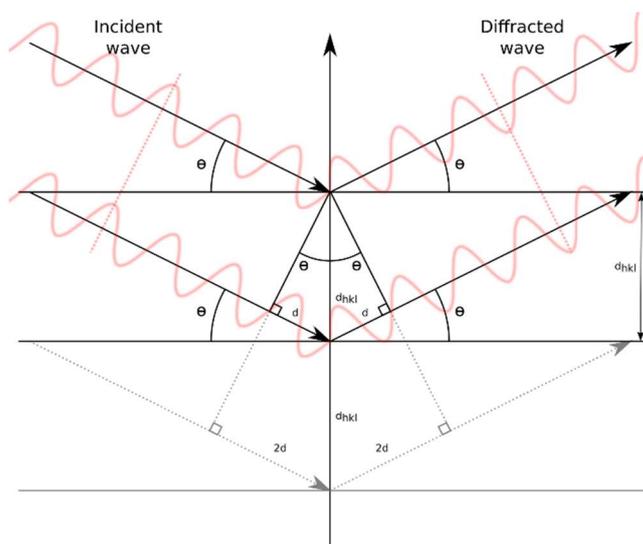


Figure 1.5. Schematic representation of Bragg's law. Three members of a family of planes diffracting a wave in an elastic manner. The incident angle θ is equal to the diffracted angle θ . If the wavelength is known, it is possible to determine geometrically the interspace distance between lattice points (d_{hkl}), using the difference in path travelled by each wave.

$$n\lambda = 2d_{hkl} \sin \theta \quad (\text{Eq. 1.3})$$

The study of x-rays interacting with atoms uses the wave nature of radiation. Taking advantage of its mathematical expression, waves can be considered as vectors. These vectors, in turn, can be conveniently handled by modern software. They can be separated into two fundamental components. The first one, the magnitude of the wave, corresponds to the electron density of each scattered wave. The second, the angle of the vector, describes the final relative phases of scattered waves. The magnitude of the diffracted wave is measured on an image detector. Structural factors can be retrieved from the images, calculated as the square root of the measured intensity ^[21]. However, the phase information of each reflection is lost because of the technical impossibility of recording phase angles at the detector point. Hence, phases should be experimentally determined. To do so, the Patterson function allows the determination of inter- and intra-molecular vectors using the square of the structural factors information, which in turn can be used to find the correct phase with the aid of heavy atoms, pre-existent related crystallographic structures or anomalous scattering ^[22].

Once a solution has been found, the next step is to refine the structure solution. This refers to the iterative process to improve the agreement between the modeled structure and the measured reflection in the experiment, aided by statistical constraints of preferred position of atoms, angles and bond distances. This process is carried out following structure quality parameters to avoid over-refinement, known as R-factor and R-Free-Factor. Here, a fraction of the reflections is not included in the refinement steps to compare the amount of bias against final refined structure ^[23].

1.1.2 Post-crystallization diffraction improvement

Protein crystallography has become the gold standard technique to determine atomic protein structures. Several technological developments have made the process more automatized and reproducible. The advent of new x-ray radiation sources and cryogenic handling of crystals successfully reduced the time of data collection ^[23] and radiation damage ^[24]. In parallel, the computing power and software development granted scientists with more and better tools to solve and refine structures, as well as validation procedures ^[25]. The aforementioned improvements rely heavily on (successfully) well diffracting crystals. This is a property not always present in crystals and certainly not trivial at all to achieve. In fact, several lines of research are dedicated to improve low quality diffracting crystals controlling water content, order, composition and more ^[26]. Crystal disorder can occur by different mechanisms. Either by substitutional disorder, where a lattice point is occupied by an isomer of the molecule forming the crystal. By random displacive disorder, which triggers the appearance of diffuse scattering, and others ^[27].

Protein crystal diffraction can be used as a parameter to determine crystal quality. Diffraction resolution is defined as the amount of interpretable data extracted from the total amount of a unique set of reflections produced by a family of planes. Sampling with different interplanar lattice distances produces diffraction at various resolutions, where a narrower family of planes renders diffraction spots of higher resolution and vice versa. Diffraction resolution is intrinsically related to crystal order. The more ordered the crystal, the brighter the intensity of high resolution diffraction spots will be on the detector. As crystal order is reduced, high resolution reflections become weaker, eventually disappearing [28].

A second parameter used to characterize crystal quality is mosaicity. The term refers to a random local misalignment of the crystal lattice. These local domains are often called mosaic blocks. Diffraction from mosaic blocks in an ideal crystal is uniform, generating discrete reflections on detectors. Real crystals, however, are not perfect, presenting several lattice distortions. These distortions can occur as soon as the crystal nucleation starts. They can also appear during crystal growing or they might happen in mature crystals, by inducing mechanical stress e.g. by crystal cooling. These crystal imperfections imply that diffracted reflections are spread with a certain width distribution to some rotational degree. Even though mosaicity and diffraction resolution are not intrinsically related, they often correlate [29].

Finally, a third useful parameter to quantify structural disorder is the Debye-Waller factor or B-factor. This value is a measure of the atomic displacement about their equilibrium position, and affects the intensity of Bragg spots, thus changing the structural factor. This term takes the following form.

$$B_{iso} = \exp \left[-8\pi^2 \langle U_h^2 \rangle \frac{\sin^2 \theta}{\lambda^2} \right] \quad (\text{Eq. 1.4})$$

Here, U_h is the projection of the atomic displacement U on the direction of diffraction vector h , θ is the scattering angle and λ is the X-ray wavelength. This expression describes an isotropic B-factor, but under specific condition of very high resolution, anisotropic B-factors can be calculated [1].

1.1.2.1 Protein crystal water content control

A search in the Protein Data Bank (PDB) sets an interval between 25% and 85% water content for protein crystals. The theoretic solvent fraction in a protein crystal can be determined using Matthews's coefficient. This coefficient was established for all space groups as the ratio of the volume of the asymmetric unit divided by the molecular weight of the molecule [30]. Using Matthews's coefficient and a virtually constant partial specific volume, it is possible to calculate the fraction of the asymmetric unit occupied by protein. Furthermore, as a difference of the total volume, the fraction of solvent can be computed. From these calculations the average amount of water per protein crystal is 43% [31].

Water content of protein crystals was known from early times on to be a crucial player in diffraction quality. Crystals were transferred and equilibrated in reservoirs of different concentrations of precipitant solutions. This would generate a lower or higher relative humidity which would change the crystal solvation [32, 33]. Most crystallization conditions have a vapor phase equilibrium of 97-99% RH, whereas high salt concentration conditions might have ~90% RH equilibrium [34]. In most of the cases it was noticed that proteins would stop diffracting below 90% RH, whereas a small minority

presented improved diffraction quality [35]. A way to study those diffraction modifications was brought by the introduction of a humidity control system. The Free Mounting System (FMS) [36] allowed crystals to be x-ray irradiated without their mother liquor and at the same time setting humidity gradients to study their diffraction quality [37]. This method bathes crystals with a laminar flow humidified air stream. The high accuracy control (0.01% RH) warrants small and controlled steps of humidity changes. Changes in crystal size are tracked by measuring two-dimensional back projections of the crystal [38]. The device is capable of setting the dew point and feedback-regulate it by gas temperature measurements. Thus, it corrects the dew point precisely. Finally, the relative humidity is calculated using equation 1.2. The crystal sits at the exit of the nuzzle on a crystal holder. Once the crystal is mounted and it has reached equilibrium with the humidified air stream, the x-ray evaluation can begin. Similar devices have pushed further the evaluation of crystal hydration states. They allow studies directly at beam-lines in synchrotrons to perform systematic experiments on different crystal systems [39, 40]. Several examples using this technology show the benefit of controlled dehydration at room temperature, exposing new order states where crystals are diffracting significantly better at higher resolutions [41, 42, 43, 44]. An outstanding example of diffraction improvement using FMS is carbon oxide dehydrogenase (CODH) protein. CODH protein crystals grown at equilibrium condition at 95% RH diffract roughly to 3.5 Å, whereas dehydration to 89% RH improves their diffraction quality substantially to 2.0 Å [35] (Fig 1.6).

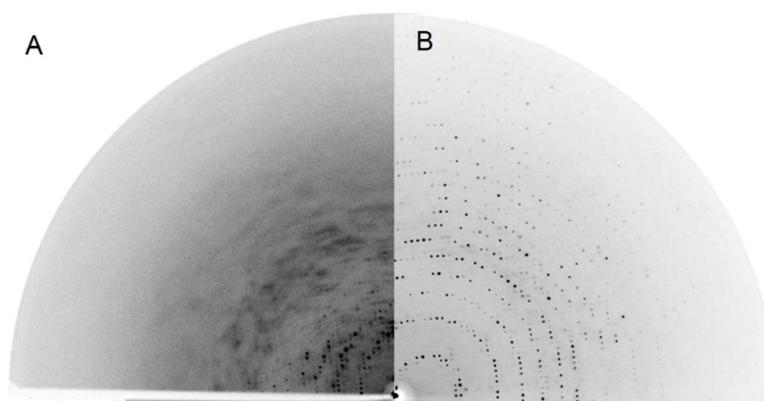


Figure 1.6. Diffraction improvement of CODH after dehydration. (A) Diffraction image of the native state of a CODH crystal grown at a humidity equilibrium of 95% RH. (B) Diffraction image of the dehydrated state of CODH crystal with improved diffraction quality at a humidity equilibrium of 89% RH. The orientation is the same in both cases. The limit of image resolution is 3.0 Å.

1.1.2.2 Protein crystal annealing

Pure water freezes at 0 °Celsius at ambient pressure. A mixture of protein and water has a lower freezing point, reducing its melting point temperature. The reduction in melting point temperature is directly related to the concentration of the solute (protein). The reason is related to the chemical potential of the solvent and increase in entropy [45]. Protein crystals, therefore, have lower freezing points than pure water, reduced even further by the presence of buffer components. If a protein solution containing a cryo-protectant like glycerol is flash-cooled to 100 K, the solution reaches a vitrified phase. This occurs

because the solution does not have enough time to generate ice crystals and remains as an amorphous material [46]. This technique is being used to prevent radiation damage while data collection [24].

Water behavior at low temperatures is a field of on-going study. If water at the amorphous state is warmed up to 160 °C, cubic ice starts to form. Further heating would render hexagonal ice at around 180 °C. These phase transitions can be followed by diffraction, since each state generates different diffraction ring patterns. There are indications for a particular behavior of water in a narrow temperature window between 130-140 °C. Namely, water can transit from an amorphous phase, rather static, to a highly viscous liquid, before becoming ice. This property enables water to diffuse through the crystal, finding lower energetic states regardless of the low temperature, assuming that confined water inside crystals behaves similarly as pure water [47].

Once crystals are cooled, annealing can be performed. Annealing is the process of heating a material and let it cool down. The process searches a new phase with different physical properties. In the case of metals to render a stronger material, or in protein crystals to find better diffraction quality. In protein crystallography cryo-annealing has been used as a post-crystallization method with some successful examples [48]. The effectiveness of cryo-annealing is determined by tracking diffraction quality, using both diffraction resolution and mosaicity as a function of thawing cycles. Cryo-annealing is usually performed only a short time, e.g. a fraction of a second. Annealing can also be studied using slow temperature ramps, measuring unit cell dimension changes. Experimental data shows a correlation between phase transitions and confined water to channels in the crystal structure, suggesting a link between a glass transition (highly viscous liquid) and the size of water channels in protein crystals. Therefore, big channels are correlated with a higher probability of glass transitions occurring [47, 49].

1.1.2.3 Protein crystal chemical modification

Protein crystals are in equilibrium with the environment. Only a narrow range of conditions allow nucleation and growing. Humidity, temperature and composition are key variables to consider when growing crystals. The buffer component sets a specific pH, to which proteins are exposed and would respond based on their average isoelectric point. The latter, directly affecting the way they interact with each other, therefore favoring or impairing crystallization [50]. The chart in figure 1.7 depicts the humidity equilibrium in a hanging drop experiment, where a drop of protein plus precipitant can freely adjust its humidity. Thus, after a period of time, the drop reaches the equilibrium with the reservoir below. Given a defined condition of protein and precipitant concentration, crystals are allowed to grow, after overcoming a Gibbs free energy threshold, whereas for an unfavorable condition, crystals might dissolve [51, 52]. Even though crystals grow in a low energetic state, they might have alternative favorable lattice arrangements [53]. This feature provides freedom to gently modify the chemical composition of crystals in order to find lower energy diffracting states. Several articles report the systematic evaluation of changing buffer components or pH by classic soaking approaches. More in detail, grown crystals are manually transferred to a new drop, containing the new buffer components. Then, crystals are

equilibrated in this condition and tested for diffraction improvements. In some cases cryo-protectants, protein substrates or inhibitors are readily included in the new condition ^[50].

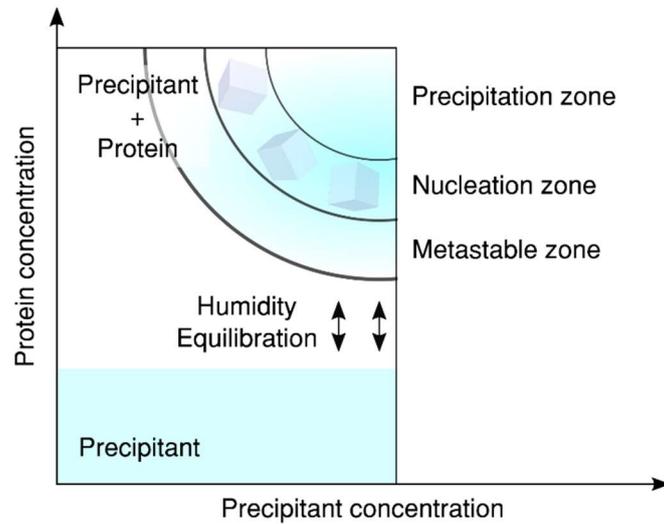


Figure 1.7. Hanging drop method equilibrium. Protein-precipitant concentration diagram of protein crystal nucleation. Crystals grow mostly in the nucleation zone. This zone is directly related to the concentration of the precipitant and protein. Too little or too much of any variable sets an unfavorable growing condition. Water moves freely from the drop towards the reservoir and vice versa, reaching humidity equilibrium.

A different soaking method uses dropping of substances onto crystals, thus changing the crystal chemical composition gently. This method relies on decreasing relative humidity to concentrate a desired substance while keeping the crystal volume constant. This approach uses FMS to set a constant dehydration gradient, so that the crystal tends to dehydrate slowly. After reaching a lower threshold of dehydration, a small drop is ejected from a piezo electrically controlled device (pico-dropper) toward the crystal surface. Thus, the crystal rehydrates meeting the higher threshold. By repeating this process several times, the crystal concentrates the dissolved substance. Eventually the vapor pressure equilibrates with the lower humidity point set by the FMS. Thereby, crystal dehydration stops without loss of volume. This method, applied systematically, can render a crystal with improved diffraction quality ^[35, 37].

1.1.3 Aim and objectives of the study: infrared radiation modifies protein crystal order

Solving protein structures at high resolution has become indispensable for both research and drug development, where diffraction of some protein crystals is a major obstacle to overcome in structure elucidation by crystallography. Dehydration of protein crystals is known to be a post-crystallization technique to improve diffraction of some crystal systems ^[41]. The exact correlation between these improvements and the mechanism behind them is not well understood. Experimental dehydration data performed on protein crystals shows how intermolecular contacts of the protein lattice shorten and additional ones are created as dehydration progresses. Given enough space to accommodate these new lattice features, protein crystals might have their diffraction improved ^[43]. In contrast, if the lattice opposes restrictions due to not having sufficient flexibility to accommodate them, it might have its diffraction impaired. Systematic and controlled modifications of the protein lattice are supposed to increase the understanding of crystal improvements and will broaden the range of protein crystals suitable for structure determination.

Protein crystal order and resulting diffraction improvements can be explored by different technical approaches. Firstly, Infrared (IR) radiation produced with a diode laser allows irradiation of protein crystals to control their water content. The Free Mounting Laser System (FML) method has been initially described as advantageous regarding the dehydration speed of crystals, among others ^[54]. Furthermore, the IR laser can be used in short pulses, carefully controlling the energy dose delivered to the crystal. Laser radiation heats crystals, thereby increasing the rate of water evaporation from them, reaching water equilibrium with the external relative humidity. Secondly, as an alternative to diffraction improvement by dehydration, protein crystal annealing by laser IR radiation will be explored. Heating crystals increases the kinetic energy of the lattice, hence taking it to a higher energetic level. Once the heating process is finished, the cooling phase takes place, possibly ending in a lower energetic state ^[48]. Thus, one of these states can have improved diffraction quality. The heating process can be performed from any ground state, either from 100 K as usually performed, or at any arbitrary temperature at constant volume and pressure. Lastly, as a third approach to induce diffraction improvements on protein crystals, the chemical composition of the crystals will be changed. FML can be employed for dropping chemicals. Substances can be concentrated to desired levels by locally increasing the temperature, therefore increasing the compound solubility ^[37].

This study seeks to explore diffraction improvements of protein crystals utilizing IR radiation to control fundamental variables of the protein crystal equilibrium such as water, order and chemical composition. IR radiation provides a gentle non-contact, fast and accurate way to control diffraction quality of a single or multiple crystals at once. Taken together, post-crystallization diffraction improvements offer a way to rescue or uncover diffraction from complicated targets, often thought to be a lost case. Therefore, more protein crystals will be able to structure determination.

1.2 Part II, study case: DPP8, DPP9 and SUMO1

1.2.1 Physiological context

1.2.1.1 Proteases

Proteins are the primary executors of functions in cells. They control cell homeostasis performing different tasks. One task is cleavage of other proteins, performed by proteases. Cleaving proteins can have several consequences, including modifying their activity, affecting their stability or location. Proteases can be classified as endo- or exo-peptidases, depending on whether they cut inside a polypeptide chain or at its end, respectively. Additionally, these enzymes can be classified according to their mechanism of action ^[55]. The protease field has a defined nomenclature for the peptidase cleavage site, with a specific name for each residue at the cutting point. The cutting occurs between P1 at the cleavage end and P1' at the remaining polypeptide (Fig 1.8, arrow) ^[56].

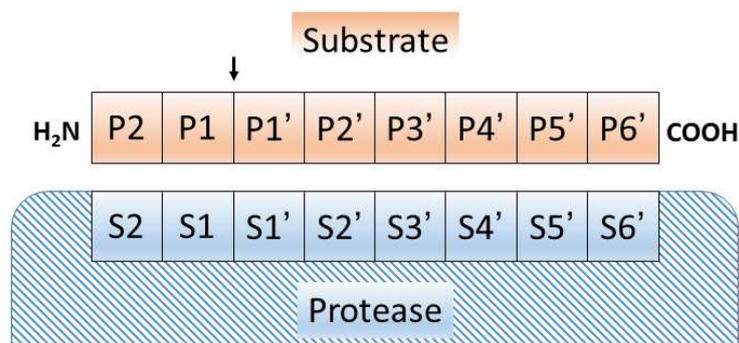


Figure 1.8. Protease cleavage site nomenclature. The substrate is cleaved between P1-P1'. P stands for peptide bond and S for subsite. Each subsite can be formed by several residues in the active site of the protease.

One class of proteases are serine proteases. They account for one third of the known proteases of the human proteome. Serine proteases are named after their proteolytic effector serine residue. Furthermore, they use water as part of the mechanism to cleave proteins, thus they are also known as hydrolases. This serine proteases family has a conserved catalytic triad contributed to by the residues serine, histidine and aspartic acid. The catalytic process starts with substrate fixation, then the formation of an oxyanion hole favors the nucleophilic attack of the carbonyl carbon of the scissile bond by the serine -OH group, generating a tetrahedral intermediate. Finally, a dipeptide is cleaved, where water helps to hydrolyze the acyl enzyme to regenerate the active site ^[57].

1.2.1.2 DPP4, DPP8 and DPP9: physiological context

Dipeptidyl peptidase (DPP) 4, DPP8 and DPP9 are members of serine exo-peptidases which cleave the last two residues at the N-terminus of the target proteins ^[58, 59]. They are members of the prolyl-oligopeptidase subfamily S9B, as classified in the MEROPS database ^[60]. DPP4 is an extracellular enzyme, often found soluble or as a transmembrane protein. It is mostly localized at the intestine and it is a key-player regulating glucose equilibrium. DPP4 is capable of controlling the half-life of insulinotropic incretins such as glucagon-like peptide-1 and glucose-dependent insulinotropic peptide.

Therefore, DPP4 plays a central role as a target of type two diabetes treatment. Several drugs aim to control DPP4 activity, in order to extend the half-life of incretins. DPP4 inhibitors have become important drugs, being developed in different chemical forms with various mechanisms and affinities [61].

In contrast, DPP8 and DPP9 are intracellular proteins, with a broader distribution in human tissues. They share ~79% of sequence identity [62]. Both proteins have a functional overlap, with similar chemical substrate specificity [63, 64]. Their functions are subject of intense study and play key roles in immunological response [65] and cancer [66]. Other publications show that DPP9 is essential for neonatal survival [67] and plays a role in cell migration and cell adhesion [68] as well as regulating B cell signaling [69]. Recently, one function associated to innate immune response against pathogens has been characterized. Inhibition of DPP8 and DPP9 induces specific cell death in monocytes and macrophage immune cell lines. The mechanism is associated with a lytic non-apoptotic cell death called pyroptosis [70]. Canonical pyroptosis mediates cell death using caspase-1, caspase-4 or caspase-5, triggered by pathogen-associated molecular patterns (PAMPS). The link between DPP8/DPP9 and pyroptosis was uncovered when using a non-specific serine protease inhibitor, Val-BoroPro (VBP). VBP treated cells showed that DPP8 and DPP9 are required to stimulate caspase-1 activation, thus pyroptosis. DPP8 or DPP9 do not directly activate or cleave caspase-1. Instead, they interact or inactivate a still unknown target, hence blocking caspase-1 repression [71]. DPP8 and DPP9 are important emerging drug development targets. Structural knowledge of both, proteins and inhibitors, would open a new chapter in the understanding and physiological significance and modulation of these two proteins.

1.2.1.3 DPP4, DPP8 and DPP9: structural relationship

DPP4 has been extensively characterized structurally due to its physiological relevance in public health. Several DPP4 crystallographic structures can be found at the protein data bank (PDB). In physiological conditions DPP4 is active only as a dimer [72]. Consistently, crystallographic structures occur as dimeric DPP4 (Fig. 1.9A), where each monomer is made of two domains. The first domain, the catalytic domain, harvest the conserved catalytic triad (Ser, His and Asp). It has a globular shape, with several central β -sheets stabilized by surrounding α -helices (Fig. 1.9B). The second domain, the propeller domain, provides the structural elements required to bind substrates. Two conserved glutamic acids and one arginine at the R-loop position bind to the N-termini of substrates. The propeller domain displays eight blades of antiparallel β -sheets, all forming a central pore (Fig. 1.9C). The active site can be reached by two ways. First, through the propeller domain pore, and second, by a side opening perpendicular to the pore. The side entry is considered as the primary substrate entrance [2]. A comparison of the liganded and non-liganded structures of DPP4 does not show any signs of structural conformational changes induced by substrate binding.

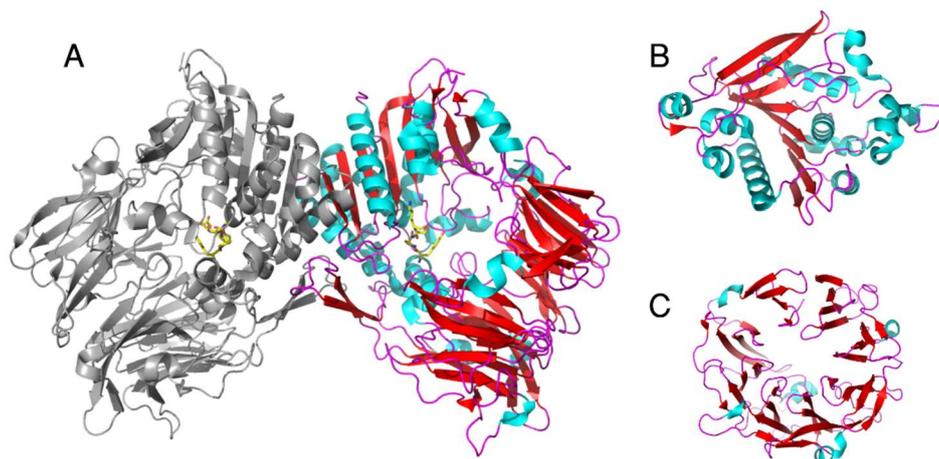


Figure 1.9. DPP4 molecular structure. (A) Dimeric DPP4. A peptidic ligand is located at the active site in yellow, between the catalytic and propeller domain. (B) Hydrolase domain. (C) Propeller domain. B and C are oriented along the central propeller pore to better display. Cyan-magenta-red color code represents helices-loops-sheets respectively (PDB 1R9N) [2].

In contrast to DPP4, little is known about DPP8 and DPP9 structures. Sequence alignments and functional assays suggest similar molecular structures and related activity, although DPP8 and DPP9 are larger molecules [73]. Mutational studies had been carried out to understand the role of different parts of the proteins and their relation to enzymatic activity [74, 75, 76]. Presumably, additional secondary structures are present compared to DPP4. In silico modeling of the three-dimensional DPP8 and DPP9 structures indicates a conserved catalytic domain and catalytic triad. However, the propeller domain structure might diverge significantly, since its sequence has several insertions and deletions compared to DPP4 [77]. Until now there is no accurate modeling available of this domain.

1.2.1.4 SUMO1

The small ubiquitin-related modifier (SUMO) 1 belongs to the ubiquitin-related protein family. It is broadly expressed in the eukaryotic kingdom with four described variants in mammals. They are used as post-translational modifications, covalently attached to lysine residues on substrates via their C-termini. One function of SUMO1 is to modify the binding partner repertoire of those sumoylated molecules. SUMO1 is 20 amino acids longer than ubiquitin and they share ~18% of their sequence identity, however they are structurally very similar (Fig. 1.10A). Furthermore, SUMO1 shares ~43% of its sequence identity with other SUMO-homologous proteins. Unlike SUMO2/3, SUMO1 is quite inefficient in forming poly-SUMO chains [78]. SUMO proteins have been implicated in several relevant cellular processes, like transcription, chromosome organization and function, nuclear transport or signal transduction pathways. Moreover, the p53 tumor suppressor protein is regulated by sumoylation, thus playing an important role in cancer biology [79].

Similar to ubiquitylation, sumoylation is performed by a sequential cascade of enzymatic reactions (Fig. 1.10B). First, SUMO1-activating enzyme (E1) generates an active SUMO-E1 complex by ATP consumption. Then, SUMO1 is transferred to SUMO1-conjugating enzyme (E2), forming a

SUMO-E2 complex. Finally, SUMO1 can be transferred to the substrate aided by a broad variety of SUMO-protein ligases (E3), conferring target specificity. The conjugation occurs to a lysine following a consensus motive Ψ KXE, where Ψ is a large hydrophobic residue and X can be any amino acid. Sumoylation is a reversible modification removed by Ulp proteins. These are also responsible for generating mature SUMO1, by cleaving a short peptide blocking the C-terminus of SUMO1 [80].

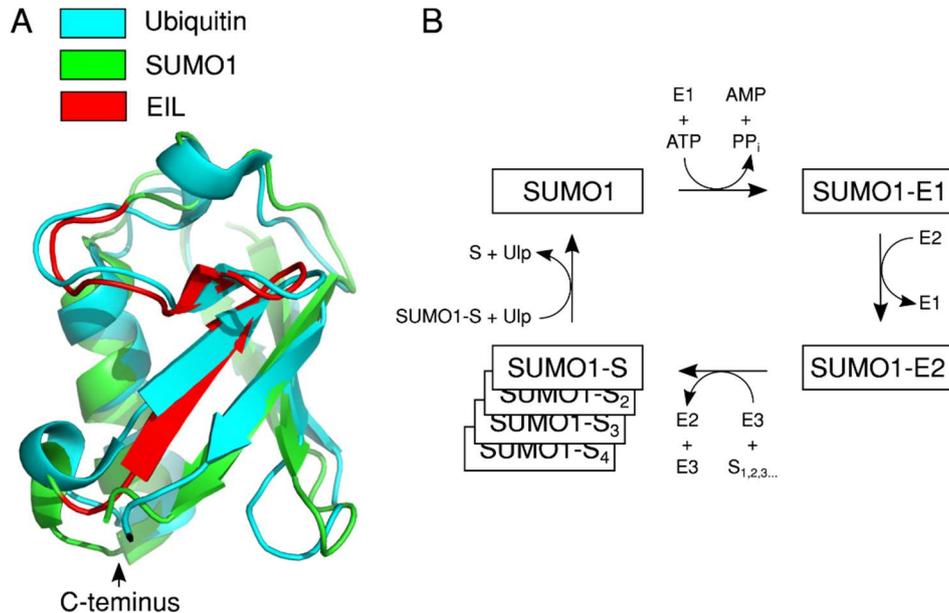


Figure 1.10. SUMO1 / ubiquitin comparison and sumoylation pathway. (A) SUMO1 and ubiquitin alpha-carbon alignment (1.51 Å rmsd). Ubiquitin and SUMO1 are colored in cyan and green, respectively. The EIL segment of SUMO1 is marked in red. (B) Sumoylation pathway including the E1, E2 and E3 enzymes required to sumoylate different substrates marked as $S_{1,2,3,\dots}$.

1.2.1.5 DPP8 and DPP9 interact with SUMO1

DPP8 and DPP9 participate in various cellular pathways and the identification of binding partners led to an appreciation of their physiological relevance. SUMO1 is one of these binding partners [81]. SUMO proteins can interact with other proteins in a non-covalent manner. The SUMO-Interacting Motive (SIM) is a hydrophobic cluster domain in SUMO proteins, known to take part in most of the protein-protein interactions of SUMO1. However, a SIM-independent interaction between SUMO1 and both DPP8 and DPP9 has been described. This interaction is restricted specifically to SUMO1, excluding SUMO2/3 [81]. The interaction region in SUMO1 is located at the opposite surface of SIM, called E67-interacting loop, in short EIL (SLRFLFEGQRIADNH). The glutamic acid 67 of EIL has been identified as one key residue in this interaction. A small peptide derived from the EIL sequence, SLRFLYEG, can efficiently compete with SUMO1 for the same binding site on DPP8 or DPP9. Strikingly, this peptide can strongly inhibit the cleavage activity of DPP8 and DPP9. The peptidic inhibitor was proposed to be the first allosteric inhibitor of DPP8 and DPP9. Surprisingly in contrast, SUMO1 binding to DPP8 or DPP9 renders a more active enzyme [82]. Although this binding has been identified, the physiological relevance of SUMO1 interaction with DPP8 or DPP9 is still unknown.

1.2.2 Aim and objectives of the study: structural and functional characterization of DPP8 and DPP9 interaction with SUMO1.

DPP8 and DPP9 are valuable proteins as pharmacological targets. Both are active serine proteases located inside cells, in contrast to DPP4, which is extracellular. They have been linked to important intracellular pathways involved in processes such as inflammation ^[71], cell migration ^[68] or cancer development ^[66, 83]. Though crucial for drug development, their crystallographic structures are unknown up to now, due to the difficulty of yielding well-diffracting DPP8 and DPP9 crystals. Therefore, the elucidation of the molecular structure of DPP8 and DPP9 will be the focus of this research. Moreover, to overcome crystallographic limitations and to enhance protein crystal diffraction, post-crystallization treatments would be implemented.

Additionally, DPP8 and DPP9 activity modulation via different inhibitors will be determined. The understanding of both the inhibitor interaction and activity modulation are important aspects to consider in drug design. Different inhibitors with different mechanisms of action will be compared. First, the peptide SLRFLYEG, a specific DPP8 and DPP9 inhibitor with allosteric properties. This inhibitor was derived from the EIL sequence of SUMO1 ^[82]. Second, 1G244 is a strong, specific, non-competitive small molecule inhibitor of DPP8 and DPP9 ^[84]. Lastly, Val-BoroPro is a non-specific covalent inhibitor of DPP4, DPP8 and DPP9 ^[85, 86]. Co-crystallization or soaking assays with each of these molecules might reveal the structural features of substrate binding as well as clarifying the structural differences leading to their different inhibition potencies and binding mechanisms.

Of particular interest is the interaction of DPP8 and DPP9 with SUMO1. This interaction was suggested by pull-down experiments with immobilized SUMO1. A weak affinity interaction was measured in the micromolar range by surface plasmon resonance ^[81]. Ambiguously, the inverse experiment with immobilized DPP9 would not confirm binding. Furthermore, the complex between DPP8 or DPP9 and SUMO1 is not present in detectable amounts in solution. Therefore, we seek to study the nature of this interaction in detail, using different protein-protein interaction assays. The determination of SUMO1 interaction with DPP8 and DPP9 can be of paramount relevance so as to understand immunological regulations or designing drugs to treat diseases like cancer.

2. Materials and Methods

2.1 Materials

2.1.1 Chemicals: Protein constructs and crystallization conditions

Different proteins were used in this study. A summary of each protein, expression constructs and their characteristics is listed in table 2.1.

Table 2.1. Proteins used in this study.

Protein Name	UniProt Access Code	Origin	Production Source	Length	Tag
DPP4	P27487	Homo sapiens	Spodoptera frugiperda	39-766	6xHis
DPP8	Q6V1X1-1	Homo sapiens	Spodoptera frugiperda	1-898	-
DPP9	Q86TI2-2	Homo sapiens	Spodoptera frugiperda	1-892	-
His-DPP9	Q86TI2-1	Homo sapiens	Spodoptera frugiperda	1-863	6xHis
DPP9-His	Q86TI2-2	Homo sapiens	Spodoptera frugiperda	1-892	6xHis
SUMO1	P63165	Homo sapiens	Escherichia coli	1-97	-
His-SUMO1	P63165	Homo sapiens	Escherichia coli	1-97	6xHis
SUMO1-His	P63165	Homo sapiens	Escherichia coli	1-97	6xHis
His-GST-SUMO1	Fusion Protein	Homo sapiens	Escherichia coli	1-97	His-GST
GST-SUMO1-(L)-SUMO1	Fusion Protein	Homo sapiens	Escherichia coli	1-97	GST
CODH	P19919	Oligotropha carboxidovorans	Oligotropha carboxidovorans	A:1-166 B:1-809 C:1-288	-
Lysozyme	P00698	Gallus gallus	Sigma L6876	1-147	-

Protein crystals were grown at 4° C or 20° C. Different protein crystals had different crystallization conditions as shown in table 2.2.

Table 2.2. Crystallization conditions of crystals used in this study.

Crystal Name	Crystallization Condition	Protein Buffer	Time of Crystallization	%RH* Equilibrium
DPP8	0.46 M Na-citrate	20 mM TRIS 8.0 150 mM NaCl 2 mM DTT	2 days	97
His-DPP9	10% PEG 8000, 25% Glycerol 0.16 M CaAc, 0.08 M Cacodilate pH 6.25	20 mM TRIS 8.0 150 mM NaCl 2 mM DTT	1 week	93
CODH	1 M KH ₂ PO ₄ /KOH pH 7.5 1 M NaH ₂ PO ₄ /NaOH pH 7.5 125 mM HEPES/NaOH pH 7.5 3% MPD	50 mM HEPES, pH 7.2	1 day	95
Lysozyme	1 M NaCl, 0.1 M NaAc pH 5.0	50 mM NaAc pH 4.5	1 day	96

*Relative humidity (RH)

2.1.2 DPP4, DPP8 and DPP9 inhibitors

Three inhibitors were used to modulate DPP4, DPP8 or DPP9 activity. SLRFLYEG (GeneScript), 1G244 (AK Scientific) and Val-BoroPro (Sigma-Aldrich). The first one, SLRFLYEG (Fig 2.1A), is an allosteric peptidic inhibitor designed by modifying the EIL segment in SUMO1. This inhibitor specifically binds to DPP8 and DPP9, blocking their activity with K_i values of 174 and 170 nM, respectively [82]. Second, 1G244 (Fig. 2.1B), is an isoindoline based inhibitor. 1G244 has reported K_i values for DPP8 and DPP9 of 13.7 and 33.7 nM, respectively [84]. Last, Val-BoroPro (Fig 2.1C), is an amino boronic dipeptide non-specific inhibitor of several serine proteases, including DPP4, DPP8 and DPP9 with 0.18, 1.5 and 0.76 nM K_i values, respectively [86].

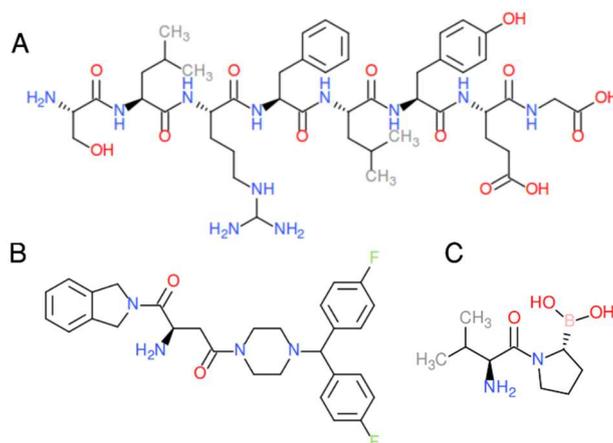


Figure 2.1. DPP4, DPP8 and DPP9 inhibitors used in this study. (A) SLRFLYEG, DPP8 and DPP9 inhibitory peptide. (B) 1G244, DPP8 and DPP9 inhibitor. (C) Val-BoroPro, DPP4, DPP8 and DPP9 inhibitor.

2.2 Methods

2.2.1 Post-crystallization improvement methods

2.2.1.1 Free Mounting Laser System (FML) hardware setup

The native humidity of protein crystals was measured using a humidifier machine FMS (Proteros Biostructures GmbH) [36]. Crystals were sampled at regular time intervals for x-ray diffraction (Rigaku RU-H3R) as well as performing two dimensional back projection measurements (Baumer camera TXG03 656x494). Fast dehydration ramps were performed using IR radiation of 938 nm. The source of radiation was a diode laser (Amtron LS453 laser system). IR light was directed towards the sample using optic fiber and focused to 200 μm diameter at the focus point using a focusing optic lens (PH101-02262 optics). A computer records the crystal area values and feeds them back to the humidity machine or laser control module, these in turn can adjust dew point or laser power, respectively, correcting values according to set parameters. Data acquisition is carried out by a dedicated board (DAQmx 662 board, National Instruments). The software was previously developed in-house using LabVIEW 10.0.1 environment (Fig 2.2).

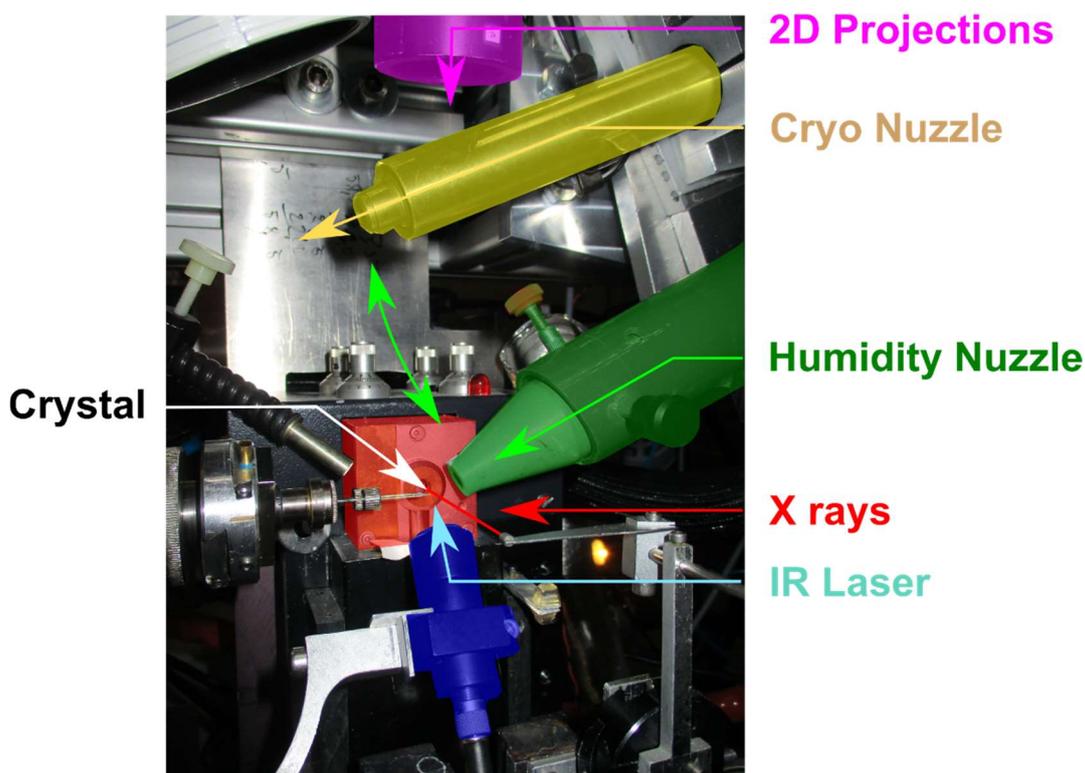


Figure 2.2. Free mounting laser (FML) setup. Parallel measurements of diffraction quality and crystal size can be performed. Each component of the FML is indicated with a different color code. The green double-headed arrow represents the possibility of changing between room temperature measurements and 100 K with a cryo-switch developed at Proteros Biostructures GmbH ^[54]. Crystals are mounted using mounting loops or capillaries depending on type of measurement. The setup is enclosed inside a security cage with an interlock system to avoid laser or x-ray leaks.

2.2.1.2 Humidity determination

Native humidity of a protein crystals is defined as the humidity of a crystallization condition at equilibrium. The first step to determine the native humidity of a protein system is to calibrate the humidifier machine with reference solutions. These solutions are saturated salt solutions. Their relative water content does not change very fast, so it can be considered constant. A set of reference solutions are used as standard solutions to calibrate the FMS (Table 2.3). A small amount of solution is placed in the laminar gas flow of the humidity nuzzle, changing the offset values until a drop reaches equilibrium, with a stable 2D area measurement (Fig. 2.3A).

Table 2.3. Relative humidity equilibrium of standard salt solutions.

Salt solution	% RH*
NaCl	75.47
KCl	85.11
KNO ₃	94.62
K ₂ SO ₄	97.59

* Vapor relative humidity at equilibrium

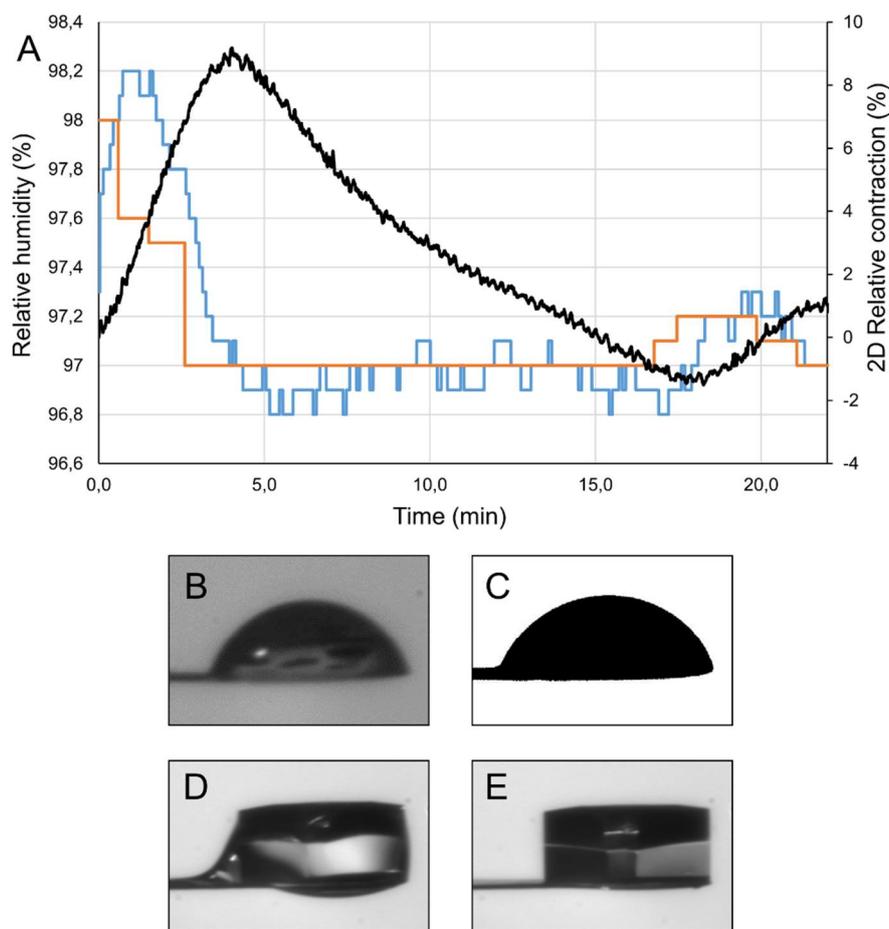


Figure 2.3. FMS calibration using K_2SO_4 . (A) FMS machine calibrated with a saturated solution of K_2SO_4 (98 %RH). The orange curve corresponds to set values of relative humidity. The blue curve is the actual relative humidity value. The black curve measures 2D-black&white drop contraction or expansion values. (B-C) K_2SO_4 drop mounted on a MiTeGen MicroMesh and 2D projection used to calculate area values, respectively. (D-E) Crystal with mother liquor and Crystal with mother liquor removed, respectively.

With a calibrated machine, a small amount of reservoir (2 μ l) from the crystal growing condition is placed on the machine to determine the native humidity. Once the 2D area measurement is stable, a crystal can be mounted (Fig. 2.3, A-C). The excess of solution around the crystal can be carefully removed using a capillary or a small piece of paper (Fig. 2.3 D-E). The crystal can now be considered to be freely mounted, thus quickly responding to FMS or FML humidity changes.

2.2.1.3 Laser power and optics

Laser treatments can be performed using a slow or fast rate. The modification of nominal power (W) or the frequency (f) affects the energy delivery on protein crystals as determined by the following relation:

$$\text{Total energy} = W \times \text{total ON pulse length (s)} = W \times s = \text{Joule} \quad (\text{Eq. 2.1})$$

Protein crystals respond to IR laser treatments by raising their temperature, leading to a new humidity equilibrium, hence the crystal contracts. The total change of temperature of a protein crystal can be calculated by the following equation assuming the absorption of a water cube:

$$\Delta T = \frac{W \times X_{\text{tal}} \% \text{absorption} \times \text{Pulse (s)}}{C \times \text{Crystal}_{\text{weight}} \text{ (g)}} = K \quad (\text{Eq. 2.2})$$

Where C is the specific heat capacity of water (4.18 J g⁻¹ K⁻¹). The laser module can be controlled to have an accurate energy dose delivered to crystals. This is performed with a pulsed mode (triggered mode). The minimum pulse length is 50 μs. The maximum frequency is 20 kHz. Figure 2.4 shows graphically the different configurations between frequency and pulse length.

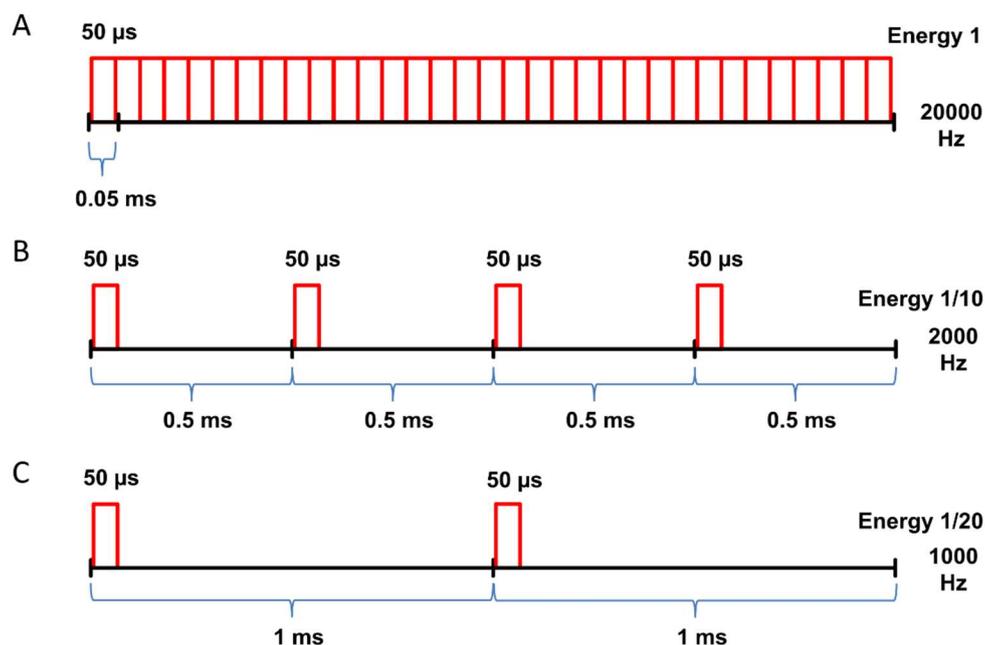


Figure 2.4. Laser energy rates. Energy produced by IR laser can be controlled by varying the frequency. (A) Continuous mode at 20 kHz. (B) Ten times less energy at 2 kHz, ON pulse of 50 μs. (C) Twenty times less energy at 1 kHz, ON pulse of 50 μs.

The laser optics is regulated so that at the focus point the IR beam width is 200 μM (Fig. 2.5A). The optics can be adjusted remotely with a micro-manipulator to set different focus point widths as shown in table 2.4. Using a capillary the laser can be visualized, this effect helps to center and adjust the position and intensity of the laser beam (Fig 2.5B).

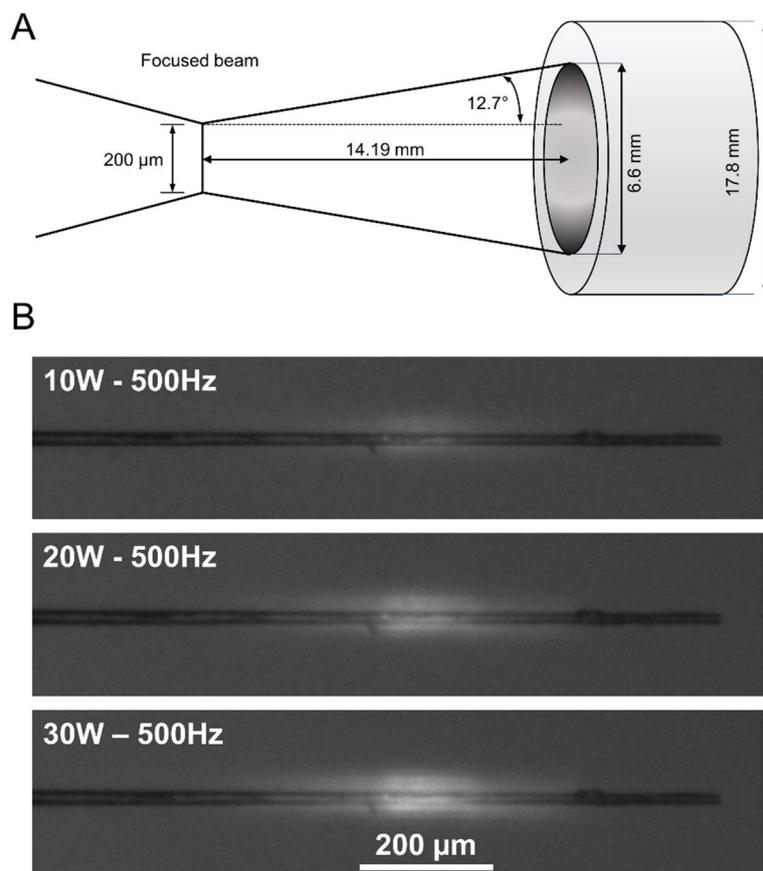


Figure 2.5. Laser optics and alignment. (A) PH 101_02262 optics used to focus IR radiation to 200 μm . (B) Laser alignment is performed using a thin glass capillary. Different laser energies are used as well as a 3-axies micro-manipulator to control remotely the optic laser head, thus the position of the laser and laser focus can be defined.

Table 2.4. Beam size and optics position correlation.

Optics position (mm)	Beam size (mm)
16	1,10231
15	0,65131
14	0,20031
13	0,65131
12	1,10231

2.2.1.4 FML dehydration experiments

FML dehydration experiments are performed using either continuous or pulse laser mode. The laser software allows to control the energy dose by setting frequencies between 1-20000 Hz and power between 0.5-30 W. Using a stable free mounted crystal, the equivalence between laser energy and relative humidity for a specific 2D area projection can be determined, aided by the FMS humidity value. Then, diffraction before and after dehydration is measured as a reference to further treatments. Dehydration experiment design consists of two major parts. First, determination of maximal

dehydration, based on relative humidity equivalence and diffraction images. And second, dehydration speed. The last one can be controlled by using the laser software module, setting watts or frequency ramps, at different increasing rates. Finally, diffraction is once again evaluated to determine changes compared to the reference images.

2.2.1.5 Dropping experiments

Experiments to explore diffraction changes caused by adding solutions or preparation for sample cooling with cryo-protectants were performed using a pico-dropper (Fig. 2.6, A-B) ^[37]. The pico-dropper is a piezo-electrically controlled device that generates small drops (~30 pL) following electric impulses. The pico-dropper works in parallel with the FMS or FML, to concentrate chemicals on crystals. Using a stable free mounted crystal, the 2D area value is used to set two thresholds (Fig. 2.7). The low threshold serves as a reference area point to start dropping. The high threshold serves as a reference area point to stop dropping. Once these parameters are set, a dehydration ramp is performed using either FMS or FML. This dehydration ramp induces crystal contraction which takes the crystal to the lower area threshold. Then, dropping starts until the higher area threshold is reached. This area oscillation implies a constant replacement of water by the newly added solution. The repeated dropping concentrates the new solute, lowering the crystal humidity equilibrium gradually. The end point is marked by a new equilibrium at the final humidity ramp point. Crystals concentrate the solution without experiencing a significant area contraction, therefore not losing diffraction. Alternatively, dropping experiments were performed using a new soaking device developed at Proteros Biostructure named Nebulizer. This device uses ultrasonic vibrations (250 kHz) to nebulize the solution to be soaked (Fig. 2.6C).

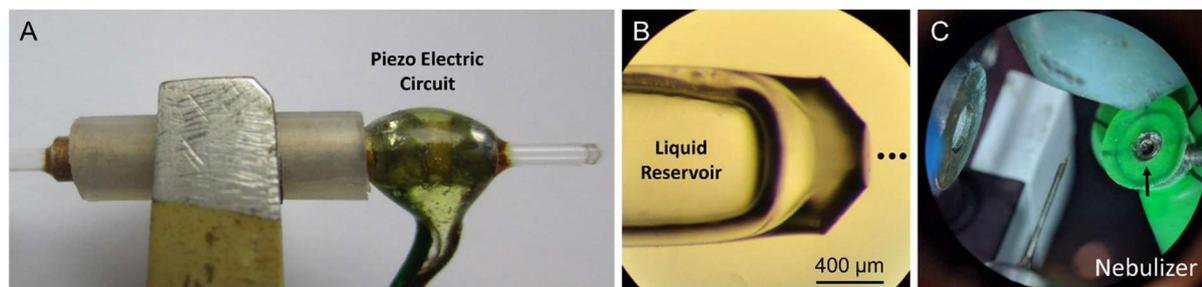


Figure 2.6. Pico-dropper devices. (A) Pico-dropper components. The center glass capillary can be filled with a liquid reservoir to be ejected following the electric impulses and mechanic contractions of the piezo component. (B) Zoom-in on the pico-dropper head and visualization of the drop exit. Black dots represent liquid drops. (C) Nebulizer setup, the black arrow points to the vibrating surface.

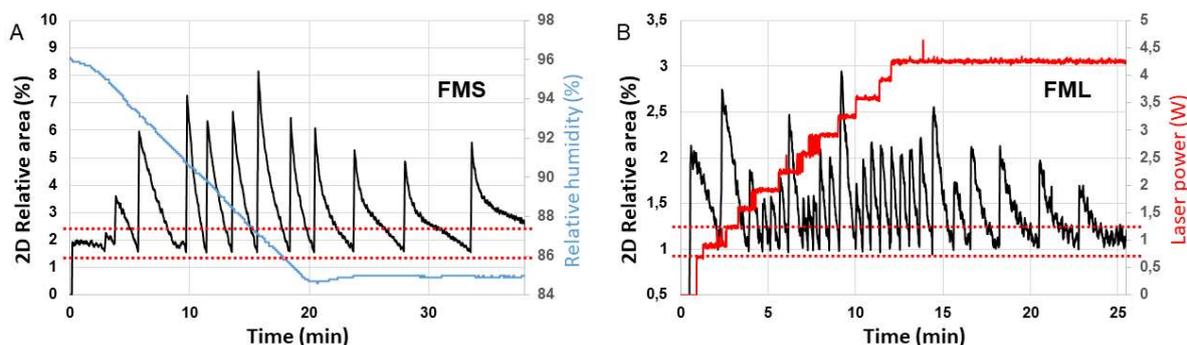


Figure 2.7. Dropping solutions using pico-dropper. (A) Dropping of solution using a FMS dehydration ramp. (B) Dropping of solution using FML dehydration ramp. Both treatments are performed with similar duration dehydration ramps. The low and high threshold are marked with a red dotted-lines.

2.2.1.6 FML annealing experiments

Experiments using protein crystal annealing are performed in two manners. Annealing at room temperature and annealing at 100 K. In the case of annealing at room temperature, crystals are first immersed in oil solution to reduce water diffusion and prevent dehydration. Paraffin oil, mineral oil and perfluoropolyalkyl ether (PFPAE) where used to cover protein crystals. A double glass layer with a spacer is used as a holder to enable 2D area projections, as shown in figure 2.8.

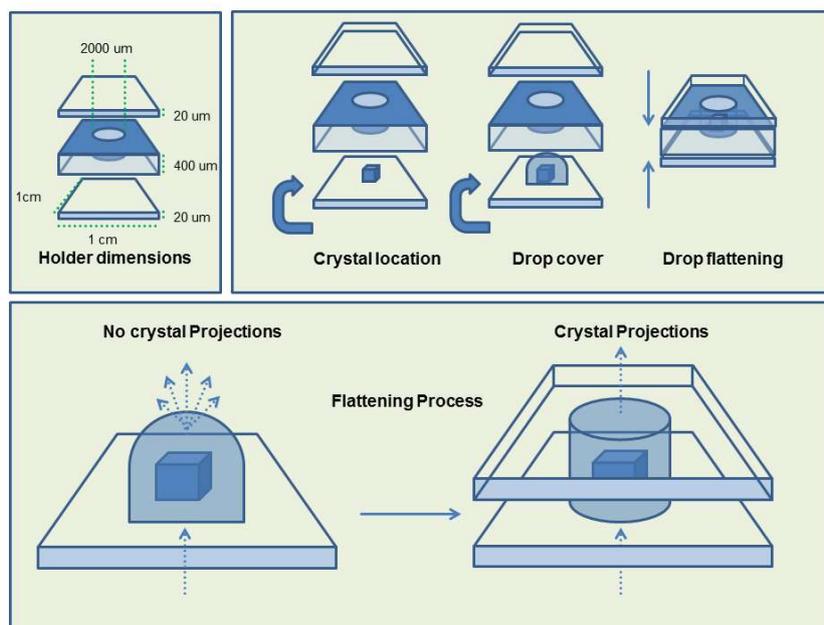


Figure 2.8. Glass-based crystal holder design. The design uses 2 layers of 20 μm glass with a plastic spacer of variable size. The setup avoids loss of water while protein crystals are IR irradiated.

To perform annealing at 100 K, first, crystals are prepared to be cooled by dropping cryo-protectants. Once crystals have concentrated the cryo-protectant, they are flash cooled using a cryo-switch (Fig. 2.1) ^[54]. Depending on the selected method, crystal absorption is modified with respect to room temperature dehydration methods. Additional factors like oil absorption have to be considered when calculating the energy dose. The following formula considers the additional absorption contribution of oil:

$$\text{Total energy} = \frac{W \times \text{Oil \% absorption} \times X_{\text{tal}} \times \text{\% absorption} \times \text{Pulse (s)} \times f}{f_{\text{max}}} = \frac{J \times s}{s} = \text{Joule} \quad (\text{Eq. 2.3})$$

Once crystals are ready to be annealed, a pulse mode is selected using the corresponding laser pulse software module (triggered mode). The pulse mode enables to set pulses as short as 1 ms, and at the same time to define frequency and power. Finally, diffraction images can be collected to determine the effect of each treatment.

2.2.2 Molecular and structure determination methods

2.2.2.1 Molecular cloning

Polymerase chain reaction (PCR) was performed in 50 µl with 1x Phusion HF buffer, 200 µM dNTPs, 0.5 µM of forward and reverse primers, 3% DMSO, 50-100 ng of DNA template and 1U Phusion DNA polymerase (New England BioLabs). The following program was used for amplification in a MasterCycler (Eppendorf): Initial denaturation: 98 °C x 30 sec. With 30 cycles of: Denaturation at 98 °C x 10 sec - annealing 55 °C x 1 min - elongation 72 °C x 30 sec. Final elongation: 72 °C x 10 min. The PCR product was then purified using the QIAquick extraction kit (Qiagen). After restriction enzyme digestion (England Biolabs), the fragments were separated and purified from agarose gels with the QIAquick extraction kit (Qiagen). Similarly, the target vector was digested and purified via gel-extraction. Ligation of the PCR product into the vector was done using T4 DNA Ligase (New England BioLabs) according to the manufacturer protocol. After transformation into *Escherichia coli* TOP10F' and growth at 37 °C overnight, plasmid DNA was purified using the QIAcube (Qiagen) with the corresponding QIAprep Spin Miniprep Kit (Qiagen). The final plasmid vector containing the target insert was checked by restriction enzyme digestion and DNA sequencing. For agarose electrophoresis and visualization of DNA, 1% (w/v) agarose was dissolved by heating in 50x TAE buffer (Applichem) and gels were prepared using a Sub-Cell GT electrophoresis system (Bio-Rad). After mixing with 6x DNA loading dye (Fermentas), the samples were loaded and electrophoresis was performed at a constant voltage of 110V in 50x TAE buffer (Applichem). 1kb or 100bp DNA ladder (peqlab) were used as a size reference. DNA was visualized with SYBR Safe DNA gel stain (Life Technologies), followed by imaging on a gel-doc system (Bio-Rad).

2.2.2.2 Protein expression and purification

Human cDNA of DPP8 isoform 1 (Uni-ProtKB Q6V1X1) was obtained from GeneArt and DPP9 isoform 2 from OriGene (UniProtKB Q86TI2-2). The genes were cloned into pFastBacHTb (Invitrogen) and viruses were generated following the Baculovirus Expression Vector System protocol (Life Technologies). The DPP8_{6His-aa1-898} protein was expressed in 5 L scale in wavebags by infecting *Spodoptera frugiperda* cells (Sf9), harvesting the cells 64 h after infection. The pellet was thawed in 20 mM phosphate buffer pH 7.4, 0.5 mM NaCl, 40 mM imidazole, 5 mM β-mercaptoethanol, and 1 mM NiSO₄. DPP8 protein was purified in three steps: Ni-nitrilotriacetic acid (NTA) (GE) affinity purification, tobacco etch virus protease cleavage of the His-tag, negative affinity on Ni-NTA column

and size-exclusion chromatography on Superdex 200 in 20 mM Tris pH 8.0, 150 mM NaCl, and 2 mM DTT buffer. DPP9_1-892-6-His protein was expressed identically. DPP9-His protein required two purification steps. First, Ni-NTA affinity chromatography purification; and second, size-exclusion chromatography on Superdex 200 in 20 mM Tris pH 8.0, 150 mM NaCl, and 2 mM DTT buffer. Both preparations yielded ~50 mg protein, with a negligible amount of contaminant proteins or aggregates. Analysis of the DPP9-His protein sample used for crystallization by LC-(electrospray ionization-TOF)-MS identified the short and long DPP9 isoforms in about equal amounts. For kinetic and pull-down assays, His-DPP9 isoform 1 (UniProtKB Q86TI2) was purified as described in Pilla et al. [81].

Human SUMO1 protein (UniProtKB P63165) was expressed using pET11 (amp), His-GST-SUMO1 using pETM30 (kan) and GST-SUMO1-linker-SUMO1 using pGEX4T1 (amp). BL21(DE3)Plus cells were transformed and grown overnight. Colonies were selected and grown in fresh medium with the corresponding antibiotic overnight. At 0.6 OD the expression was induced with 1 mM IPTG and grown for 3-4 h at 37 °C. The harvested bacteria was suspended and sonicated in 80 ml 50 mM Tris-Cl pH 8.0, 50 mM NaCl, 0.1 mM PMSF, 1 µM leupeptine, pepstatine and aprotinin plus 2 mM DTT. The lysate was treated with 50 µg/ml DNase I plus 1 mM MgCl₂, 1 h at 4 °C. Finally, the lysate was ultracentrifuged at 100000 g for 1 h at 4 °C. SUMO1 lysate was precleared using a 5 ml Q-Sepharose anion exchange equilibrated in 50 mM Tris/Cl pH 8.0, 50 mM NaCl, protease inhibitors and 2mM DTT. Q-sepharose was removed by centrifugation for 15 min at 250 g, collecting the supernatant. The sample was concentrated and loaded to a Superdex 75 in transport buffer. Purified SUMO1 fractions were collected. The purification yielded 5-10 mg/L of protein. GST proteins were purified using one step GST-affinity chromatography and Superdex 200 size exclusion chromatography in transport buffer.

2.2.2.3 Crystallization and x-ray measurements

Crystallization was performed by the hanging drop method. DPP8 crystals were grown at 4 °C, setting drops in a 1:1 ratio of 10 mg/mL protein and 0.46 M Na-citrate pH 6.75 precipitant solution. Crystals appeared after 1–2 d, mostly in a P2₁2₁2₁ space group. After seeding, crystals with the space group C222₁ prevailed, offering the best diffracting crystal form. Soaks with SLRFLYEG peptide powder were done from 1 h to overnight incubations. This method produced the first set of structures at 3.0 Å resolution. In a further method development, we treated C222₁ crystals with 1 M trimethylamine N-oxide (TMAO) as a cryo-protectant and lattice stabilizer using a free mounting system soaking-based method [87]. We found a diffraction improvement of these crystals up to 2.4 Å. DPP9-His crystals appeared at 20 °C, setting drops in a 1:1 ratio of 20 mg/mL protein and 10% PEG 8000, 25% glycerol, 0.16 M calcium acetate, and 0.08 M cacodilate pH 6.25 as precipitant solution. Crystals were fully grown after 1 wk in the presence of 0-2 mM 1G244. DPP9-His with 1G244 crystals occurred as small 5 µm needles and larger stacks with multiple splits. X-ray measurements were performed in different facilities. In-house rotation data set collection was performed using Rigaku RU-H3R with a MAR345 detector, for measurements with FML or FMS. Also, in-house Rigaku FR-E+ with a MARdtb detector, to perform rotation data set collection at room temperature or at 100 K. Synchrotron measurements of rotation and

serial crystallography data sets were performed at Swiss Lightsource (SLS, Villigen, Switzerland) and Deutsches Elektronen-Synchrotron (DESY, Hamburg, Germany).

2.2.2.4 Structure solution and data analysis

Solution of protein structures can be described in 5 steps. First, reflection finding. Second, image indexation, possibly determining space group. Third, merging and scaling. Fourth, phasing. And finally, structure refinement. Structure solution of protein crystals was done with different software packages. The first three steps were carried out using *XDS suite* [88]. XDS produces a set of hkl coordinates from the reflections after space group identification. Then, using *XSCALE*, the data is merged and scaled. The resulting file can be transformed to an mtz file with *Xdsconv* for further processing, separating a set 5% of reflections to calculate R_{free} . The fourth step uses *Phaser* to determine the correct phase of the structure by molecular replacement [89]. Phaser uses maximum likelihood (ML) to determine the rotation and translation functions. Finally, structure refinement is performed by *REFMAC5* [90]. Likewise, REFMAC uses ML statistic method to maximize the agreement between the model and the x-ray data. REFMAC ML uses two principal variables, previous knowledge of atom position and x-ray data [91], assigning a defined weight to each term depending on the data quality. Each structure has been deposited in the PDB with the following access codes: DPP8-SLRFLYEG C222₁ (PDB: 6EOP), DPP8 unliganded C222₁ (PDB: 6EOO), DPP8-SLRFLYEG P2₁2₁2₁ (PDB: 6EOT), DPP8 unliganded P2₁2₁2₁ (PDB: 6EOS), DPP9-1G244 P12₁1 (PDB: 6EOR), DPP9 unliganded P12₁1 (PDB: 6EOQ) and DPP8-Val-BoroPro C222₁ (PDB: 6HP8).

2.2.3 Functional and biophysical methods

2.2.3.1 SUMO1 oligomer formation

SUMO1 dimers, tetramers and high molecular weight oligomers were generated by crosslinking protein monomeric SUMO1 samples. Bis[sulfosuccinimidyl] suberate (BS3) (Thermo Fisher 2158) was selected as crosslinking agent due to its lysine specificity and water solubility. Incubation times varied from 1 h to 24 h. The selected concentration was 20 mM, due to efficiency producing discrete oligomer forms. The experiment render the best results using 16 μM of SUMO1 protein.

A second approach was to generate recombinant SUMO1 dimeric proteins, cloned and expressed in *E. coli* cells in the pGEX-4T-1 vector (Amersham). First, HIS-GST-SUMO1 protein was designed to form a dimer in solution, where the GST provides the interface to make a stable homodimeric complex. Similarly, the GST-(thr)-(tev)-SUMO1-linker-SUMO1 (~53 kDa) construct was purified. Alternatively, this protein can be cleaved using thrombin to have a SUMO1-linker-SUMO1 homodimer protein, where a 27 aa linker connects both SUMO1 proteins.

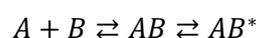
2.2.3.2 Surface plasmon resonance

Protein-protein interaction experiments between DPP9 and SUMO1 or SUMO1 oligomers were done using a BIACORE 8K (GE, Healthcare). CM5 chips were coupled with 200 nM of DPP9 or SUMO1. pH 6.0 and 5.5 were used for sample accumulation on the surface, respectively. NTA chips were coupled

with 0.1-5 μM of DPP9 using the provided NTA activation and loading protocol, including 0.5 mM NiCl_2 to charge NTA groups. Ligand immobilization and chip regeneration was done using 350 mM EDTA. In all cases the reference surface on chips was equally treated with analyte but not having ligand bound to it. The analyte interaction with the surface was 60 sec. The dissociation curve was measured 300 sec or more, depending on the stability of the formed complex. Kinetics and affinity constants were calculated by fitting models against the experimental data. Data fitting was performed using the integral form of the rate equation describing Langmuir adsorption reaction for a 1:1 interaction ^[92]:

$$R_t = R_{eq} \left(1 - e^{-(k_a \cdot C + k_d)(t-t_0)} \right) \quad (\text{Eq. 2.4})$$

Where R_t are the response units at time t , R_{eq} are the response units at equilibrium, k_a and k_d are the association and dissociation constants, respectively, and C is the analyte concentration. Furthermore, the model can be complemented as a “two state reaction model”:



Where A and B are the ligand and analyte, AB is the complex formed following a 1:1 interaction, and AB^* is a conformational change which occurs after the analyte is bound. The quality of fit is determined with the chi-square parameter, which is a measure of the average square residuals and the significance of such values is obtained with the T-value.

2.2.3.3 DPP8 and DPP9 kinetics measurements

Activities of DPP4, DPP8 and DPP9 were determined with a fluorimetric assay detecting the amount of fluorescent 7-Amino-4-methylcoumarin (AMC) produced after cleaving Gly-Pro-AMC (GP-AMC). The reaction buffer was “transport buffer” (20 mM Hepes/KOH pH 7.3, 110 mM potassium acetate, 2 mM Mg acetate, 0.5 mM EGTA) supplemented with 0.02% Tween 20. 12 nM of DPP4, DPP8 or DPP9 were mixed with increasing concentrations of substrate in a total volume of 20 μl per reaction. Serial dilutions of GP-AMC from 0.01 to 0.5 mM were used to measure enzyme activity and normalized with non-cleaved substrate fluorescence. The measurements were performed with a PHERAstar FSX microplate reader (BMG LABTECH) with emission and absorption filters of 350/440 nm. All assays were pipetted simultaneously by an INTERGRA VIAFLO 0383 automatic pipet (INTEGA Bioscience). Fluorescence rate values were obtained from the first 5 min of data of non-saturated points for each substrate concentration. Each treatment was compared to the activity of apo DPP8/9. For inhibitor comparison, a 1% DMSO mock was used for comparison. Kinetics parameters were calculated, $K_{0.5}$ is the concentration of substrate that produces a half-maximal enzyme velocity. The Hill coefficient “ h ” is used to quantify the degree of cooperativity between two sites. If $h > 1$, then there is a positive cooperativity. Fitting of different inhibition models was performed against the fluorescence produced after proteolytic cleavage as function of the substrate concentration in presence of 150 nM of SLRLYEG, 1G244 or Val-BoroPro. Each inhibition model has a defined equation fitted to the data as follows:

Competitive inhibition model:

$$K_m^{observed} = K_m \times (1 + [I]/K_i), Y = V_{max} \times X / (K_m^{observed} + X) \quad (\text{Eq. 2.5})$$

Non-competitive inhibition model:

$$V_{maxin} = V_{max} / (1 + I/K_i), Y = V_{maxin} \times X / (K_m + X) \quad (\text{Eq. 2.6})$$

Allosteric sigmoidal model:

$$Y = V_{max} \times X^h / (K_{0.5}^h + X^h) \quad (\text{Eq. 2.7})$$

Where K_m is the concentration of substrates that gives half of the maximal velocity (Michaelis-Menten constant), K_i is the concentration of inhibitor required to produce half maximum inhibition, $[I]$ is the inhibitor concentration, V_{max} is the maximum velocity and $K_{0.5}$ is half of the maximum velocity for a non-Michaelis-Menten equation (cooperative system). Y is velocity and X is substrate concentration.

2.2.3.4 Pull-down assays

Two hundred nanograms of recombinant DPP8 or His-DPP9 in transport buffer supplemented with 0.05% Tween 20 and 0.2 mg/mL ovalbumin were incubated with SLRFLYEG (110 μ M) or with 1G244 (40 μ M) to allow saturation. Control reactions included His-DPP9 or DPP8 alone. Following a 1 h incubation at 4 °C, bead-immobilized SUMO1 was added to the reactions. The reactions were incubated for 2 h at 4 °C. Next, beads were washed in transport buffer containing 0.05% Tween 20, and proteins were eluted with loading buffer. Finally, the samples were loaded to 10% SDS-PAGE and transferred to a nitrocellulose membrane for western blot. Pulled down DPP8 and DPP9 were detected with respective antibodies.

2.2.3.5 Molecular dynamics

Both proteins, DPP8 and DPP9, were simulated in monomeric states under two different conditions; one considering the co-crystallized inhibitor and second one where the coordinates of the bound inhibitor were removed. Consequently, 4 different systems were simulated; DPP8-lig-removed, DPP8-lig, DPP9-lig-removed, DPP9-lig. All systems were simulated with GROMACS 5.0.^{3 [93, 94]}

3. Results

3.1 Crystal order and post-crystallization treatments

In this study post-crystallization methods are used to reveal diffraction improvements in protein crystals. The mechanistic behind these improvements hereafter are referred to as “transformations” instead of “phase transitions”, to include those improvements which technically are not phase transitions, like dehydration of thermal expansion. Moreover, crystalline states separated by a transformation are referred to as isomorphs.

3.1.1 Response of protein crystals to dehydration

The broad diversity of proteins sets a challenging scenario to study protein crystals. Therefore, is important to address how different protein crystals respond to changes in relative humidity. The dimension of the crystal size and unit cell changes differently depending on the crystal system. Our results indicate that some crystal systems like lysozyme or CODH are quite robust. They can contract their lattice constants by several % points, keeping indexable diffraction patterns. Furthermore, crystal lattice contraction is reversible, in the same way as diffraction recovers after rehydration. Figure 3.1 A and B shows the unit cell dimension of a lysozyme and a CODH crystal, respectively. In both cases the lattice parameters are completely recovered after one dehydration-rehydration cycle. The maximum dehydration achievable before the diffraction pattern is not indexable is roughly 80 %RH in both cases. However, DPP8 crystals, in both crystal forms ($C222_1$ and $P2_12_12_1$), are extremely sensitive to humidity changes. Interestingly, changing RH by 3% does not induce a significant change in the lattice constants, although further reduction of RH abruptly impairs diffraction in a non-reversible manner (Fig. 3.1 C-D).

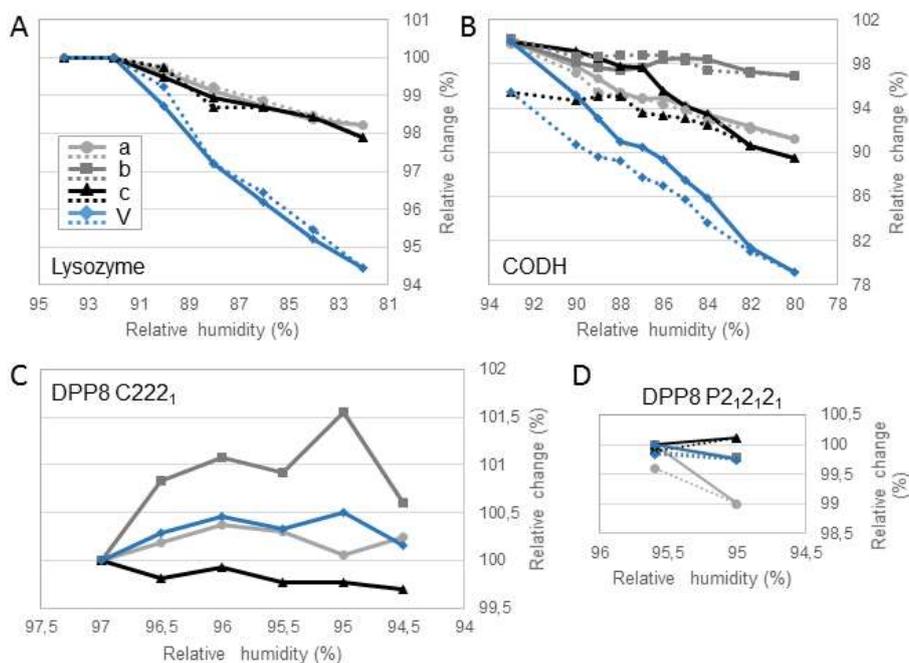


Figure 3.1. Unit cell changes as a function of relative humidity. (A) A tetragonal lysozyme crystal oriented with the c crystallographic axis along the rotation axis. Relative values of “a” axis (circles), “c” axis

(triangles) and calculated unit cell volume (diamond) for the dehydration and rehydration using FMS (94 %RH to 82 %RH) are shown. Diffraction images below 82 %RH did not provide consistent indexing parameters. (B) Unit cell dimensions of a CODH crystal shown as a function of relative humidity. (C-D) Unit cell dimensions as a function of relative humidity for two DPP8 crystals with different space groups. The solid line marks the dehydration from 93% RH to 80% RH. The dotted line corresponds to the rehydration process.

The change in size of the unit cell correlates to the macroscopic crystal size change. 2D projections of a CODH crystal, visualizing both the a-b or a-c crystallographic axis cross-section, show the characteristic hysteresis behavior of this protein crystal system, as the lattice contracts or expands. The dehydration-rehydration is anisotropic depending on the rate of change of each axis. It is also noticeable that the transformation happens between 90 and 89 %RH, as a jump in the continuity of the dehydration-rehydration curve (Fig. 3.2).

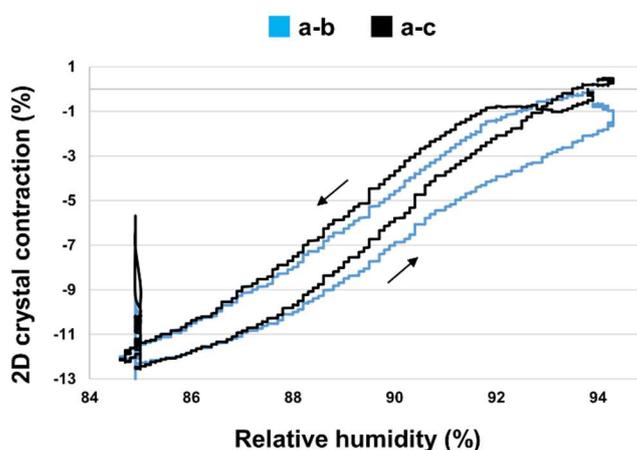


Figure 3.2. CODH anisotropic dehydration. CODH crystal contraction of a-b plane (black line) and a-c plane (blue line) were measured using a FMS ramp from 94 %RH to 85 %RH. Dehydration or rehydration is marked with black arrows. Both, dehydration and rehydration were performed using the FMS. Dehydration and rehydration ramps were run for 20 min.

3.1.2 Laser IR-radiation water absorption

The control of water content of protein crystals using IR radiation might induce changes in the structure of proteins itself. Thus, we sought to determine the absorption ratio between water and protein, as an approximation of a protein crystal. At near-IR wavelength, water has different strong absorption peaks [15]. Figure 3.3A shows an absorption spectrum carried out from 900 to 1000 nm. The absorption contribution of water and protein can be determined by measuring the amount of transferred light through the sample at each wavelength. The relative comparison of absorption shows that pure water absorbs almost 10% of the radiation between 960 and 980 nm. At 938 nm, the IR laser wavelength used in this study to control humidity, water still accounts for the major absorption, but somewhat reduced to 5%. Addition of lysozyme protein, at increasing concentration up to 700 μ M, does not increase absorption at any measured wavelength. Moreover, sample dilution in lysozyme crystallization buffer

further reduces the absorption. This effect can be explained by water remotion. Taking these results into consideration, we assume water as the major absorbing component of protein crystals.

As an additional tool to control crystal radiation absorption, molecules like Epolight 2735 (EPOLIN) can be added to crystals. This dye has a strong IR absorption at 938 nm. The absorption is concentration dependent. A solution of 100 $\mu\text{g/ml}$ Epolight absorbs twice as strong as pure water does (Fig. 3.3B). Therefore, heating of protein crystals is quicker.

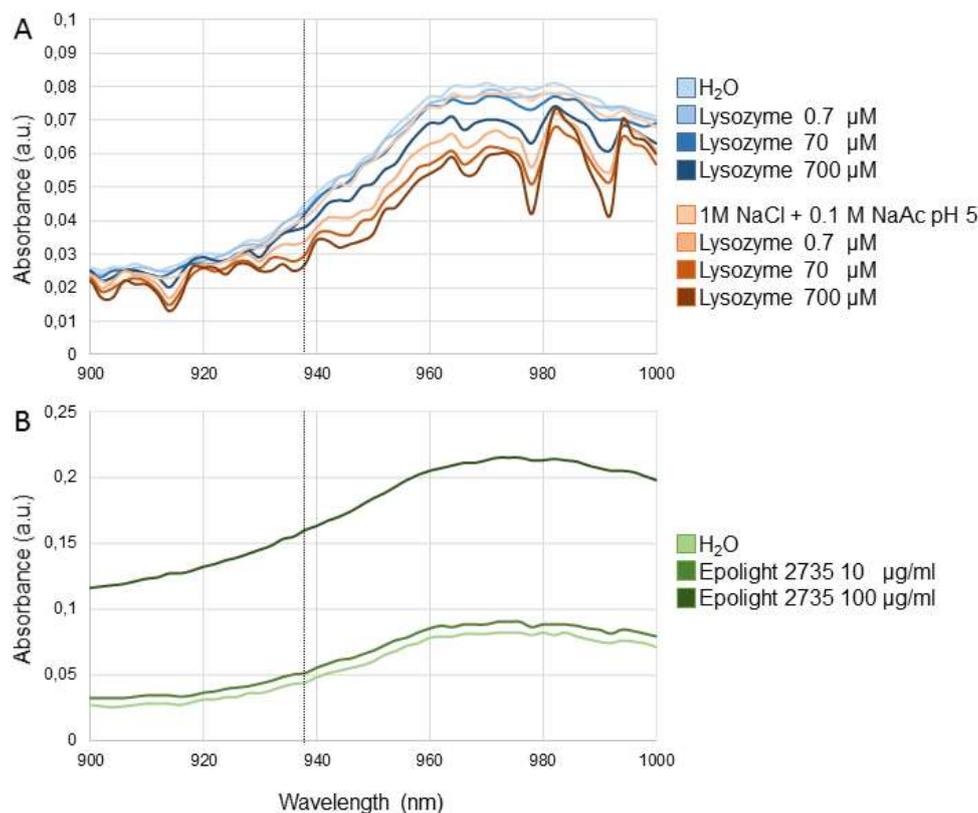


Figure 3.3. Near-infrared absorption of molecules. (A) Absorption spectra from 900-1000 nm of lysozyme protein dilutions in pure water (blue) and crystallization buffer (brown). (B) Absorption spectra from 900-1000 nm of Epolight 2735 IR dye. The dotted line marks 938 nm as the IR-laser wavelength. Absorption spectra were measured using a LUMIstar OMEGA from 220-1000 nm with 2 nm period (BMG LABTECH).

3.1.3 Laser IR radiation increases temperature of protein crystals

Non-contact temperature measurements using an OPTRIS-PI-450 IR camera were performed to experimentally determine the degree of temperature change induced by IR laser treatments on protein crystals. Figure 3.4 shows a significant increment of temperature in the crystal with a certain degree of heterogeneity, most likely due to a not perfect match between laser and crystal size. The measurements determined a temperature change from 25 $^{\circ}\text{C}$ to 34.6 $^{\circ}\text{C}$ within the applied energy range. There are several limitations of the measurement setup, mostly related to spatial resolution. The optics used to zoom-in on the crystal were not sufficient to resolve temperature changes within a protein crystal with confidence. Crystals smaller than 500 μm are not big enough to detect temperature variations above the background levels.

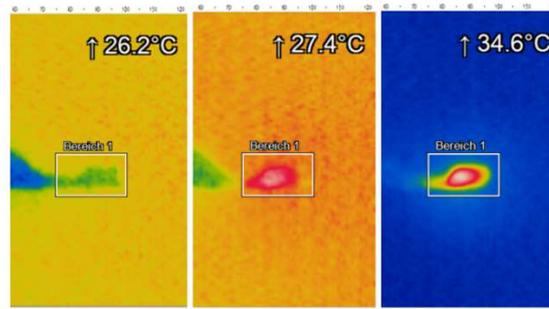


Figure 3.4. CODH crystal temperature change during IR treatment. IR temperature measurements were done with an OPTRIS-PI-450 IR camera. The dimension of the square around the crystal is 2x1 mm. Consecutive exposures of 1, 10 and 30 W at 1000 Hz are shown from left to right. The camera has a spectrum range of temperature sensitivity between 7.5-13 μm and a thermal sensitivity of 0.1 K. 13° optic objective (ACPIO13 OPTRIS).

3.1.4 Dehydration response of protein crystals to IR treatments

Protein crystal dehydration can be performed in diverse manners. Only some features of protein crystal dehydration have been explored using FMS technique, but the limits of crystal dehydration in terms of maximum dehydration, dehydration speed limits and how these correlate to diffraction quality are not well understood. Therefore, to determine the characteristics of protein crystals response to IR radiation, a 200 μm CODH crystal was irradiated with different energies at constant laser focus diameter (200 μm). Figure 3.5A shows the dependency of crystal contraction on applied energy. The measurements were performed for a-b (circles) and a-c crystallographic axes (triangles). A linear correlation was found in both cross-section areas with non-isomorphic characteristics, where a-b contracts the most. Furthermore, calculation of contraction speed of CODH crystals pointed at a speed limit reached with higher energy (Fig. 3.5B). Raising IR laser energy induces higher dehydration speeds, as well as increasing the maxima of the speed limit.

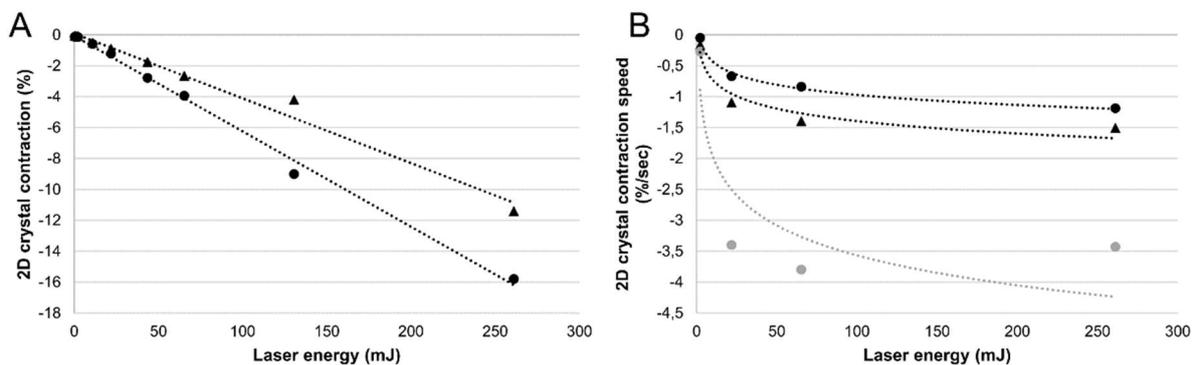


Figure 3.5. CODH dehydration and its dependence on laser power. A CODH 200 μm crystal was oriented with the “a” crystallographic axis along the rotation axis. (A) Values of 2D crystal relative area contraction and (B) speed of contraction for a-b plane (circles) and a-c plane (triangles). Speed calculation was done considering the time interval after laser went on until crystal equilibrium was reached (first derivative). Laser power was 13.6 W (black) and 27.2 W (grey). Crystals were mounted at 94 %RH as initial humidity at FMS. The laser diameter was 200 μm (focus point).

The maximum dehydration speed tolerated by a crystal before its diffraction is damaged and how crystal size influence this parameter was measured on lysozyme crystals. The assay tested diffraction quality before and after dehydration-rehydration cycles. Protein crystal dehydration speed was increased by raising the laser frequency and shortening total pulse time duration, thus the total energy delivered was kept constant (Fig. 3.6B). To ensure equal treatments between different crystal sizes, the laser energy was normalized with respect to the crystal size, considering the square of the crystal cross-section area at constant beam focus diameter. Therefore, the laser energy applied to a crystal with half of cross-section area was four times bigger (Fig 3.6A). Two crystals, 100 and 200 μm in diameter, have their diffraction impaired at the same maximum dehydration speed, with a maximum total energy rate of 10.8 J/s. All treatments, regardless of the energy rate, reached 4 % contraction. This contraction corresponds to a conservative dehydration with good diffracting quality, except for the critical dehydration point (Fig. 3.6C).

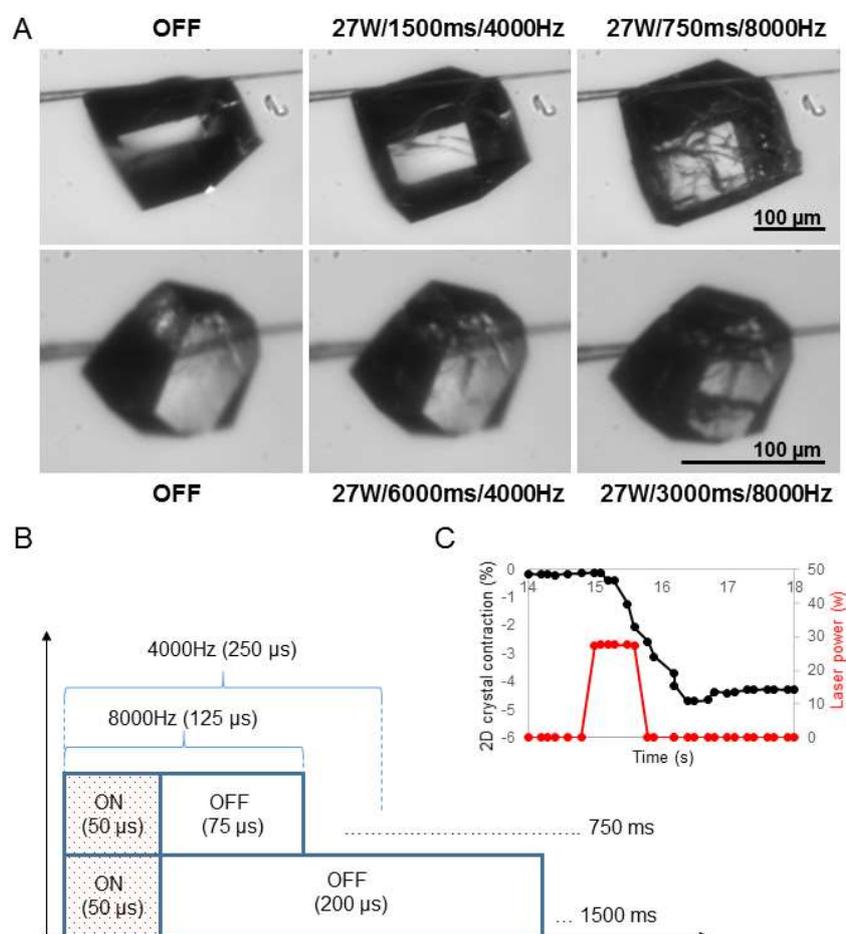


Figure 3.6. Lysozyme dehydration speed limit. (A) Two lysozyme crystals were exposed to IR radiation, at increasing dose rates. The upper panel shows a 200 μm crystal and the lower panel a 100 μm crystal. The laser energy was proportionally adjusted to the crystal cross-section area. Thus, 8.1 J and 32.4 J of total energy were used, respectively. (B) Experimental setup of laser energy dose regarding time pulses for each frequency (4 kHz and 8 kHz). (C) 2D lysozyme contraction response at 27 W/750ms/8000Hz laser treatment. In all cases, crystals contracted to the same extent, with non-disturbed diffraction pattern at 4% contraction of the total 2D area.

3.1.5 Crystal annealing

Two variants of crystal annealing were tested to explore protein crystal response and lattice order. The aim was to establish an experimental methodology and set out control experiments to lay the basis of rational annealing experiment design. The first approach, annealing under oil, seeks to induce lattice shaping without loss of water at room temperature. The second approach, annealing at 100 K, intends to improve diffraction by alternating the temperature by transiently heating a cooled crystal.

3.1.5.1 Crystal annealing under oil

Annealing under oil was employed to improve diffraction of protein crystals at reduced water evaporation conditions. Since water evaporation is minimized, raising crystal temperature becomes the only variable to take into account. Hence, diffraction improvements might be triggered purely by annealing without the influence of dehydration. Crystal response was compared under different oils, following the unit cell dimensions as a parameter to pick an oil with the lowest water diffusion rate (Fig. 3.7). Dehydration was significantly reduced by all tested oils, allowing room temperature measurements without the use of humidity devices over long periods of time (more than 1 h). Perfluoropolyalkyl ether was the mildest oil cover, with a minimum influence on crystal morphology, thus was selected for further annealing measurements with the IR laser (Fig. 3.7C).

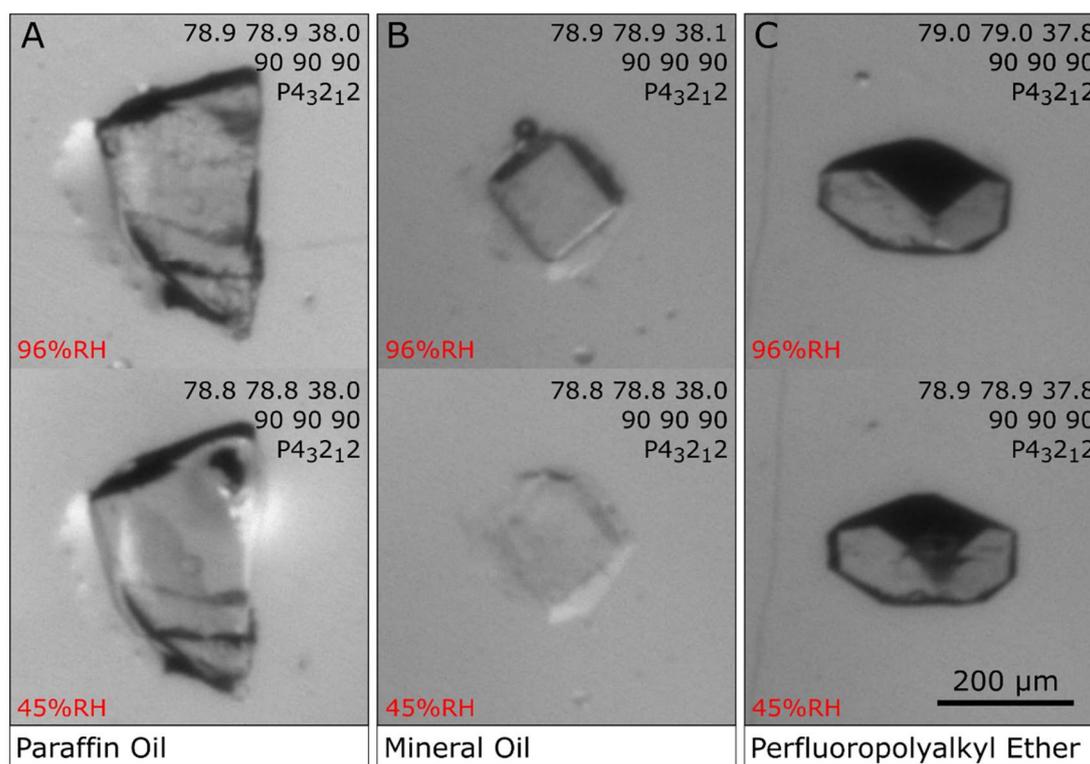


Figure 3.7. Unit cell dimensions of lysozyme crystals under oil. (A) Lysozyme crystal mounted under Paraffin Oil. (B) Lysozyme crystal mounted under Mineral Oil (C) Lysozyme crystal mounted under PFPPE. The crystal size used in all cases was approximately 200 μm . Unit cell measurements were performed after every treatment.

Then, using the experimental set up previously prepared, a single lysozyme crystal was mounted on a customized holder. This holder allowed IR annealing and measurement of crystal projections while immersed in oil, thus reducing water loss. First, as a control, the unit cell parameters of a naked crystal at 96 %RH were collected (Fig 3.8A). Then, the IR laser was turned on to dehydrate the crystal and measure the final dehydration state of a non-protected crystal (Fig 3.8C). Subsequently, the rehydrated crystal was covered with PFP AE, keeping the same orientation, and measuring the unit cell parameters. It is noticeable how good PFP AE preserves both crystal morphology as well as the hydration state, reflected in the invariance of the unit cell dimensions (Fig. 3.8B). Furthermore, treating the covered crystal with IR radiation did not induce dehydration, even at 10 fold higher energy (Fig 3.8D).

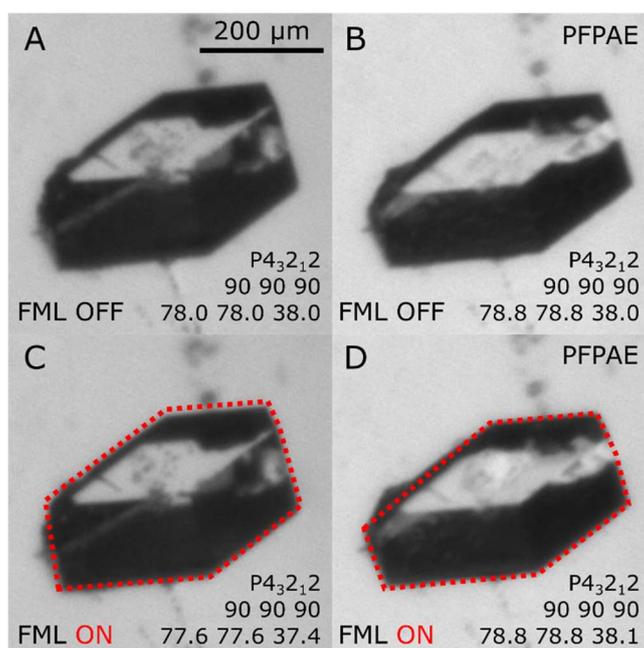


Figure 3.8. Lysozyme crystal IR irradiated under PFP AE oil cover. (A-B) Unit cell dimension of a naked and oil covered crystal, respectively, at 96 %RH. (C-D) Unit cell dimension of a naked and oil covered crystal, respectively, under IR radiation. The laser was setup to 200 μm focus beam, 5 kHz, 10 W and 10 s exposure.

3.1.5.2 Crystal annealing at 100 K: ice phase transition

Annealing at 100 K implies that water is in a frozen state, therefore studying the behavior of pure water at low temperatures can be used as a reference point to understand transformation of confined water in protein crystals. Thus, a methodology to apply IR annealing at 100 K can be better prepared. Ice phase transitions from 100 to 273 K were achieved by raising the temperature with a cryo-nuzzle (Fig 3.9A). The area measurement of a drop of 25% glycerol solution mounted on a crystal holder shows a gradual expansion between 100 and 150 K. The total expansion was approximately 1%. Shortly after 150 K, a sudden expansion of almost 4% marks the appearance of cubic ice. Then, after a plateau, hexagonal ice starts to form until the thawing point is reached at 273 K, where the drop dehydrates rapidly. A similar response can be observed when raising the drop temperature by laser IR radiation (Fig. 3.9B). First a flat phase, where the drop size does not change significantly, with a fast second phase of expansion to

finally dehydration. Interestingly, there are two spikes before reaching cubic ice formation. X-ray diffraction of each phase shows how ice rings appear: From the absence of ice rings in the glass phase (Fig. 3.9C), to a single diffuse ring in the cubic phase (Fig. 3.9D), and three well defined rings in the hexagonal phase (Fig. 3.9E).

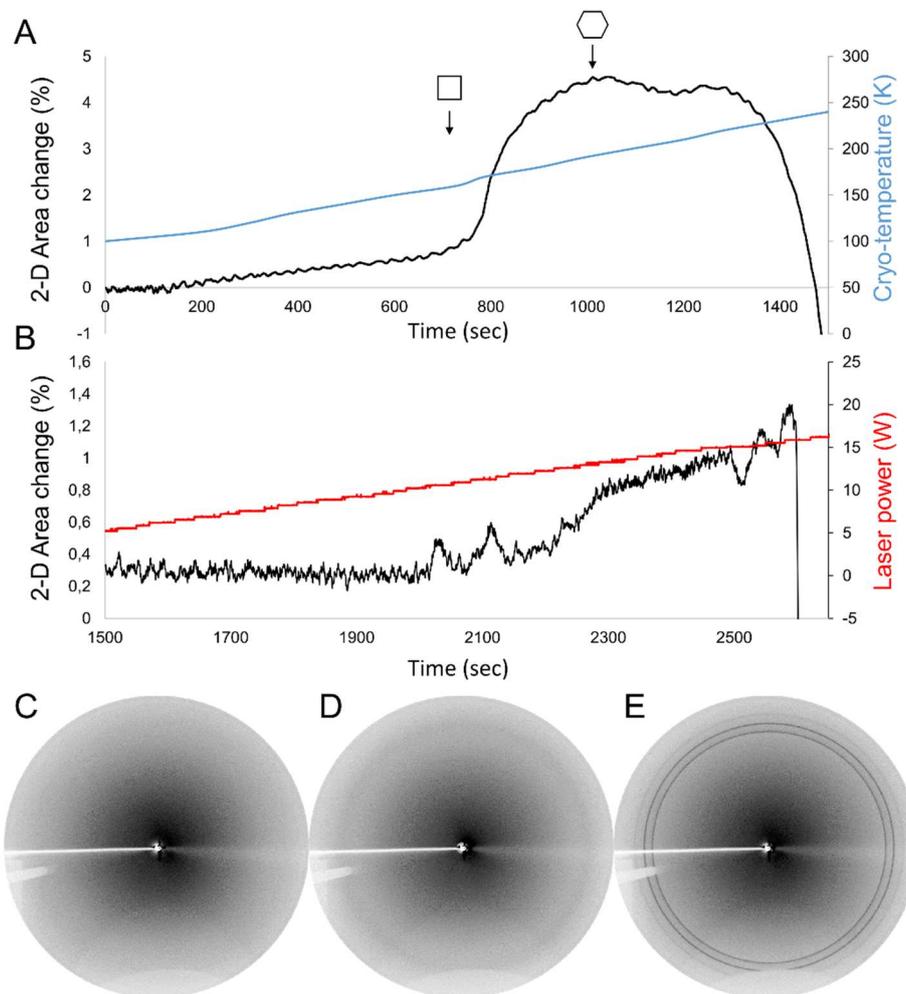


Figure 3.9. Cryogenic water phase transition induced by FML. A cooled glycerol solution (25%, 100 K) was heated up to reveal water phase transitions. (A) The heating ramp from 100 K to 273 K was performed using an Oxford Cryosystems device. (B) The heating ramp was performed using FML system ramping from 0.5 W/1KHz to 20 W/1KHz. (C-E) Diffraction images at different temperatures. (C) Glass state. (D) Cubic ice. (E) Hexagonal ice.

3.1.5.3 DPP9 crystal annealing at 100 K using FML

To establish a controlled protein crystal annealing method aimed to improve diffraction of protein crystals we used FML to heat up DPP9 crystals from 100 K, as base line temperature. A slow IR laser ramp was set to raise crystal temperature. The 2D response of DPP9 crystals was measured as function of laser energy. DPP9 crystals were mounted at 93 %RH, corresponding to their native humidity. Laser energy was raised gradually from 0 to 25 W at 1 KHz. Between 0 and 2000 seconds the 2D area measurement increased due to holder absorption, thus increasing the observable portion of holder in the region of interest. After reaching a stable plateau, the crystal reacts by increasing its 2D area (0.4% Δ area).

Diffraction quality also changed, from 2.99 Å to 2.67 Å after the 22.6 W point (Fig. 3.10). A critical point in this experiment is to find the point where hexagonal ice starts to form. Thus, it is important to check diffraction to determine the presence of ice rings as an indicator of overheating.

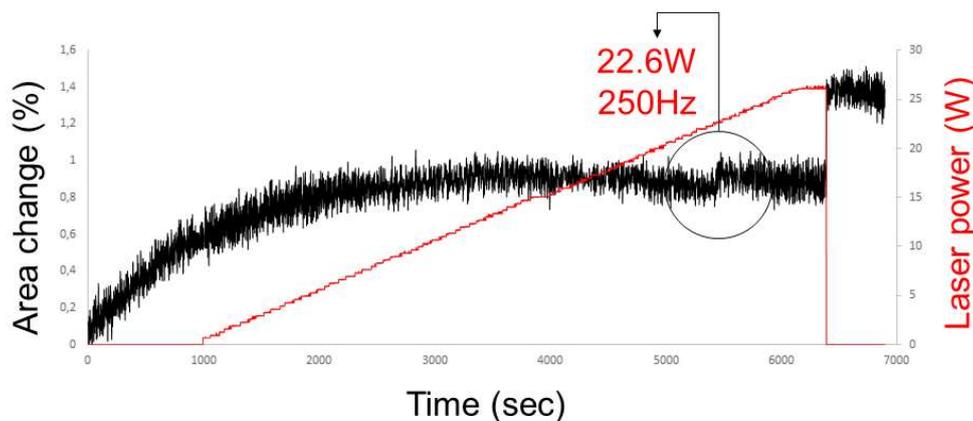


Figure 3.10. FML annealing of DPP9 crystal at 100K. A DPP9 crystal was mounted at the FML at 93 %RH and 2D projections were collected to determine discontinuities in the area curve. The laser treatment was set to ramp from 0 to 25 W at 1 kHz. The crystal was previously cryo-prepared by dropping 1M TMAO.

3.1.6 Chemical modifications

Chemical modification of protein crystals is an avenue to improve diffraction quality. This technique can be performed in several ways, including transferring crystals to a new reservoir. Often this method causes mechanical stress due to excessive manipulation and can ultimately dissolve the crystal. Therefore, mild non-contact techniques are needed for changing the chemical environment of protein crystals, inducing diffraction improvements as well as preparing crystals to be cryo-cooled.

3.1.6.1 DPP8 and DPP9 chemical treatment

The chemical environment of DPP8 and DPP9 crystals was changed to improve their diffraction. Instead of transferring the crystals to a new reservoir, the chemicals to be tested were dropped onto the crystals using either a pico-dropper or a nebulizer device. DPP8 crystallizes in two fashions, $P2_12_12_1$ and $C222_1$. The last one has significantly better diffraction quality, reaching to 2.4 Å in synchrotron datasets. Interestingly, poorly diffracting $P2_12_12_1$ crystals exhibit a transformation after dropping the crystal with 100% DMSO, improving substantially their diffraction quality (Fig. 3.11, A and B). This transformation was not stable in time as DMSO evaporates out of the crystal. The transformation was characterized by a change in the unit cell dimension (Fig. 3.12 C-D).

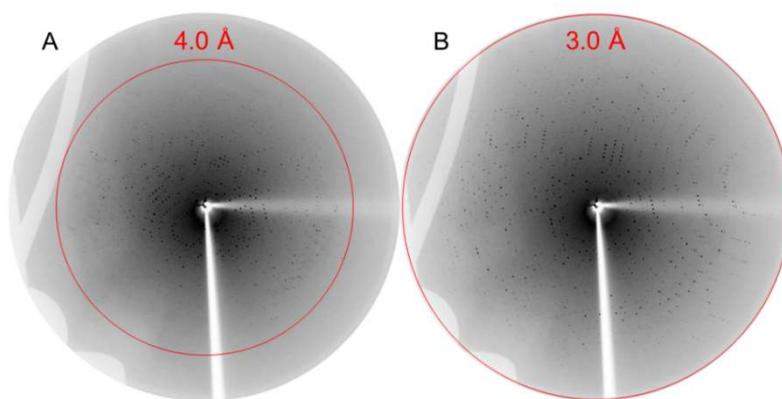


Figure 3.11. DPP8 crystal transformation induced by dropping DMSO. (A) Native X-ray diffraction of a DPP8 ($P2_12_12_1$) crystal mounted at 97 %RH (4.0 Å, collected in-house). (B) Improved X-ray diffraction of a DPP8 crystal mounted at 97 %RH (3.0 Å, collected in-house). The crystal was treated with one drop of 100% DMSO. The transformation occurred in seconds, but was stable for some minutes. The diffraction improvement can be restored adding an additional DMSO drop. Further dropping dissolved crystals. Data was collected using the same crystal and orientation.

In a similar manner, soaking of DMSO can be performed using the newly developed “nebulizer” device. This technique allows the formation of drops of micrometers in diameter, thus significantly reducing the mechanical stress on protein crystals. This method uses ultrasound frequencies to generate vibration, which in turn nebulize drops of solution deposited on the vibrating surface (Fig. 3.12A). In addition to the diffraction improvement observed in DPP8 crystals ($P2_12_12_1$), crystals swell after addition of DMSO, possibly due to its high hygroscopicity. (Fig. 3.12, C-D).

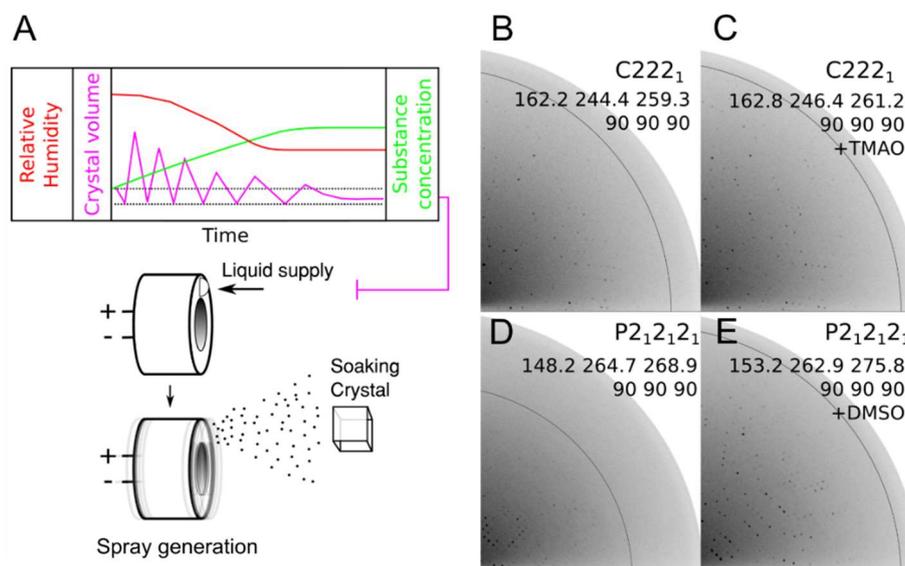


Figure 3.12. Nebulizer soaking method. (A) The upper panel is a trace of parameters during soaking. The magenta line indicates that the crystal area controls dropping of solution. The lower panel shows the nebulizer device and the production of “spray drops” by vibration at 250 kHz. (B-C) DPP8 crystal of space group $C222_1$ soaked with TMAO to prepare the crystal for cryo-cooling. The resolution edge limits are 2.6 Å and 2.4 Å, respectively. (D-E) DPP8 crystal of space group $P2_12_12_1$ soaked with DMSO to expose the transformation and improve diffraction. The resolution edge limits are 3.4 Å and 2.4 Å

The pico-dropping method was then applied to DPP9 crystals to preserve diffraction after cooling crystals. They grew as stacks of rods with high values of mosaicity. Good diffracting crystals were obtained only after carefully soaking crystals with 1M TMAO as cryo-protectant (Fig. 3.13). Similarly, DPP8 crystals ($C222_1$) were treated with 1M TMAO to cryo-protect them and improve diffraction quality from 3.0 to 2.8 Å (in-house data collection) (Fig. 3.12, A-B). TMAO often was observed to induce a slight improvement of diffraction e.g. 0.4 Å, besides protecting crystals from ice formation in almost 80% of all cases of water soluble proteins tested. FMS or FML using 1M TMAO requires a final humidity equilibrium of at least 85 %RH. Each peak in the chart is a dropping event, reflected in the crystal area increment.

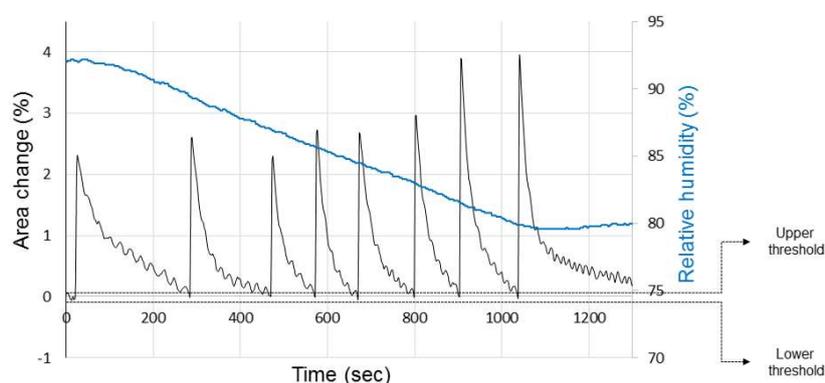


Figure 3.13. Dropping of DPP9 crystal with TMAO as cryo-protectant. FMS was used to set a dehydration ramp from 93 to 80 %RH while dropping of 1 M TMAO was performed using a pico-dropper.

3.1.6.2 Lysozyme crystal structures after FMS and FML treatments

To assess if treating crystals with FML induces changes in the molecular structure of protein crystals, we performed comparative experiments. We dropped TMAO to three different lysozyme crystals of similar size and shape using either FMS, FML or a combination of both. All treatments ensure an equal final humidity, i. e. the concentration of TMAO added to crystals was roughly the same. Crystals were cooled to 100 K and datasets were collected in-house. All structures were refined to similar R-factors with 1.56 Å (Table 3.1). Readily from the data collection it was noticeable that treatments with FML alone rendered broader diffraction spots. This effect might be attributed to higher B-factors. After data processing, the B-factors showed a significantly increased average value (Figure 3.14, B and D). Cooling crystals while the IR laser is ON raises the temperature, ending up with “hot-cooled” structures. Therefore, to avoid this unwanted effect, the humidity had to be FMS adjusted to match the new humidity equilibrium set while TMAO FML dropping. The laser module is able to gradually reduce the power as crystals contract, thus keeping crystal volume constant. After 85 %RH was reached, the crystal was cooled to 100 K. This method showed similar B-factor values as the FMS treated structure, without molecular structural differences (Fig. 3.14, C and D).

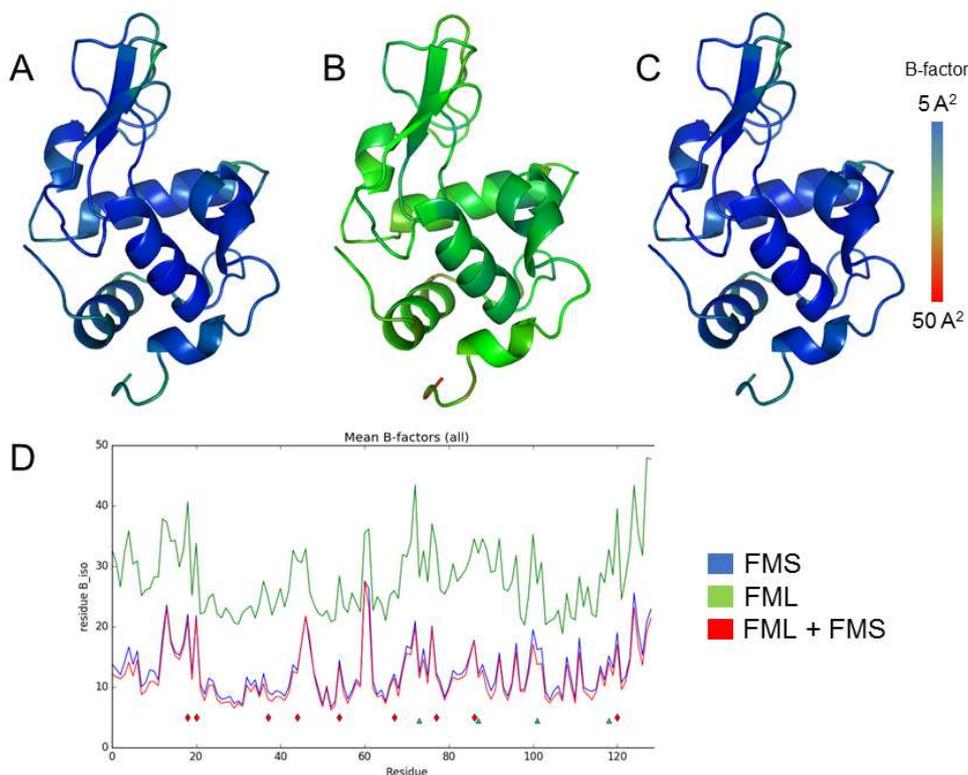


Figure 3.14. FMS and FML dropping effect comparison. (A) Molecular structure of lysozyme after soaking with 1M TMAO, using a dehydration ramp from 96 %RH to 85 %RH by FMS method. (B) Molecular structure of lysozyme after soaking with 1M TMAO, using a dehydration ramp from 96 %RH to 85 %RH by FML method. (C) Molecular structure of lysozyme after soaking with 1M TMAO, using a dehydration ramp from 96 %RH to 85 %RH by FML method. After the equilibrium was reached, the laser power was gradually reduced as humidity was lowered from 96 %RH to 85 %RH using FMS. (D) Comparative diagram for each treatment displaying isotropic B-factors per residue. All data sets were collected with the same parameters and refined to 1.56 Å. From A-C the color code displays isotropic B-factors from 5 to 50 Å². The αC alignment are B > A = 0.129 Å and C > A = 0.136 Å

Table 3.1. Crystallographic parameters for lysozyme treated structures.

	FMS lysozyme	FML lysozyme	FML-FMS lysozyme
<i>Data Collection</i>			
Space Group	P4 ₃ 2 ₁ 2	P4 ₃ 2 ₁ 2	P4 ₃ 2 ₁ 2
Resolution (Å)	33.46-1.58 (1.626-1.58) ^a	31.02-1.59 (1.636-1.59)	39.21-1.59 (1.63-1.59)
<i>Cell dimensions</i>			
a, b, c (Å)	78.3, 78.3, 37.0	77.7 77.7 37.5	78.4 78.4 36.9
α, β, γ (Å)	90, 90, 90	90, 90, 90	90, 90, 90
R-meas	3.8 (9.6)	6.7 (120)	3.1 (9.8)
I/σ	39.69 (17.48)	17.3 (1.83)	47.90 (16.15)
Completeness (%)	97.6 (87.8)	98.3 (92.9)	99.5 (97.3)
Total observations	125259	124293	127904

Total unique observations	15927	15732	15856
<i>Refinement</i>			
R _{cryst} /R _{free}	17.2/22.2	25.7/31.2	16.1/20.4
Number of reflections	15130 (797) ^b	14944 (787)	15062 (793)
RMS bonds (°)	0.020	0.019	0.022
RMS angles (°)	2.030	1.974	2.061
Number of atoms	2337	2166	2412
Average B-factor	13.75	28.32	12.976
Ramachandran plot (%)			
Preferred region	124	123	122
Allowed region	3	4	5

^a Values in parentheses correspond to the highest-resolution shell

^b Values in parentheses correspond to free R-value test set

3.2 Structure and mechanism of dipeptidyl peptidases 8 and 9

To understand in detail the nature of interaction between DPP8/9 and the 1G244, SLRFLYEG or Val-BoroPro inhibitors, we performed crystallographic structural studies of both proteins and their complexes with those ligands. Post-crystallization treatments were applied to improve diffraction quality of DPP8 and DPP9 crystals. Furthermore, we determined kinetics of complex formation to explore the substrate binding mode. DPP8 and DPP9 sequence comparisons against DPP4 suggest the existence of new structural features, which are possible candidates for specific inhibition of these important pharmacological drug targets. Part of the results shown in this section were published separately, describing for the first time the structure and mechanism of DPP8 and DPP9 ^[94].

3.2.1 DPP8 and DPP9 structure solution as homodimer and monomer

Using the structure information and multiple sequence alignment, we mapped and compared the secondary structures of DPP4, DPP8, and DPP9, suggesting the sequence alignment presented in Fig. 3.15. The alignment shows the general similarity between DPP4, 8 and 9, as well as characteristic structural features.

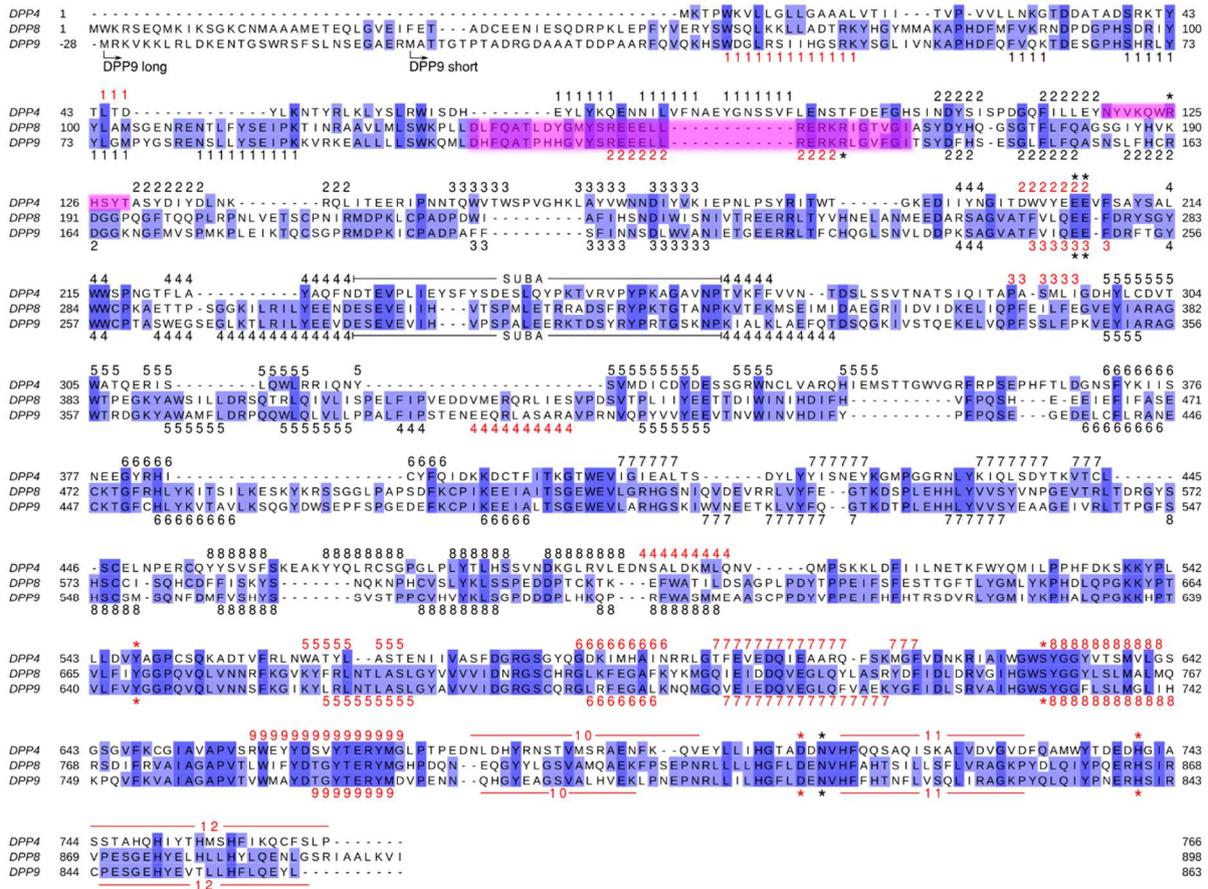


Figure 3.15. Multiple sequence alignment of DPP4, 8 and 9 highlighting structural features. Human sequences of DPP4, DPP8 isoform 1 and DPP9 isoform 2 (long) are shown. Secondary structural elements are marked above (for DPP4) and below (for DPP8/9) the sequence indicating the helix number in red and the blade number in black. The R-loop (DPP4) and the R-segment (DPP8/9) are highlighted in magenta, including the relevant R125/R160/133 marked with asterisks respectively. The SUMO1 binding arm (SUBA) is labeled. The blue gradient represents degree of residue conservation. Asterisks indicate important residues for binding substrates (black) and protease activity (red). The numbering of DPP9 long follows the numbering of the DPP9 short, thus the first 28th residues are negative.

The crystal structure of DPP8 was determined in space group C222₁, refined to 2.5 Å and 2.4 Å for the unliganded and liganded form, respectively. The structures R-factors are 22.9% (R_{free} 25.4%) and 21.4% (R_{free} 23.7%), respectively. The structures were solved using molecular replacement with a DPP4 model (PDB: 1ORV) [95]. There are three polypeptide chains in the asymmetric unit. Two form a non-crystallographic dimer and the remaining molecule forms a crystallographic dimer with a twofold rotation at the “a” axis (0.17 Å average α -carbon rmsd for the three molecules). DPP8 also crystallized in space group P2₁2₁2₁, both in unliganded and liganded form, with six polypeptides (three non-crystallographic dimers) in the asymmetric unit (on average 0.24 Å α -carbon rmsd). All liganded forms show full occupation of the ligand sites, indicating that ligand binding is not influenced by crystal packing.

While DPP8 produced well-ordered single crystals, DPP9 tended to form clusters of crystals whose diffraction images could be reliably processed, but gave high-symmetry R-factors. The unliganded and liganded structures were solved using molecular replacement with the DPP8 model. The space group is P12₁1. The structures were refined to 3.0 Å and 2.9 Å with R-factor 27.3% (R_{free} 33.4%) and 26.5% (R_{free} 33.2%), respectively. There are four polypeptides in the asymmetric unit forming two non-crystallographic dimers with an average α -carbon rmsd of 0.18 Å. A summary of all structure statistics is presented in table 3.2. DPP4, DPP8, and DPP9 are active dimers in solution. Using a PISA server ^[96], we determined similar interface areas for each dimer of ~2,200 Å² with a complexation significance score of 1. The DPP4, DPP8, and DPP9 dimers are compared in Figure 3.16. As the best-defined structure in the series of DPP8 and DPP9 crystals, the DPP8 liganded form was used for comparison with DPP4. The DPP8 unliganded form has interpretable electron density from residues 48–70, 77–105, 109–137, and 165–897, while in DPP8 in complex with SLRFLYEG, residues 48–105, 109–139, and 148–897 are well ordered. DPP9 is less well defined than DPP8 with several additional loops missing electron density in the β -propeller domain. The residues with interpretable electron density are 20–43, 48–79, 82–93, 101–229, 232–266, 270–581, 583–599, and 604–836. Two residues of the His-tag are visible in the DPP9 unliganded structure. The overall DPP-family structure is conserved. The monomer consists of two domains: the C-terminus α/β globular domain, harboring the catalytic triad in DPP4/8/9 (S630/755/730, H740/864/840, and D708/833/708), and the N-terminus β -propeller domain, providing most of the elements required for ligand binding.

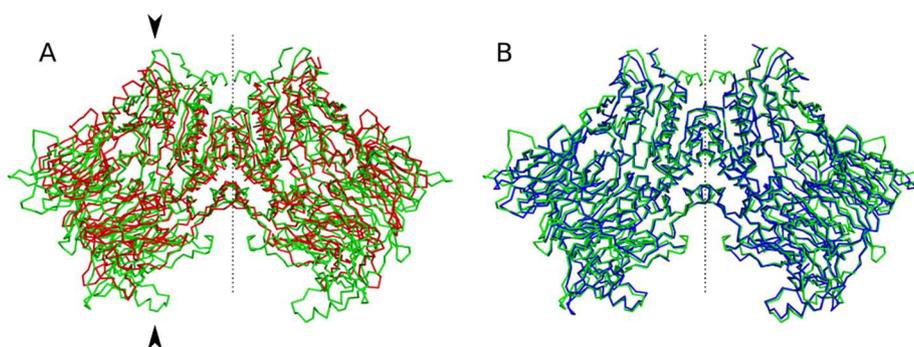


Figure. 3.16. Homodimer alignment and comparison of DPP4/8 and DPP8/9. (A) DPP4 (red) and DPP8 (green). The catalytic and propeller domains of one monomer are marked with arrows at the top and bottom, respectively. (B) DPP8 (green) and DPP9 (blue). The α -carbon rmsd for DPP4/8 is 2.9 Å and 1.0 Å for DPP8/9. The dotted lines represent the homodimer diad axis.

Table 3.2. Crystallographic parameters for DPP8 and DPP9 structures.

	DPP8 / apo			DPP8 / SLRFLYEG			DPP8 / apo			DPP8 / SLRFLYEG			DPP9 / apo			DPP9 / 1G244		
Space group	C222 ₁ (20)			C222 ₁ (20)			P2 ₁ 2 ₁ 2 ₁ (19)			P2 ₁ 2 ₁ 2 ₁ (19)			P12 ₁ 1 (4)			P12 ₁ 1 (4)		
Cell constants	162.84Å	247.06Å	260.85Å	162.83Å	246.37Å	261.19Å	148.12Å	266.78Å	268.09Å	148.16Å	264.68Å	268.94Å	120.37Å	118.02Å	164.46Å	119.36Å	117.21Å	163.40Å
a, b, c, α, β, γ	90.00	90.00	90.00	90.00	90.00	90.00	90.00	90.00	90.00	90.00	90.00	90.00	90.00	105.49	90.00	90.00	105.57	90.00
Resolution (Å)	44.65 – 2.50			44.62 – 2.40			44.46 – 3.10			43.86 – 3.50			49.26 – 3.00			43.52 – 2.90		
% Data completeness (in resolution range)	100.0 (44.65-2.50)			100.0 (44.62 – 2.40)			99.9 (44.46-3.10)			100.0 (43.86-3.50)			99.9 (49.26 – 3.00)			98.9 (43.52-2.90)		
R _{sym}	0.12			0.07			0.18			0.35			0.55			0.16		
< I=σ(I) > ^a	1.3 (at 2.50Å)			2.16 (at 2.40Å)			1.71 (at 3.12Å)			2.01 (at 3.48Å)			1.73 (at 3.00Å)			4.00 (at 2.90Å)		
< I=σ(I) > data set	13.15			22.8			13.7			6.3			5.3			9.52		
Total observations	1535055			1711684			1604806			1118402			1205386			406807		
Total unique observations	180252			203220			191919			133413			88858			95096		
Redundancy	8.96			8.42			8.36			8.38			13.57			4.28		
R _{cryst} , R _{free}	0.229 , 0.254			0.214 , 0.237			0.247 , 0.283			0.205 , 0.262			0.273 , 0.334			0.265 , 0.332		
R _{free} test set	9013 reflections (5%)			10161 reflections (5%)			9593 reflections (5%)			6671 reflections (5%)			4443 reflections (5%)			4755 reflections (5%)		
Wilson B-factor (Å ²)	60.9			61.9			83.2			68.8			56.8			39.1		
Anisotropy	0.434			0.298			0.550			0.442			0.672			0.601		
Bulk solvent k _{sol} (e/Å ³), B _{sol} (Å ²)	0.31 , 31.0			0.31 , 31.2			0.30 , 63.5			0.31 , 63.3			0.33 , 57.3			0.36 , 47.8		
L-test for twinning ^b	< L > = 0.50, < L ² > = 0.33			< L > = 0.49, < L ² > = 0.33			< L > = 0.47, < L ² > = 0.30			< L > = 0.47, < L ² > = 0.30			< L > = 0.52, < L ² > = 0.35			< L > = 0.54, < L ² > = 0.38		
F _o ,F _c correlation	0.93			0.95			0.93			0.91			0.90			0.91		
Total number of atoms ^c	20380			21351			39998			41124			24794			26650		
Average B, all atoms (Å ²)	61.0			62.0			98.0			73.0			56.0			38.0		

^aIntensities estimated from amplitudes. ^bTheoretical values of < |L| >, < L² > for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets. ^cValues calculated considering the amount of molecules in the asymmetric unit. REFMAC 7.0 was used with at least 50 cycles of refinement.

3.2.2 DPP8 and DPP9 β -propeller domain

The β -propeller domain, similar to DPP4, consists of eight blades, which enlase a central round pore. The blades are arranged in two subdomains (blades 3–6 and 1, 2, 7–9) with a β -strand average length of six and eight residues for DPP4 and DPP8/9, respectively. Blades 4 and 5 are longer in DPP8/9 and span up to 13 residues. In DPP8/9, the EE-helix is inserted in blade 4, which arches with a sharp turn at G267/240 toward the active site, forming a helical turn harboring both E275/248 and E276/249, corresponding to the primary binding sites for the substrate N-termini. While blades in DPP4 are invariably four-stranded, the number of strands varies in DPP8 and DPP9. Blade 4 and 5 are tightly packed and intertwined. The fifth β -strand of blade 4 is formed by residues from a loop of blade 5. In DPP8/9, immediately after blade 4, a three-turn helix is observed, not present in DPP4. This helix blocks a surface, which, in DPP4, is the consensus binding region of adenosine deaminase (ADA), offering an explanation for the lack of ADA binding to DPP8/9 [97]. Blade 4 shares with DPP4 a characteristic conserved arm of similar size and position (~34 residues) protruding toward the side opening, named SUBA. This arm structure remains fixed upon substrate binding in DPP4 [81]. R125 in DPP4 is fundamental for substrate fixation, located in a loop of the second blade linking β -strands 2 and 3, named the R-loop (Fig. 3.17A). Due to a general low homology in the propeller domain, this residue aligns with K190 and R163 in DPP8/9 (Fig. 3.15). Instead, the molecular structure shows R160 and R133, respectively, adopting the same structural and functional role. Interestingly, they are provided by a different region of the propeller domain located in the R-segment, at the interconnecting loop between blades 1 and 2 (DPP8: 137–165; DPP9: 110–138) (Fig. 3.17, C and F). Part of this segment folds into the R-helix, which harbors the arginine residues at its C-terminus. The R-helix becomes ordered upon substrate binding, but is mostly disordered in the unliganded forms (Fig. 3.17, B, D, and E).

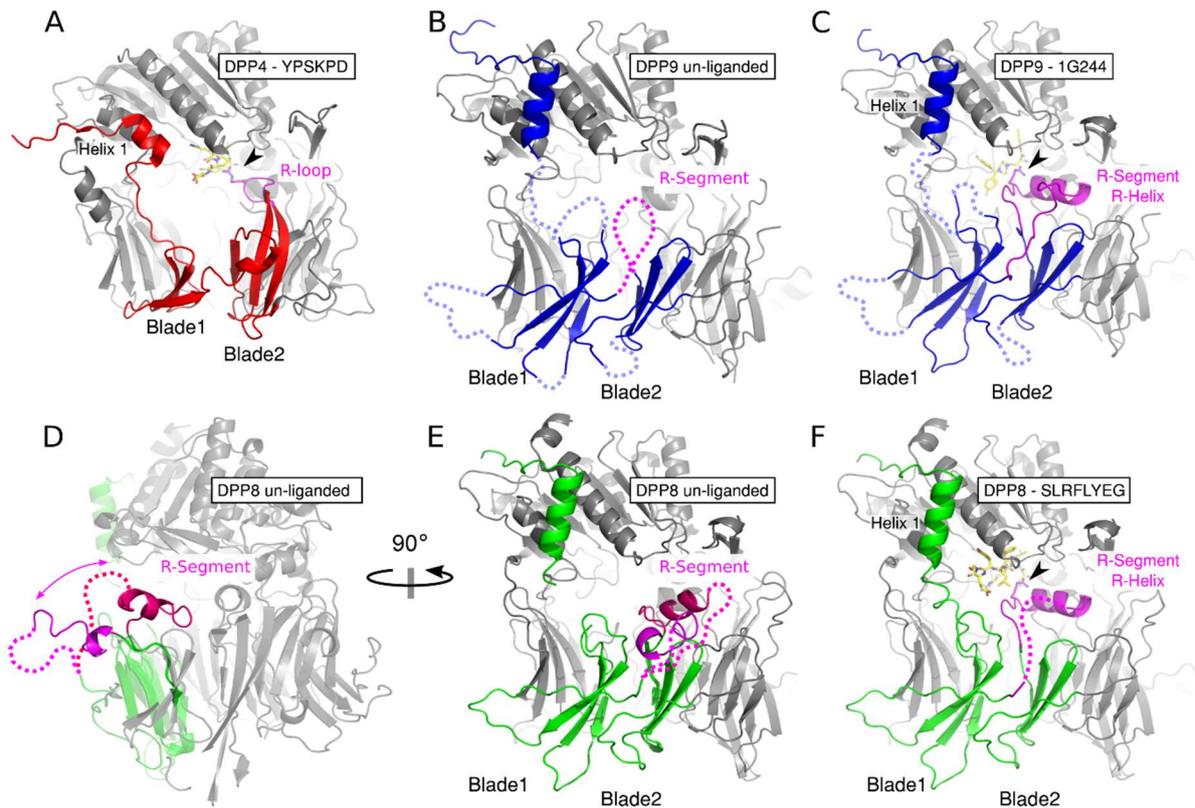


Figure. 3.17. R-segment order/disorder transition in DPP8 and DPP9. The R-loop and R-segment (including R-helix) are highlighted in magenta. (A) Structure of YPSKPD-liganded DPP4 [PDB: 1R9N ^[2]]. (B and C) Unliganded and 1G244-liganded DPP9, respectively. Dotted lines indicate undefined segments. (D and E) Unliganded DPP8 in two orientations favoring the visualization of the opened R-segment conformation. Unliganded DPP8 displays two hypothetical conformations partially adopted by the R-segment, as observed in different molecules of different unliganded DPP8 structures. (F) SLRFLYEG-liganded DPP8. The black arrowheads indicate the position of the relevant R125/160/133. All panels, except for D, have the same orientation. The monomeric structure is presented for simplification.

Disconnected electron density for this helix in unliganded DPP8/9 is visible in some subunits in the asymmetric units, where it adopts a wide range of conformations, suggesting partial order in the unliganded form (Fig. 3.17, D-F; Fig. 3.18). In sharp contrast, in DPP4 such structural change upon ligand/substrate binding has not been observed ^[95].

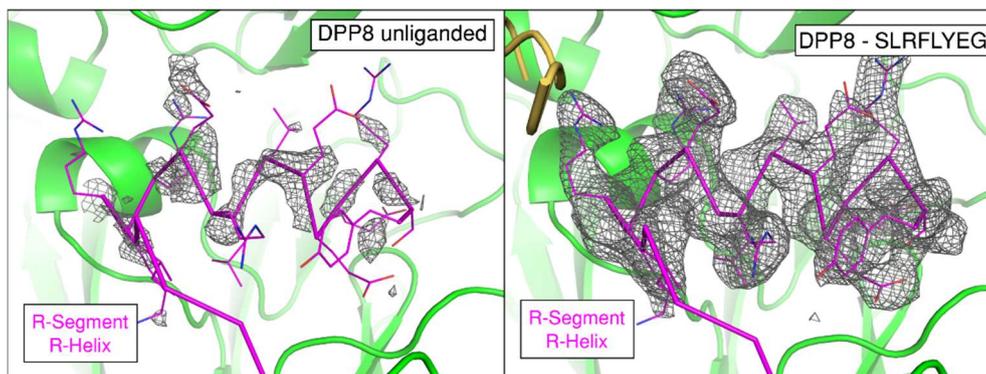


Figure 3.18. Ordering of R-helix upon substrate binding. The left panel corresponds to the unliganded DPP8 structure, zoomed in at the R-helix position. Discontinuous electron density (Fo-Fc - 3 sigma; grey) is attributed to a disordered R-helix. The ordered R-segment of liganded DPP8 structure (magenta) is overlaid as a reference. The right panel shows the omit map (Fo-Fc - 3 sigma; grey) of the ordered R-helix in DPP8 structure interacting with SLRFLYEG (yellow). The monomer structures were aligned with an alpha-carbon rmsd of 0.25 Å.

3.2.3 DPP8 and DPP9 α/β hydrolase domain

This domain is the most conserved region in DPP4, DPP8, and DPP9. In DPP8, it encompasses the C-terminus residues 629–897 with the contribution of an α -helix from the N-terminus of residues 48–70. It is composed of eight parallel twisted β -strands flanked by five close α -helices and three additional more distant helices. These appear to stabilize and link the hydrolase and propeller domains. An interesting observation in DPP8 and DPP9 is the different orientation of the first α -helix in the hydrolase domain. This change causes a shift in the side opening relative to DPP4, thus explaining the different paths followed by the peptides bound in the active site of DPP4 and DPP8 (Fig. 3.19). We observed a strong elongated electron density in a hydrophobic cavity of the α/β catalytic domain, accessible via the side entry. A pentadecanoic acid was modeled with its acidic group fixing the side chain of -R- at P1' in the DPP8 liganded structure. The unliganded DPP8 also has visible, albeit lower, electron density in this cavity.

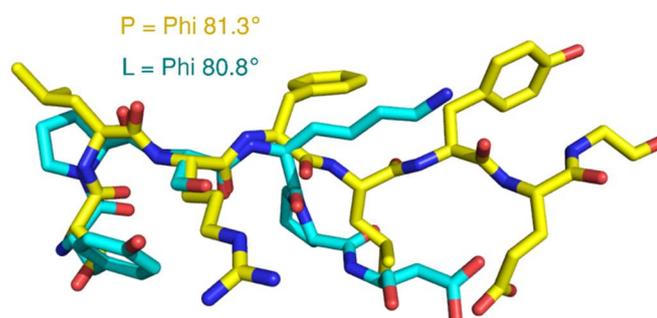


Figure 3.19. Comparison of peptides bound in the active site of DPP4 and DPP8. SLRFLYEG bound by DPP8 in yellow and YPSKPD bound by DPP4 in cyan. Note the difference in relative position caused by the first α -helix of the hydrolase domain. α -carbon alignment of liganded DPP8 and liganded DPP4 monomers with 3.04 Å rmsd. DPP4 PDB: 19RN.

3.2.4 Inhibitor binding exposes active site architecture of DPP8 and DPP9

Next, we analyzed the interaction of DPP8 and DPP9 with two well-characterized inhibitors. 1G244 was developed as a competitive inhibitor of DPP8 and DPP9 [84]. The peptide SLRFLYEG was designed as an allosteric inhibitor of these peptidases. This inhibitory peptide was previously described and developed based on amino acids 61–67 of SUMO1, corresponding to a fraction of the SUMO1 E67 interacting loop (EIL: SLRFLFEGQRIADNH). Furthermore, ELI as synthetic peptide competed with SUMO1 for binding to DPP9 and acted as a DPP9 inhibitor, with a K_i of 5.4 μM (5.6 μM for DPP8) when analyzed with a non-competitive fit. SLRFLYEG shows K_i values in the nanomolar range for DPP8 and DPP9 when analyzed with a non-competitive fit [82]. Surprisingly, the crystal structure of DPP8 in complex with SLRFLYEG disclosed the peptide bound in a substrate-like manner. Crystals soaked overnight with SLRFLYEG displayed a clear difference electron density with well-defined amino acid side chains (Fig. 3.20A). Further refinements including a link between O_γ of S755 in DPP8 and the carbonyl carbon of the scissile -L-R-peptide bond resulted in negative electron density between these two atoms, suggesting a tight non-covalent interaction rather than a tetrahedral intermediate, as had been observed in a peptide complex of serine protease trypsin [98]. The presence of an oxyanion hole and the polarization of the carbonyl oxygen of the scissile bond by a hydrogen bond with the side-chain hydroxyl group of Y669 is a precondition for enzymatic activity. This residue is embedded in a fully conserved segment in all DPPs. The S1 subsite, which accommodates the side chain of the scissile peptide, is the most conserved region among all three proteins. It possesses a conserved particular arrangement of four residues perpendicular to each other (T-shaped), starting with W353/446/420 to Y662/787/762, endowing it with a hydrophobic character (Fig. 3.20C).

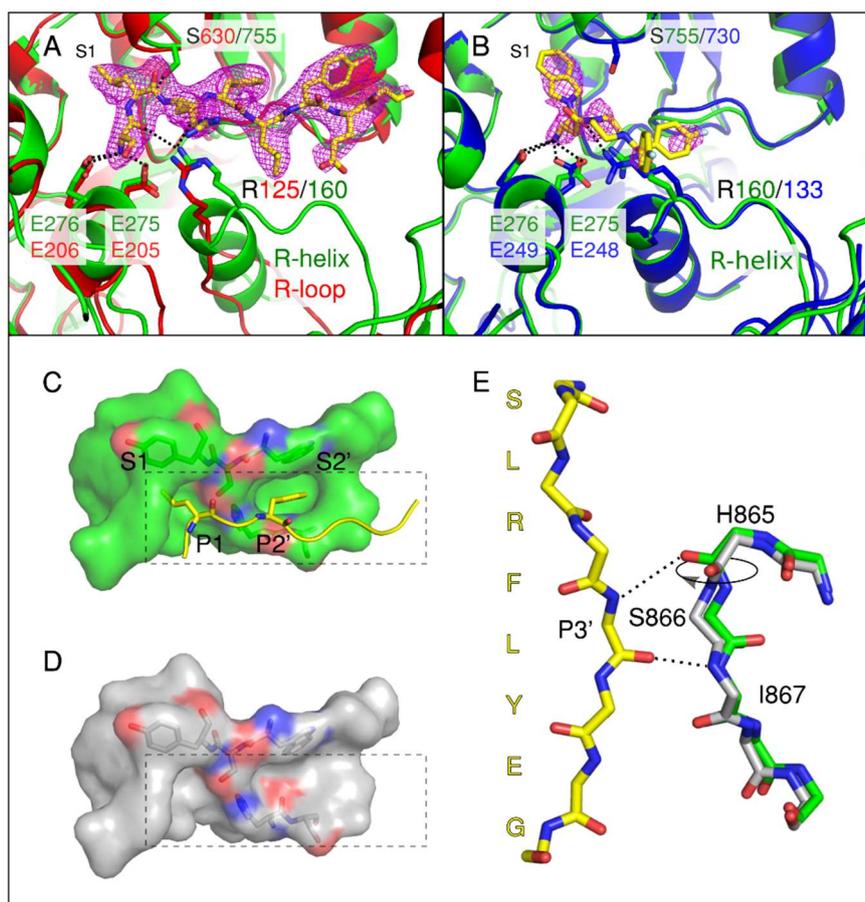


Figure 3.20. SLRFLYEG and 1G244 active site binding and induced fit. (A) Overlay of liganded DPP4 (red) and liganded DPP8 (green). The omit map difference electron density (Fo-Fc) for SLRFLYEG is displayed at 3σ . (B) Overlay of liganded DPP8 (green) and DPP9 (blue). The omit map difference electron density (Fo-Fc) for 1G244 is displayed at 3σ . (C and D) Surface representations of residues forming the S1, S1', S2', and S3' subsites of liganded DPP8 (green) and unliganded DPP8 (grey), respectively. The side chains of P1 and P2' are represented as sticks; the rest are omitted for simplification. Dashed rectangles correspond to the peptide binding region. (E) Parallel β -strand arrangement of SLRFLYEG with the residues H865 and I867 in DPP8. The arrowed circle highlights the psi angle change of H865 upon peptide binding.

Regardless of the high homology of S1, the comparison of unliganded and liganded structures of DPP8 highlights significant differences induced by peptide binding in other subsites with respect to DPP4. One major change is the reorganization of the amino acid sequence H865-S866-I867, not observed in DPP4 (Fig. 3.20, C-E). The H865 psi torsion angle changes from -54° to $+49^\circ$ upon peptide binding. This remodeling generates a parallel β -sheet interaction with the incoming peptide formed between H865/I867 and the P3' residue -L-. It allows the formation of a hydrophobic pocket of S2', where -F- fits (Fig. 3.20, C-E). It is noticeable that the SLRFLYEG peptide is also involved in a β -sheet in native SUMO1 (Fig. 3.21) ^[99].

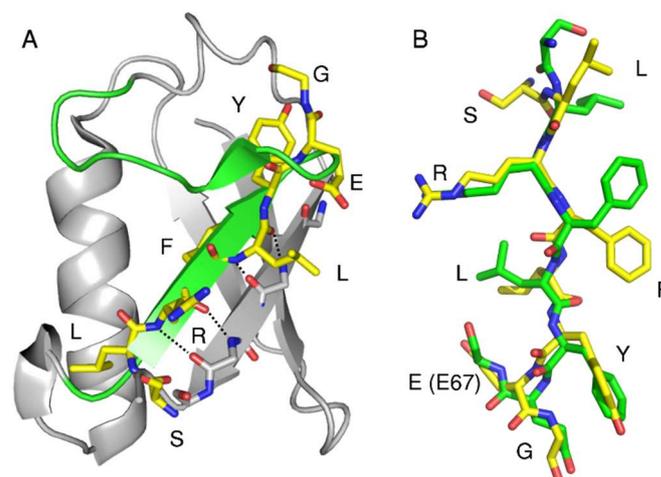


Figure 3.21. DPP8 bound SLRFLYEG peptide aligned with SUMO1. (A) α -carbon alignment of SLRFLYEG (yellow) with EIL (green) of SUMO1. The antiparallel β -sheet hydrogen bonds of SLRFLYEG interacting with the fourth sheet of SUMO1 are shown in discontinuous lines. (B) Alignment of SLRFLYEG with a fraction of EIL (SLRFLFEG). The side chains in both peptides are displayed for visualization. PDB: 2PE6 was used as SUMO1 model [99].

1G244 bound to DPP9 provides further information regarding the active site. Strong electron density for the R-helix is observed in all four monomers in the asymmetric unit, similar to liganded DPP8; this feature is attributed to the ligand bound state. The isoindoline group in 1G244 with a clear difference electron density fills the S1 subsite and hydrophobic pocket. Its amino substituent binds the E248, E249, and R133 side chains. The 1-(4-4'-difluorbenzhydryl)-piperazine substituent is not defined in electron density (Fig. 3.20B). A modeling study based on DPP4 had predicted that the S2 subsite is more voluminous in DPP8/9 [77]. Our results confirm these findings. Three loops in DPP8 and DPP9 present significant differences compared with DPP4 forming the S2 subsite (Fig. 3.22). First, enforced by the sequence change of G355 (DPP4) to N448/422 (DPP8/9), the main and side chain of residue H450/424 is displaced by 7 Å, in an opposite orientation with respect to F357 in DPP4, generating a more spacious S2 subsite (Fig. 3.22A). Second, an additional new feature of this subsite is an extra loop, which buds from the first β -strand of the seventh blade of the β -propeller domain. This loop is absent in DPP4, with the residue H525/500 in DPP8/9 lining the pocket (Fig. 3.22B). A further difference is provoked by the sequence exchange of C551 (DPP4) to Q673/648 (DPP8/9), offering additional contact fixing the SLRFYLEG peptide, here interacting with serine at P2 (Fig. 3.22C).

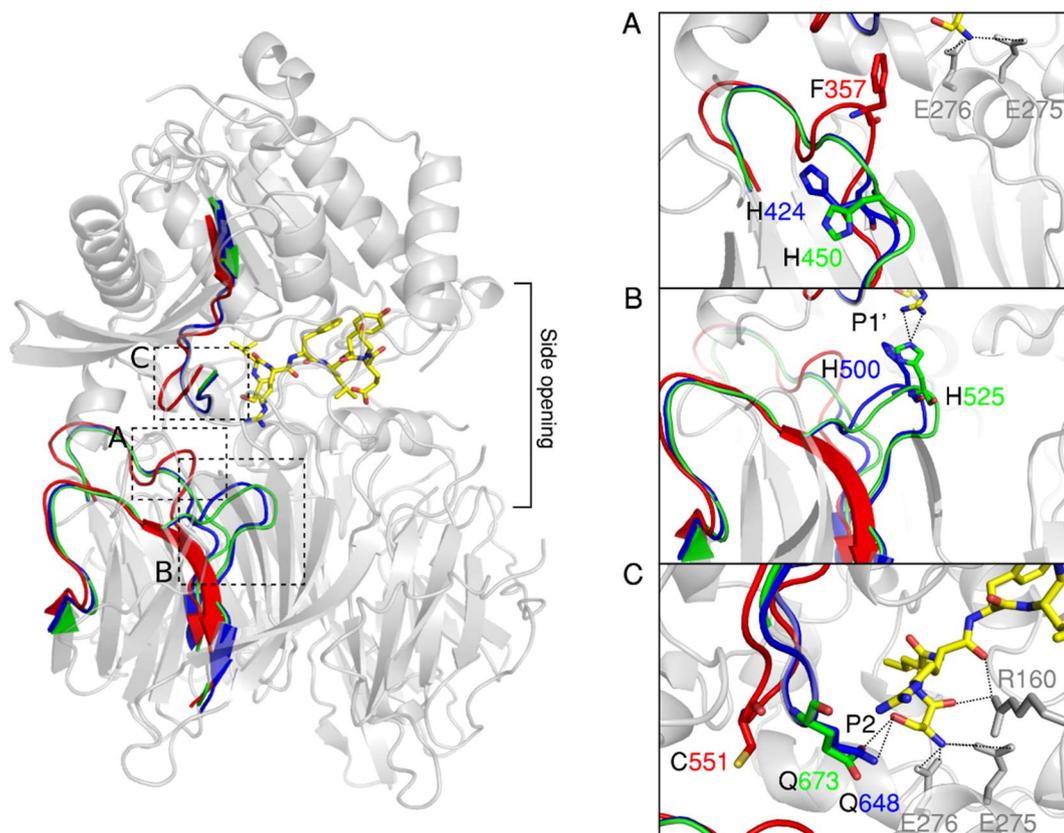


Figure 3.22. S2 subsite loop comparison between DPP4, DPP8, and DPP9. The monomer α -carbon alignment of DPP4 (red), DPP8 (green), and DPP9 (blue) is shown. In the background the liganded DPP8 secondary structure serves as a loop position reference. The main loop differences contributing to the S2 subsite are highlighted and zoomed in at the discontinuous line boxes. (A) F357 in DPP4 is exchanged for the equivalent residues H450/424 in DPP8 and DPP9, respectively. (B) The most different loop of all has an H525/500 in DPP8 and DPP9, whereas this loop does not exist in DPP4. (C) The loop bearing C551 in DPP4 is exchanged for Q673/648 in DPP8 and DPP9, respectively.

3.2.5 Structure of DPP8 interacting with Val-BoroPro

Generating inhibitory specificity against DPP8 or DPP9 proved to be quite challenging. Only small differences in inhibitory power have been found in numerous attempts^[100]. This limitation drove us to try to understand more about DPP8 and DPP9 binding mode. We setup soaking experiments using Val-BoroPro and DPP8 crystals, to expose the details related to binding of small molecules in the active site of this protein. Val-BoroPro was successfully soaked in C222₁ DPP8 crystals diffracting to 2.5 Å (Table 3.3). Interestingly, the binding mechanism for this small molecule is the same as the one observed in SLRFLYEG-liganded DPP8 structure, where the R-segment undergoes a disorder-order transition (Fig. 3.23). Val-BoroPro uses the empty P-orbital centered at boron atom to form a stable covalent bond with the catalytic serine, which resembles the transition state of amide hydrolysis^[101]. This correlates with literature reports, suggesting Val-BoroPro as a strong competitive inhibitor of serine proteases^[102].

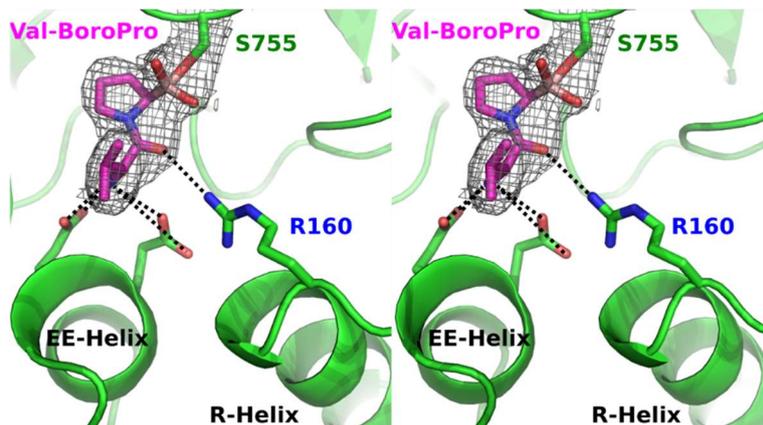


Figure 3.23. Val-BoroPro bound to DPP8. Stereo image of Val-BoroPro electron density (omit map, $F_o - F_c - 3\sigma$) and ordered R-helix interacting with the inhibitor via the R160. The glutamic acids from the EE-helix interacting with the N-terminus of Val-BoroPro are also shown.

Table 3.3. Crystallographic parameters for DPP8-Val-BoroPro structure.

	DPP8-Val-BoroPro
<i>Data Collection</i>	
Space Group	C222 ₁
Resolution (Å)	44.62-2.50 (2.56-2.50) ^a
<i>Cell dimensions</i>	
a, b, c (Å)	163.16 245.32 261.59
α, β, γ (Å)	90, 90, 90
R-meas	14.5 (156.3)
I/ σ	15.64 (1.94)
Completeness (%)	99.9 (99.6)
Total observations	1518649
Total unique observations	179844
<i>Refinement</i>	
R _{cryst} /R _{free}	20.4/22.7
Number of reflections	170840 (8992) ^b
RMS bonds (°)	0.002
RMS angles (°)	0.916
Number of atoms	21264
Average B-factor	48.577
Ramachandran plot (%)	
Preferred region	2380
Allowed region	92
Outliers	9

^a Values in parentheses correspond to the highest-resolution shell

^b Values in parentheses correspond to free R-value test set

3.2.6 DPP8 and DPP9 β -propeller tunnel, active site cavity, and side opening

In DPP4, the active site cavity is connected to the exterior via two pores: a tunnel of ~ 6 Å along the center of the β -propeller domain and a wide side opening of 8 Å (Fig. 3.24A). Based on homology models and sequence comparisons, the existence of a side opening in DPP8/9 was not clear [77]. Interestingly, we see both structures: a tunnel of similar proportions as in DPP4 and a side opening. The latter has variable dimensions, depending on whether a substrate is bound or not. In the unliganded form, the R-segment is not ordered, leaving a wide side opening of ~ 7 Å, close to the values observed for DPP4. In turn, after binding of a substrate, the side opening tightens to a narrow tunnel (Fig. 3.24B). Our data show the 8-residue polypeptide SLRFLYEG bound in the active site of DPP8, pointing to the side opening as the primary access of unprocessed substrates, similar to a DPP4 bound decapeptide (Fig. 3.17, A and F) [2].

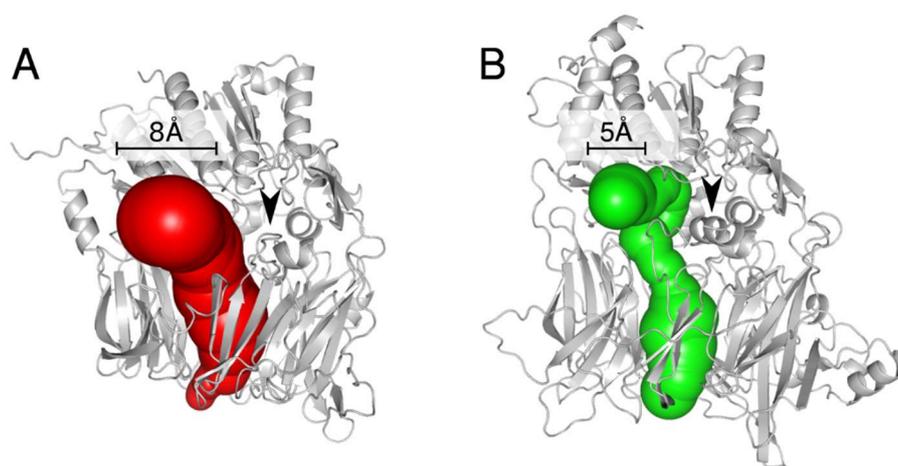


Figure 3.24. DPP4 and DPP8 pore size comparison. (A) DPP4 in complex with YPSKPD [PDB: 1R9N [2]]. The peptide has been omitted to calculate the pore size void volume represented by red spheres. (B) Void volume of DPP8 in complex with SLRFLYEG (omitted for calculation) is shown with green spheres. In both cases, the size of the side exit is indicated above it. The R-loop (DPP4) and R-helix (DPP8) are marked with arrowheads. The structures are displayed in the same orientation.

3.2.7 Inhibitor binding reveals allosteric and cooperative inhibition

The binding of SLRFLYEG to the active site of DPP8 suggests that it acts as a competitive inhibitor. This finding was unexpected since SLRFLYEG is a variant of the EIL SUMO1 peptide, which acted as an inhibitor of DPP8 and DPP9, and competed with SUMO1 for binding to DPP9, suggesting a non-competitive inhibition [81, 82]. Indeed, we find that, similar to the EIL, incubation of DPP8 or DPP9 with SLRFLYEG reduces their interaction with SUMO1, suggesting that SLRFLYEG also competes with the interaction of DPP8 and DPP9 with SUMO1. Strikingly however, incubation of DPP8 or DPP9 with 1G244 leads to a similar effect (Fig. 3.25A). To further study the inhibitory effect of 1G244 and SLRFLYEG, we performed enzyme kinetic assays in the presence of these inhibitors and analyzed the data using non-linear regression. We assumed a non-competitive inhibition [null hypothesis: $V_{\max}^{\text{inh}} = V_{\max}/(1 + I/K_i)$, $Y = V_{\max}^{\text{inh}} \times X/(K_m + X)$] and compared the fitting with a competitive model

[$K_m^{\text{Observed}} = K_m \times (1 + [I]/K_i)$, $Y = V_{\text{max}} \times X/(K_m^{\text{Observed}} + X)$], and vice versa. The extra sum of squares F test was used to compare two equations at a time. This analysis revealed no preference to either model. On the other hand, the interaction of the substrate with the enzyme in the presence of the inhibitor fits with an allosteric model [$Y = V_{\text{max}} \times Xh/(K_{0.5h} + Xh)$], showing a sigmoid behavior (Fig. 3.25B). Furthermore, the average value of the Hill coefficient for DPP8 in the presence of SLRFLYEG was 1.49, suggesting cooperative substrate binding (Table 3.4). Similar observations were made for DPP9 inhibition by SLRFLYEG. The Hill coefficient in the presence of SLRFLYEG also points to a cooperative interaction of DPP9 with its substrate, with calculated average Hill values of 1.28, and an R2 average value of 0.98 (Table 3.4). Consistently, cooperative binding of DPP8 to its substrate was also revealed in the presence of 1G244 (Fig. 3.25C), with a maximal Hill coefficient value of 3.55 (Table 3.4).

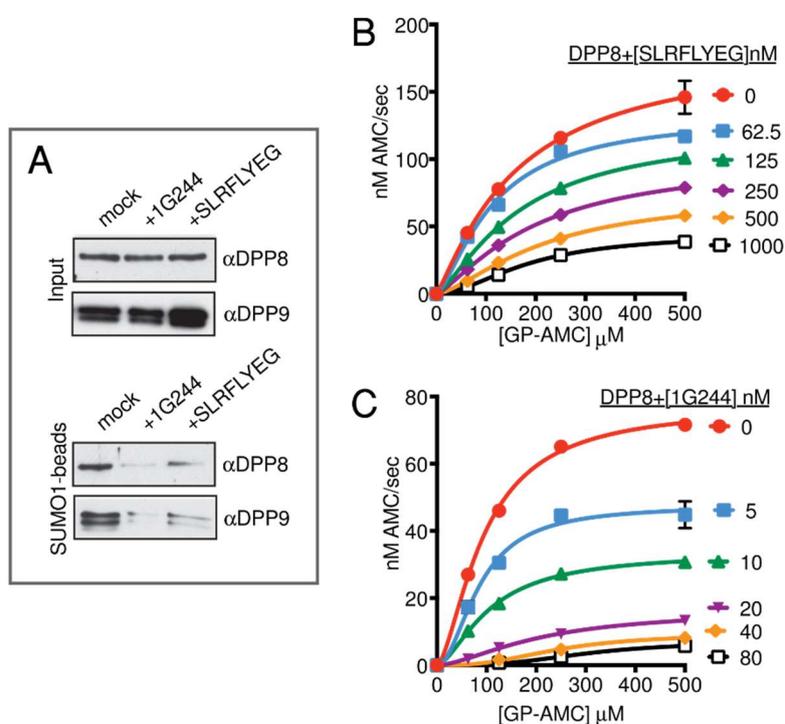


Figure 3.25. DPP8 and DPP9 interaction and enzyme kinetics with SLRFLYEG and 1G244. (A) Pull-down assays with immobilized SUMO1 showing that interaction of DPP8 and DPP9 with SUMO1 is strongly reduced in the presence of 1G244 or SLRFLYEG. (B) Inhibition of DPP8 by SLRFLYEG showing an allosteric fit. The experiment was performed three times in triplicates; shown are the results from one experiment, with error bars within one experiment. (C) Data analysis as above but for inhibition of DPP8 by 1G244.

Table 3.4. Inhibition data for DPP8 and DPP9. Best fit values calculated from a non-linear fit analysis to an allosteric sigmoidal model ($Y=V_{max} \cdot X^h / (K_{0.5}^h + X^h)$). $K_{0.5}$ is the concentration of substrate that produces a half-maximal enzyme velocity. h is the Hill slope, if $h > 1$ positive cooperativity.

SLRFLYEG	0 nM	62.5 nM	125 nM	250 nM	500 nM	1000 nM
Nonlin 1 DPP8						
Vmax nM/sec	183,5±33,5	130,78±5,349	114±7,425	113,7±23,62	64,61±3,55	53,53±4,756
h	1,115±0,3363	1,311±0,07831	1,556±0,2355	1,037±0,1548	1,884±0,2243	1,356±0,1118
$K_{0.5}$ μ M	132,5±51,47	158,7±11,68	116±13,26	335,9±142,2	146,8±12,74	274,4±41,36
R square	0,9676	0,9981	0,9853	0,9917	0,9881	0,9962
Nonlin 2 DPP8						
Vmax nM/sec	156,2±31,05	114±5,089	110,2±6,98	99,44±8,099	59,19±7,285	49,6±4,03
h	1,301±0,4189	1,437±0,1152	1,339±0,112	1,237±0,151	1,581±0,295	1,478±0,164
$K_{0.5}$ μ M	149,4±56,79	143,6±11,59	189,7±22,03	157,2±25,23	183,7±37,44	195,4±27,26
R square	0,948	0,996	0,996	0,9927	0,9755	0,9918
Nonlin 3 DPP8						
Vmax nM/sec	184,5±29,9	133,2±8,84	123,5±7,37	102,3±5,731	74,11±7,6	44,83±4,83
h	1,186±0,247	1,435±0,212	1,365±0,123	1,318±0,092	1,514±0,196	1,866±0,347
$K_{0.5}$ μ M	162,1±53,27	113,9±13,73	166,6±18,31	199,1±20,54	214,9±36,4	185,4±30
R square	0,9763	0,9878	0,9952	0,9972	0,9889	0,9721
Summary of 9 repetitions						
Vmax nM/sec	174,6±21,4	127,5±9,69	114,2±9,7	101,8±12,7	65,1±4,74	48,18±3,37
h	1,191±0,226	1,393±0,196	1,419±0,200	1,201±0,174	1,653±0,197	1,562±0,146
$K_{0.5}$ μ M	148±36,84	135,3±18,9	148,9±23,16	204,9±50,16	175,9±20,82	207,1±23,63
R square	0,9467	0,9613	0,9554	0,9655	0,9637	0,979
SLRFLYEG	0 nM	31.25 nM	62.5 nM	125 nM	250 nM	500 nM
Nonlin 1 DPP9						
Vmax nM/sec	266,4±34	212,7±19,8	228,6±17,7	189,9±53,6	211,9±49,6	217±53,5
h	1,487±0,167	1,262±0,087	1,202±0,05	1,37±0,26	1,278±0,12	1,145±0,07
$K_{0.5}$ μ M	141,9±29,1	168±26,85	212,8±27,84	175,6±78,47	269,2±93,75	428,3±154
R square	0,981	0,9966	0,9987	0,9601	0,994	0,9982
Nonlin 2 DPP9						
Vmax nM/sec	404,4±122,5	254,3±52,11	219,3±26,5	186,3±19,6	216,6±33,4	166,4±35,7
h	1,197±0,172	1,284±0,179	1,319±0,125	1,419±0,129	1,16±0,089	1,234±0,14
$K_{0.5}$ μ M	245,2±121	165±56,3	150,7±30,96	135,7±23,1	234,4±61,3	212,8±75,75
R square	0,981	0,9966	0,9987	0,9601	0,994	0,9982
Nonlin 3 DPP9						
Vmax nM / sec	400,8±128,5	216±37,8	423±184,5	182,2±17,04	190,8±29,8	117,8±12,4
h	1,228±0,186	1,36±0,215	1,046±0,122	1,259±0,08	1,277±0,11	1,388±0,11
$K_{0.5}$ μ M	246,6±125	132,9±39,8	450±312	163,7±25,99	213,5±53,19	153,2±26,29
R square	0,9852	0,9778	0,9913	0,9964	0,9949	0,9943
Summary of 9 repetitions						
Vmax nM/sec	384,2±67,41	218,7±25,39	250,9±38,17	210,2±34,9	207,1±32,76	155,1±30,75
h	1,226±0,105	1,319±0,125	1,197±0,09	1,296±0,137	1,229±0,09	1,244±0,126
$K_{0.5}$ μ M	238,9±67,19	145,9±29,02	210,1±54,01	184,4±49,97	240,2±60,67	225±71,87
R square	0,9838	0,9754	0,983	0,9717	0,987	0,979
1G244	0 nM	5 nM	10 nM	20 nM	40 nM	80 nM
Nonlin DPP8_1						
Vmax nM/sec	145,5±9,40	141,5±15,34	99,72±1,87	72,58±4,62	45,04±10,2	22,26±1,53
h	1,487±0,218	1,211±0,199	1,564±0,0675	1,347±0,127	1,223±0,200	1,843±0,138
$K_{0.5}$ μ M	115±13,3	153,1±33,22	117,4±3,851	168,8±19,93	339,3±134,6	282,1±25,97
R square	0,9874	0,9876	0,9988	0,9949	0,9875	0,9965
Nonlin DPP8_2						
1G244	0 nM	5 nM	10 nM	20 nM	40 nM	80 nM
Vmax nM/sec	153,2±8,53	106,7±2,62	82,39±2,06	50,31±3,50	25,29±1,95	16,47±0,934
h	1,291±0,127	1,434±0,0773	1,502±0,0865	1,404±0,166	1,799±0,212	2,151±0,185
$K_{0.5}$ μ M	135,3±14,57	115,8±5,17	114,3±5,10	149,9±19,07	202,5±23,41	240,1±17,37
R square	0,9951	0,9984	0,998	0,9916	0,9893	0,9946
Nonlin fit DPP8_4						
1G244	0 nM	31,25 nM	62,5 nM	125 nM	250 nM	
Vmax nM/sec	177,6±38,6	51,54±5,2	29,49±3,18	15,98±1,8	8,329±0,5	
h	1,067±0,323	1,45±0,243	1,373±0,135	1,956±0,266	3,559±0,662	
$K_{0.5}$ μ M	156,2±74,21	157,9±28,72	279,6±50,54	270,4±39,76	250,8±14,21	
R square	0,9682	0,9822	0,9945	0,9883	0,9855	

3.3 DPP8 and DPP9 interact non-covalently with SUMO1

Experimental data shows that DPP8 and DPP9 interact with SUMO1 in SUMO-immobilized beads pull-down assays. However the interaction is in the μM affinity range. The complex between DPP9 and SUMO1 monomers was not observed in size-exclusion chromatography ^[81]. Pull-down with beads ensures a high local concentration of SUMO1 molecules, possibly allowing avidity to take place, thus increasing binding affinity. Therefore two molecules of SUMO1 can interact with DPP9, which is a dimer in solution. Furthermore, this interaction can be blocked by pre-incubations with DPP8/9 inhibitors (Fig. 3.25A). To determine the existence of avidity, and to obtain a stable complex formation between DPP9 and SUMO1, we explored the effect of SUMO1 oligomerization on binding affinity and complex formation.

3.3.1 SUMO1 oligomers purification

The first experimental approach to obtain SUMO1 oligomers was to use the bi-functional lysine-specific crosslinker BS3. SUMO1 and BS3 ratio was crucial to obtain controlled oligomerization (Fig. 3.26A). 16 μM of SUMO1 and 20 mM of BS3 yielded the best results. Incubation times longer than 2 h at room temperature did not increase the amount of oligomers. The reaction was stopped adding an excess of tris buffer, quenching the free BS3 molecules. After the sample was crosslinked, the monomeric SUMO1 had an increased electrophoretic mobility on SDS-PAGE gels (Fig. 3.26A). This effect might be explained by intra-molecular crosslinks, thus rendering a more compact SUMO1. Separation of crosslinked SUMO1 by size-exclusion chromatography generates a discreet monomer peak. However, dimers and tetramers are somewhat overlapped (Fig 3.26, B and C). Interestingly, monomers of SUMO1 elute together with dimers and tetramers, probably due to a rather heterogeneous monomer sample, generating a broad peak. Therefore, using this chromatography approach, dimer and tetramer of SUMO1 were always mixed with SUMO1 monomeric species. However, the separation of enriched samples of each specie was successful (Fig. 3.26D).

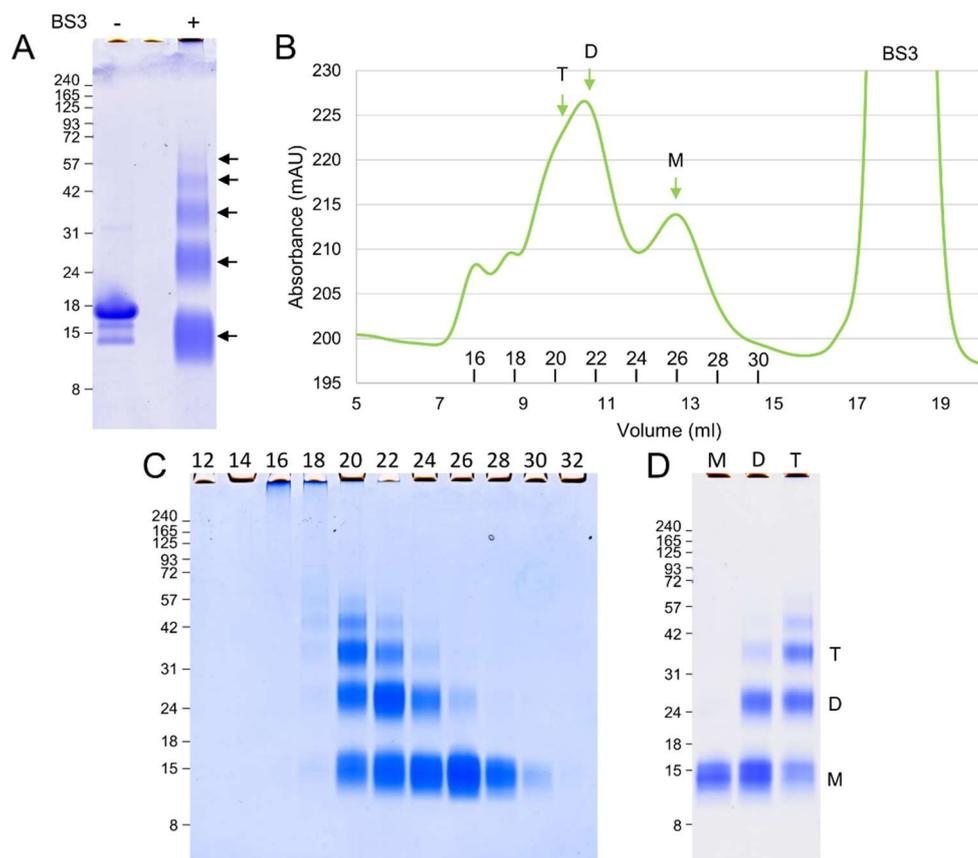


Figure 3.26. SUMO1 oligomers formation. (A) 16 μ M SUMO1 were treated with 20 mM BS3 during 2 h at room temperature. Sample was quenched with 200 mM TRIS pH 7.0. The arrows mark different oligomer sizes of SUMO1. (B) Size separation of different oligomeric forms of crosslinked SUMO1 sample. Monomer, dimer and tetramer retention volume are marked with arrows. (C) 10% SDS-PAGE of SEC fractions from 12 to 32. (D) Pooled fractions of monomers (26-30), dimers (22-25) and tetramer (19-21).

3.3.2 DPP9 and SUMO1 complex formation

Then, we sought to determine if complex formation between SUMO1 and DPP9 depends on the oligomeric state of SUMO1. To do so, we determined the retention volume by size exclusion chromatography of SUMO1 oligomers and DPP9 (Fig. 3.27, A, C-D). Then, these retention volumes were compared with a mixed sample of SUMO1 oligomers plus DPP9 (Fig. 3.27, A and B). A shift in DPP9 retention volume was observed. Furthermore, an increased amount of SUMO1 co-elutes with DPP9 in comparison to the respective fractions of SUMO1 alone (Fig. 3.27B, asterisks). Altogether, the data suggest a complex formation between DPP9 and oligomeric forms of SUMO1. Possibly this complex becomes more stable as SUMO1 oligomers increase in size. Otherwise this complex is unstable with a very short half-life, offering an explanation to the difficulties to obtain the complex in solution using monomeric SUMO1.

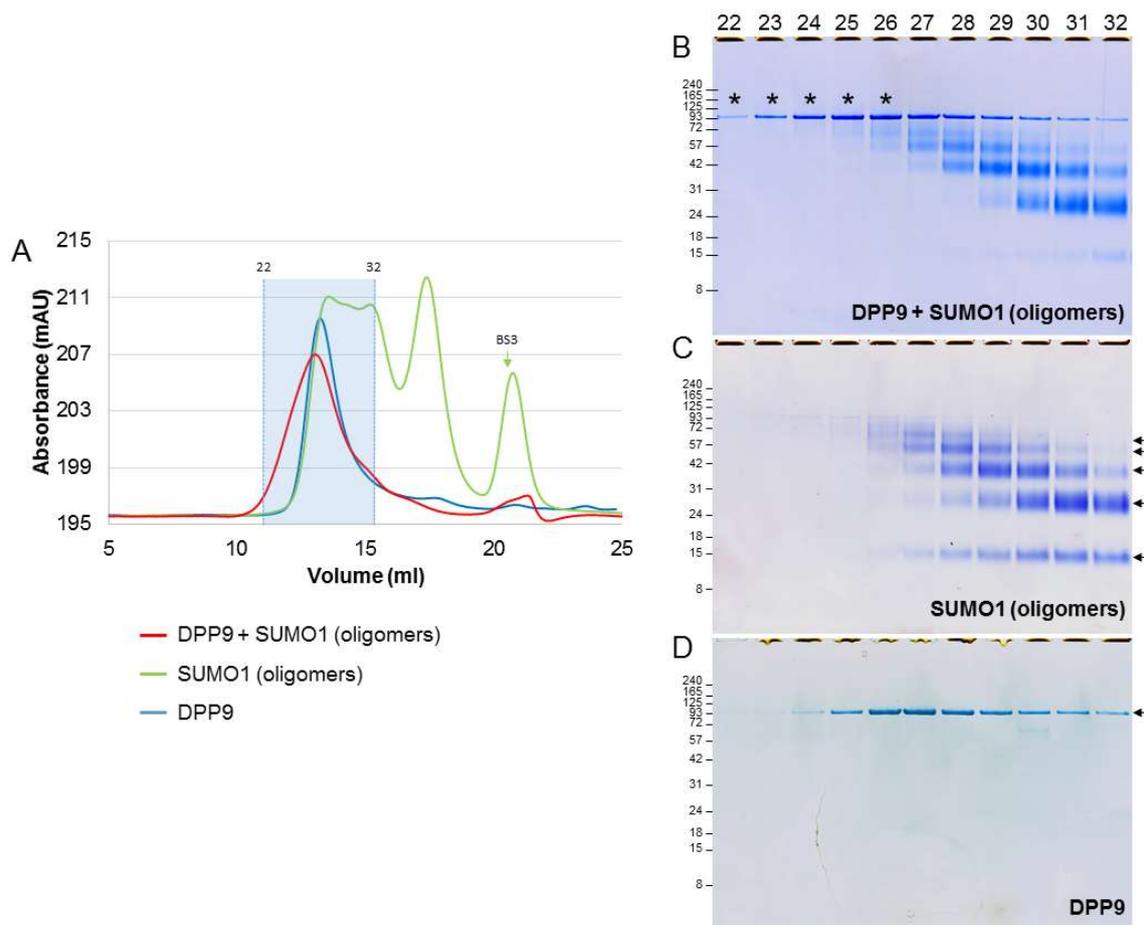


Figure 3.27. DPP9 forms a stable complex with SUMO1 oligomers. (A) SUMO1 protein was crosslinked with 20 mM BS3 and subsequently BS3 was quenched and removed. Samples including DPP9 (blue) and oligomer of SUMO1 (green) were ran independently in size-exclusion chromatography, as reference to compare their retention volumes. Then, after sample pooling from volume 12 to 16 ml, the mix of DPP9 plus SUMO1 was run (red). (B) Denaturing gel of DPP9 plus SUMO1 (oligomers) sample. The asterisks mark the fractions attributed to a complex. (C) Denaturing gel of SUMO1 (oligomers) sample. The arrows mark the different SUMO1 oligomers, starting with the monomer at ~15 kDa. (D) Denaturing gel with DPP9 sample. The arrow marks the monomer of DPP9 at ~95 kDa.

3.3.3 Binding of crosslinked SUMO1 to DPP9

To quantify the affinity of complex formation between oligomers of SUMO1 and DPP9, we measured binding kinetics using SPR. As a first approach, monomeric SUMO1 was used as analyte to measure interaction on DPP9 crosslinked lysine N-coupling chips. The results did not show binding, thus the experimental setup was inverted, using SUMO1 as ligand, to mimic the experimental conditions of the pull-down experiments. In this case a concentration-dependent binding was observed. The later result was repeated using NTA-Chips. Once again, specific binding of SUMO1 was measured (Fig. 3.28A). The result showed a very fast dissociation rate constant (in the order of a second), accounting for an unstable complex, in the micro molar range of binding affinity (KD of 28 μ M).

Remarkably, a crosslinked sample of SUMO1 with BS3 increased significantly the complex stability. All previously purified crosslinked fractions showed strong stable complex formation, where the tetramer or aggregated fractions bind the strongest, with an increased affinity and K_D of ~ 173 nM (Fig. 3.28B). The two state reaction model fits best the measured data, and correlates to the proposed structural mechanism for DPP8/9, indicating that after a conformational change, the K_{off} decreases from $1.9e^{-2}$ to $5.8e^{-4}$. Moreover, as expected from literature reports, parallel experiments using crosslinked SUMO2 did not show binding, regardless of the oligomeric state of the sample, supporting the notion of specific binding between SUMO1 and DPP8/9 (Fig. 3.28B).

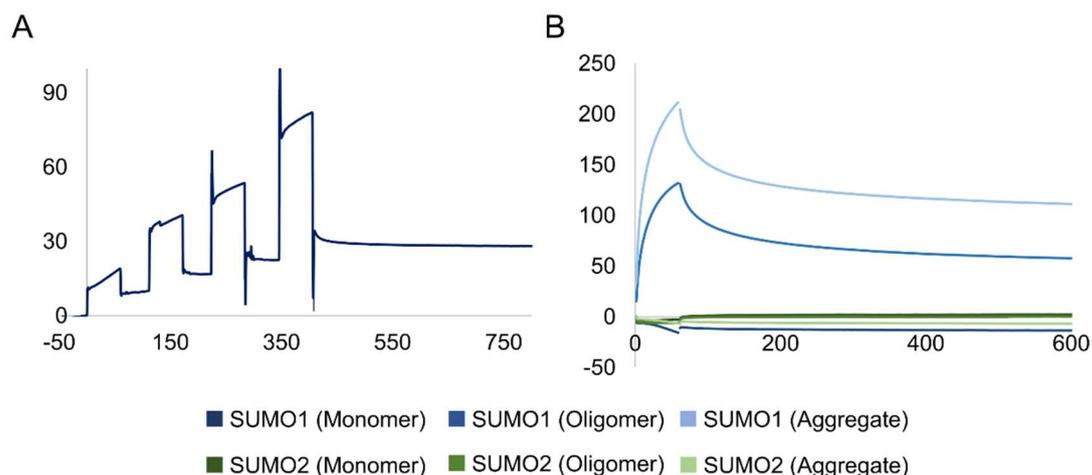


Figure 3.28. Increased binding of DPP9 to oligomeric SUMO1. (A) SUMO1 binding titration. 0.5 μ M DPP9-His was immobilized on a NTA-Chip. Increasing concentration of SUMO1 (20, 40, 80 and 160 μ M) were tested. (B) Crosslinked samples of SUMO1 or SUMO2 were used to measure interaction with DPP9-His. All treatments were performed using the same protein concentration (12.5 μ M).

To further characterize if the observed SUMO1 binding in these and previously reported assays is affected by molecules binding to the active site, we measured DPP9 binding to immobilized SUMO1 in presence or absence of three well known DPP8 and DPP9 small molecule inhibitors. The binding of DPP9 dimer is impaired by all three molecules to different extent, in agreement to what it was observed in pull-down experiments (Fig. 3.29).

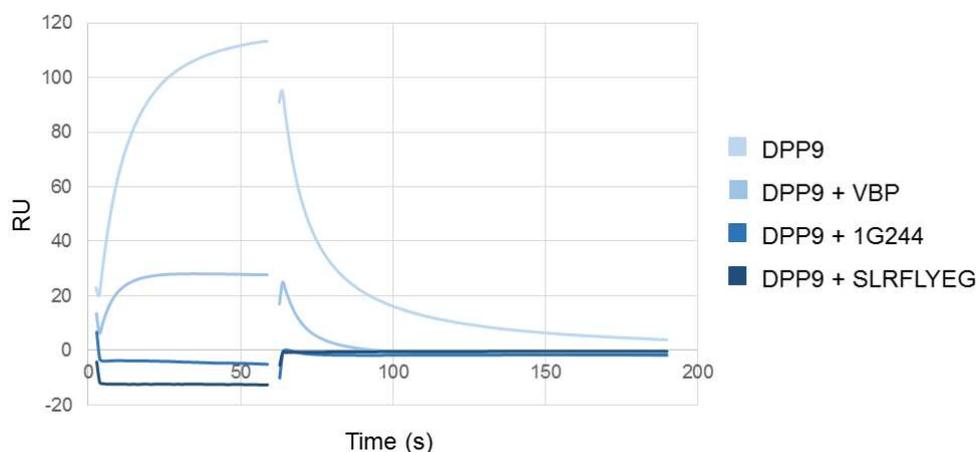


Figure 3.29. DPP9-SUMO1 complex is disrupted by small molecules binding. The interaction of DPP9 with immobilized SUMO1 was measured by SPR in presence of 100 μ M of small molecule inhibitors VBP, 1G244 and SLRFLYEG, respectively.

Next, as an attempt to use a sample that resembles better a physiological probe, the recombinant protein GST-SUMO1-linker-SUMO1 was expressed in *E. coli* and purified. GST serves as dimerization surface, generating a SUMO1 “tetramer”, thus increasing the probability of interaction with DPP9. Alternatively, there is a Thrombin cleavage site between the GST and SUMO1 to generate a SUMO1-linker-SUMO1 recombinant dimer. The linker between SUMO1 molecules was designed based on the DPP9 dimeric crystallographic structure to have an adequate length, namely to cover the distance from one DPP9 monomer to the opposite site of the non-crystallographic dimer (27 aa). SPR measurements using GST-SUMO1-linker-SUMO1 did not show an increased binding to DPP9 in comparison to monomeric SUMO1 (Fig. 3.30). Furthermore, cleavage of the GST tag did not change binding either.

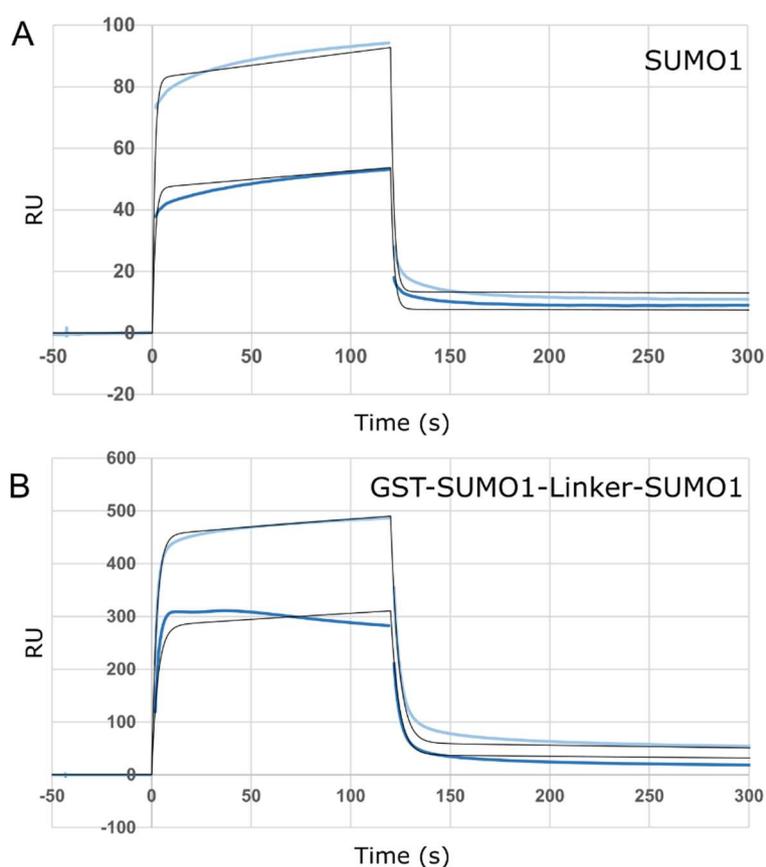


Figure 3.30. SUMO1-linker-SUMO1 oligomer does not increase binding to DPP9. Binding of SUMO1 and GST-SUMO1-Linker-SUMO1 to DPP9 were measured using SPR. 100 nM of DPP9 were immobilized on a NTA chip. Two concentration were used for each analyte, 35 μ M and 70 μ M. 80% of the expected RU were observed in both cases, calculated using 12 kDa and 53 kDa, respectively. A two state reaction binding model was used to fit the raw data. The black line correspond to the fitted curve at each sample concentration.

3.3.4 SUMO1 does not affect the activity of DPP8 or DPP9.

The functional effect of SUMO1 interaction on DPP8 and DPP9 was determined by measuring peptidase enzymatic activity for each protein. DPP8 and DPP9 enzymatic activities were compared at increasing substrate concentration after a pre-incubation time of 1 h with each SUMO1 forms at room temperature. The fluorimetric assay detects the cleaved AMC molecule signal at 440 nm. At each point the respective non-cleaved GP-AMC fluorescence was subtracted. Val-BoroPro was used as a control to inhibit DPP8 and DPP9, in both cases, strongly inhibiting (Fig. 3.31, A and B). At low concentration (1 μ M), SUMO1 did not induce a significant effect on the activity of either proteins (Fig. 3.31, A and B). As SUMO1 concentration raised to 5 μ M, there seemed to be an increased enzymatic activity in both, DPP8 and DPP9 (Fig. 3.31, C and D). However, this effect might not be specific, as a control experiment using similar amount of ovalbumin rendered enzymes with similar activities. Treatments with GST-SUMO1-linker-SUMO1 produced similar results, not inducing a significant change in enzymatic activity (Fig. 3.31, C and D).

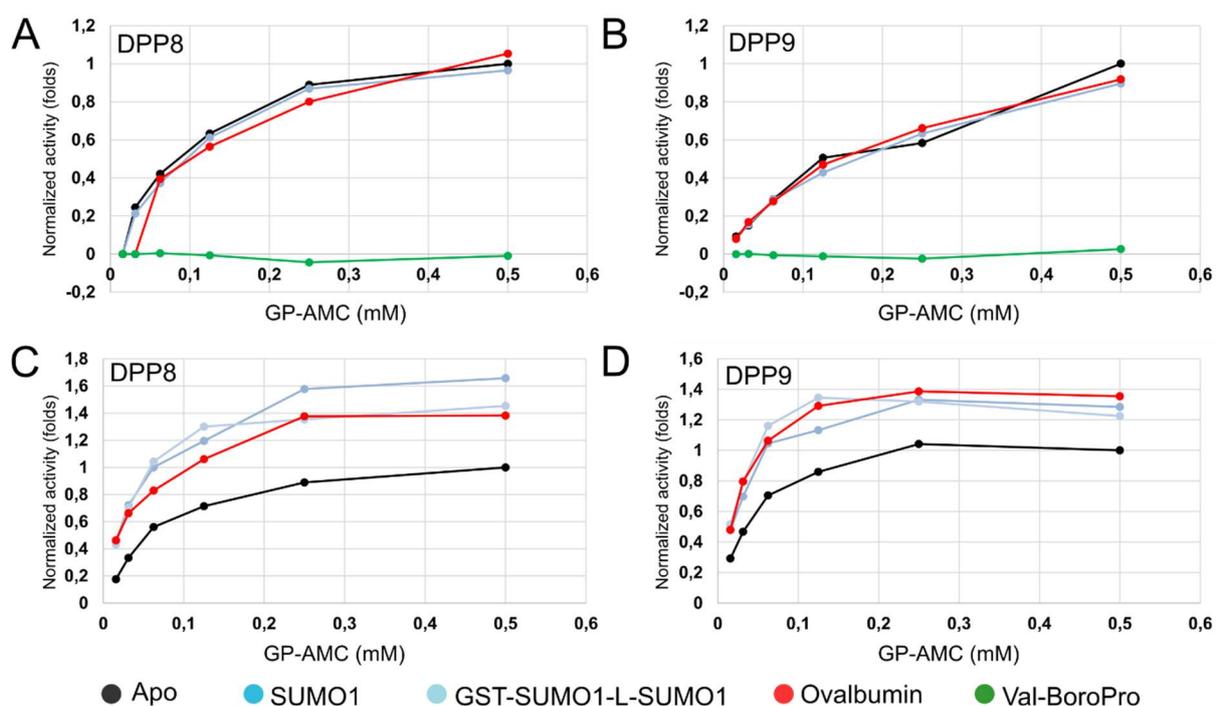


Figure 3.31. DPP8 and DPP9 activity in presence of SUMO1. (A-B) Activity of 12.5 nM of DPP8/9 in presence of 1 μ M of SUMO1. (C-D) Activity of 12.5 nM of DPP8/9 in presence of 5 μ M or SUMO1 or GST-SUMO1-L-SUMO1. 200 nM of Val-BoroPro was added as a control of inhibition (green). 1 or 5 μ M of ovalbumin was used as negative control of interaction (red). The experiments were repeated two times showing a similar trend.

4. Discussion

Protein crystallography is the technique that contributes with the vast majority of molecular structures deposited in Protein Data Banks, accounting roughly to 85% of all structures deposited until now (Fig. 4.1) [103]. It is therefore clear the chief relevance of each step towards solving structures using crystals. Furthermore, crystallography can be applied to solve structures using all kind of crystals, from inorganic materials to high molecular weight protein complexes.

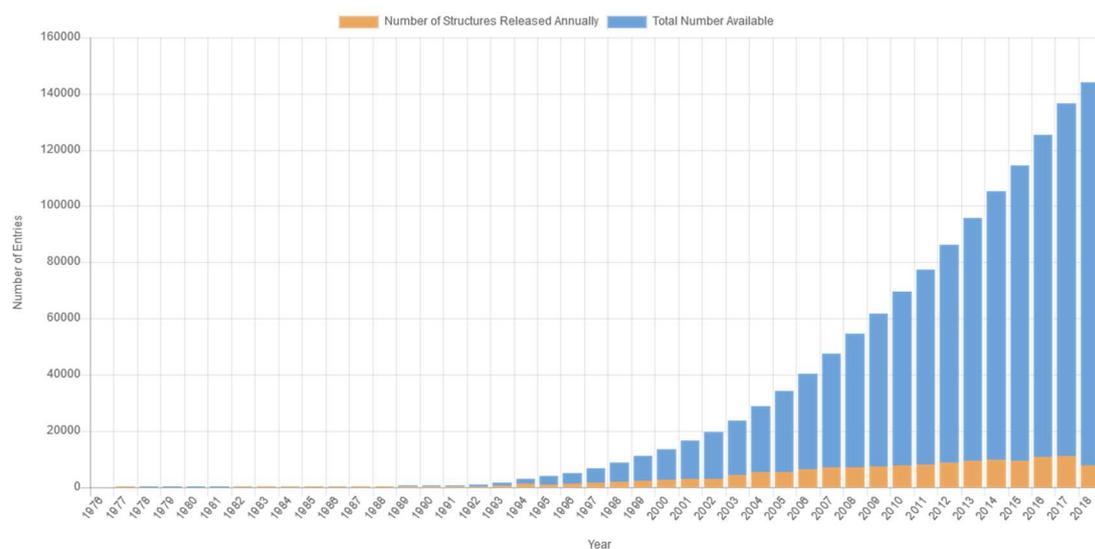


Figure 4.1. Protein data bank deposited structures. All structures deposited until September 2018. 92.8% of entries correspond to protein structures, 4.9% to protein-nucleic acids complexes and 3.2% to nucleic acids. Crystallography accounts to 83.8% of structures, 7.5% are NMR structures and 1.2% are electron microscopy structures [103].

Protein crystals are almost entirely a product of serendipity. Many efforts have been made to control crystallization [104]. However, crystallization still relies on big screens to find a favorable condition where crystals can grow. Thus, once a crystal condition is found, crystals are expected to be exhaustively measured in order to extract the most information out of them. Additionally, ideal experimental design strategies help at synchrotrons to maximize the structural data retrieved from protein crystals [21].

Unfortunately, protein crystals are far from being perfect, and very often they do not diffract as required to obtain meaningful structural data. This oppose limitations in different fields, e.g., towards structural rational drug design, delaying fundamental research in order to solve health issues. Therefore development of strategies to recover or improve “bad diffracting crystals” is desired [105].

This research sought to stablish and develop new post-crystallization methods to improve diffraction of protein crystals, successfully applying these techniques on DPP8 and DPP9 drug targets. As a result, two novel protein structures were solved, involved in immune response [65] and cancer development [66].

4.1 Post-crystallization diffraction improvement: dehydration

Water is a fundamental member of protein crystals, therefore dehydration has played a central role in protein crystallography from the very beginning [33]. Many tools to control humidity of protein crystals have been developed [36, 106]. Recently a new approach developed in our laboratory, using IR radiation, proved to efficiently induce transformations in some crystal systems by dehydration [54]. In this study we established the basic properties of applying IR radiation to protein crystals, to induce dehydration, annealing or perform soaking.

Dehydration and annealing are two different processes by which diffraction of protein crystals can be improved. Several reports account for either one or the other as responsible for the observed improved diffraction, when in reality, in many cases both variables are present [26]. This makes the interpretation of results cumbersome and not straightforward. Therefore we sought to design our experiments to isolate these two variables.

Bulk water is free to move within the crystal lattice using long water channels. This water is restructured as humidity is modified, thus reshaping the lattice simultaneously. As suggested by our experiments, IR radiation can effectively control crystal humidity without affecting the structure of the protein forming the crystals. This is further supported by data indicating that IR radiation at 938 nm stimulates water motion only, and is not absorbed by proteins. We observed a broad diversity of responses to IR radiation. Some crystal systems are extremely robust, flexible and resilient to humidity changes. Cycling between dehydration and rehydration several times did not induce loss of diffraction quality. However, crystals like DPP8 or DPP9 are not as stable, losing diffraction quality quickly after dehydration, and not recovering back. This fragility might be related to the amount of effective crystal contacts gained after dehydration without losing order. A comparative experiment was not possible due to the impossibility of collecting data on dehydrated DPP8 or DPP9 crystals.

Protein crystals respond to IR radiation linearly within a certain energy and speed range. The limit is most likely related to mechanical limitations imposed mainly by crystal lattice rigidity and intra-crystalline water diffusion (Fig. 3.5). As a reference, the self-diffusion coefficient of water is $2.4 \times 10^{-9} \text{ m}^2 \text{ s}^{-1}$ [107]. The diffusion coefficient of water within a lysozyme crystal along the (100, 010) plane is 1.8×10^{-10} and $4.6 \times 10^{-11} \text{ m}^2 \text{ s}^{-1}$ for that in (001) plane. This means a reduction of 15 and 5 times within a crystal lattice, respectively [108]. If the diffusion coefficient of water is forced to higher values, the lattice order will be mechanically damaged, and the crystal will crack. This effect was observed systematically above 10.8 J s^{-1} of laser energy rate (Fig. 3.6A). The diffusion coefficient is related to the specific permeability in a certain material, and the specific permeability is directly proportional to the thickness of the samples. Therefore it is reasonable to assume a linear correlation between the diffusion coefficient of water and crystal size. This was also observed in our experiments, where changing proportionally the energy with respect to crystal size kept the damage-energy-limit constant, regardless of the crystal size (Fig 3.6).

Taking into account the molar heat capacity of water ($80.9 \text{ J mol}^{-1} \text{ K}^{-1}$), the expected energy required to raise the temperature of a water cube of $100 \text{ }\mu\text{m}^3$ in one degree is $4.15 \text{ }\mu\text{J}$. Furthermore, experimental data shows that a lysozyme protein crystal of similar dimensions requires only 0.36 nJ and 0.12 nJ for 45.7 and 31.0 percent water content, respectively^[109]. Comparatively, we found 843 mJ K^{-1} as the energy required to dehydrate a lysozyme protein crystal at the maximum speed, achieved in one second approximately (assuming a total temperature raise of 9.6 K). This value is several orders of magnitude higher than the minimum energy required to raise the temperature of a lysozyme crystals in one degree. Possibly due to a reduced energy absorption of protein crystals (0.16%) as calculated by the Beer-Lambert law^[54], as well as additional factors reducing energy absorption like laser focusing errors and temperature raise time frame. Even though the laser proved to be an efficient tool to accelerate dehydration in comparison to other devices, we did not observe a positive correlation between fast-speed dehydration and diffraction improvements in the surveyed crystal systems. This is not a conclusive statement, since crystal heterogeneity is as diverse as protein diversity. More crystals systems need to be measured to have a significant sample spectra. However, the FML introduces significant new improvements, related to time-efficient sample processing and non-contact accurate humidity control.

4.2 Post-crystallization diffraction improvement: annealing

Then, as a second approach to improve diffraction quality, we used lysozyme crystals to develop a new annealing method, where annealing would serve as the major lattice-changing driving force. This was attained by immersing crystals under different oils, using a special crystal holder. The holder was designed to allow parallel measurements of 2D crystal area and diffraction quality. We found the oil perfluoropolyalkyl ether to be the most crystal friendly and water impermeable. Crystal's morphology was preserved and humidity was fixed, allowing room temperature measurements without use of a humidifier. The setup fulfills all requirements to perform systematic experiments on different crystal systems, following both, 2D area and diffraction to determine the existence of diffraction improvements after annealing treatments at room temperature.

Depending on the sample, room temperature measurements might not be the best option, due to radiation damage and other effects. Hence our efforts to setup experimental conditions to perform controlled annealing at 100 K . The theoretic reason to perform these experiments is based on translational motions of water at low temperature when undergoing phase transitions. Therefore, cycling temperature up and down may expose those transition reflected in the crystal volume^[110], in particular those crystals with big water channels (bigger than 60 angstroms)^[47]. We were able to see crystal size changes associated with cubic and hexagonal ice formation, as well as characteristic ice rings linked to those water phases, respectively. However, using lysozyme and DPP9 crystals was not possible to detect crystal size or diffractions changes within 100 to 150 K . This might be due to technical limitations, like camera resolution, or because these crystals systems do not present such phase transitions.

Even though there are 65 possible non-enantiogenic space groups, only a handful is highly represented in macromolecular structures, where $P12_11$ and $P2_12_12_1$ are the most prevalent, accounting

for one third of deposited structures. This fact is explained by an increased rigid-body degrees of freedom of molecules forming the lattice ^[111]. This purely statistical explanation makes of P2₁2₁2₁ the space group with the greatest degree of freedom to accommodate isomorphs (CODH is an example of it). Moreover, in addition to a dependency on crystal water channel size, as suggested in literature ^[112], rational experiment design can be planned to address those crystals matching such characteristics. More importantly, those crystals of space group P2₁2₁2₁ with big water channels may show diffraction improvements more frequently.

4.3 Post-crystallization diffraction improvement: chemical modification

Changing chemical composition of protein crystals has proven to successfully improve crystal diffraction ^[113, 114]. Therefore is especially important to have complexation or soaking methods which complement the success rate of chemicals alone. The non-contact soaking methods presented in this study proved to efficiently concentrate solutions without inducing major effects on the molecular structure of protein crystals. Furthermore, they were instrumental to keep and improve diffraction on DPP8 and DPP9 drug targets.

The methylamine naturally occurring osmolyte trimethylamine N-oxide (TMAO) is known to work as a protein stabilizer, opposite to urea ^[115]. Different reports link this molecule to diffraction quality improvements of protein crystals ^[116]. Moreover, a cryo-protectant function further extends the usage of TMAO in macromolecular crystallography ^[87]. We successfully used TMAO to improve diffraction of DPP8 and DPP9 by soaking crystals using both FMS and FML. Even though a diffraction improvement was present, we were not able to observe positive electron density of TMAO in either structures.

DPP8 protein crystals have 24 molecules per unit cell, therefore 24 possible active sites to be occupied in a total unit cell volume of 10.487.185 Å³. This corresponds to a concentration of 4 mM of active sites for C222₁. As a rule of thumb, ligands should be soaked using a concentration 5-times higher than their K_m affinity values, or to reach a concentration as high as active sites concentration in the crystal. FML offers the possibility to increase temperature of protein crystals, thus increasing ligand solubility. However, previous knowledge of how ligand solubility changes with temperature is rarely available, particularly in experiments with new chemicals.

4.4 DPP8 and DPP9 as successful targets of post-crystallization treatments

4.4.1 DPP8 and DPP9 molecular structure

DPP8 and DPP9 are intracellular serine dipeptidyl peptidases that modify in a non-reversible manner the N-terminus of their substrates. The outcome of this processing and formation of a neo N-terminus may alter the life span or activity of a variety of proteins ^[69, 74]. They are a focus of attention because of their relevance in immune response and cancer ^[58, 65, 66, 69, 71, 83]. Therefore, molecular structures of both targets are a valuable basis for development of specific inhibitors. The DPP4 Activity Structure Homolog (DASH) family of proteases with its members DPP10, DPP6, DPP4, FAP, DPP8, and DPP9

share a common modular structure, consisting of the N-terminus β -propeller domain and the C-terminus α/β hydrolase domain, despite a very low sequence homology in the former module. The last four members commonly occur as active functional homodimers, whereby the association is mediated by the α/β hydrolase domain.

The first crystal structure of a member of the protein family was published in 2003 for DPP4 [95]. Here we report the structures of DPP8 and DPP9 and extend earlier studies of in vitro functional investigations. These new structures were obtained by application of post-crystallization methods to obtain optimal crystal diffraction. Furthermore, only using a DPP4 modified molecular model allowed phasing by molecular replacement of DPP8 and DPP9 crystal forms. All crystal forms contain multiple copies of the polypeptide chain in the asymmetric unit. The comparison of DPP8 and DPP9 with DPP4 disclosed extensive variations in the β -propeller domain by additional secondary structures, strand exchanges, and loop alterations.

In contrast to DPP4, where ligand binding does not significantly alter the protein structure, the binding of the inhibitory SUMO1- derived peptide SLRFLYEG to DPP8 induces ordering of the R-helix, which is part of the R-segment, shaping the substrate binding site. The unliganded structure of DPP8 shows no or disconnected electron density, which may be traced as pieces of the R-helix, albeit differently positioned. These observations hint at induced fit and/or conformational selection for ligand binding. The unexpected discovery of SLRFLYEG binding in the active site was instrumental in revealing the essential structural features of substrate binding. Although it has the canonical proline residue replaced by leucine at P1, its ϕ angle is compatible with proline. Discontinuous electron density between O γ of S755 and the carbonyl carbon of the scissile peptide bond -L-R- indicates a tight non-covalent binding. The peptide displays tight interactions at P1', P2', and P3'. In particular, the phenylalanine's role might be underestimated in defining enzyme specificity, fitting in an additional hydrophobic pocket not existent in DPP4. The octa-peptide extends toward the surface occupying the side entry/exit tunnel similar to DPP4 but following a somewhat different path [2].

The R-helix plays a major role in ligand binding by providing R160/133 at its C-terminus. The arginine side chain anchors the peptide through hydrogen bonds to the carbonyl oxygens of the P2 and P1' residues, thereby stabilizing the proline turn conformation at P1. R125 plays this role in DPP4, but emanates from a different structural segment of the protein, the R-loop (Fig. 3.20 A and B). The active sites of DPP4, DPP8, and DPP9 exhibit a conserved characteristic S1 subsite, with similar dimensions in the three species. The site is almost fully occupied by 1G244 in DPP9, offering little room for expansion. On the other hand, the S2 subsite in DPP8 and DPP9 diverges significantly from DPP4, presenting different features, most remarkably the positions of two loops. First, the H450/424 loop in DPP8/9, with the side chain pointing away from the active site, increasing the size of S2 subsite compared with DPP4. Second, the loop H525/500 contributing to the S2 subsite in DPP8 and DPP9 does not exist in DPP4. The H525/500 loop is a possible candidate to interact with large P2 side chains as in 1G244 the 1-(4-4'-difluor-benzhydryl)-piperazine. Furthermore, the significant size expansion of the P2

subsite in DPP8 and DPP9 causes an overlap of S2 and S1'. This feature provides options for specific ligand generation. The S2 subsite can accept a variety of residues, with a preference for voluminous hydrophobic groups [74]. A significant difference between DPP8 and DPP9 is a region contained within the R-segment. This solvent-exposed loop possesses two consecutive histidines, H117 and H118 in DPP9, ordered in the liganded form. DPP8 has D134 and Y135 in the same positions, which are disordered and not visible either in the unliganded crystal structure or in the peptide-complex structure. This segment offers itself as an epitope for antibodies with specific inhibitory properties in a similar approach as for DPP4 [117].

Solving the structure of DPP8 with Val-BoroPro supports the conformational change observed in previous liganded structures. This results suggest that the disorder-order transition of the R-segment is independent of the ligand size, highlighting the fundamental role of this segment in enzymatic activity. We did not observed any additional particular difference in the molecular structure of DPP8 bound to Val-BoroPro which might serve to generate specificity. However, extending the knowledge of how molecules bind to dipeptidyl peptidases is the way to find new inhibitors of DPP8 and DPP9 [118].

4.4.2 DPP8 and DPP9 display allosteric and cooperative binding

1G244 was designed for specificity against DPP8 and DPP9 and discriminating against DPP4. The analysis of the enzymatic binding mode of 1G244 had indicated a small difference between both proteins, with competitive and slow-tight competitive inhibition for DPP9 and DPP8, respectively [119]. However, the kinetic data presented here are consistent with an allosteric interaction between the two subunits of DPP8 and DPP9, resulting in a cooperativity in their substrate binding. The allosteric effect of both 1G244 and SLRFLYEG on substrate turnover is supported by the observation that although both inhibitors bind in the active site, they have little effect on $K_{0.5}$ (Table 3.3). The structural features described, specifically the ligand-induced rearrangements and formation of the substrate binding site and the strap of contacts between the active sites in the dimer formed by “ligand-[R-helix]-SUBA-SUBA-[R-helix]-ligand,” suggest a tentative molecular interpretation of these data, whereby ligands first bind to the partially disordered unliganded conformation or, alternatively, select competent conformers, ensued by active site stabilization, which is signaled to the other subunit. Fast kinetic measurements would need to be performed to further study the substrate binding mode and conformational selection mechanism associated with partially defined R-helices in unliganded structures. The discovery of communication between the subunits in the dimer and the putative transduction signal pathway offers opportunities for specific functional interference.

4.4.3 Molecular dynamics simulation

To assess the stabilization that the bound ligand provides to the overall structure and the R-helix, respectively, molecular dynamics simulations were carried out. Hence, experimental DPP8 and DPP9 structures, crystallized with their respective ligands, were modeled under two different conditions: first, the ligand bound protein structure with the ligand being present and second, the ligand bound protein structure with the ligand removed. Four independent simulations, on each of the four modeled systems,

were carried out at 310 K for 200 ns to provide statistical robustness to the observations. The rmsd showed deviations of 1.8 Å and 1.9 Å of the overall structures for DPP8 and DPP9, respectively, independent of the presence of the ligand, while the R-helix deviates by 1.2 Å and 2.5 Å for DPP8 and 0.4 Å and 2.3 Å for DPP9 under the same conditions. Interestingly, while the global protein fold is preserved during the dynamics runs, the R-helix structure is highly sensitive to the presence of the ligand, in agreement with the crystallographic structure observations (Fig. 4.2).

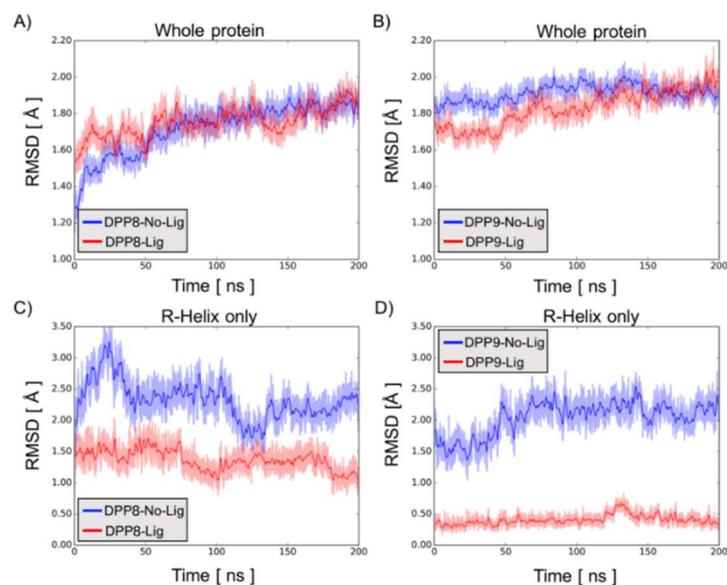


Figure 4.2. RMSD of DPP8 and DPP9 simulated systems. In the panels are shown the average rmsd curves (shaded region) with a running average with 50 frame window (solid line). Average curves were computed out of 4 independent simulation runs for each system; DPP8 (panels A and C) and DPP9 (Panel B and D) with and without inhibitor, red and blue curves, respectively. The rmsd was computed considering only alpha-carbons of the whole protein structure (panels A and B) and alpha-carbons of the R-helix (panels C and D; DPP8 150-160 and DPP9 123-133). In all cases the corresponding experimental structure was used as reference.

4.4.4 Bacterial DPP4 structural diversity

Screening the PDB database and comparing human DPP4/8/9 with bacterial DPP4 reveals several features ^[103]. Some bacterial DPP4s (e.g. *Porphyromonas gingivalis*) display an R-loop and are structurally related to human DPP4 ^[120] (Fig. 4.3 A and F), whereas DPP4 from other bacterial species has an R-segment, lacks the R-loop, and is closer to DPP8/9. The R-segment in DPP4 of *Stenotrophomonas maltophilia* is disordered as in human DPP8/9 ^[121], but data of a liganded structure are not available (Fig. 4.3 B–D). Furthermore, a third case is compared for *Bacteroides ovatus*. The R-loop does not exist, and the R-segment is ordered in an open conformation, having a lysine instead of an arginine (Fig. 4.3E). Very interestingly and simultaneously to our findings in human DPP8 and DPP9, a non-related laboratory has reported an identical mechanism (disorder-order transition) in one DPP4 orthologous protein of *Pseudoxanthomonas mexicana* ^[122]. This result support our hypothesis and suggest that the diversity observed in bacteria is quite complex, reflecting a broad function and flexibility of these proteins across species.

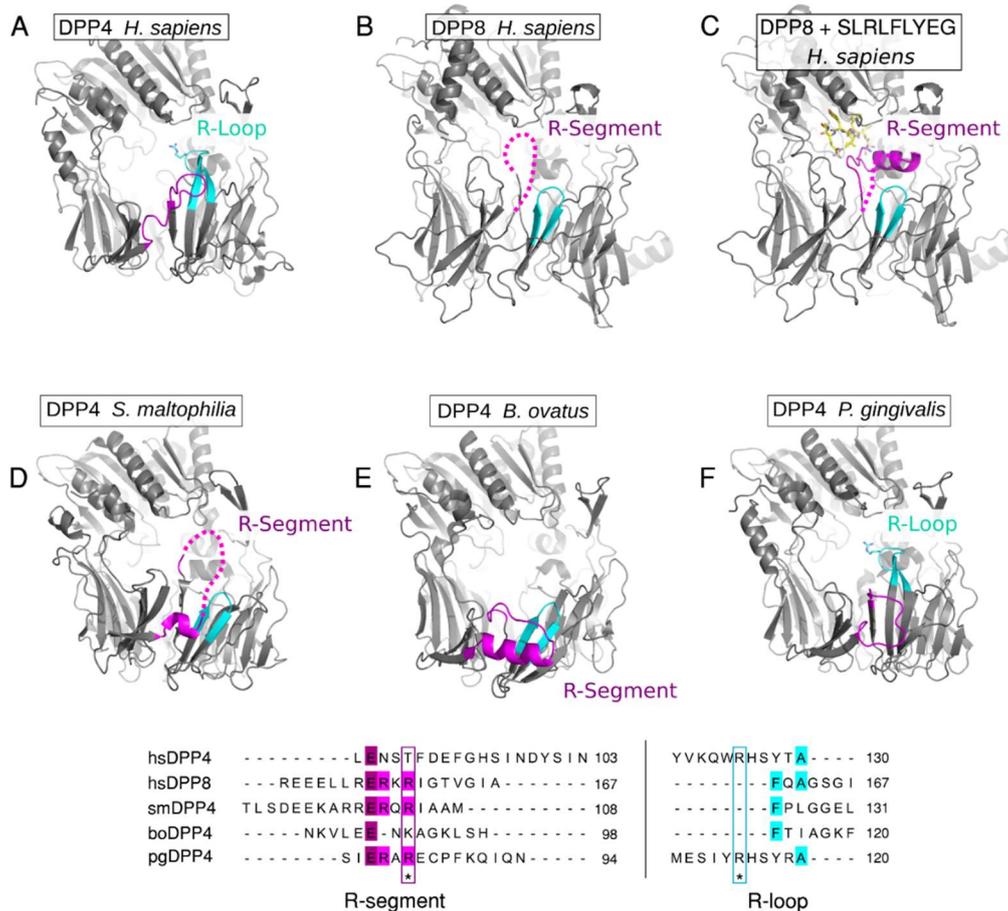


Figure 4.3. Structural comparative analysis of human DPP4/8 versus bacterial DPP4. The R-loop (hsDPP4 - cyan) and R-segment (hsDPP8 - magenta) counterparts of each bacterial DPP4 were identified by multiple sequence alignment and structural alpha-carbon alignment. (A) Human DPP4 (PDB: 1ORV). (B-C) Human DPP8. (D-F) Bacterial DPP4 (PDB: 2ECF; 4Q1V; 5OLJ). The dotted-line correspond to a disordered segment. The asterisk in multiple sequence alignment marks the conserved arginine residue used to fix substrates.

4.5 SUMO1 interacts with DPP8 and DPP9

A specific interaction between DPP8/9 and SUMO1 has been described based on pull-down experiments with bead-immobilized SUMO1. We therefore set up co-crystallization experiments with DPP8/9 and SUMO1, which, however, were not successful. Also, we did not observe the complex in solution using size exclusion chromatography. These observations denote a transient and low-affinity interaction. SUBA has been characterized as the interaction region of SUMO1, and a single mutation, V285A, in this subdomain abolishes binding^[81, 82]. SLRFLYEG, a peptide derived from SUMO1, where it is a terminal strand of the central β -sheet in the molecular structure, was found to displace SUMO1 from its complex. It was tempting to assume that the peptide mirrors SUMO1 binding. However, the structural data, described here, present SLRFLYEG at the active site of DPP8, which is a narrow crevice and unfit to receive SUMO1. Extensive unfolding of either ligand or receptor is unlikely and not supported by experimental data. Modeling by docking of SUMO1 to the DPP8 dimer was pursued and demonstrated that there is sufficient space to allow the approach of the EIL strand to SUBAs, such that the specific

contacts defined earlier (L321 and F329) by mutations in SUBA and sequence variations of the peptide can be satisfied. Nonetheless, the physiological role of the internal SUMO1 peptide found in the active site of DPP8 remains to be solved (Fig. 4.4).

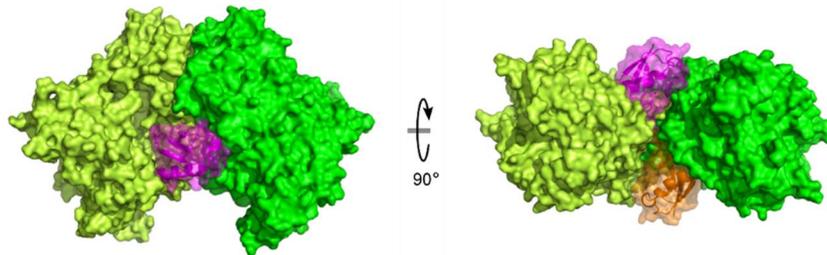


Figure 4.4. Model of interaction between SUMO1 and DPP8/9. Rigid body docking model using SUMO1 (PDB 2PE6). DPP8 dimer interacting with two molecules of SUMO1 docked to SUBA at the dimer interface.

We were able to confirm the non-covalent interaction between DPP8/9 and SUMO1 by using in vitro experiments with recombinant protein. However, the complex formation affinity is quite weak (28 μM), falling apart in seconds. This offers an explanation to the impossibility to recover a protein complex by size exclusion chromatography. Based on pull-down experiments, we hypothesized that multi-sumoylated targets might interact strongly with DPP8 or DPP9, thus we performed SUMO1 crosslinking to generate oligomeric forms. Even though the SUMO1 crosslinked sample had an increased affinity to DPP9, purification of the complex was inefficient. Moreover, recombinant SUMO1 oligomeric forms (dimers or tetramers) did not enhance SUMO1 binding. This result is restricted to a set of particular constructs, which might not reflect the physiological interactions accurately. Various avenues might be used to obtain a stable complex, either by expressing constructs with different linker lengths, using different orientations of SUMO1, using sumoylated peptidic substrates or using multi-sumoylated binding partners like p53 [79].

Finally, our functional experiments contradict literature reports, where a DPP8/9 activating role has been attributed to SUMO1 [81]. We observed that incubations of DPP9 with SUMO1 do not induce a significant change in enzymatic activity. This result seems to not correlate the SUMO1 binding region in DPP9 to the R-segment, however, we have observed that molecules binding in the active site induce a “closed” conformation of DPP9, releasing SUMO1. This should theoretically be translated into an inhibitory function of SUMO1, favoring an “opened” DPP9 conformation. This apparent lack of correlation could be due to experimental conditions, which makes use of a small substrate and might not be a good analog of a physiological substrate. Nevertheless, to explain the competition experiments by inhibitors, we propose an essential role of the R-segment, which undergo a profound change and structural fixation upon ligand binding at the active site and suggest SUMO1 binding to the unliganded enzyme. When ligands bind, the ensuing rearrangement of the R-segment disrupts these interactions, leading to the dissociation of the complex.

5. Conclusion

In this study we explored how to improve diffraction of protein crystals. Protein crystals were irradiated with infrared light to induce isomorphous lattice conformations, with the aim of finding an improved state of diffraction quality. We were able to establish the experimental bases of controlled dehydration using IR radiation, as well as annealing and chemical exchange. These three methods can modify crystal order, as quantified by diffraction resolution improvements or B-factors reductions. The FML presents several advantages, chiefly that crystals can be heat up in a quick manner, thus inducing fast dehydration or short annealing pulses. Moreover, the chemical exchange proved to be a powerful technique, which when performed in a careful way it can induce significant diffraction improvements. The major disadvantage of this technique relates to humidity calculation, since the FML relies on the FMS to set indirectly a humidity value. Furthermore, this method measures 2D projections, sometimes missing significant crystal changes occurring in the hidden axis to the camera.

The second part of the study, sought to determine the crystallographic structure of DPP8 and DPP9. These are relevant drug targets in the field of immunology and cancer biology. Approximately 18.000 different conditions were tested per protein construct to obtain a crystallization condition, altogether spanning a period of 6 months. After a challenging crystal optimization, in particular for DPP9, we applied the previously developed techniques to improve diffraction of those crystals successfully. While dehydration rendered negative effects and annealing was neither negative nor positive, chemical exchange improved crystal order and allowed effectively cooling of crystals.

Solving the crystallographic structure of DPP8 and DPP9 revealed new structural features. The disorder/order transition of the R-segment accounts as a novel substrate binding mechanism. Furthermore, additional induced structural changes in DPP9 by substrate binding broadens the possible substrate specificity of this protein. Subsequently, we characterized the binding of SUMO1 to DPP8 and DPP9. This was not straightforward, since the low binding affinity of these complexes rendered short half-life times, thus the complexes exist in solution only few seconds before they disassemble. After diverse cross-linking tests and high concentration titrations, the measured binding K_D supports the physiological relevance of the DPP8-SUMO1 and DPP9-SUMO1 complexes.

While crystal diffraction improvement is a parameter tightly related to crystal order, is quite complicated to predict how crystal order might respond to post-crystallization treatments. However, a treatment that hinders Bragg spots, might still contribute to induce random disorder, which in turn can be used to extract structural meaningful information.

6. Outlook

The development of post-crystallization techniques to improve diffraction quality, including those presented here, will allow to screen and reveal transformation of complicated cases, where protein crystals are available but do not diffract sufficiently enough to provide relevant structural data. In this manner, saving time and resources, but also speeding up the development of pharmacological tools.

Post-crystallization techniques will move towards the different approaches of coherent diffracting imaging. Thus, improvement of diffraction can be aimed to serial crystallography as well as to micro-electron diffraction, both dealing with crystalline materials. In particular, IR laser treatments might come in handy when seeking for continuous diffraction, as a mean to address the phase problem or data refinements. Moreover, improvements in single crystals visualization towards a 3D reconstruction might simplify the detection of transformation points, as well as enabling real time x-ray absorption corrections by accurate crystal volume measurements.

The crystallographic structures presented here, solved with application of post-crystallization methods, are of enormous value in different human health fields. They contribute to clarify the structural basis of their inhibition, as well as providing possible ways to modify their activities by future drug design. In this regard, the interactome between these and regulator proteins like SUMO1 will have a central role in the future. Most importantly, the determination of the DPP8/9-SUMO1 complex might reveal new cues to regulate natural or aberrant signaling pathways leading to health improvement.

7. References

- [1] B. Rupp, *Biomolecular Crystallography: Principles, Practice, and Application to Structural Biology*, First ed., NY: Garland Science, 2009.
- [2] K. Aertgeerts, M. G. Tennant, M. L. Kraus, J. Rogers, B.-c. Sang, R. J. Skene, D. R. Webb and G. Sridhar Prasad, "Crystal structure of human dipeptidyl peptidase IV in complex with a decapeptide reveals details on substrate specificity and tetrahedral intermediate formation," *Protein Science*, vol. 13, no. 2, pp. 412-21, 2003.
- [3] X. Chen, I. Weber and R. W. Harrison, "Hydration water and bulk water in proteins have distinct properties in radial distributions calculated from 105 atomic resolution crystal structures," *Journal of Physical Chemistry B*, vol. 112, p. 12073–12080, 2008.
- [4] E. K. H. Salje, "Crystallography and Structural Phase Transitions, an Introduction*," *Lead Article Acta Cryst*, vol. 47, pp. 453-469, 1991.
- [5] C. J. Howard and H. T. Stokes, "Structures and phase transitions in perovskites - A group-theoretical approach," *Acta Crystallographica Section A: Foundations of Crystallography*, vol. A61, pp. 93-111, 2005.
- [6] G. Jaeger, "The Ehrenfest Classification of Phase Transitions: Introduction and Evolution," *Archive for History of Exact Sciences*, vol. 53, pp. 51-81, 1998.
- [7] V. Pokrovsky, "Landau Theory and Modern Physics," *Physics Uspekhi*, vol. 52, no. 11, pp. 1169-1172, 2009.
- [8] S. Salinas, *Introduction to Statistical Physics*, Springer ed., NY, 2001, pp. 235-256.
- [9] M. Baazaoui, S. Hcini, M. Boudard, S. Zemni and M. Oumezzine, "Critical behavior near the ferromagnetic-paramagnetic phase transition temperature of $\text{Pr}_{0.67}\text{Ba}_{0.33}\text{Mn}_{1-x}\text{Fe}_x\text{O}_3$ ($x=0$ and 0.05) manganite," *Journal of Magnetism and Magnetic Materials*, vol. 401, pp. 323-232, 2016.
- [10] M. Sakoh and M. Shimizu, "Thermodynamics of Phase Transitions in Iron," *phys. stat. sol. (a)*, vol. 441, no. 211, pp. 511-16, 1977.
- [11] K. Harata and T. Akiba, "Phase transition of triclinic hen egg-white lysozyme crystal associated with sodium binding," *Acta Crystallographica Section D: Biological Crystallography*, vol. D60, pp. 630-37, 2004.
- [12] K. Harata and T. Akiba, "Structural phase transition of monoclinic crystals of hen egg-white lysozyme," *Acta Crystallographica Section D: Biological Crystallography*, vol. D62, pp. 375-382, 2006.
- [13] M. G. Lawrence, "The Relationship between Relative Humidity and the Dewpoint Temperature in Moist Air," *American Meteorological Society*, vol. 86, p. 225–234, 2004.

- [14] D. F. Swinehart, "The Beer-Lambert Law," *Journal of Chemical Education*, vol. 39, no. 7, pp. 333-35, 1962.
- [15] H. Horvath, "Atmospheric Light Absorption: A Review," *Atmospheric Environment*, vol. 27, no. 3, pp. 293-317, 1993.
- [16] J. A. Curcio and C. C. Petty, "The Near Infrared Absorption Spectrum of Liquid Water," *Journal of the Optical Society of America*, vol. 41, no. 5, pp. 302-4, 1951.
- [17] K. F. Palmer and D. Williams, "Optical properties of water in the near infrared*," *Journal of the Optical Society of America*, vol. 64, no. 8, pp. 1107-10, 1974.
- [18] J. K. Gregory, D. C. Clary, K. Liu, M. G. Brown and R. J. Saykally, "The Water Dipole Moment in Water Clusters," *Science*, vol. 275, no. 5301, pp. 814-817, 1997.
- [19] B. Wozniak and J. Dera, *Light Absorption in Sea Water*, First ed., Springer, 2007, p. 452.
- [20] C. G. Pope, "X-Ray Diffraction and the Bragg Equation," *Journal of Chemical Education*, vol. 74, no. 1, pp. 129-131, 1997.
- [21] T. Krojer, A. C. W. Pike and F. Von Delft, "Squeezing the most from every crystal: The fine details of data collection," *Acta Crystallographica Section D: Biological Crystallography*, no. D69, pp. 1303-13, 2013.
- [22] G. L. Taylor, "Introduction to phasing," *Acta Crystallographica Section D: Biological Crystallography*, vol. D66, no. 1, pp. 325-338, 2010.
- [23] R. L. Owen, J. Juanhuix and M. Fuchs, "Current advances in synchrotron radiation instrumentation for MX experiments," *Archives of Biochemistry and Biophysics*, vol. 602, pp. 21-31, 2016.
- [24] E. F. Garman and R. L. Owen, "Cryocooling and radiation damage in macromolecular crystallography," *Acta Crystallographica Section D: Biological Crystallography*, vol. D62, pp. 32-47, 2006.
- [25] G. J. Kleywegt, "Validation of protein crystal structures," *Acta Cryst*, vol. D56, no. 1, pp. 249-265, 2000.
- [26] B. Heras and J. L. Martin, "Post-crystallization treatments for improving diffraction quality of protein crystals," *Acta Crystallographica Section D: Biological Crystallography*, vol. D61, pp. 1173-1180, 2005.
- [27] K. Ayyer, O. M. Yefanov, D. Oberthür, S. Roy-Chowdhury, L. Galli, V. Mariani, S. Basu, J. Coe, C. E. Conrad, R. Fromme, A. Schaffer, K. Dörner, D. James, C. Kupitz, M. Metz, G. Nelson, P. L. Xavier, K. R. Beyerlein, M. Schmidt, I. Sarrou, J. C. H. Spence, U. Weierstall, T. A. White, J.-H. Yang, Y. Zhao, M. Liang, A. Aquila, M. S. Hunter, J. S. Robinson, J. E. Koglin, S. Boutet, P. Fromme, A. Barty and H. N. Chapman, "Macromolecular diffractive imaging using imperfect crystals," *Nature*, vol. 530, pp. 202-220, 2016.

- [28] C. Giacovazzo, *Fundamentals of crystallography*, Second ed., Michigan: Oxford University Press, 2002, p. 825.
- [29] S. Kriminski, C. L. Caylor, M. C. Nonato, K. D. Finkelstein and R. E. Thorne, "Flash-cooling and annealing of protein crystals," *Acta Crystallographica Section D: Biological Crystallography*, vol. D58, pp. 459-471, 2002.
- [30] C. X. Weichenberger, P. V. Afonine, K. Kantardjieff and B. Rupp, "The solvent component of macromolecular crystals," *Acta Crystallographica Section D: Biological Crystallography*, vol. D71, pp. 1023-38, 2015.
- [31] B. W. Matthews, "Solvent content of protein crystals," *J. Mol. Biol.*, vol. 33, pp. 491-497, 1968.
- [32] J. D. Bernal and D. Crowfoot, "X-ray Photographs of Crystalline Pepsin," *Nature*, vol. 133, p. 794-5, 1934.
- [33] M. F. Perutz, "The Composition and Swelling Properties of Haemoglobin Crystals," *Transactions of the Faraday Society*, vol. 42, pp. 187-195, 1946.
- [34] P. W. Winston and D. H. Bates, "Saturated solutions for the control of humidity in biological research," *Ecology*, vol. 41, no. 1, pp. 232-237, 1960.
- [35] H. Dobbek, L. Gremer, R. Kiefersauer, R. Huber and O. Meyer, "Catalysis at a dinuclear [CuSMo(AO)OH] cluster in a CO dehydrogenase resolved at 1.1-Å resolution," *PNAS*, vol. 99, no. 25, pp. 15971-76, 2002.
- [36] R. Kiefersauer, M. E. Than, H. Dobbek, L. Gremer, M. Melero, S. Strobl, J. M. Dias, T. Soulimane and R. Huber, "A novel free-mounting system for protein crystals: Transformation and improvement of diffraction power by accurately controlled humidity changes," *Journal of Applied Crystallography*, vol. 33, no. 1, pp. 1223-1230, 2000.
- [37] J. Bttcher, A. Jestel, R. Kiefersauer, S. Krapp, S. Nagel, S. Steinbacher and H. Steuber, "Key factors for successful generation of proteinfragment structures: Requirement on protein, crystals, and technology," *Methods in Enzymology*, vol. 493, pp. 61-89, 2011.
- [38] R. Kiefersauer, J. Stetefeld, F. Xaver Gomis-Roth, M. J. Romao, F. Lottspeich and R. Huber, "Protein-Crystal Density by Volume Measurement and Amino-Acid Analysis," *J. Appl. Crvst*, vol. 29, no. 1, pp. 311-317, 1996.
- [39] S. Russi, D. H. Juers, J. Sanchez-Weatherby, E. Pellegrini, E. Mossou, V. T. Forsyth, J. Huet, A. Gobbo, F. Felisaz, R. Moya, S. M. McSweeney, S. Cusack, F. Cipriani and M. W. Bowler, "Inducing phase changes in crystals of macromolecules: Status and perspectives for controlled crystal dehydration," *Journal of Structural Biology*, vol. 175, no. 1, pp. 236-243, 2011.
- [40] W. Awad, G. S. Birkedal, M. M. G. M. Thunnissen, K. Mani and D. T. Logan, "Improvements in the order, isotropy and electron density of glypican-1 crystals by controlled dehydration," *Acta Crystallographica Section D: Biological Crystallography*, vol. D69, pp. 2524-33, 2013.

- [41] I. K. Russo, F. Sica, C. A. Mattia and A. Merlino, "Increasing the X-ray diffraction power of protein crystals by dehydration: The case of bovine serum albumin and a survey of literature data," *International Journal of Molecular Sciences*, vol. 13, pp. 3782-3800, 2012.
- [42] G. Hagelueken, H. Huang, K. Harlos, B. R. Clarke, C. Whitfield and J. H. Naismith, "Crystallization, dehydration and experimental phasing of WbdD, a bifunctional kinase and methyltransferase from *Escherichia coli* O9a," *Acta Crystallographica Section D: Biological Crystallography*, vol. D68, pp. 1371-1379, 2012.
- [43] R. Oliete, J. Pous, S. Rodriguez-Puente, C. Abad-Zapatero and A. Guasch, "Elastic and inelastic diffraction changes upon variation of the relative humidity environment of PurE crystals," *Acta Crystallographica Section D: Biological Crystallography*, vol. D69, pp. 194-212, 2013.
- [44] S. Klingl, M. Scherer, T. Stamminger and Y. A. Muller, "Controlled crystal dehydration triggers a space-group switch and shapes the tertiary structure of cytomegalovirus immediate-early 1 (IE1) protein," *Acta Crystallographica Section D: Biological Crystallography*, no. D71, pp. 1493-1504, 2015.
- [45] P. Mazur, "A biologist's view of the relevance of thermodynamics and physical chemistry to cryobiology," *Cryobiology*, vol. 60, pp. 4-10, 2010.
- [46] D. R. Macfarlane, "Physical Aspects of Vitrification in Aqueous Solutions," *Cryobiology*, vol. 24, pp. 181-195, 1987.
- [47] M. Weik, G. Kryger, A. M. M. Schreurs, B. Bouma, I. Silman, J. L. Sussman, P. Gros and J. Kroon, "Solvent behaviour in flash-cooled protein crystals at cryogenic temperatures," *Acta Crystallographica Section D: Biological Crystallography*, vol. D57, pp. 566-573, 2001.
- [48] B. L. Hanson, J. M. Harp and G. J. Bunick, "The Well-Tempered Protein Crystal: Annealing Macromolecular Crystals," *Methods in Enzymology*, vol. 368, pp. 217-35, 2003.
- [49] S. Parkin and H. Hope, "Low-temperature water reconstruction in concanavalin A, with implications for controlled protein crystal annealing," *Acta Crystallographica - Section D Biological Crystallography*, vol. D59, pp. 2229-36, 2003.
- [50] I. K. Russo, A. Merlino, A. Vergara and F. Sica, "An overview of biological macromolecule crystallization," *International Journal of Molecular Sciences*, vol. 14, pp. 11643-91, 2013.
- [51] P. G. Vekilov, "Nucleation," vol. 10, pp. 5007-5019, 2010.
- [52] M. Dimova and Y. D. Devedjiev, "Protein crystal lattices are dynamic assemblies: The role of conformational entropy in the protein condensed phase," *IUCrJ*, vol. 5, pp. 130-140, 2018.
- [53] P. Ma, Y. Xue, N. Coquelle, J. D. Haller, T. Yuwen, I. Ayala, O. Mikhailovskii, D. Willbold, J. P. Colletier, N. R. Skrynnikov and P. Schanda, "Observing the overall rocking motion of a protein in a crystal," *Nature Communications*, vol. 6, no. 8361, 2015.

- [54] R. Kiefersauer, B. Grandl, S. Krapp and R. Huber, "IR laser-induced protein crystal transformation," *Acta Crystallographica Section D: Biological Crystallography*, vol. D70, pp. 1224-32, 2014.
- [55] N. D. Rawlings and A. J. Barrett, "Evolutionary Families of Peptidases," *Methods in Enzymology*, vol. 290, pp. 205-18, 1995.
- [56] E. Harper and A. Berger, "On the size of the active site in proteases: Papain," *Biochemical and Biophysical Research Communications*, vol. 27, no. 2, pp. 157-62, 1972.
- [57] Y. Zhou and Y. Zhang, "Serine protease acylation proceeds with a subtle re-orientation of the histidine ring at the tetrahedral intermediate," *Chemical Communications*, vol. 47, pp. 1577-1579, 2011.
- [58] Y. Waumans, L. Baerts, K. Kehoe, A.-M. Lambeir and I. De Meester, "The dipeptidyl peptidase family, prolyl oligopeptidase, and prolyl carboxypeptidase in the immune system and inflammatory disease, including atherosclerosis," *Frontiers in Immunology*, vol. 6, no. 387, 2015.
- [59] L. Wagner, C. Klemann, M. Stephan and S. Von Hörsten, "Unravelling the immunological roles of dipeptidyl peptidase 4 (DPP4) activity and/or structure homologue (DASH) proteins," *Clinical and Experimental Immunology*, vol. 184, no. 3, p. 265–83, 2016.
- [60] N. Rawlings, A. Barrett, P. Thomas, X. Huang, A. Bateman and R. Finn, "The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database.," *Nucleic Acids Res*, vol. 46, pp. D624-D632, 2018.
- [61] A. M. Lambeir, S. Scharpé and I. De Meester, "DPP4 inhibitors for diabetes-What next?," *Biochemical Pharmacology*, vol. 76, no. 12, pp. 1637-43, 2008.
- [62] K. Ajami, C. A. Abbott, G. W. McCaughan and M. D. Gorrell, "Dipeptidyl peptidase 9 has two forms, a broad tissue distribution, cytoplasmic localization and DPIV-like peptidase activity," *Biochimica et Biophysica Acta - Gene Structure and Expression*, vol. 1679, pp. 18-28, 2004.
- [63] C. A. Abbott, D. M. T. Yu, E. Woollatt, G. R. Sutherland, G. W. McCaughan and M. D. Gorrell, "Cloning, expression and chromosomal localization of a novel human dipeptidyl peptidase (DPP) IV homolog, DPP8," *European Journal of Biochemistry*, vol. 267, no. 20, pp. 6140-50, 2000.
- [64] H. Zhang, S. Maqsudi, A. Rainczuk, N. Duffield, J. Lawrence, F. M. Keane, D. Justa-Schuch, R. Geiss-Friedlander, M. D. Gorrell and A. N. Stephens, "Identification of novel dipeptidyl peptidase 9 substrates by two-dimensional differential in-gel electrophoresis," *FEBS Journal*, vol. 282, p. 3737–57, 2015.
- [65] Spagnuolo P, Hurren R, Gronda M, MacLean N, Datti A, Basheer A, Lin F-H, Wang X, Wrana J and Schimmer A, "Inhibition of intracellular dipeptidyl peptidases 8 and 9 enhances parthenolide's anti-leukemic activity," *Leukemia*, vol. 27, no. 6, pp. 1236-44, 2013.

- [66] Z. Tang, J. Li, Q. Shen, J. Feng, H. Liu, W. Wang, L. Xu, G. Shi, X. Ye, M. Ge, X. Zhou and S. Ni, "Contribution of upregulated dipeptidyl peptidase 9 (DPP9) in promoting tumorigenicity, metastasis and the prediction of poor prognosis in non-small cell lung cancer (NSCLC)," *International Journal of Cancer*, vol. 140, no. 7, pp. 1620-1632, 2017.
- [67] Y. Chen, M. G. Gall, H. Zhang, F. M. Keane, G. W. McCaughan, D. M. Yu and M. D. Gorrell, "Dipeptidyl peptidase 9 enzymatic activity influences the expression of neonatal metabolic genes," *Experimental Cell Research*, vol. 342, no. 1, pp. 72-82, 2016.
- [68] H. Zhang, Y. Chen, C. Wadham, G. W. McCaughan, F. M. Keane and M. D. Gorrell, "Dipeptidyl peptidase 9 subcellular localization and a role in cell adhesion involving focal adhesion kinase and paxillin," *Biochimica et Biophysica Acta - Molecular Cell Research*, vol. 1853, no. 8, pp. 470-480, 2015.
- [69] D. Justa-Schuch, M. Silva-Garcia, E. Pilla, M. Engelke, M. Kilisch, C. Lenz, U. Mö Ller, F. Nakamura, H. Urlaub and R. Geiss-Friedlander, "DPP9 is a novel component of the N-end rule pathway targeting the tyrosine kinase Syk," *Elife*, vol. 10, no. 5, p. e16370, 2016.
- [70] C. Y. Taabazuing, M. C. Okondo and D. A. Bachovchin, "Pyroptosis and Apoptosis Pathways Engage in Bidirectional Crosstalk in Monocytes and Macrophages," *Cell Chemical Biology*, vol. 24, no. 4, pp. 507-514, 2017.
- [71] M. C. Okondo, D. C. Johnson, R. Sridharan, E. B. Go, A. J. Chui, M. S. Wang, S. E. Poplawski, W. Wu, Y. Liu, J. H. Lai, D. G. Sanford, M. O. Arciprete, T. R. Golub, W. W. Bachovchin and D. A. Bachovchin, "DPP8 and DPP9 inhibition induces pro-caspase-1-dependent monocyte and macrophage pyroptosis," *Nature Chemical Biology*, vol. 13, no. 1, pp. 46-53, 2016.
- [72] H. B. Rasmussen, S. Branner, F. C. Wiberg and N. Wagtmann, "Crystal structure of human dipeptidyl peptidase IV/CD26 in complex with a substrate analog," *nature structural biology*, vol. 10, no. 1, pp. 19-25, 2003.
- [73] H. J. Lee, Y. S. Chen, C. Y. Chou, C. H. Chien, C. H. Lin, G. G. Chang and X. Chen, "Investigation of the dimer interface and substrate specificity of prolyl dipeptidase DPP8," *Journal of Biological Chemistry*, vol. 281, no. 50, pp. 38653-62, 2006.
- [74] R. Geiss-Friedlander, N. Parmentier, U. Möller, H. Urlaub, B. J. Van den Eynde and F. Melchoir, "The cytoplasmic peptidase DPP9 is rate-limiting for degradation of proline-containing peptides," *Journal of Biological Chemistry*, vol. 284, no. 40, pp. 27211-9, 2009.
- [75] K. Ajami, C. A. Abbott, M. Obradovic, V. Gysbers, T. Kähne, G. W. McCaughan and M. D. Gorrell, "Structural requirements for catalysis, expression, and dimerization in the CD26/DPIV gene family," *Biochemistry*, vol. 42, no. 3, pp. 694-701, 2003.
- [76] H. K. Tang, K. C. Chen, G. G. Liou, S. C. Cheng, C. H. Chien, H. Y. Tang, L. H. Huang, H. P. Chang, C. Y. Chou and X. Chen, "Role of a propeller loop in the quaternary structure and

- enzymatic activity of prolyl dipeptidases DPP-IV and DPP9," *FEBS Letters*, vol. 585, no. 21, pp. 3409-14, 2011.
- [77] C. Rummey and G. Metz, "Homology models of dipeptidyl peptidases 8 and 9 with a focus on loop predictions near the active site," *Proteins*, vol. 66, no. 1, pp. 160-71, 2007.
- [78] H. D. Ulrich, "The Fast-Growing Business of SUMO Chains," *Molecular Cell*, vol. 32, no. 3, pp. 301-5, 2008.
- [79] F. Melchior and L. Hengst, "SUMO-1 and p53.," *Cell cycle (Georgetown, Tex.)*, vol. 1, no. 4, pp. 243-47, 2002.
- [80] R. T. Hay, "SUMO: A history of modification," *Molecular Cell*, vol. 18, pp. 1-12, 2005.
- [81] E. Pilla, U. Möller, G. Sauer, F. Mattioli, F. Melchior and R. Geiss-Friedlander, "A novel SUMO1-specific interacting motif in dipeptidyl peptidase 9 (DPP9) that is important for enzymatic regulation," *Journal of Biological Chemistry*, vol. 287, no. 53, pp. 44320-9, 2012.
- [82] E. Pilla, M. Kilisch, C. Lenz, H. Urlaub and R. Geiss-Friedlander, "The SUMO1-E67 interacting loop peptide is an allosteric inhibitor of the dipeptidyl peptidases 8 and 9," *Journal of Biological Chemistry*, vol. 288, no. 45, pp. 32787-96, 2013.
- [83] M. L. Smebye, A. Agostini, B. Johannessen, J. Thorsen, B. Davidson, C. G. Tropé, S. Heim, R. I. Skotheim and F. Micci, "Involvement of DPP9 in gene fusions in serous ovarian carcinoma," *BMC Cancer*, vol. 17, no. 1, p. 642, 2017.
- [84] W. T. Jiaang, Y. S. Chen, T. Hsu, S. H. Wu, C. H. Chien, C. N. Chang, S. P. Chang, S. J. Lee and X. Chen, "Novel isoindoline compounds for potent and selective inhibition of prolyl dipeptidase DPP8," *Bioorganic and Medicinal Chemistry Letters*, vol. 15, no. 3, pp. 687-91, 2005.
- [85] G. R. Flentke, E. Munozt, B. T. Hubert, A. G. Plautt, C. A. Kettner and W. W. Bachovchin, "Inhibition of dipeptidyl aminopeptidase IV (DP-IV) by Xaa-boroPro dipeptides and use of these inhibitors to examine the role of DP-IV in T-cell function," *Proc. Nati. Acad. Sci. USA*, vol. 88, pp. 1556-1559, 1991.
- [86] S. J. Coutts, T. A. Kelly, R. J. Snow, C. A. Kennedy, R. W. Barton, J. Adams, D. A. Krolikowski, D. M. Freeman, S. J. Campbell, J. F. Ksiazek and W. W. Bachovchin, "Structure-Activity Relationships of Boronic Acid Inhibitors of Dipeptidyl Peptidase IV. 1. Variation of the P 2 Position of X aa-boroPro Dipeptides," *J. Med. Chem.*, vol. 39, pp. 2087-94, 1996.
- [87] H. Marshall, M. Venkat, N. S. Hti Lar Seng, J. Cahn and D. H. Juers, "The use of trimethylamine N-oxide as a primary precipitating agent and related methylamine osmolytes as cryoprotective agents for macromolecular crystallography," *Acta Crystallographica Section D: Biological Crystallography*, vol. D68, pp. 69-81, 2012.
- [88] Kabsch W, "XDS," *Acta Cryst*, vol. D66, no. 1, pp. 125-132, 2009.

- [89] A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni and R. J. Read, "Phaser crystallographic software," *Journal of Applied Crystallography*, vol. 40, no. 1, pp. 658-674, 2007.
- [90] G. N. Murshudov, A. I. A. Vagin and E. J. Dodson, "Refinement of Macromolecular Structures by the Maximum-Likelihood Method," *Acta Cryst*, vol. 53, no. 3, pp. 240-255, 1997.
- [91] R. A. Engh and R. Huber, "Accurate Bond and Angle Parameters for X-ray Protein Structure Refinement," *Acta Cryst.*, vol. A47, pp. 392-400, 1991.
- [92] I. Langmuir, "The adsorption of gases on plane surfaces of glass, mica and platinum," *Journal of the American Chemical Society*, vol. 40, no. 9, p. 1361-1403, 1918.
- [93] S. Pronk, S. Páll, R. Schulz, P. Larsson, P. Bjelkmar, R. Apostolov, M. R. Shirts, J. C. Smith, P. M. Kasson, D. Van Der Spoel, B. Hess and E. Lindahl, "GROMACS 4.5: A high-throughput and highly parallel open source molecular simulation toolkit," *Bioinformatics*, vol. 29, no. 7, pp. 845-854, 2013.
- [94] B. Ross, S. Krapp, M. Augustin, R. Kierfersauer, M. Arciniega, R. Geiss-Friedlander and R. Huber, "Structures and mechanism of dipeptidyl peptidases 8 and 9, important players in cellular homeostasis and cancer," *Proceedings of the National Academy of Sciences*, vol. 115, no. 7, p. E1437-E1445, 2018.
- [95] M. Engel, T. Hoffmann, L. Wagner, M. Wermann, U. Heiser, R. Kiefersauer, R. Huber, W. Bode, H.-U. Demuth and H. Brandstetter, "The crystal structure of dipeptidyl peptidase IV (CD26) reveals its functional regulation and enzymatic mechanism," *PNAS*, vol. 100, no. 9, p. 5063-5068, 2003.
- [96] E. Krissinel and K. Henrick, "Inference of Macromolecular Assemblies from Crystalline State," *Journal of Molecular Biology*, 2007.
- [97] W. A. Weihofen, J. Liu, W. Reutter, W. Saenger and H. Fan, "Crystal structure of CD26/dipeptidyl-peptidase IV in complex with adenosine deaminase reveals a highly amphiphilic interface," *Journal of Biological Chemistry*, vol. 279, no. 41, pp. 43330-5, 2004.
- [98] M. Marquart, J. Walter, J. Deisenhofer, W. Bode and R. Huber, "The geometry of the reactive site and of the peptide groups in trypsin, trypsinogen and its complexes with inhibitors," *Acta Crystallographica Section B*, 1983.
- [99] A. D. Capili and C. D. Lima, "Structure and Analysis of a Complex between SUMO and Ubc9 Illustrates Features of a Conserved E2-Ubl Interaction," *Journal of Molecular Biology*, vol. 369, no. 3, pp. 608-18, 2007.
- [100] S. Van Goethem, V. Matheeußen, J. Joossens, A. M. Lambeir, X. Chen, I. De Meester, A. Haemers, K. Augustyns and P. Van Der Veken, "Structure-activity relationship studies on

- isoindoline inhibitors of dipeptidyl peptidases 8 and 9 (DPP8, DPP9): Is DPP8-selectivity an attainable goal?," *Journal of Medicinal Chemistry*, vol. 54, no. 16, pp. 5737-46, 2011.
- [101] H. Fu, H. Fang, J. Sun, H. Wang, A. Liu, J. Sun and Z. Wu, "Boronic Acid-based Enzyme Inhibitors: A Review of Recent Progress," *Current Medicinal Chemistry*, vol. 21, pp. 3271-3280, 2014.
- [102] E. G. Van Schaftingen, E. Walsh, B. Walsh and A. B. Shenvi, "α-Aminoboronic acid derivatives: Effective inhibitors of aminopeptidases," *Adv. Enzymol. Relat. Areas Mol. Biol.*, vol. 94, pp. 1286-1291, 1986.
- [103] H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov and P. E. Bourne, "The Protein Data Bank," *Nucleic Acids Research*, vol. 28, no. 1, pp. 235-242, 2000.
- [104] Z. S. Derewenda, "Rational protein crystallization by mutational surface engineering," *Structure*, vol. 12, no. 1, pp. 529-35, 2004.
- [105] T. L. Blundell, "Protein crystallography and drug discovery: Recollections of knowledge exchange between academia and industry," *IUCrJ*, vol. 4, pp. 308-21, 2017.
- [106] C. M. Lobley, J. Sandy, J. Sanchez-Weatherby, M. Mazzorana, T. Krojer, R. P. Nowak and T. L. Sorensen, "A generic protocol for protein crystal dehydration using the HC1b humidity controller," *Acta Crystallographica Section D: Structural Biology*, vol. D72, pp. 629-40, 2016.
- [107] J. H. Wang, "Self-diffusion coefficients of water," *J. Phys. Chem.*, vol. 69, no. 12, p. 4412, 1965.
- [108] V. N. Morozov, G. S. Kachalova, V. U. Evtodienko, N. F. Lanina and T. Y. Morozova, "Permeability of lysozyme tetragonal crystals to water," *European Biophysics Journal*, vol. 24, no. 1, pp. 93-98, 1995.
- [109] Y. Miyazaki, T. Matsuo and H. Suga, "Low-Temperature Heat Capacity and Glassy Behavior of Lysozyme Crystal," *The Journal of Physical Chemistry B*, vol. 104, no. 33, p. 8044-52, 2000.
- [110] R. S. Smith and B. D. Kay, "The existence of supercooled liquid water at 150 K," *Nature*, vol. 398, pp. 788-791, 1999.
- [111] S. W. Wukovitz and T. O. Yeates, "Why protein crystals favour some space-groups over others," *Nature Structural Biology*, vol. 2, pp. 1062-67, 1995.
- [112] M. Weik, A. M. M. Schreurs, H. K. S. Leiros, G. Zaccai, R. B. G. Ravelli and P. Gros, "Supercooled liquid-like solvent in trypsin crystals: Implications for crystal annealing and temperature-controlled X-ray radiation damage studies," *Journal of Synchrotron Radiation*, vol. 12, pp. 310-317, 2005.
- [113] V. R. Samygina, S. V. Antonyuk, V. S. Lamzin and A. N. Popov, "Improving the X-ray resolution by reversible flash-cooling combined with concentration screening, as exemplified with PPase," *Acta Crystallographica Section D: Biological Crystallography*, vol. D56, pp. 595-603, 2000.

- [114] Z. Q. Fu, G. C. Du Bois, S. P. Song, R. W. Harrison and I. T. Weber, "Improving the diffraction quality of MTCP-1 crystals by post-crystallization soaking," *Acta Crystallographica Section D: Biological Crystallography*, vol. D55, pp. 5-7, 1999.
- [115] Q. Zou, B. J. Bennion, V. Daggett and K. P. Murphy, "The molecular mechanism of stabilization of proteins by TMAO and its ability to counteract the effects of urea," *Journal of the American Chemical Society*, vol. 124, no. 7, pp. 1192-1202, 2002.
- [116] J. Jiang, E. M. Lafer and R. Sousa, "Crystallization of a functionally intact Hsc70 chaperone," *Acta Crystallographica Section F: Structural Biology and Crystallization Communications*, vol. F62, pp. 39-43, 2006.
- [117] J. Tang, J. Majeti, A. Sudom, Y. Xiong, M. Lu, Q. Liu, J. Higbee, Y. Zhang, Y. Wang, W. Wang, P. Cao, Z. Xia, S. Johnstone, X. Min, X. Yang, H. Shao, T. Yu, N. Sharkov, N. Walker, H. Tu, W. Shen and Z. Wang, "An inhibitory antibody against dipeptidyl peptidase IV improves glucose tolerance in vivo," *Journal of Biological Chemistry*, vol. 288, no. 2, pp. 1307-16, 2013.
- [118] L. Heirbaut, S. Van Goethem, K. Jansen, H. De Winter, N. Lamoen, J. Joossens, J. Cheng, X. Chen, A. M. Lambeir, I. De Meester, K. Augustyns and P. Van Der Veken, "Probing for improved selectivity with dipeptide-derived inhibitors of dipeptidyl peptidases 8 and 9: The impact of P1-variation," *MedChemComm*, vol. 7, no. 3, pp. 433-438, 2016.
- [119] J. J. Wu, H. K. Tang, T. K. Yeh, C. M. Chen, H. S. Shy, Y. R. Chu, C. H. Chien, T. Y. Tsai, Y. C. Huang, Y. L. Huang, C. H. Huang, H. Y. Tseng, W. T. Jiaang, Y. S. Chao and X. Chen, "Biochemistry, pharmacokinetics, and toxicology of a potent and selective DPP8/9 inhibitor," *Biochemical Pharmacology*, vol. 78, no. 2, pp. 203-10, 2009.
- [120] D. Rea, R. Van Elzen, H. De Winter, S. Van Goethem, B. Landuyt, W. Luyten, L. Schoofs, P. Van Der Veken, K. Augustyns, I. De Meester, V. Fülöp and A. M. Lambeir, "Crystal structure of *Porphyromonas gingivalis* dipeptidyl peptidase 4 and structure-activity relationships based on inhibitor profiling," *European Journal of Medicinal Chemistry*, vol. 139, pp. 482-91, 2017.
- [121] Y. Nakajima, K. Ito, T. Toshima, T. Egawa, H. Zheng, H. Oyama, Y. F. Wu, E. Takahashi, K. Kyono and T. Yoshimoto, "Dipeptidyl aminopeptidase IV from *Stenotrophomonas maltophilia* exhibits activity against a substrate containing a 4-hydroxyproline residue," *Journal of Bacteriology*, vol. 190, no. 23, p. 7819-29, 2008.
- [122] S. Roppongi, Y. Suzuki, C. Tateoka, M. Fujimoto, S. Morisawa, I. Iizuka, A. Nakamura, N. Honma, Y. Shida, W. Ogasawara, N. Tanaka, Y. Sakamoto and T. Nonaka, "Crystal structures of a bacterial dipeptidyl peptidase IV reveal a novel substrate recognition mechanism distinct from that of mammalian orthologues," *Scientific Reports*, vol. 8, no. 2714, pp. 1-18, 2018.

