Neutron spectroscopy and crystallography to understand the structure-dynamics relationship in enzymes

Submitted in accordance with the requirements for the degree of
Doktor der Naturwissenschaften

Tushar Bhim Raskar
Universität Hamburg
Fachbereich Chemie

November 2020

This thesis was reviewed by:

Dr. Nils Huse (Universität Hamburg, Fachbereich Physik)

Dr. Henning Tidow (Universität Hamburg, Fachbereich Chemie)

Date of Disputation: 18th of December 2020

Date of release for publication: 15th of January 2021

Declaration/Erklärung

I hereby declare that this doctoral dissertation is my own work and that I have not used any sources other than those listed.

Hiermit erkläre ich an Eides statt, die vorliegende Dissertation selbst verfasst und keine anderen als die angegebenen Hilfsmittel benutzt zu haben.

Place, Date, Signature: Tushar Raskar

Preamble

The work in this thesis was carried out in Center for Free Electron Laser Science, Universität Hamburg, Hamburg, jointly with Institut Laue-Langevin (ILL), Grenoble, between January 2017 and February 2021. The work was supervised by Prof. Dr. Arwen Pearson (Universität Hamburg), Dr. Tilo Seydel (ILL), Prof. Dr. Trevor Forsyth (ILL).

Acknowledgements

First, I would like to thank my supervisor Dr. Arwen Pearson for giving me the opportunity to work on this highly interdisciplinary project. Her constant support, motivation and supervision have helped me tremendously throughout the entire period of this project and leaving her group to work elsewhere feels very sad.

The help and support from our collaborators has also been invaluable. It would have been impossible to reach any of the milestones of this thesis without them. The two years of my PhD which I spent in the European Photon and Neutron science campus in Grenoble were intellectually stimulating. It was a sheer pleasure to work, discuss and live science while I was in Grenoble. I thank Dr. Trevor Forsyth for all the support that he has provided for the experimental work that has gone into this thesis. I would like to give a special thanks to Dr. Michael Härtlein for his critical comments on proposal and his overall contribution to this project. Working in the Life Science Group was a great learning experience and gave me opportunity to work in one of the best perdeuteration facilities in the world. Next, I cannot thank Dr. Tilo Seydel enough who has been my mentor throughout this project. The long discussions with him and his thorough explanations have helped me understand the fundamental principles of different neutron spectroscopy techniques and specifically, Quasi-Elastic Neutron Scattering from scratch. I would like to thank Dr. Juliette Devos in Life Sciences Group who was my friend and colleague during my tenure in ILL. She has helped me with the experiments, sample preparations and in sorting out many aspects related with the general logistics in France. I will forever be grateful to her for being so nice. I thank Susanne Meier in the group in Hamburg for being so kind and helpful, especially in setting up the crystallization trials when I was in Grenoble. I thank Dr. Michael Marek Koza (IN5 beamline, ILL) and Dr. Monica Jimenez-Ruiz (IN1 beamline, ILL) for their support during beamtimes and discussions which helped me in data analysis and interpretation. I was privileged to have an extremely supportive friend and colleague like Dr. Stephan Niebling and I thank him for all his help. He has been a great support for running and analyzing Molecular Dynamics simulations of ADC. I also thank Dr. Briony Yorke for her

contribution to the project and for giving me access to the X-ray diffraction data which was eventually analyzed and included in the thesis. I would like to thank Dr. Nils Huse for all the discussions and his overall contribution to this project.

I thank all my labmates in the groups in Hamburg and in Grenoble with whom I had a great time. The dinners we had, the coffee breaks full of diverse range of discussions and the lab fun were all really enjoyable. I would also like to extend my thanks to Ms. Shilpa Sharma in Indian Institute of Technology Delhi for providing the support for running the molecular dynamics simulations of *E.coli* copper amine oxidase in a short time frame. We hope to continue our collaboration in future as well in order to keep working on the project that we have started.

I thank Centre for Ultrafast Imaging, University of Hamburg and ILL PhD program for funding my work. The work was carried out in Centre for Free Electron Laser Science, University of Hamburg in my first year. My second and third years were spent in Institut Laue-Langevin, Grenoble where Dr. Tilo Seydel and Prof. Dr. Trevor Forsyth were my local supervisors. The ILL PhD seminars were very helpful in engaging in critical discussions on various neutron based methods and to get to know about the projects which the fellow PhD students were working on. My last year was spent in Hamburg and was funded by CUI, University of Hamburg. Last but not the least, I would like to thank my parents who have always stood by me and I would not be who I am today without them.

Contents

List of Figures

List of Tables

List of Abbreviations and Symbols

$\langle \cdots \rangle$	Denotes average over indicated quantity in the text (usually time origins)
ADC	Aspartate α-decarboxylase
ECAO	*E.coli* copper amine oxidase
EISF	Elastic Incoherent Structure Factor
DLS	Dynamic light scattering
(E/I)FWS	(Elastic/Inelastic)Fixed-window scan
EINS	Elastic Incoherent neutron scattering
EISF	Elastic incoherent structure factor
FWHM	Full width at half maximum
GDOS	Generalized Density of States
ILL	Institut Laue-Langevin (Grenoble - France)
INS	Inelastic Neutron Scattering
k_{B}	Boltzmann constant, $k_{\mathrm{B}} = 8.6173303\ eV/K$
LB	Luria Bertani
MD	Molecular dynamics
MSD	Mean squared displacement
NMR	Nuclear magnetic resonance

NPT	constant number (N), pressure (P), and temperature (T)
NVT	constant number (N), volume (V), and temperature (T)
PDT	Protein dynamical transition
PST	Phase transformation chopper
QENS	Quasi-elastic neutron scattering
TEMED	Tetramethylene diamine
THz-TDS	Terehertz time domain spectroscopy
TPQ	2,4,5-trihydroxyphenylalanine quinone

Abstract

Proteins are key macromolecules required for a plethora of critical biological processes and hence, are called the workhorses of the cell. Various techniques have been used to probe the structure and dynamics of proteins over the past several decades. These techniques cover different time and length scales associated with structure and dynamics of proteins. When these techniques are employed at different temperatures, they can help establish the structure-function-dynamics relationship in proteins. However, it is a challenge to correlate the results from solution, powder and crystalline states of these molecules. In this thesis, we have used several complementary crystallographic and spectroscopic techniques which cover a broad energy range to cross-correlate the results between these three properties. The experiments presented here were acquired at Institut Laue-Langevin (ILL), Grenoble and Center for Free Electron Laser Science (CFEL), Hamburg. We have used two well studied enzyme systems to demonstrate the complementarity between these techniques: *Escherichia coli* copper amine oxidase (ECAO) and aspartate α-decarboxylase (ADC).

In case of ADC, we have extensively studied the effect of binding of D-Serine which is a substrate analogue of its natural substrate (L-aspartate), on its structure and dynamics. The structural change induced following the binding of D-Serine was studied using X-ray diffraction. The effect of binding of D-Serine on the dynamics of ADC is also explored. For this purpose, we studied the diffusive dynamics of ADC using neutron backscattering and dynamic light scattering techniques. The existence of higher order oligomers of ADC tetramer on nanosecond to millisecond time scales was observed. Binding of D-Serine appears to further increase the size of these higher order oligomers. Hence, the binding of the ligand is found to affect both the internal and global dynamics of ADC.

We also probed the vibrational dynamics and the evolution of dynamics at atomic level as a function of temperature for both apo (ADC-APO) and D-Serine complexed (ADC-LIG) states of ADC using inelastic neutron scattering. These studies showed that the change in vibrational dynamics results in sharper peaks for ADC-LIG complex than for ADC-APO and hence are in general, consistent with the results from

solution backscattering experiments which suggest stiffening of the protein as a result of binding of the ligand. In addition, it was observed that binding of D-Serine is associated with stabilization of ADC which makes it harder than its apo state.

We explored the effect of two mutations: E573Q and I342F on the dynamics of ECAO in solution using neutron backscattering and dynamic light scattering. Despite the fact that these mutations are $\approx$ 20-30 Å away from the active site of ECAO, they are observed to influence the internal and global dynamics of ECAO. The (E573Q/I342F) double mutant of ECAO is observed to be more flexible than wild-type ECAO. The dynamics are affected to a higher extent on the nanosecond time scale than on the millisecond time scale.

Zusammenfassung

Proteine sind Makromoleküle, die für eine Vielzahl kritischer biologischer Prozesse benötigt werden, und werden daher als die Arbeitspferde der Zelle bezeichnet. In den letzten Jahrzehnten wurden verschiedene Techniken eingesetzt, um die Struktur und Dynamik von Proteinen zu untersuchen. Diese Techniken decken verschiedene Zeit- und Längenskalen ab, die mit der Struktur und Dynamik von Proteinen verbunden sind. Wenn diese Techniken bei unterschiedlichen Temperaturen eingesetzt werden, können sie dazu beitragen, die Struktur-Funktions-Dynamik-Beziehung in Proteinen zu ermitteln. Es ist jedoch eine Herausforderung, die Ergebnisse von flüssigen, pulverförmigen und kristallinen Zuständen dieser Moleküle zu korrelieren. In dieser Arbeit haben wir mehrere komplementäre kristallographische und spektroskopische Techniken verwendet, die einen breiten Energiebereich abdecken, um die Ergebnisse zwischen diesen drei Zuständen miteinander zu korrelieren. Die hier vorgestellten Experimente wurden am Institut Laue-Langevin (ILL), Grenoble, und am Center for Free Electron Laser Science (CFEL), Hamburg, durchgeführt. Wir haben zwei gut untersuchte Enzymsysteme verwendet, um die Komplementarität zwischen diesen Techniken nachzuweisen: Escherichia coli Kupfer-Amine-Oxidase (ECAO) und Aspartat α-Decarboxylase (ADC).

Im Falle der ADC haben wir die Wirkung der Bindung von D-Serin, das ein Substratanalog seines natürlichen Substrats (L-Aspartat) ist, auf seine Struktur und Dynamik eingehend untersucht. Die nach der Bindung von D-Serin induzierte Strukturveränderung wurde mittels Röntgenbeugung untersucht. Die Wirkung der Bindung von D-Serin auf die Dynamik der ADC wurde ebenfalls untersucht. Zu diesem Zweck untersuchten wir die Diffusionsdynamik von ADC mit Neutronenrückstreuung und dynamischen Lichtstreuungstechniken. Es wurde die Existenz von Oligomeren höherer Ordnung als ADC-Tetramer auf Zeitskalen von Nanosekunden bis Millisekunden beobachtet. Die Bindung von D-Serin scheint die Größe dieser Oligomere höherer Ordnung weiter zu erhöhen.

Wir untersuchten auch die Schwingungsdynamik und die Entwicklung der Dynamik auf atomarer Ebene als Funktion der Temperatur sowohl für apo (ADC-

APO) als auch für D-Serin-komplexierte (ADC-LIG) Zustände von ADC unter Verwendung inelastischer Neutronenstreuung. Diese Studien zeigten, dass die Änderung der Schwingungsdynamik für den D-Serin-Komplex zu schärferen Peaks führt als für das Apoenzyme und daher im Allgemeinen mit den Ergebnissen von Lösungsrückstreuexperimenten übereinstimmt, die eine Versteifung des Proteins als Folge der Bindung des Liganden vermuten lassen. Darüber hinaus wurde beobachtet, dass die Bindung von D-Serin mit einer Stabilisierung von ADC assoziiert ist, was es steifer macht als seinen apo-Zustand.

Wir untersuchten den Effekt zweier Mutationen (E573Q und I342F) auf die Dynamik der ECAO in Lösung mit Neutronenrückstreuung und dynamischer Lichtstreuung. Trotz der Tatsache, dass diese Mutationen ≈20-30 Å von dem aktiven Zentrum der ECAO entfernt sind, wird beobachtet, dass sie die interne und globale Dynamik der ECAO beeinflussen. Es wird beobachtet, dass die (E573Q/I342F)-Doppelmutante der ECAO flexibler ist als die Wildtyp-ECAO. Die Dynamik wird auf der Nanosekunden-Zeitskala stärker beeinflusst als auf der Millisekunden-Zeitskala.

Chapter 1

Introduction and background

1.1 Introduction

Proteins are ubiquitous entities in living systems that play a vital role in a variety of natural processes. These processes include catalysis of a plethora of chemical reactions which are required for some of the most important biological processes e.g. proteasomal degradation, fatty acid metabolism, transport of nutrients across the membrane *etc*. This functional role of proteins renders them essential for life. Understanding the structural and functional details of proteins at molecular level is therefore important to:

1. Gain an insight into the fundamental molecular phenomena associated with function of proteins in general given that they are key to understanding the governing principles of the biological events at molecular level.

2. Using the gained knowledge to inhibit and possibly control these chemical reactions which maybe of importance for the survival of disease causing pathogens [1].

While the first objective is important for an understanding of the physico-chemical principles which govern the behavior of biomolecules, the second is important for application of this knowledge base in therapeutics or drug discovery [2]. For both these interrelated objectives, it is extremely important to access both structure and dynamics of biological macromolecules as they govern the function of these biomolecules.

Single crystal X-ray diffraction has been used as a primary technique in structural biology. In this technique, the scattered X-rays are measured at increasing scattering angles, with the higher scattering angle corresponding to the structural

information at a higher resolution. In a single crystal diffraction experiment, the changes in energy are typically not measured. Since the scattered X-ray intensity is measured as a function of the scattering angle and not as a function of energy, the resolution mainly reflects spatial resolution [3]. X-ray diffraction experiments are therefore normally used to derive information on the structure of the molecule at a high resolution. However, changes in dynamics cannot be directly inferred from these experiments.

X-ray crystallography has been used to explore the structural change in biomolecules as a function of several parameters. These include room temperature experiments that are used to understand the effect of temperature on the conformational subspace and flexibility of specific residues [4], soaking protein crystals or co-crystallizing proteins with ligands to decipher the key interactions between the protein and ligand [5, 6], using specific conditions (e.g. mutations, changes in pH, soaking conditions) to capture transient intermediates during the course of a chemical reaction [7, 8]. One of the most common ways this is achieved is by using substrate analogues which arrest the reaction at a specific stage of the catalytic cycle [8]. These methods exploit the property of highly ordered crystals to diffract to high resolution. Typically, this 'high resolution' is taken to be $\approx$ 2.0-1.5 Å or higher. It is less common to obtain data higher than 1.1 Å which in some cases can make it possible to resolve a fraction of the hydrogen atom coordinates along with the heavy atoms, C, O and N. On the other hand, resolution less than 1.5 Å can only be used to resolve the heavy atoms as their scattering power is several fold higher than hydrogen due to a higher atomic number.

Neutron diffraction can be used to reliably resolve the positions of deuterium atoms after successful crystallization of a perdeuterated biomolecule [9]. This technique exploits the high coherent scattering length of deuterium (+6.67 fm) which is comparable to that of heavy atoms of the protein (C, N, O) [10] and hence can be used to resolve the positions of deuterium atoms even at a resolution as low as 2.5-2.7 Å [11, 12]. Moreover, the difference in the sign of the coherent scattering lengths of D and H (positive for D and negative for H) allow D/H replacement to be exploited to selectively change the visibility of specific regions of the protein [11, 13]. This method is known as 'contrast variation' [13]. This selection is not possible in case of X-rays as the visibility of atoms depends on their electron density.

Although X-ray and neutron diffraction are very important to understand the structure of biomolecules, there are some limitations to these methods. These are mainly due to the spatiotemporally averaged information that is obtained from the crystals. It is also important to note that there exists a fundamental difference in the thermodynamics of crystalline and physiologically relevant solution states.

There are several works which investigate the fundamental differences between the thermodynamic parameters of a protein in crystalline form and in solution [14–16]. Murakawa and others studied the equilibrium mixture between two conformations of the topaquinone cofactor and compared the thermodynamic parameters for this equilibrium in solution and in crystal [14]. Interestingly, the ΔH° and ΔS° for transition between the two conformations were found to be ≈ 1.5 fold higher in crystal than those in solution [14]. These changes in thermodynamic parameters indicate that the observables related with the structure and dynamics of protein are significantly affected by whether the experiment is carried out using solutions, crystals or amorphous powder samples [17–19].

Pérez and others systematically studied the evolution of internal dynamics as the hydration level of Lysozyme and Myoglobin powders is increased in order to capture the evolution of internal dynamics from dry powder to solution. They reported that when the protein is covered by a monolayer of water, the protein side chains diffuse locally. As the hydration level is increased, the major effect is seen in the rate at which these diffusive motions occur [18]. The average amplitude of these motions was found to be ≈ 3 times higher in solution compared to that in hydrated powders whereas the relaxation time is found to be approximately double for the hydrated powder as compared to the solution [18]. Another study focused on the calculation of the mean squared displacements (MSD) of human butyrylcholinesterase at different length and time scales using elastic incoherent neutron scattering (EINS) [19]. The fundamental concepts related with EINS will be explained in chapter 2. The authors observed the onset of translational diffusion in solution. It was also inferred that as the powder state changes to the solution state, the drop in the recorded elastic intensity is faster [19]. To summarize, these studies indicate that as the powder state changes into a solution state, the rate of diffusion of the more disordered parts of the protein increases. It is therefore clear that it is challenging to combine the information obtained from solution, powder and crystalline states as the fundamental physical properties of these states are different which can also affect the observables associated with a biological phenomenon under study. Hence, it is evident from the aforementioned studies that there are different experimental windows which exist in terms of the length and the time scales into which the molecular motions of the protein fall.

Figure 1.1 shows the different time and length scales which can be probed using inelastic neutron scattering, dynamic light scattering, Brillouin and Raman spectroscopy and other related techniques [20]. It can be seen that combining different methods can be useful in resolving different dynamical modes. For example, a combination of inelastic X-ray and neutron scattering can be used to elucidate dynamical behavior ranging from 1 Å to greater than 100 nm [20]. While inelastic X-ray scattering can probe dynamics at large length scales (smaller scattering vector Q), inelastic neutron scattering has proved to be an important tool in probing changes on smaller length scales (larger Q). On the other hand, quasi-elastic neutron scattering (QENS), is a special case of inelastic neutron scattering in which the incoherent part of the scattered neutrons can be used to access length scales from Ångströms to nanometers and time scales from picoseconds to nanoseconds. This is exactly the dynamic range that is relevant for proteins. These motions, which correspond to an energy range between 30 and 100 μeV are the 'relaxation motions' which are associated with the diffusive and subdiffusive dynamics of biomolecules [21,22]. Similarly, dynamic light scattering (DLS) can be used to investigate into the diffusive dynamics of molecules in solution but at a millisecond time scale as opposed to the nanosecond time scale in case of QENS [23]. DLS involves the use of a laser having a wavelength of ≈ 650 nm which can be used to calculate translational diffusion coefficients of macromolecules and can be used as a complementary technique to QENS [24].

QENS can also be used to probe the dynamics of protein in a powdered state. This approach is especially useful in the study of internal diffusive dynamics of proteins since the global center-of-mass diffusion is switched off in a powdered state [25]. Before performing a classical QENS experiment, it is a routine practice to perform an elastic fixed window scan experiment (EFWS). This experiment is analogous to differential scanning calorimetry and involves recording elastic incoherent intensity of the samples within a specific range of temperatures [26]. EFWS can be used to derive 'vibrational mean squared displacements' [25]. Hartmann and coworkers, using X-ray crystallography and Mössbauer spectroscopy, measured these mean-square displacements, of atoms in metmyoglobin [27]. They could observe both harmonic and anharmonic motions as a function of temperature. They found that at 80K, most of the modes of intramolecular motion are frozen out suggesting well defined modes at energies larger than $kT\ (80\ K) = 7\ meV$. However, even at this temperature, a fraction of atoms have mean squared displacements greater than 0.1 Å^2 which indicates the presence of conformational substates with energy barriers below the available thermal energy at 80 K.

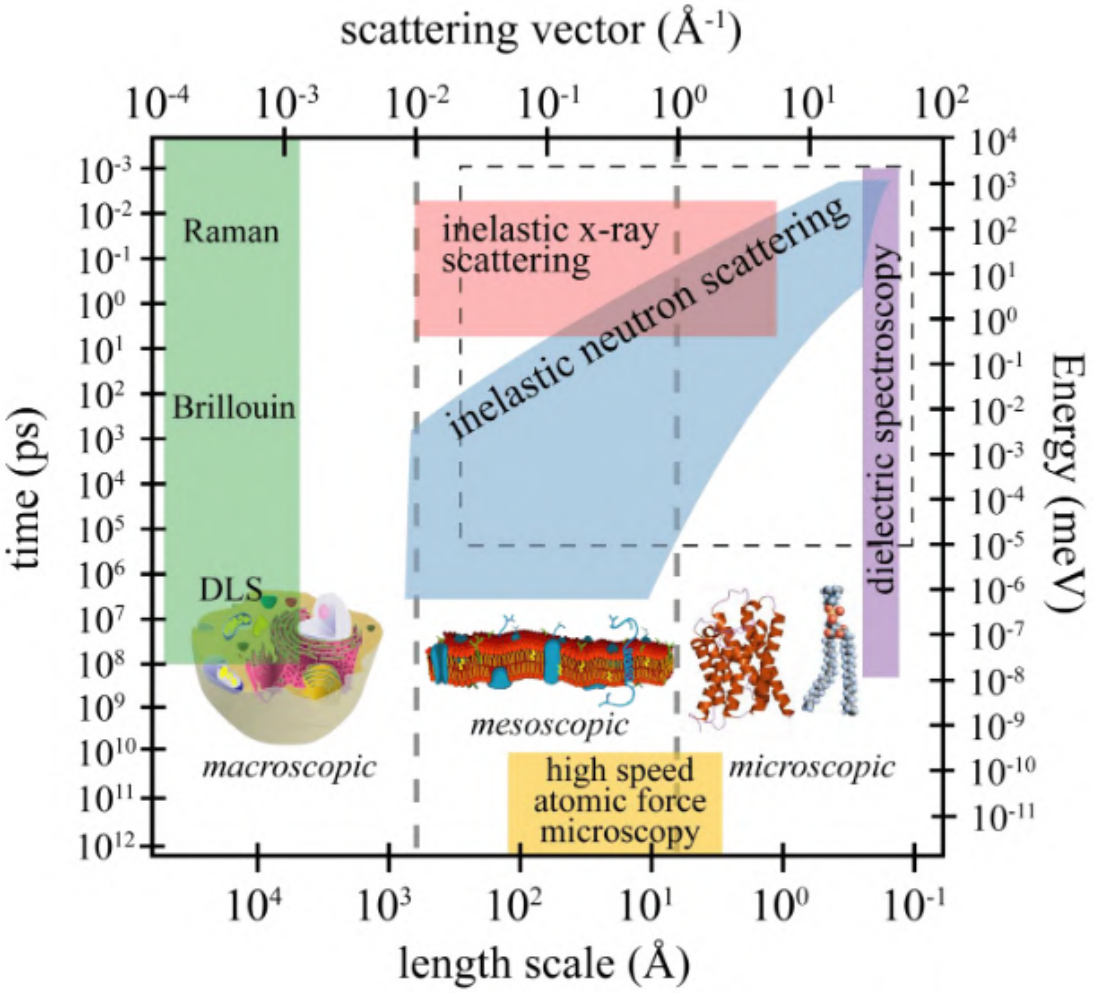

Figure 1.1: An overview of the length and time scales and the corresponding energy and momentum transfers for the spectroscopic techniques covering a broad range of dynamics. The dotted lines forming the top-right quadrilateral enclose the area which represent the dynamical range which can be accessed by MD simulations. Reprinted from [20]

Doster and coworkers in their work found that that harmonic and vibrational motions can be well described by a Debye-Waller factor using inelastic neutron scattering (INS), with mean squared displacement proportional to temperature, except for very low temperatures, where quantum tunnelling effects dominate [28]. The mean squared displacement can be used to calculate the B-factor or temperature factor using the following equation:

$$B = 8\pi^2\langle u^2\rangle \tag{1.1}$$

where, $\langle u^2\rangle$ is the mean squared displacement. The B-factor depends on the scattering contribution from a particular atom which reduces with an increase in the temperature causing an increase in the B-factor. In Doster's work, a transition is observed, around 200 K, which could be explained by considering an atom jump between two sites separated by a given distance d, and by an energy ΔG, and was termed the 'protein dynamical transition (PDT)'.

EFWS can therefore be used to determine the temperatures at which different dynamical modes in the protein are activated and hence the temperature at which a classical QENS spectra are to be recorded depending on the type of dynamics to be observed (e.g. methyl rotations, larger internal motions *etc*). Using high-resolution neutron backscattering (cf. section 2.2), a first transition typically becomes visible around 135 K figure 1.2), and corresponds to the onset of methyl group rotation [30–33]. This is followed by a second transition, corresponding to the aforementioned PDT, associated with side-chain motions and backbone fluctuations. The temperature dependence of the mean squared displacement after this second transition is highly dependent on amino acid composition, as demonstrated in a neutron scattering study [34].

Interestingly, the mean squared displacement presents a relatively small contribution from the backbone compared to the contribution from amino acid composition, as it was shown by Schirò and co-workers [34]. The backbone and amino acid sequence might mostly influence function-related structure and slow motions. We can consider that fast dynamics, less than the nanosecond time-scale - provides the most significant contribution to conformational stability, except for some proteins that perform electron transport for instance. This is due to the fact that the motions which fall within sub-nanosecond regime, are generally related to conformational entropy [35]. Slow, collective dynamics are mostly related to function [36]. Fast dynamics contributions can be understood in terms of free energy, using

$$\mathrm{G} = \mathrm{H} - \mathrm{TS} \tag{1.2}$$

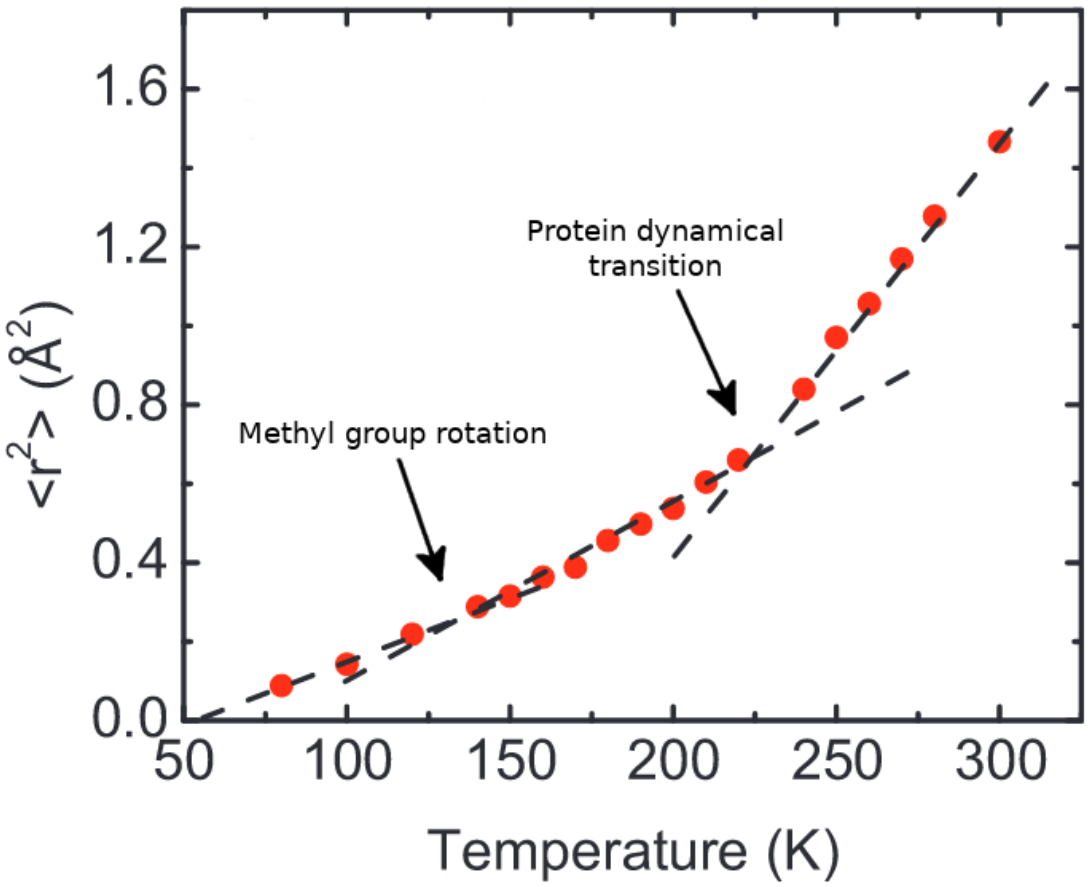

Figure 1.2: Protein dynamical transition (PDT) resulting from anharmonic motions activation in Lysozyme, seen in the apparent mean-squared displacement ($\langle r^2 \rangle$). Typical experimental result obtained with elastic fixed windows scans (EFWS) obtained on a neutron backscattering spectrometer are shown in blue - here, Lysozyme. Dashed lines are added for better visualisation of the transitions. Adapted from [29]

with G, the free energy, H the enthalpy, T the temperature and S the entropy. EFWS and the analysis involved is described in detail in chapter 3. It is therefore clear that it is possible to extract information on collective and single particle diffusive dynamics using solution techniques, on internal dynamics using powder scattering techniques and on the detailed structure using crystallographic methods. As all these methods are essential but insufficient to cover a broad range of dynamical modes, there is a need to cross-correlate the results to understand and correlate structure with internal and global dynamics of biomolecules as each of these techniques can be used to explore a different dynamic range of biological macromolecules. In the present work, I have attempted to integrate neutron and X-ray crystallography with QENS and DLS to access the internal and global diffusive dynamics and molecular dynamics simulations to corroborate the results from the aforementioned methods. In this project, I have used two exemplar enzyme systems to demonstrate that a combination of these methods can be used to gain a comprehensive understanding of the structure-function-dynamics relationship in these biomolecules. The two enzymes are: aspartate α-decarboxylase and *E.coli* copper amine oxidase.

1.2 Aspartate α-decarboxylase

Aspartate α-decarboxylase (ADC) is a homotetrameric enzyme which catalyzes the oxidative dcarboxylation of L-aspartate into β-alanine [37]. β-alanine is required for pantothenate biosynthesis pathway [38] after which pantothenate is converted to its biologically active form, Coenzyme A (CoA) [1] which is essential to all life forms. Even in prokaryotes, CoA is used for diverse purposes. These range from acting as a carrier for naturally occurring or xenobiotic weak organic acids of varied lengths, to maintaining redox homeostasis in cells [39–41]. In bacteria, the pantothenate pathway is a branched pathway in which β-alanine is one of the intermediates which along with pantoate, forms pantothenate [42]. β-alanine is derived from different sources in the case of bacteria, yeast and plants. In case of bacteria, it is derived from L-aspartate by the action of aspartate α-decarboxylase enzyme (ADC) which is encoded by panD gene [37]. Hence, ADC is an essential source of β-alanine in *E.coli*.

1.2.1 Structure of ADC

The first structure of ADC was reported in 1998 by Albert and others [43]. The tetramer is mushroom shaped and has a pseudo four-fold axis which corresponds

to the two-fold crystallographic axis leading to two molecules per asymmetric unit. The molecule is ≈ 35 Å in height and 50 Å in diameter. The tetramer has a solvent filled cavity which runs across the four-fold axis. There exists an acidic region at the base of the tetramer mainly due to ASP-95, GLU-96 and GLU-97 on helix $\alpha 3$ of the base. This region can potentially mediate the assembly on the cell surface. Figure 1.3 depicts the three dimensional structure of mature ADC. The core of the molecule comprises a six-stranded β-barrel capped by small α-helices at each end. The β-barrel has a local two-fold axis relating strands $\beta 1$-$\beta 1$', $\beta 2$, $\beta 3$ and helix $\alpha 1$ with strands $\beta 4$, $\beta 5$, $\beta 6$ and helix $\alpha 2$. The pairs of antiparallel strands $\beta 1$-$\beta 1$'$\beta 2$ and $\beta 4\beta 5$ are interdigitated and related by a two-fold axis to form two ψ-structures, in which $\beta 1$-$\beta 1$'$\beta 2$ and $\beta 4\beta 5$ lie over the central strands of the ψ-structures [43]. The strands $\beta 1$ and $\beta 6$ at the N- and the C-termini respectively form a β ribbon. This forms the β cylinder at the base of the tetramer. The inner surface of the cylinder contains a pattern of hydrophobic (MET-1, ILE-2, PRO-94), polar (THR-4, THR-92) and hydrophobic (MET-5, VAL-91) residues. The interface between the β barrels is lined by polar residues with fewer hydrogen bonds, indicating a weaker interaction, hence allowing the cap to open up so that the substrate can enter the active site [43]. This particular structural change is very important for the turnover of the substrate since this is the very first step after which the conversion of L-aspartate to β-alanine is initiated. In terms of molecular dynamics, this is the major functionally relevant motion in ADC which influences the catalysis. To this end, we have explored the changes in structure and dynamics of ADC associated with binding of a substrate inhibitor (D-Serine) using X-ray crystallography, quasi-elastic neutron scattering (QENS) and molecular dynamics simulations.

1.2.2 Activation of PanD

Pyruvoyl-dependent enzymes are a small subset of enzymes which also includes decarboxylases e.g. S-adenosylmethionine decarboxylases [45, 46]. ADC is is synthesized as a 14 kDa zymogen pro-protein called PanD. The zymogen form contains a single chain, the π-chain. This zymogen form undergoes self-processing or self-catalysed proteolysis leading to the formation of active enzyme [38]. The post-translational modification is initiated with an N-O acyl shift between GLY-24 and SER-25 (figure 1.4). The SER-25 hydroxyl group then attacks the carbonyl backbone of GLY-24 (I), resulting in the formation of oxyoxazolidine ring (II). The ring opens to form an ester intermediate (III) [43]. This is cleaved through β-elimination reaction to a C-terminal Glycine (GLY-24) and dehyhdroalanine (IV) which is hydrolyzed to form the N-terminal pyruvoyl group (V). This step involving self processing

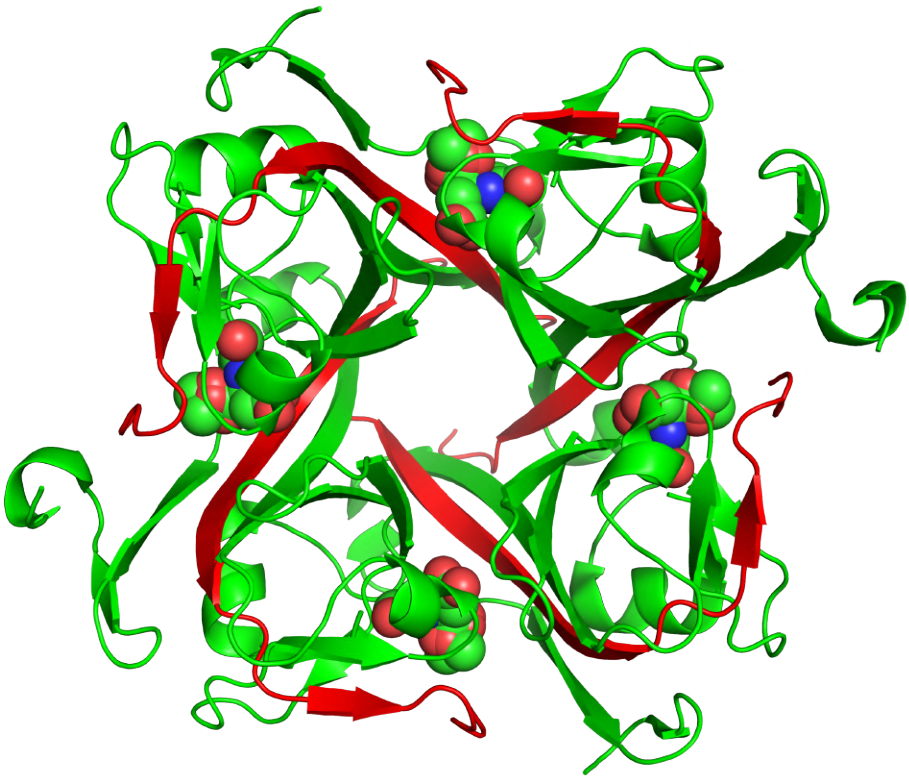

Figure 1.3: Aspartate α-decarboxylase tetramer (PDBID 4AON). β chain is shown in red and α chain is shown in green. Pyruvoyl co-factor is shown as spheres.
Image was rendered using PyMol [44].

yields an 11 kDa α-subunit (PYR-25 to ALA-126) and a 3 kDa β-subunit (MET-1 to GLY-24) which also includes the formation of pyrovoyl moiety at N-terminus of the α-subunit [47, 48]. The covalently linked, protein derived pyruvoyl cofactor is also essential for the catalysis [38] and plays an important role in initial binding of the ligand to the active site.

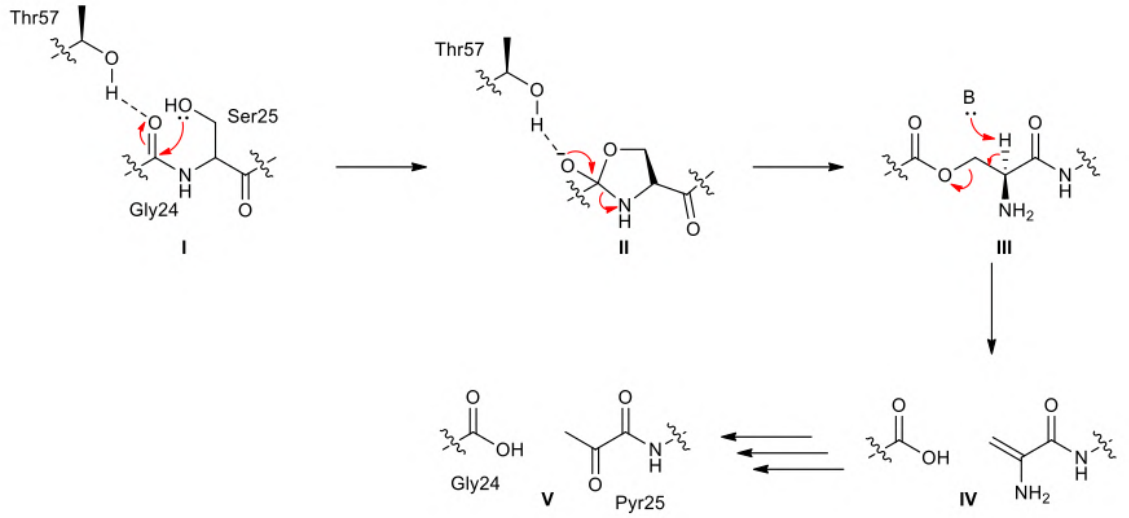

Figure 1.4: Mechanism of conversion of Serine 25 to pyruvoyl involving the hydroxyl residue of SER-25 and the backbone carbonyl of GLY-24 which is originally hydrogen bonded to THR-57 (PDBID 1PPY).
Image reprinted from [49]

The residues in PanD which are essential for its activation and subsequent conversion to ADC have been extensively studied [50, 51]. The mutation of Serine 25 to Alanine results in a mutant which cannot be activated to form ADC. However, other mutations along with the activation peptide significantly reduce the processing ability of ADC, most probably as an effect of the distortions in the loop. In the crystal structure of PanD, GLY-24 is seen to be hydrogen bonding to the side chain of the highly conserved THR-57 residue [50]. This interaction may polarise the carbonyl group, leading to nucleophilic attack by SER-25. Mutation of THR-57 to Valine leads to a PanD mutant which cannot be activated [51]. Apart from GLY-24 and SER-25, THR-57 is the only proximal residue which is essential for the activation of PanD. Other proximal residues like THR-58, HIS-11, ASN-72 have also been tested but none of these affected the activation of PanD to ADC [51]. Activation of PanD

in overexpressed *E.coli* cells occurs slowly at 37°C. Hence, the protein is required to be incubated overnight in order to get sufficient yield of the active protein [52]. It should be noted that the thermal activation of PanD is never 100 % which leads to the presence of both the desired pyruvoyl cofactor as well as the SER-25 byproduct from the hydrolysis of the ester intermediate (figure 1.4, II). The presence of this byproduct can be clearly seen in both the published structures (PDBID 1AW8, [43]) as well as the structure which was solved in the present work (PDBID 7A8Y) as a SER-25 conformer with partial occupancy along with the processed pyruvoyl moiety.

1.3 *E.coli* copper amine oxidase

E.coli copper amine oxidase (ECAO) is a periplasmic, soluble, homodimeric enzyme that catalyzes the oxidation of primary amines into aldehydes. This enzyme is important for slightly different functions in procaryotes, mammals and plants [53, 54]. Amine oxidases are becoming increasingly important drug targets because of reports which conclude that they mediate the interactions between lymphocytes and endothelial cells [55]. ECAO is encoded by the *tynA* gene [56]. The molecular weight of the functional ECAO homodimer is 161kDa. *E.coli* copper amine oxidase contains a cofactor along with a divalent copper ion in each of the two active sites [57]. This 2,4,5-trihydroxyphenylalanine quinone (TPQ) cofactor is generated after the modification of the endogenous TYR-466 residue by a series of oxidative reactions in presence of molecular oxygen [58].

1.3.1 Structure of *E.coli* copper amine oxidase

So far, the structures of copper amine oxidases from human, *Arthrobacter globiformis* (AGAO), *Hansenula polymorpha* (HPAO), *pea seedling* (PSAO), *Pichia pastoris* (PPAO) and *Escherichia coli* (ECAO) have been solved [59]. The tertiary structures of these enzymes are very similar, making their functional domains very similar as well [59]. However, the major difference between ECAO and other CuAOs is that ECAO contains an additional amino-terminal domain which is not present in HPAO, AGAO or PSAO. This enzyme is predominantly a β structure with a large C-terminal catalytic domain and two or three peripheral domains (figure 1.5). This C-terminal domain forms a large part of the intimate dimer interface and contains two β hairpins which reach across from one monomer to the other and also contains the active site. The residues which interact with the copper ion in the active site, as well as the conserved endogenous Tyrosine 466 residue are parts of the adjacent β sheets. The active site consists of a Cu(II) ion which coordinates the Nitrogen atoms from the three Histidines (524, 526, 689) at a distance of $\approx$ 2.0 Å . The coordination occurs in a distorted square pyramidal arrangement wherein the Nitrogens from Histidines occupy the equatorial positions [58, 59]. Along with the Nitrogen atoms, catalytic water molecules also coordinate with the copper ion (figure 1.5). The axial water which interacts with TPQ is at a distance of $\approx$ 2.4 Å while the labile equatorial water is at a distance of $\approx$ 2.0 Å if present. The axial water in figure 1.5 interacts with the dioxygen species during the turnover. It can also be seen in the figure that in case of active ECAO, the TPQ is in an off-copper conformation which means it is close to the copper ion but doesn't directly interact with it. The O2 position of

TPQ forms a hydrogen bond with the axial water while the O4 is hydrogen bonded to the hydroxyl group of the side chain of the conserved Tyrosine 369 residue. The active site also consists of ASP-383 which acts as a base during the turnover.

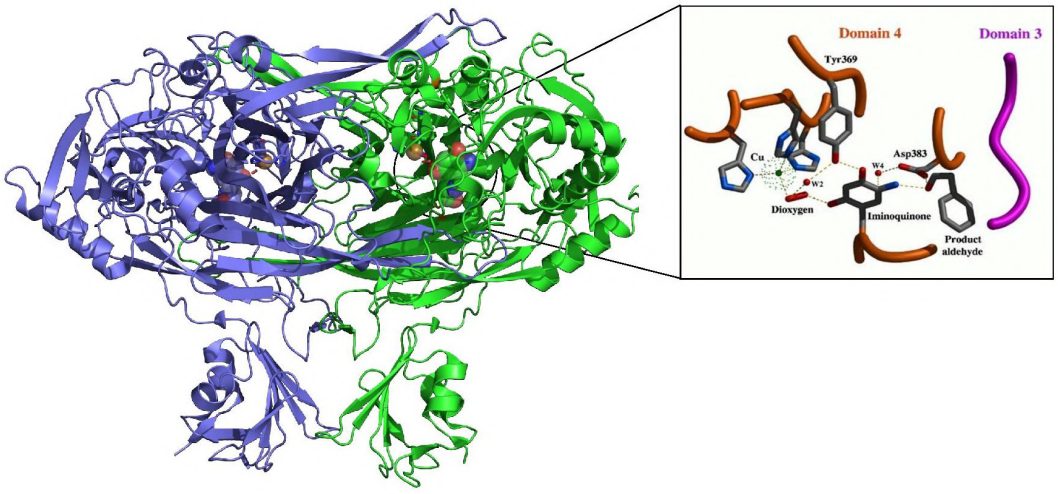

Figure 1.5: *E.coli* copper amine oxidase enzyme (PDBID 1D6Z). Inset: active site of ECAO with the catalytic waters, Cu(II), the released product aldehyde and the freeze trapped catalytic intermediate iminoquinone (figure 1.6)
Schematic illustration inset is adapted from [57]. Image was rendered using PyMol [44]

1.3.2 TPQ cofactor regeneration and catalysis

Copper amine oxidases catalyze the oxidative deamination of amines to corresponding aldehydes followed by the reduction of oxygen to hydrogen peroxide:

$$RCH_2NH_3 + O_2 + H_2O \longrightarrow RCHO + {NH_4}^+ + H_2O_2$$

The catalysis involves a ping-pong bi ter mechanism with reductive and oxidative half cycles [60, 61] (figure 1.6).

$$E_{ox} + {RNH_3}^+ \longrightarrow RCHO + E_{red}$$

$$E_{red} + O_2 \longrightarrow E_{ox} + H_2O_2 + {NH_4}^+$$

During reductive half reaction, amine is oxidized to aldehyde resulting in the reduced state of the enzyme through the formation of two separate Schiff base intermediates (figure 1.6, steps II and III). It has been shown that the Schiff base formation occurs at the O5 position of TPQ. During the oxidative half reaction, the enzyme reacts with oxygen resulting in the release of an ammonium ion (figure 1.6, steps IV to V). It has been proposed that oxidative half reaction occurs via a semiquinone of TPQ and reduced Cu^{+2} [61].

It has been an interesting problem to decipher the mechanism of *in situ* generation of TPQ in copper amine oxidase which is fully loaded (containing oxygen and copper). This process was initially thought to be autocatalytic, requiring the catalytic copper and the molecular oxygen but not necessarily the reducing equivalents [62, 63]. Two moles of O_2 are consumed for each mole of TPQ and H_2O_2 produced [64]. Kinetic and structural investigations on AGAO and HPAO have revealed the mechanism of generation of TPQ. The fact that these are the only CuAOs for which the precursors are metal free has been exploited to propose this mechanism. This process in HPAO is illustrated in figure 1.7. In case of metal free precursors, both mononuclear copper and the TPQ are absent. In this state, the conserved TYR side chain which is a part of the sequence THR-X-X-ASN-TYR-ASP/GLU [65] is unmodified (A in figure 1.7) [66, 67]. After this, copper binds to the three conserved Histidine residues in the active site (state B in figure 1.7). In kinetic studies on apoHPAO with and without copper, it was found that the rate of TPQ formation was unchanged pre-binding of Cu(II) which indicates that the binding of copper is a fast process relative to the overall rate of biogenesis [67]. After this step, the molecular oxygen binds to an off-copper hydrophobic binding pocket which induces a conformational change in the precursor Tyrosine. In this new conformer, the hydroxyl becomes oriented towards the copper (state B to C, figure 1.7). The structural work suggests that at this stage, the precursor Tyrosine is in its protonated form [58, 67]. On the other hand, the evidence for oxygen dependent formation of a tyrosine/copper complex comes from spectroscopic studies carried out on apoHPAO [47]. The conformational change of Tyrosine is followed by its deprotonation which results in a ligand-metal charge transfer (LMCT) species (C in figure 1.7) which absorbs at 350 nm.

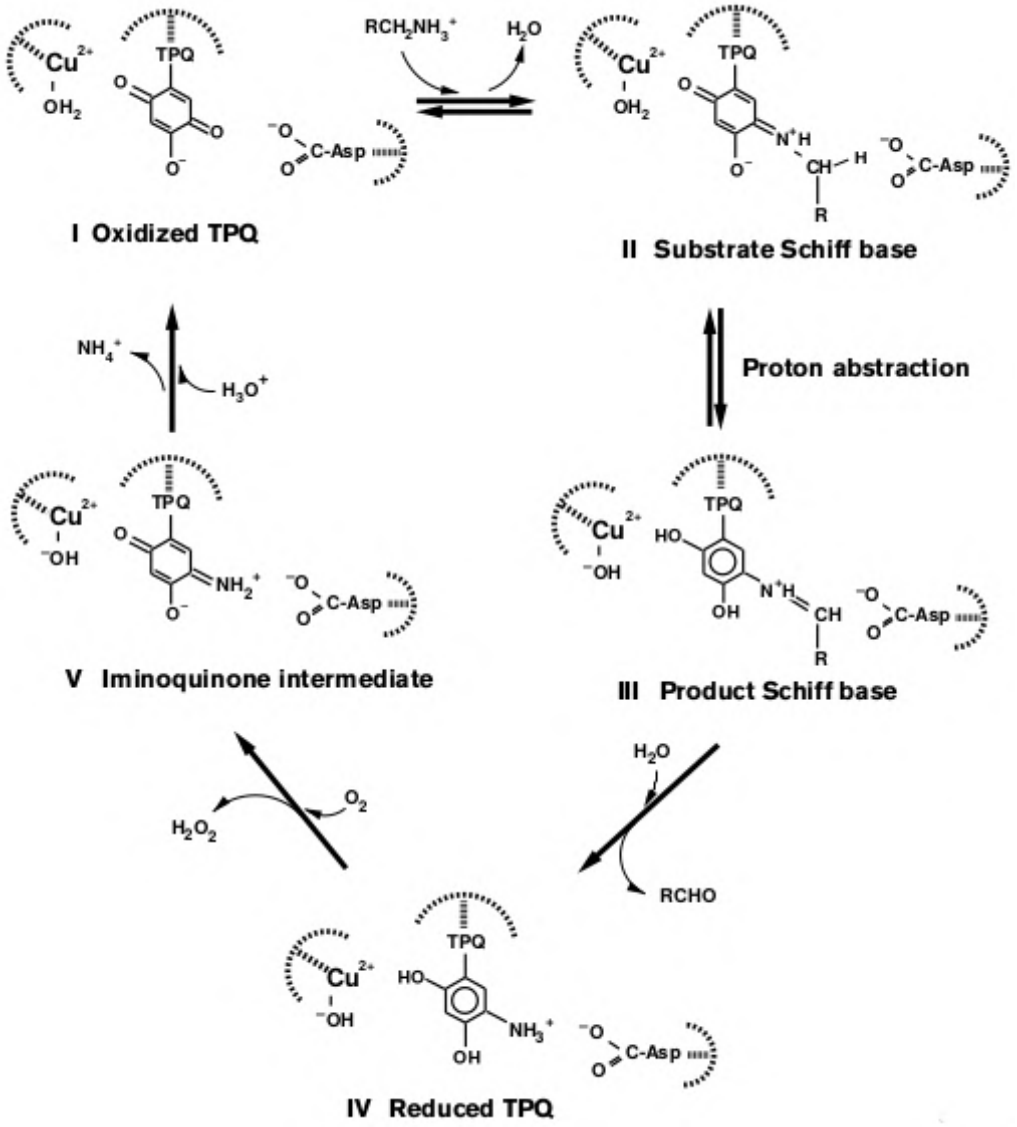

Figure 1.6: Oxidative and reductive half reactions involving the topaquinone (TPQ) cofactor.
Reprinted from [61]

The successive steps after this result in formation of dopaquinone (DPQ) after which the second oxygen atom is incorporated at position 6 of the ring (G to H, figure 1.7). This step leads to formation of reduced form of TPQ which is then oxidized to mautre TPQ which is followed by the reduction of molecular oxygen to hydrogen peroxide (H to I, figure 1.7). This mechanism for AGAO which is one of the most extentively studied CuAOs is strongly supported by crystallographic studies [67]. However, it should be noted that the axial and the equatorial water molecules which are observed in ECAO are absent in AGAO. Hence, these along with the additional domain in ECAO may also play a role in the catalysis. Another important aspect is that tyrosyl/copper complex is important in activating the precursor Tyrosine for the initial oxygenation reaction [68]. On the other hand, zinc is known to bind tightly at the CuAO active site and resists replacement by copper ions [69]. Binding of zinc renders the enzyme inert [70].

The region between the two subunits described in previous section is filled with solvent. This cavity is known as the 'inland lake'. All the CuAOs except HPAO are found to have two metal binding sites distinct from the copper binding site, one of which is solvent exposed and the other is buried. In case of ECAO, these are occupied by calcium ions. These calcium ions are not essential for the activity but their removal results in reduction of activity by $\approx$ 60-90 %. This removal can be a result of a mutation, E573Q in case of ECAO. This particular mutation results in a reduction in the activity by $\approx$ 90 %. However, a second mutation I342F results in partial rescue. In case of I342F mutation without the E573Q mutation, the activity is almost the same as wild-type ECAO. The partial rescue can also be a result of addition of another exogenous divalent cation in the E573Q mutant [71]. This loss of activity has been proposed to be due to a long range structural changes which alter the conformation of TPQ. It has also been a question as to which specific residues are involved in forming the pathway of molecular oxygen to the active site.

Johnson and others explored the oxygen entry pathways in HPAO using a combination of crystallographic, kinetic and computational approaches. They used Xenon as an oxygen probe as its size and hydrophobicity are analogous to those of oxygen [74]. Distinct Xenon binding sites were identified. Two of these which are $\approx$ 7 and 7.5 Å away from TPQ, were found to be consistent with the other reported copper amine oxidase structures in complex with Xenon [75]. *In silico* implicit ligand binding studies on HPAO, AGAO, PPLO and PSAO indicated two minimum energy pathways for entry of oxygen. One of these started from the amine entry channel close to the 5 position of TPQ and the other one from the hydrophobic β-sheet sandwich towards the active site [76].

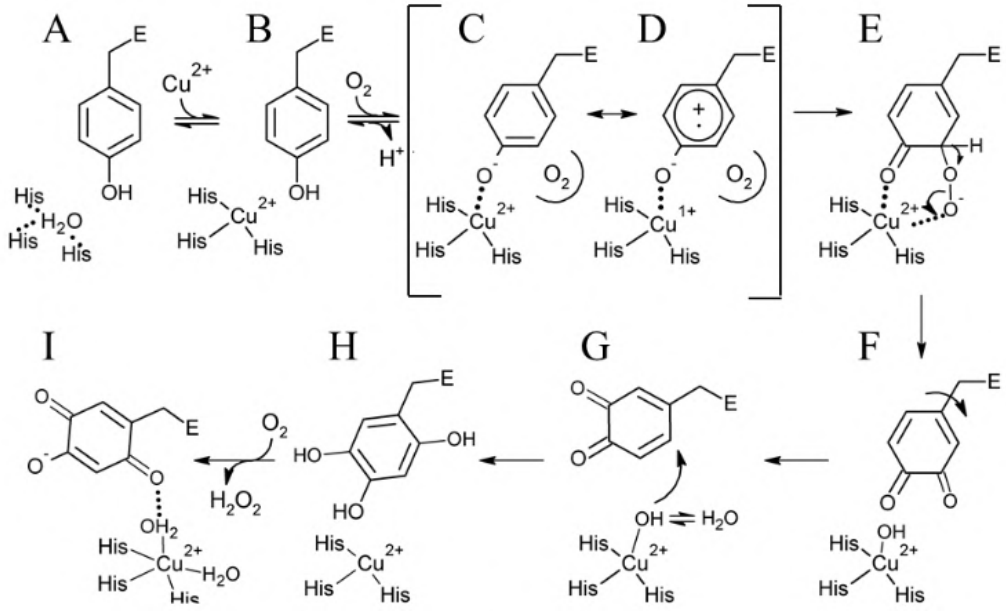

Figure 1.7: Proposed mechanism of regeneration of TPQ in HPAO.
Figure reprinted from [58]

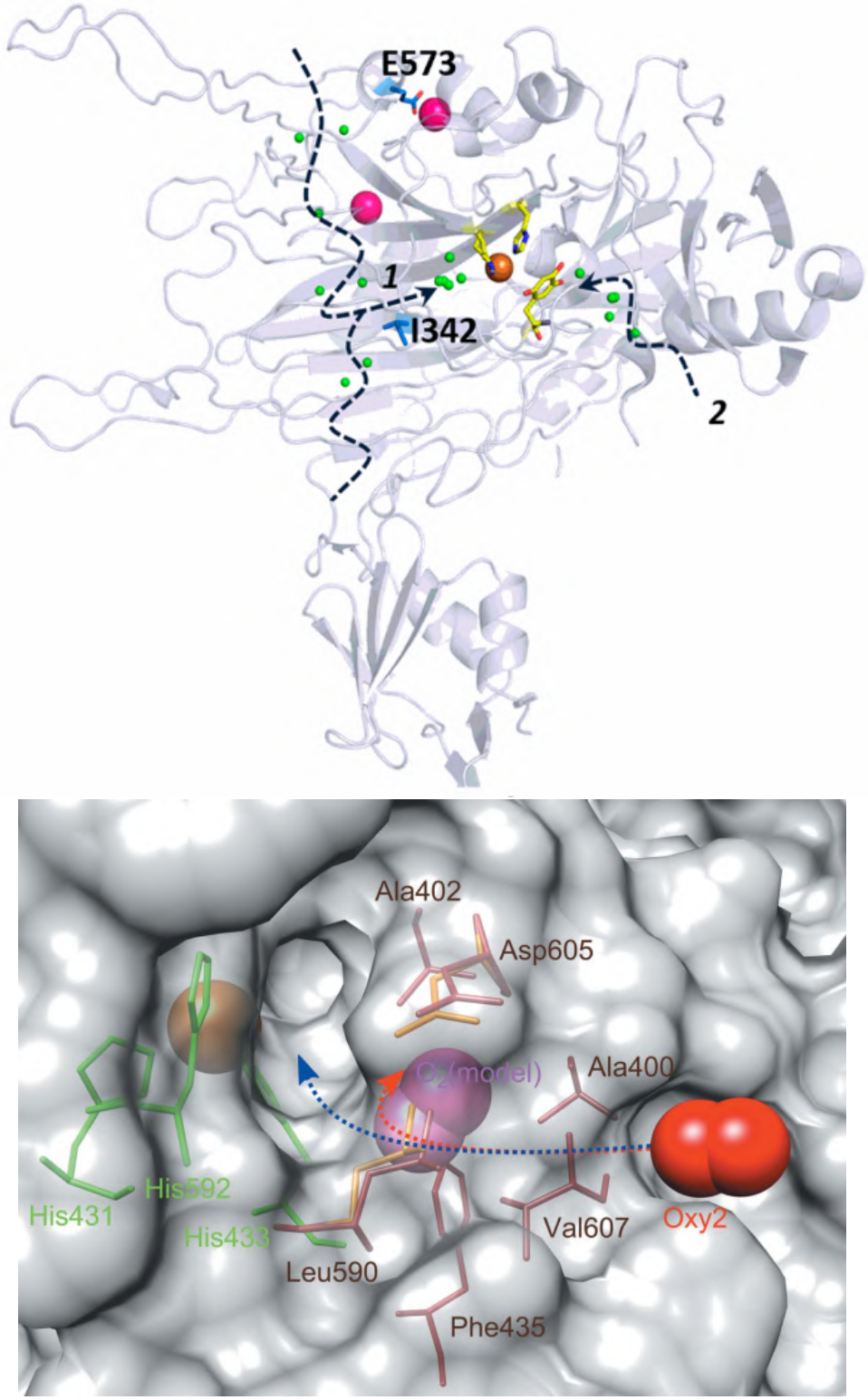

Figure 1.8: Proposed oxygen binding sites and pathways in ECAO (top) from the Xenon derivative structures and HPAO (bottom) from the high resolution O_2 bound structures. Figure adapted from [72] and [73]

In order to confirm these oxygen entry pathways, another study was performed by Pirrat and others which revealed very similar results [72]. High resolution structures of AGAO have been solved by growing the crystals in O_2 and N_2 atmospheres to compare the structures with and without molecular oxygen [73]. These structures revealed potential pathways of entry of oxygen. A total of three distinct oxygen entry pathways were identified as shown in the top panel of figure 1.8. Two of these merge close to the active site where an energetically favorable O_2 docking site was identified in the *in silico* studies [76]. As seen in figure 1.8, these oxygen entry pathways are close to the active site. Hence, apart from the long range structural changes, another possible reason for the effect of the two mutations (E573Q, I342F) on the activity of ECAO in spite of these being 20-30 Å from the active site can be that these mutations affect the dynamics of the enzyme. This change in dynamics may affect the trafficking of molecular oxygen in E573Q mutant in such a way that its entry channel is blocked leading to inactivation of ECAO. In order to understand the influence of the two mutations on the structure which can then be correlated with the function and hence the change in the activity of ECAO, the X-ray structures of single (E573Q) and double (E573Q/I342F) mutants of ECAO were solved [77]. Surprisingly, no major structural changes were observed in the two mutants. Some of the subtle changes include the change in the side chain conformation of the residue MET-699 which maybe related to hindrance of the pathway of molecular oxygen entry as MET-699 is in close proximity to one of the oxygen entry channels. It was also found that the water which replaces the calcium ion bound in wild-type enzyme is present only in subunit A and not in B. To summarize, the subtle structural changes in the two mutants were insufficient to explain the effect of the mutations on the activity of ECAO. This effect can be attributed either to the changes in long-range electrostatic interactions resulting in the alteration of the protonation state of TPQ or the associated residues, hence inaccessible to X-ray crystallography or to the change in structural dynamics of the enzyme resulting in altered conformational sampling either of the residues involved in O_2 or amine entry channel or those involved in the catalysis.

The former hypothesis can be addressed using neutron crystallography which can be used to visualize the positions of deuterium atoms in the crystal structure of a per-deuterated molecule. It has been previously used to address biologically important questions associated with the catalysis of enzymes. HIV-1 protease is one such system which is extensively studied but the questions related with its catalysis remain ambiguous. The two proton transfer from catalytic aspartates to the hydroxyl group of the substrate or the drug molecule is one such question that is of interest to chemists and biologists especially due to the two-fold symmetry which raises the question

as to whether the proton transfer occurs from aspartate in subunit-1 or 2 [78]. Adachi and others found out that ASP-25 is protonated and ASP-125 is deprotonated when HIV-1 protease is complexed with KNI-272 [79]. Another study by Fukuda and others on copper-containing nitrite reductase indicates that the catalytic copper is coordinated by a hydroxide ion and not by water, thereby giving new insights into the mechanism of nitrite reduction [80]. It is therefore clear that neutron diffraction can be used for probing even minute changes in protonation states of amino acid residues and hence in their key interactions with other residues.

On the other hand, the changes in protein dynamics within picosecond-nanosecond time scales in the latter hypothesis can be accessed using QENS. It has been used in the past to reveal key changes in dynamics of proteins as a function of multiple parameters e.g. concentration, pressure, temperature and the presence of ions [81, 82]. Grimaldo and others have explored crucial details regarding the changes in internal and global dynamics of proteins in crowded environment [83]. They could identify the relationship between the size of the macromolecules and the degree to which their diffusive dynamics are slowed down. Recent studies also reveal the effect of aggregation on the internal and global dynamics of protein. They conclude that the internal dynamics of Lysozyme remain unchanged whereas the self-diffusion is significantly affected in the aggregated state [84]. In another study of the photosynthetic reaction center of purple bacterium, *Rhodobacter sphaeroides*, site-specific mutations were found to significantly affect the overall dynamics of the protein [85]. These results are a major step towards understanding the relationship between fast dynamics and activity of the protein [85]. Hence, it is clear that QENS can be used to probe changes in the diffusive dynamics of proteins.

1.4 Objectives

1.4.1 ADC

We will study the effect of binding of D-Serine ligand on structure and dynamics of ADC. We will use a range of complementary techniques to address this objective:

1. The structural change in ADC associated with binding of the ligand will be elucidated using X-ray crystallography
2. QENS will be used to study the changes in global and internal dynamics of ADC upon binding of D-Serine
3. we will also use elastic fixed window scans (EFWS) and neutron vibrational spectroscopy to understand the changes in dynamics upon binding of the ligand at atomic level.
4. MD simulations will be used to corroborate the observations from X-ray crystallography and QENS.

1.4.2 ECAO

We will focus on elucidation of the mechanism of inactivation and partial reactivation of ECAO. For this purpose, we will explore both of the aforementioned hypotheses:

1. Neutron diffraction will reveal the possible long range structural effects either on the protonation state of TPQ or of the other active site residues. In case of E573Q mutant of ECAO, neutron diffraction will be used to probe any possible changes in the protonation states of TPQ or of the residues which are essential for catalysis in ECAO. On the other hand, QENS will be used to assess the changes in protein dynamics.
2. QENS will be used to assess whether the internal or global dynamics or both are affected as a result of the two mutations (E573Q, I342F) in case of ECAO.
3. MD simulations will be used to corroborate and correlate the observations from QENS with the residue level changes in dynamics that can be used to explain the possible reason of inactivation and partial reactivation of ECAO.

Chapter 2

Theoretical framework

This chapter will be used to explain the key concepts of diffusion and the theory behind some of the major techniques which were used to address the objectives of this thesis. Concepts of diffusion are explained, followed by fundamental theory of neutron spectroscopy, neutron diffraction and molecular dynamics simulations. This section follows the book chapter [86] to explain the concepts of diffusion, textbooks [87] and [88] to explain the theory of neutron spectroscopy and [89] to explain the theory of molecular dynamics simulations.

2.1 Fundamentals and concepts of diffusion

Diffusion plays a very important role in life and is of fundamental importance to different cellular processes [90–92]. This includes the transport of molecules within the cell, nutrient transport across the membrane and also influences the signal transduction mechanism at the molecular level [93]. There are two fundamental types of diffusion: self-diffusion and collective diffusion. The former refers to diffusion of a tagged colloidal particle or 'a tracer particle' in suspension with other particles [86]. The latter refers to the isothermal relaxation of inhomegeneities through correlated and collective motions of many particles. At the infinite dilution limit, both self and collective diffusion converge to the same value.

2.1.1 Self and Brownian diffusion

In an infinitely diluted solution, diffusion can be simply described as the random motion of a particle, as the solvent molecules collide with the diffusing particle [86].

The force exerted by the solvent molecules on the diffusing particle acts as a Gaussian distributed fluctuating force $\xi(t)$ such that,

$$\langle \xi(t) \rangle = 0 \qquad \langle \xi(t), \xi(t') \rangle = 2d\beta\delta(t - t') \tag{2.1}$$

wherein, d is the dimension of the system, $\beta = k_B T\gamma$ is the strength of the fluctuating force with γ being the frictional coefficient, brackets represent the thermal average over the fast solvent collisions. The translational displacement of a tagged colloidal particle can be considered as the diffusive mean-squared displacement (W) which signifies the average distance over which the particle has displaced itself from the initial position over time:

$$W(t) = \frac{1}{2d}\langle |r(t) - r(0)|^2 \rangle \tag{2.2}$$

wherein $r(t)$ is the position vector of the center-of-mass of the diffusing particle at time t, d is the dimension of the system and $\langle ... \rangle$ denotes the ensemble average.

When the evolution of $W(t)$ is studied with time, different time scales for the diffusing colloidal particle can be identified (figure 2.1).

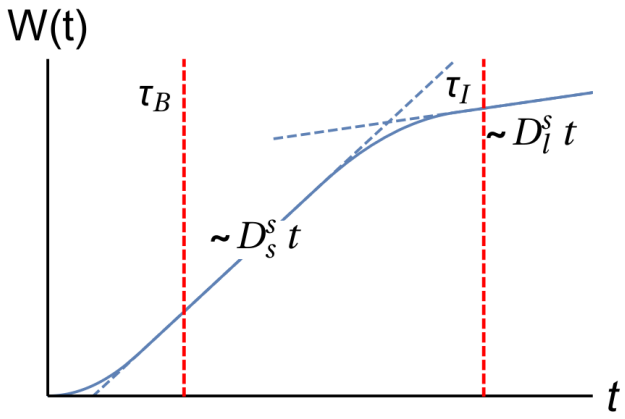

Figure 2.1: Schematic of the diffusive mean squared displacement $W(t)$ as a function of time. For very short times, $W(t) \propto t^2$. For $\tau_B \ll t \ll \tau_I$, $W(t) \approx D_s^s t$ and for $t \gg \tau_I, W(t) \approx D_l^s t$
Figure taken from [94]

If a particle is diffusing at time $t = 0$ with a velocity $\mathbf{v}_0$, the velocity doesn not change significantly due to the collision of the solvent molecules for very short time

scales i.e. $\mathbf{r}(t) - \mathbf{r}(0) \approx \mathbf{v}_0 t$ In this case, the diffusive mean-squared displacement is proportional to t^2,

$$W(t) \propto t^2 \qquad\qquad t \ll \tau_B \tag{2.3}$$

Here, τ_B signifies the interaction time scale over which the velocity of the diffusing particle remains constant.

For times much larger than τ_B but smaller than the time required for the diffusing solute molecule to collide, the particle experiences several collisions from the solvent molecules [86]. The trajectory of such a particle can be defined as Brownian random motion. Hence,

$$W(t) \approx D_{\mathrm{s}}^{\mathrm{s}} t \qquad\qquad \tau_B \ll t \ll \tau_I \tag{2.4}$$

with the short-time self-diffusion coefficient $D_{\mathrm{s}}^{\mathrm{s}}$ and the interaction time scale τ_I after which a particle experiences a substantial change of the potential interactions through a perceptible change in its next neighbor particle configuration [95]. As opposed to τ_B, τ_I signifies the interaction time scale over which the diffusing particle experiences several collisions from the solvent molecules which results in a change in its velocity. In this context, Brownian diffusion can be described as an overdamped motion driven by thermal fluctuations in a solvent medium [86]. In modern physics, the concept of Brownian diffusion has also been extended to anomalous diffusion processes such as sub and superdiffusion [92]. There has also been investigation into the so called 'hydrodynamic memory' arising from the diffusing particle acting back on its solvent [96]. According to Fick's second law [97], diffusion can be written analogous to the heat equation as a concentration field which changes with time as,

$$\frac{\partial \rho}{\partial t} = D \nabla^2 \rho \tag{2.5}$$

wherein, ρ is the particle concentration and D is the diffusion coefficient. While equation 2.5 refers to the theoretical definition, a scattering experiment gives access directly to the center-of-mass diffusion and follows a Lorentzian distribution where, the linewidth γ of the Lorentzian $\mathcal{L}(\gamma, \omega)$ reads as,

$$\gamma(q) = D q^2 \tag{2.6}$$

It should be noted here that the diffusion coefficient D of a colloidal particle in an ***infinitely diluted*** solution is described by the Stokes-Einstein relation [91]:

$$D_0 = \frac{\xi}{2d\alpha^2} = \frac{k_B T}{\kappa} \tag{2.7}$$

wherein, κ is the hydrodynamic friction coefficient [98] and ξ is the strength of the fluctuating force, d represents the dimensions of the system and α is a phenomenological parameter which typically lies between 0 and 1 [86]. However, in case of ***dense colloidal suspension***, equation 2.7 is generalized by a scalar function $f(\varphi, t)$ [81, 173, 185] depending on the volume fraction φ occupied by the colloidal particles in the suspension and on the observation time t:

$$D = \frac{k_B T}{6\pi\eta R} f(\varphi, t) \tag{2.8}$$

where $\kappa = 6\pi\eta R$ is the hydrodynamic friction coefficient with the solvent viscosity η and the colloidal particle hydrodynamic radius R.

2.1.2 Jump diffusion

When a large tracer molecule follows a simple diffusion and the solvent is considered as a continuous implicit medium, the diffusing particle follows Brownian diffusion (equation 2.6). However, when we take into account the fact that the solvent itself constitutes explicit tracer particles e.g. water molecules, this usually follows a more complicated diffusion [86]. One such frequently used model is 'jump diffusion' which constitutes a generalization of the simple Brownian diffusion by introducing the residence time τ between the molecular diffusive jumps. In this case, the Lorentzian linewidth associated with the solvent diffusive motion is,

$$\gamma(q) = \frac{Dq^2}{1 + Dq^2\tau} \tag{2.9}$$

As soon as more than one large particle of the solute (e.g. protein) suspended in a solution containing small solvent molecules (e.g. water) is being observed, the terms of 'self' and 'collective' diffusion need to be distinguished [86]. It should also be noted that the volume fraction φ occupied by the particles in the solution crucially influences the diffusion of the individual particles. Concepts more relevant for the biological molecules will be discussed in subsection 2.2.3.

2.2 Quasi-elastic neutron scattering (QENS)

Neutrons are fermions that can be produced using either nuclear fission or spallation. The former makes use of fuel elements consisting of heavy fissile atoms. There are quite a few sources operating on this principle. Institut Laue-Langevin in France, the

Moderator	Energy [meV]	Temperature [K]	Wavelength λ [10^{-10}m]
cold	0.1-10	1-120	30-3
thermal	5-100	60-1000	4-1
hot	100-500	1000-6000	1-0.4

Table 2.1: Energy, temperature and wavelength ranges from three types of sources. Adapted from [87]

.

FRM-II in Germany are two of the major ones. Spallation consists in hitting a target with accelerated protons. The impact allows disintegration of the nucleus, which produces several particles, including neutrons. Sources operating on this principle are the Spallation neutron source (SNS) in the USA, ISIS in the UK, and in the near future the European Spallation Source in Sweden.
Neutrons possess a magnetic moment and are electrically neutral, which allows them to penetrate deeply into materials, and probe various properties such as dynamics and structures of atoms or spins. Moreover, neutrons can be produced at various energies, depending on the moderator used (table 2.1). The use of cold neutrons ($\simeq 2$ meV energy) allows us to match the energies encountered in proteins, providing access to both spatial and temporal information. Cold neutrons also have the advantage of preserving chemical bonds due to their low energy.

2.2.1 Scattering cross-section

It can be seen in the book by Squires [87] or that by Bée [88] that a beam of neutrons passing through a sample, which is a collection of atoms, from which the neutron wave can be scattered in a given direction with a given energy exchange between the neutron and the atom. The general principle of neutron scattering is to count the number of neutrons being scattered into a solid angle Ω with energy E'. This gives the partial differential cross-section $\frac{d^2\sigma}{d\Omega dE'}$ and the total number of neutrons is given by,

$$\frac{d^2\sigma}{d\Omega dE'}\ \Phi\ d\Omega dE' = \begin{array}{l}\text{number of neutrons scattered per second} \\ \text{into solid angle d}\Omega,\ \text{and with} \\ \text{energy in the interval } [E',\ E'+dE'].\end{array} \tag{2.10}$$

where Φ is the flux of incident neutrons, as the number passing through a flat area, perpendicular to the beam propagation vector.

We can then obtain the total scattering cross-section by integrating over energy, which is the total number of neutrons scattered per second, divided by incident flux. Assuming isotropic scattering for angle ϕ, this quantity is given by

$$\sigma_{\text{tot}}\ \Phi = \int_{\text{all space}} d\Omega \int_0^\infty\ dE'\ \frac{d^2\sigma}{d\Omega dE'} = \int_0^\pi d\theta\ 2\pi\sin(\theta)\ \frac{d\sigma}{d\Omega} \tag{2.11}$$

All atoms in the sample will contribute to this quantity. However, each of the atoms or a group of atoms will scatter the neutrons with different efficiencies i.e. for a given nucleus, if we consider an incoming wave propagating along z given by

$$\Psi_{\text{in}}\ =\ e^{ikz} \tag{2.12}$$

then, the scattered wave will be spherical with a scaling factor b such that

$$\Psi_{\text{out}}\ =\ -\frac{b}{r}e^{ikr} \tag{2.13}$$

The factor b is known as the scattering length and is determined experimentally. It varies erratically from one nucleus to another. If we consider elastic scattering, the number of neutrons passing through unit area per unit time is

$$v\ dS|\Psi_{\text{out}}|^2 = v\ dS\ \frac{b^2}{r^2} = v\ b^2\ d\Omega \tag{2.14}$$

Using (2.10), with $\Phi\ =\ v|\Psi_{\text{in}}|^2$, we have

$$\frac{d\sigma}{d\Omega}\ =\ \frac{vb^2d\Omega}{vd\Omega} = b^2 \tag{2.15}$$

and the total cross-section $\sigma_{\text{tot}}\ =\ 4\pi b^2$. Thus, the scattering length is directly related to elastic scattering probability for a given nucleus.

2.2.2 Coherent and incoherent scattering function

The system under study can exchange kinetic energy with the incoming neutron. This can be expressed by the transition rate for a given state $|k,\lambda\rangle$ to turn into the final state $|k',\lambda'\rangle$, with k being the neutron state, λ the nucleus state and $\rho_{\mathbf{k}'}$ is the density of the final states. We use the so-called *Fermi's golden rule*, which states that

$$\sum_{\mathbf{k}'\ \text{in}\ \Omega} W_{\mathbf{k},\lambda\to\mathbf{k}',\lambda'}\ =\ \frac{2\pi}{\hbar}\rho_{\mathbf{k}'}|\left\langle\mathbf{k}'\lambda'\right|V\left|\mathbf{k}\lambda\right\rangle|^2 \tag{2.16}$$

It can be shown that this expression yields the following for the differential cross-section [87],

$$\left(\frac{d^2\sigma}{d\Omega dE'}\right)_{\lambda\to\lambda'} = \frac{k'}{k}\left(\frac{m}{2\pi\hbar^2}\right)^2 |\left\langle \mathbf{k}'\lambda'\right| V \left|\mathbf{k}\lambda\right\rangle|^2 \delta(E_\lambda - E_{\lambda'} + E - E') \tag{2.17}$$

Here, the nucleus is considered as a point in space, therefore, the potential V takes the the form of the Fermi pseudopotential

$$V(\mathbf{r}) = \frac{2\pi\hbar^2}{m} b\delta(\mathbf{r}) \tag{2.18}$$

with r being the distance between the neutron and the nucleus.
Using this potential, representing the delta on energy by an integral on a complex exponential, and taking the averaged scattering length for the system, it can be shown that the double differential cross-section for coherent scattering is given by [87]

$$\left(\frac{d^2\sigma}{d\Omega dE'}\right)_{\text{coh}} = \frac{\sigma_{\text{coh}}}{4\pi}\frac{k'}{k}\frac{1}{2\pi\hbar}\sum_{\text{j}\neq\text{j}'}\int_{-\infty}^{\infty}\langle e^{i\boldsymbol{q}(r_{\text{j}}(t)-r_{\text{j}'}(0))}\rangle e^{i\omega t}dt \tag{2.19}$$

and is related to collective dynamics and mostly used with neutron crystallography for example. While incoherent scattering can be described by

$$\left(\frac{d^2\sigma}{d\Omega dE'}\right)_{\text{inc}} = S(\boldsymbol{q},\ \omega) = \frac{\sigma_{\text{inc}}}{4\pi}\frac{k'}{k}\frac{1}{2\pi\hbar}\sum_{\text{i}}\int_{-\infty}^{\infty}\langle e^{i\boldsymbol{q}(r_{\text{i}}(t)-r_{\text{i}}(0))}\rangle e^{i\omega t}dt \tag{2.20}$$

and is related to self dynamics, with $\sigma_{\text{coh}} = 4\pi\bar{b}^2$, $\sigma_{\text{inc}} = 4\pi(\overline{b^2} - \bar{b}^2)$, **q** = **k'** - **k**, and r_{i} denotes the position of nucleus i. It should be noted here that equation 2.19 is valid for a pair of particles whereas equation 2.20 is for the same atom.
Therefore, incoherent scattering will be the interesting quantity for us since self dynamics is probed using quasi-elastic neutron scattering.
Some values for scattering cross-sections are given in table 2.2. We can readily see that hydrogen has a significantly higher cross-section than other nuclei. This feature can be used to probe specifically some parts of the system using site-specific deuteration.

Nucleus	Z	σ_{coh} $[10^{-28} m\check{s}]$	σ_{inc} $[10^{-28} m\check{s}]$
^{1}H	1	1.8	80.2
^{2}H	2	5.6	2.0
C	6	5.6	0.0
N	7	11.01	0.5
O	8	4.2	0.0
Al	13	1.5	0.0

Table 2.2: Scattering cross-section for experimentally relevant nuclei Adapted from [87].

.

2.2.3 QENS and its significance in the study of biological systems

Living cells are generally considered a soft matter. However, there is usually an absence of long range order in biological molecules. The degree of ordering may depend on the specific role that the molecule plays and the type of molecule under consideration [86]. For example, in case of enzymes which catalyze a particular chemical reaction, the substrate binds at the active site which results in a conformational change or reordering [99–101]. Hence, in such cases, the underlying molecular motion which is associated with a particular stimulus plays an important role in the biological function of a molecule. In the aforementioned example, this stimulus is binding of the substrate at the active site. Hence, the experimental methods used should be sensitive to the length and the time scales corresponding to these molecular rearrangements. Neutron spectroscopy methods which involve scattering of cold and thermal neutrons are suitable to detect these changes and in addition are also non-invasive to the samples [86]. Since most of the structural and in some cases even the dynamical investigations into biochemical processes involves the use of coherent scattering, incoherence has a somewhat negative connotation. In case of QENS, the incoherence purely refers to the spatial incoherence which arises due to the nuclear spin statistics [87]. It should be noted at this point that although the neutron spectroscopy experiments are spatially incoherent, there indeed exists a strong coherence in time i.e. an energy resolution in order to access the time-correlations in the sample. As seen from table 2.2, the incoherent cross section for hydrogen is very high compared to the other atoms which implies that these methods are also highly sensitive to the hydrogen atoms. Since proteins have a large number of hydrogen atoms ($\approx 50\%$ of the total number of atoms), they can provide a very strong scattering signal. Hence, there is no need for labels in case of neutron spectroscopy in contrast to fluorescence spectroscopy, as hydrogens act as natural 'labels' and is also a non-destructive technique unlike X-ray scattering or IR spectroscopy techniques which are often associated with radiation damage in the sample [102–104]. This enables us to follow the changes in the dynamics of proteins and other biomolecules.

The classical time-dependent van Hove correlation function is obtained by applying the inverse Fourier transform twice to the incoherent scattering function $S(\boldsymbol{q},\ \omega)$ from ω to t and from q to r (equation 2.24). The same follows for the coherent scattering function as well (equation 2.23). In case of IN16B spectrometer, the accessible momentum transfer is between 0.2 and 2.0 Å^{-1}. The corresponding accessible length scale can be calculated using the following relation:

$$l = \frac{2\pi}{q} \tag{2.21}$$

where, q is the momentum transfer and l is the accessible length scale. Hence, the accessible length scale in case of IN16B corresponding to a q-range of 0.2 - 2.0 Å $^{-1}$ is 3.0 - 30 Å .

Similarly, the accessible time scale can be calculated as follows:

$$T = \frac{2\pi}{\omega} \tag{2.22}$$

This indicates that the slowest time scale which is accessible to IN16B is ≈ 4 ns corresponding to an instrumental resolution of ≈ 0.75 μeV. The time-dependent van Hove correlation functions are of fundamental importance to access biomolecular dynamics using neutrons [105]:

$$G(\mathbf{r}, t) = \frac{1}{N} \sum_{i,j=1}^{N} \langle (\delta(\mathbf{r} - \mathbf{R}_{\mathrm{j}}(t) + \mathbf{R}_{\mathrm{i}}(0)) \rangle \tag{2.23}$$

$$G_{\mathrm{s}}(\mathbf{r}, t) = \frac{1}{N} \sum_{j=1}^{N} \langle (\delta(\mathbf{r} - \mathbf{R}_{\mathrm{j}}(t) + \mathbf{R}_{\mathrm{j}}(0)) \rangle \tag{2.24}$$

where, t denotes time, $\mathbf{r}$ is the instantaneous position of the neutron and $\mathbf{R}_{i,j}$ are the time dependent positions of N nuclei which constitute the sample. The correlation function $G(\mathbf{r}, t)$ in equation 2.23 contains the information about the distinct pairs of nuclei $\mathbf{R}_{i,j}$ and hence includes both temporal and spatial correlations. The self-correlation function $G_s(\mathbf{r}, t)$ in equation 2.24 includes the time-correlation of nuclei with identical index j only. Hence, it constitutes the ensemble-averaged single-particle self-correlation function [86].

The two different correlation functions represented by equations 2.23 and 2.24 are schematically presented in figure 2.2 for the case of randomly moving particles which are being observed at two arbitrary points in time. Equation 2.24, is important in case of the ligand induced loop motion in aspartate α-decarboxylase since it enables us to unambiguously access the internal and global self-diffusion. There is no other method than neutrons that can access self-diffusion on a similar short time and length scale on biological samples. However, one of the drawbacks is that the sample concentrations and volumes required are higher (50-300 mg/mL, >1 mL). This is because the neutron beam size is at least a few centimeters and shows an angular divergence of at least a few degrees unlike X-rays which are highly parallel. Another advantage in case of QENS is that the energy resolution is several orders of magnitude

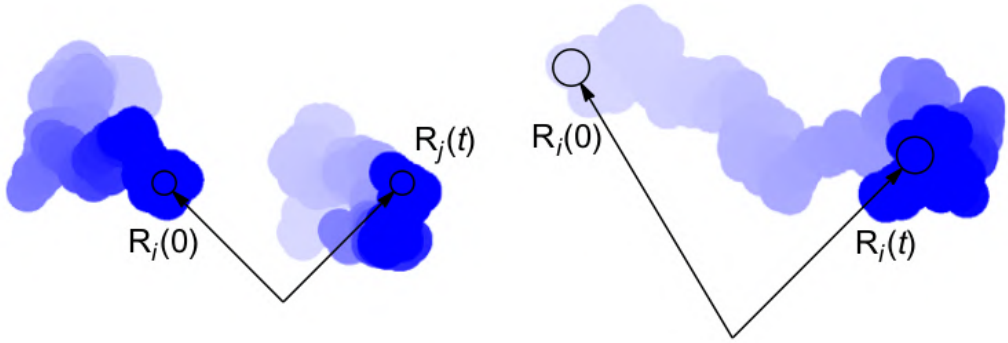

Figure 2.2: Schematic illustration for the interpretation of the van Hove correlation functions for coherent scattering (left, equation 2.23) and incoherent scattering (right, equation 2.24), respectively. Here, coherent scattering is typically employed to probe the ensemble-average over pairs of particles with coordinates $\mathbf{R}_{i,j}(t)$, marked in the schematic by the time-dependent coordinates $\mathbf{R}_{i,j}(t)$ and shadows. This indicates the random paths followed by a pair of particles that are within the observation time scale. Incoherent scattering on the other hand is used to probe the ensemble-average over single particles with a tag. This is indicated in the schematic by the path followed by a single particle and being observed at the time points 0 and t. Reprinted from [86]

higher than that offered by X-rays. It is now possible to reach an energy resolution of up to at least 1 μeV at 1 Å $^{-1}$ momentum transfer, whereas in case of X-rays it is impossible to reach close to 1 μeV and difficult to get to an order of magtinude of meV. In case of an incoherent neutron spectroscopy experiment in which the vibrational and the diffusive motions are uncoupled, we can write the scattering function as follows:

$$S(q,\omega) = exp\left(\frac{-1}{d}q^2\langle u^2_{\text{vib}}\rangle\right) S_{\text{diff}}(q,\omega) \tag{2.25}$$

In equation 2.25, the term $\langle u^2\rangle$ is the ensemble-averaged mean-squared displacement of the atomic vibrations and the term $e^{\left(\frac{-1}{d}q^2\langle u^2\rangle\right)}$ is called the vibrational Debye-Waller factor of the system with d dimensions (typically, $d = 3$). $S_{\text{diff}}(q,\omega)$ contains all the contributions from the diffusive and the relaxational molecular motions which are characteristic in soft matter in general and in biomacromolecules in particular. $S_{\text{diff}}(q,\omega)$ is a symmetric function which is centered at $\omega = 0$ at least at physiolo-

gical temperatures. It is called the 'Quasi elastic scattering function' because it is established experimentally in terms of the symmetric line broadening relative to the experimental resolution function.

The Debye-Waller factor hence results in a q-dependent reduction in the elastic scattering signal. The scattering function only due to the diffusive motions can be written as [106]:

$$S_{\text{diff}}(q,\omega) = \mathcal{L}(\gamma(q),\omega) \otimes \{A(q)\delta(\omega) + [1 - A(q)]\mathcal{L}_\alpha(\Gamma(q),\omega)\} \tag{2.26}$$

wherein, $\mathcal{L}(\gamma(q),\omega)$ is the first Lorentzian function with the linewitdh $\gamma(q)$ which accounts for the global translational and rotational diffusive motions of a macromolecule in solution. The first Lorentzian is convoluted ($\otimes$) with the second term containing the second Lorentzian $\mathcal{L}_\alpha(\Gamma(q),\omega)$ which describes the ensemble-averaged internal diffusive molecular motions of the macromolecule. Here, $A(q)$ denotes the internal molecular motion within a confined geometry and is known as 'Elastic Incoherent Structure Factor (EISF)'. It is explained further in subsubsection 4.1.4.2 of chapter 4.

2.2.4 Modeling the elastic scattering signal

If we consider only the elastic scattering signal, useful information can already be extracted. Using equation 2.20, the correlation function in the integrand can be expanded, considering the actual measured signal by convolution with the resolution function $S_{\text{exp}}(q,\ \omega) = S(q,\ \omega) \otimes R(q,\ \omega)$, to obtain [107]

$$\begin{aligned} S(q,\ \omega)_{\text{exp}} = \hat{R}(\omega) & \\ & - \frac{q^2}{2\sqrt{2\pi}} \int dt e^{-i\omega t} \langle \Delta r^2 \rangle(t) R(t) \\ & + \frac{q^4}{24\sqrt{2\pi}} \int dt e^{-i\omega t} \langle \Delta r^4 \rangle(t) R(t) \\ & + \mathcal{O}(r^6) \end{aligned} \tag{2.27}$$

In line with the Debye-Waller factor $e^{-\langle u^2\rangle q^2/3}$, the mean-squared displacement can be obtained by

$$\begin{aligned}\langle u^2\rangle_\omega &= -\lim_{\mathrm{q}\to 0}\frac{3}{q^2}\log\left[S_{\exp}(q,\omega)/\hat{R}(\omega)\right] \\ &= \frac{3}{2\sqrt{2\pi}}\int dt e^{i\omega t}\langle\Delta r^2\rangle(t)R(t)/\hat{R}(\omega)\end{aligned} \tag{2.28}$$

This is the most general mean-squared displacement (calculated for ADC in section 5.2 of chapter 5). In the case where we are interested only in the elastic signal, $\omega = 0$, and the equation becomes

$$\langle u^2\rangle_0 = \frac{3}{2\sqrt{2\pi}}\frac{1}{\hat{R}(0)}\int dt\langle\Delta r^2\rangle(t)R(t) \tag{2.29}$$

We see that the mean squared displacement provides a model-free indicator of the type of dynamics that strongly depends on the time window given by the resolution function. This indicator can be used with both elastic and inelastic fixed-window scans.
Other methods were proposed [108] [109], the simplest of which makes use of a simple Gaussian, being valid only on a limited q-range. Using cumulant expansion on the auto-correlation function, it is possible to derive either a simple Gaussian expression for the Mean squared displacement, or a more complicated one that includes also higher terms in the expansion [108], thereby giving the following,

$$S(q,\ \omega=0) \simeq e^{\frac{1}{6}q^2\langle\Delta r^2\rangle}\left(1+\frac{q^4}{72}\sigma^2\right) \tag{2.30}$$

where σ^2 is the variance of $\langle\Delta r^2\rangle$ over all atoms. This model was successfully applied to α-lactalbumin to highlight the heterogeneity of motions in hydrated powder samples [110]. Also, a gamma distribution can be used to take into account the heterogeneity of motions in the sample [111] [112]. In practice, the data to be fitted can be noisy and exhibit few features. Therefore the fitting process can quickly become unstable when the model gets more complicated. The Gaussian model presents the advantage to give a simple, and quick way to compare dynamics of different samples, but it is limited to momentum transfer $\mathrm{q}^2\langle\mathrm{r}^2\rangle < 1$, where $\langle r^2\rangle$ is the mean squared displacement.

2.3 Single crystal X-ray diffraction

Single crystal X-ray diffraction is one of the oldest method of structure determination at atomic level which was pioneered by William Henry and his son William Lawrence Bragg in 1912. This was followed by some important discoveries by Dorothy Hodgkin between 1937 and 1945. She solved the structures of penicillin, vitamin B12 and cholesterol. This was followed by some of the pioneering work on unraveling the structural features of biomolecules. From 1950s, the structures of many biomolecules like DNA by Francis, Crick and Wilkins in 1953, Myoglobin by Max Perutz in 1957 and later on Lysozyme by Blake in 1965 were solved. X-ray crystallography has come a long way ever since and has reached a level at which we are able to calculate precise 3-dimensional coordinates of molecules as large as Ribosomes [113]. This has resulted into X-ray crystallography being a major technique which has contributed to 89% of the structures which are deposited in the Protein Data Bank (RCSB-PDB) [114].

In a typical single crystal X-ray crystallography experiment, a bright X-ray light from synchrotron or home source X-ray generator is diffracted by millions of identical biomolecules. In this case, these molecules are those which form a protein crystal which is better visualized as a crystalline grid [115] with a grid spacing of 'd' (figure 2.3). The outgoing X-ray waves which are in phase then interfere constructively which results in a sharp diffraction pattern. The resulting peaks are called 'Bragg peaks' or 'Bragg reflections' (figure 2.3). Bragg peaks lie in those directions in which the scattered waves interfere constructively.

Assuming that the path difference between the two waves is $2dsin\theta$, *Bragg's law* can be stated as:

$$2d \cdot sin\theta = n\lambda \tag{2.31}$$

where, λ is the wavelength of the incident X-ray beam, θ is the angle of scattering and 'd' is distance between the lattice planes. The crystallographic lattice planes are described by Miller notation *'hkl'*.

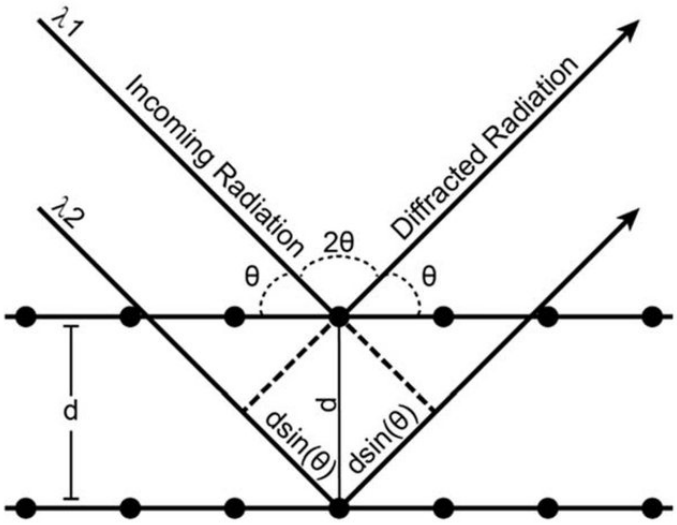

Figure 2.3: 2 dimensional representation of incident and diffracted waves with a path difference of 2d sin θ [116]

The lattice planes can be arranged in different ways and hence can have different geometries. These geometries are termed as *Bravais lattices*. These can be generated by discrete translation operations in 3-dimensional space. The information about structure of the molecule which forms the asymmetric unit of the crystal is contained by the intensities of the peaks. These intensities are related to a fundamental quantity called 'structure factors (F_{hkl})'. The structure factors are associated with the amplitudes of the Fourier transform of the electron density of the molecule. The structure factors are a convolution of the Fourier transform of the lattice and the Fourier transform of the electron density of the molecule: $\mathcal{F}_{lattice} \otimes \mathcal{F}_{molecule}$.

$$\begin{aligned} F_{hkl} &= \sum_{j=1}^{N} f_j e^{[-2\pi i \mathbf{k}.\mathbf{x}]} \\ &= \sum_{j=1}^{N} f_j e^{[-2\pi i (hx_j + ky_j + lz_j)]} \end{aligned} \tag{2.32}$$

In equation 2.32, the summation is over N atoms whereas $\mathbf{x}$ represents the real space coordinates for these N atoms. f_j is the *atomic form factor* of the j'th atom, f_j(k) $= \int \rho(\mathrm{r})\ e^{i\mathbf{k}.\mathbf{r}}\ \mathrm{d}^3\mathbf{r}$. Here, the wavevector k is expressed in the basis ($\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$) of the lattice as $\mathbf{k} = h\mathbf{a} + k\mathbf{b} + l\mathbf{c}$ and the Miller indices (hkl) define a point in the reciprocal space which corresponds to the crystallographic plane in the real space at which scattering/diffraction has occurred. The scattering factors are defined as the measure of the scattering amplitude of a wave by an atom and depend on the scattering power of an individual atom. The scattered intensity at reflection $\mathbf{k}(hkl)$ on the 2-dimensional detector depends on the simple mathematical correlation:

$$I(\mathbf{k}) \propto |F_{hkl}|^2 \tag{2.33}$$

where, F_{hkl} represents the crystallographic structure factors.

A major drawback in single crystal X-ray diffraction measurements is that the phases of the scattered waves are not measured during the experiment. Hence, the phases either have to be obtained experimentally or they have to be obtained from the previously obtained phases of either an identical or a similar molecule.

There are a number of methods which can be used for obtaining the phases. For example, the phases can be obtained from the crystal structure of a similar molecule which has already been solved. This method is called molecular replacement and is currently the most widely used method for obtaining the phases [117]. Positivity and atomicity of the electron density which leads to relationships between (normalized) structure factors which leads to a strong correlation between the phases of multiple

reflections which can be used to obtain the phases [118]. This is called the direct or the *ab initio* method. Apart from these methods, there are the experimental methods of calculating the phases which includes the use of anomalous X-ray scattering (MAD and SAD phasing) [119, 120]. The phased electron density map is then used to build the atomic model of the biomolecule under consideration. In the process of model building, the backbone atoms are fitted first since their scattering contribution is larger than the side chain atoms as they are relatively rigid in a crystal.

2.3.1 Crystallographic refinement

The process of statistical adjustments in the atomic coordinates so that the model parameters represent the experimental data as well as possible is known as 'crystallographic refinement'. This process includes further refinement of the model by including the side-chain orientations, occupancies, the B-factor refinement (which follows the Debye-Waller equation) and the ordered solvent. During refinement, restraints are applied to the movements of individual atoms based on the stereochemical configuration, other conformational parameters and B-factors. Overall parameters such as bulk solvent corrections, anisotropy corrections, overall B-factors and scale factors are also applied. All these parameters are refined against the experimental data to obtain the best agreement between the experimental or the observed and the calculated or the model structure factor amplitudes.

In this process, incorrect phases may dominate ultimately leading to a biased model. These incorrect phases may result if the molecule used for phasing is similar but not identical. To overcome this problem, $2F_{obs} - F_{calc}$ maps are created which reflect the improvement in the model (subsubsection 6.4.2). The more widely used *Sigma A* figure of merit weighted map ($2mF_o - DF_c$) is the maximum likelihood of the best possible map, where 'm' represents the maximum likelihood estimates for each model phase and 'D' represents the errors in the atomic coordinates of the current model. To know the ambiguities arising due to the possible difference between the experimental and the modeled structure factor amplitudes, difference maps are calculated. These maps are typically calculated by subtracting the calculated structure factor amplitudes ($|F_{calc}|$) from an integral multiple of the observed structure factor amplitudes ($|F_{obs}|$) in each Fourier term. This difference is then highlighted in the difference electron density maps ($F_{obs} - F_{calc}$). This results in either positive or negative peaks in the electron density map. A positive peak corresponds to regions of missing information in the model. On the other hand, a negative peak corresponds to a region where the atoms are falsely modeled.

Using the information obtained from the difference peaks, the atomic coordinates

in the model are iteratively changed with a final aim to converge the difference between ($|F_{calc}|$) and ($|F_{obs}|$). Two statistical parameters are used to assess the quality of the model after each iteration of refinement. These parameters are R_{free} and R_{work} which are calculated using the following equation:

$$R = \frac{\sum_{hkl} |F_{obs}(hkl) - F_{calc}(hkl)|}{\sum_{hkl} |F_{obs}(hkl)} \tag{2.34}$$

Hence, the R values are used to assess the overall agreement between experimental and the calculated structure factor amplitudes. As the agreement between the experimental and the calculated data improves, the model becomes more and more accurate and hence the R factor decreases. As seen from equation 2.34, the R factors can iteratively attain lower values if there are parameters which can be used to overfit the data. To overcome this problem, Brunger and colleagues introduced the R_{free} statistic in 1992 [121]. R_{free} is calculated in the same way as shown in equation 2.34 but is based on a small set of randomly chosen intensities. This small set is usually 5% of the total number of reflections and are not used during refinement. R_{free} is used to assess how accurately the model can predict the randomly chosen subset of the reflected intensities. If the changes implemented in the atomic coordinates are correct then the change in R_{free} would be in the same direction as that of R_{work} which uses the remaining 95% of the data. The absolute R values are strongly dependent on the overall resolution of the data. For example the R values for a data with the highest resolution of 1.9 Å should be ≈ 0.17. The acceptable difference between R_{free} and R_{work} is usually 5 % unless the overall resolution of the data is below 2.5 Å. For crystallographic data analysis, we have used XDS for data integration and scaling [122], PHASER of CCP4 suite for structure solution using molecular replacement [123], REFMAC5 for crystallographic structure refinement [124] and the computer graphics software COOT [125] for the interpretation of mF_o-DF_c and $2mF_o$-DF_c electron density maps.

2.4 Single crystal neutron diffraction

This method measures elastic, coherent scattering from the neutrons which get scattered by a single crystal. The fundamental principles and the interpretation of the data is the same as that for single crystal X-ray diffraction (section 2.3). However, the difference between X-ray and neutron diffraction is that the X-rays are scattered by the electrons whereas the neutrons are scattered by the nuclei. The nuclear radius is $\approx 10^4$ smaller than the standard neutron wavelengths which are

used in diffraction experiments. Therefore, the neutron scattering amplitude is a constant independent of the momentum transfer, $Q = 4\pi(sin\theta/\lambda)$. For X-rays on the other hand, the scattering amplitude, also called as 'the form factor', falls off with the increase in Q [126]. This is due to the destructive interference of the scattering caused by different parts of the electron density cloud.

As seen from table 2.10, each of the isotopes have different coherent scattering cross sections. It can be seen that the cross section for deuterium is very high among these, which makes it possible to resolve the positions of deuterium atoms with almost the same accuracy as that for the other heavy atoms (C, N, O). Hence, if hydrogen atoms in a molecule are replaced with deuterium atoms, a process called 'perdeuteration', it is possible to determine the positions of these deuterium atoms. This means that it is possible to indirectly locate the hydrogen atoms in e.g. a protein molecule. This is especially important in structural biology since hydrogen is known to play a crucial role in the catalysis of many biologically important chemical reactions, in hydrogen bonding interactions and also in understanding the orientation of the ordered solvent molecules [127–129]. A recent example of this is the peptide cleavage catalyzed by HIV-1 protease [130] for which neutron scattering has been extensively used to generate critical information about the hydrogen atom positions of the catalytic aspartates [131]. Neutron diffraction datasets on an average result in a highest resolution between 1.5 and 3 Å. When combined with prior information, stereochemical arguments and model building, such datasets can yield quasi-atomic resolution in the model [132]. Such a resolution also allows the determination and refinement of the positions of deuterium atoms.

Since X-rays and neutrons interact differently with the atoms, the information obtained from these techniques can be used in a complementary manner [133–135]. The large incoherent cross section of the hydrogen atoms results in significant background scattering [11]. Hence, it is always helpful to use perdeuterated or at least partially deuterated samples rather than completely hydrogenated ones although the hydrogenated ones can be used in case if the solvent is D_2O. This is usually done to understand the changes in orientational preferences or hydrogen bonding patterns of water molecules. In the case of partially deuterated molecules, there can be either a partial or a complete cancellation of the hydrogen electron density by adjacent heavy atoms or exchangeable D atoms which sets a limit on the visibility and interpretation of H-atoms [136]. This is known as 'the cancellation effect'.

Since X-ray and neutron diffraction are complementary methods, the model generated during the data processing can be refined against both X-ray and neutron datasets simultaneously [136]. The corresponding new total refinement target is then as follows:

$$T_{total} = w_{nxc}(w_{xc\_scale}T_{xray} + w_{nc\_scale}T_{neutron}) + w_cT_{geom} \tag{2.35}$$

where w_{nxc} is determined automatically [137], T_{xray} and $T_{neutron}$ can be any of the crystallographic targets computed with X-ray and neutron data, respectively [least-squares, maximum-likelihood, phased maximum-likelihood (mlhl) or twinned least-squares] and w_{xc_scale} and w_{nc_scale} are additional scales (set to 1 by default) that allow the user to adjust the automatically determined weights. By default w_c is set to 1, but setting it to zero turns off the geometry restraints, allowing unrestrained refinement. The equation 2.35 is similar for both restrained coordinate and ADP refinement. Either or both reflection data sets can be treated for twinning in refinement using a twinned least-squares target. The parameters of the H and D atoms are refined in the same way as the other heavier atoms in the structure.

2.5 Dynamic light scattering

2.5.0.1 General theory

Dynamic light scattering is used routinely in order to access the translational diffusion of biomolecules. The fundamental principle of this technique involves generation of a laser beam which is scattered by the sample. The intensity fluctuations are then monitored in the outgoing scattered signal at a given angle to yield 'an autocorrelation function'. In case of DLS, the outgoing signal is described in terms of two autocorrelation functions. Among these, the second order autocorrelation function describes the measured scattered intensity which depends on motions of the particle and is given by [138]:

$$g^{(2)}(\tau) = \frac{\langle I(t)I(t+\tau)\rangle}{\langle I(t)\rangle^2} \tag{2.36}$$

Equation 2.36 above also includes several instrumental artefacts. For example, there is an assumption that the properties of the particles are averaged over time which is not completely true, especially in case of solution. Equation 2.36 can be related with the first order correlation function using Siegert relation:

$$g^{(2)}(\tau) = b + \beta|g_{(1)}(\tau)|^2 \tag{2.37}$$

where b is a baseline term, usually close to unity, and β is a coherence factor, that depends on instrumental parameters, and also on some of the studied system parameters. The first order correlation function is usually represented by a decaying

exponential for diffusive motions, with a parameter Γ, related to the diffusion coefficient by $\Gamma = Dq^2$, with q^2 being the momentum transfer.

The first order correlation function can be related to the scattered electric field correlation function - also known as first order correlation function - which directly depends on coherent scattering from particles moving relative to each other. This relation is represented by the following equation:

$$g^{(1)}(\tau) = \frac{\langle E(t)E(t+\tau)\rangle}{\langle E(t)\rangle^2} \tag{2.38}$$

The correlation function for monodisperse systems can be represented by a single exponential as seen in subsection 4.1.5 of chapter 4. However, for polydisperse systems, $g^{(1)}$ is represented as an integral over the decay rate distribution and is represented by the following equation:

$$g^{(1)}(\tau) = \int_0^\infty G(\Gamma)\exp\{-\Gamma_\tau\}d\Gamma \tag{2.39}$$

2.6 Molecular dynamics simulations

The principles governing the dynamics and electronic structure of atoms have been known since the beginning of the 20^{th} century. However, for atoms bigger than helium, the problem becomes too complex to be solved exactly. With the advent of computers, it has been possible to calculate, under some simplifying assumptions, the electronic structure of various small molecules. This was done using quantum mechanical principles within the framework of the Hartree-Fock matrix and Roothan-Hall equation [89]. More recently, density functional theory has gained attention as a good alternative way to compute electronic structures. However, these quantum mechanical methods are computationally intensive especially in the case of macromolecules which typically contain a few thousand to a few hundred thousand atoms. It was found that even in absence of the electronic structure it is possible to develop reliable models, considering only the nuclei positions, and making use of classical physics to compute energy and motions.

2.6.1 Equation of motion

The basis of molecular dynamics simulations is Newton's second law,

$$\frac{d^2}{dt^2}\boldsymbol{r}_{\mathrm{i}} = \frac{F_{\mathrm{i}}}{m_{\mathrm{i}}} \tag{2.40}$$

where $\boldsymbol{r}_{\mathrm{i}}$ is the i^{th} atom position, F_{i} the force acting on it and m_{i} its mass. Using a proper potential function, the force can be easily obtained by derivation with respect to coordinates. Thus, the strategy is to use small time steps to integrate the second law. Given an atom at a certain position at time t, we can, using the computed force, determine the acceleration to be applied on, and obtain its position and velocity at time $t + \delta t$. The most widely used algorithm for this purpose is the Verlet algorithm [139]. This algorithm makes use of the following expansions,

$$\begin{aligned} \boldsymbol{r}(t+\delta t) &= \boldsymbol{r}(t) + \delta t\dot{\boldsymbol{r}}(t) + \frac{\delta t^2}{2}\ddot{\boldsymbol{r}}(t) + \mathcal{O}(\delta t^3) \\ \boldsymbol{r}(t-\delta t) &= \boldsymbol{r}(t) - \delta t\dot{\boldsymbol{r}}(t) + \frac{\delta t^2}{2}\ddot{\boldsymbol{r}}(t) + \mathcal{O}(\delta t^3) \end{aligned} \tag{2.41}$$

Adding these therefore leads to an expression giving the new position in terms of known quantities,

$$\boldsymbol{r}(t+\delta t) = 2\boldsymbol{r}(t) - \boldsymbol{r}(t-\delta t) + \delta t^2\ddot{\boldsymbol{r}}(t) \tag{2.42}$$

Then, velocities can be obtained using

$$\dot{\boldsymbol{r}}(t) = \boldsymbol{r}(t+\delta t) - \boldsymbol{r}(t-\delta t)/\delta t \tag{2.43}$$

Various types of algorithms exist but are beyond the scope of this work. An ideal or error-free algorithm would be able to accurately predict the energy conservation throughout the simulation. However, because of the truncated expansion, the computation gives rise to an erroneous result. A solution to this problem is to compare the predicted acceleration from the previous step using Taylor expansion with the computed one at the current step. The $\Delta\ddot{\boldsymbol{r}}$ can then be used to apply a simple linear correction for current values of position, velocity and acceleration.
Furthermore, various motions, with different time scales are involved in biological molecules, many of which do not require the usual 1 or 2 femtosecond time steps, but could be very well approximated using larger time steps. To address this point, a commonly used algorithm is r-RESPA [140], which consists in a reversible application of time translation operators on the system in the form

$$\Gamma(t) = e^{i(L_1+L_2+L_3+L_4)t}\Gamma(0) \tag{2.44}$$

where $\Gamma(t)$ is the system state at time t, and L_{i} denotes the Liouville operator associated with a given term i in the potential energy function,

$$iL = \sum_{\mathrm{i}=1}^{3N}\left[\frac{\partial x_{\mathrm{i}}}{\partial t}\frac{\partial}{\partial x_{\mathrm{i}}} + F_{\mathrm{i}}(x)\frac{\partial}{\partial p_{\mathrm{i}}}\right] \tag{2.45}$$

with x_{i} and p_{i} denoting position and conjugate momentum of atom i, respectively, and F_{i} the force acting on it. Each type of force comes with a given time step, being a multiple of the smallest one. The standard Verlet algorithm is retrieved if all time steps are equal. This method substantially improves simulation performances, and shows a good stability, even for long runs.
In all the aforementioned algorithms, we see that we always require to compute the force. This is done by means of a defined force field, which is discussed in the next subsection.

2.6.2 Force field

In case of the assessment of macromolecular dynamics, several parameters, such as bond lengths, angles, torsions can be taken into account. These parameters usually come with an equilibrium value and the atoms or groups of atoms oscillate around

it. There are a plethora of force fields available, but they usually contain some of the fundamental terms as follows,

$$\begin{aligned} V(\boldsymbol{r}) = &\sum_{\text{bonds}} \frac{k_\text{i}}{2}(l_\text{i} - l_{\text{i},0})^2 + \sum_{\text{angles}} \frac{k_\text{i}}{2}(\theta_\text{i} - \theta_{\text{i},0})^2 + \sum_{\text{torsions}} \frac{V_\text{n}}{2}(1 + \cos(n\omega - \gamma)) \\ &+ \sum_{\text{i,j=i+1}}^{N} \left(4\epsilon_\text{ij} \left[(\sigma_\text{ij}/r_\text{ij})^{12} - (\sigma_\text{ij}/r_\text{ij})^{6} \right] + \frac{q_\text{i} q_\text{j}}{4\pi\epsilon_0 r_\text{ij}} \right) \end{aligned} \tag{2.46}$$

The first term corresponds to the bond length, with an equilibrium length $l_{\text{i},0}$ for an i atom pair. The second term is exactly the same, but for bond angles. The third term denotes rotations of a group around a bond axis, and the last one contains Coulomb and Lennard-Jones potentials for electrostatic and van der Waals interactions, respectively. Cross-terms can be used as well. As an example, stretching of two bonds with a finite angle between the two bond axes can be controlled using an expression of the form

$$V(l_1, l_2, \theta) = \frac{k_{\text{l}_1,\text{l}_2,\theta}}{2} \left[(l_1 - l_{1,0})(l_2 - l_{2,0}) \right] (\theta - \theta_0) \tag{2.47}$$

Electrostatic interactions are usually computed using point-charge models. The idea is to place partial charges on the molecule so that its properties can be correctly reproduced. In the case of the N_2 molecule, placing a charge -q at each nucleus center and a +2q charge at the center of mass permits to retrieve the neutrality, absence of dipole moment and presence of quadrupole moment of the molecule. For proteins, partial charges might be calculated and optimized for each amino acid separately, using e.g. quantum mechanical calculations. In addition, various models have been proposed to compute effects such as anisotropy, polarisation, van der Waals forces, and others. Several of these are reviewed in Leach's book [89]. It is important to keep in mind that all force fields are always empirical, a general form is tried and parameters are optimized by comparison with experiment. Therefore, some might be more appropriate depending on the system being studied or the conditions used, and over-fitting might be a problem to take care of. Once the equation of motion and force field are set, it can be necessary to minimize the initial structure to avoid atom contacts and large energy terms due to bad geometry.

2.6.3 Running the simulation

2.6.3.1 Minimization

Several algorithms exist to minimize the energy. The simplest of these is the steepest-descent which consists in using derivatives of energy with respect to parameters using

the following expression,

$$\boldsymbol{v}_{\mathrm{k}} = -\boldsymbol{g}_{\mathrm{k}} + \gamma_{\mathrm{k}}\boldsymbol{v}_{\mathrm{k}-1} \tag{2.48}$$

with $\boldsymbol{g}_{\mathrm{k}}$ being the gradient in direction k, $\boldsymbol{v}_{\mathrm{k}-1}$ the displacement from the previous iteration in the direction k-1, perpendicular to k. The weighting factor γ is given by

$$\gamma_{\mathrm{k}} = \frac{\boldsymbol{g}_{\mathrm{k}}\boldsymbol{g}_{\mathrm{k}}}{\boldsymbol{g}_{\mathrm{k}-1}\boldsymbol{g}_{\mathrm{k}-1}} \tag{2.49}$$

Other methods make use of the second derivative. The best known example is the Newton-Raphson algorithm, where the derivative of the potential function is used along with its expansion

$$V'(x) = V'(x_{\mathrm{k}}) + (x - x_{\mathrm{k}})V''(x_{\mathrm{k}}) \tag{2.50}$$

such that the minimum can be found at x* using

$$x^* = x_{\mathrm{k}} - \frac{V'(x_{\mathrm{k}})}{V''(x_{\mathrm{k}})} \tag{2.51}$$

Notwithstanding its accuracy, this method is computationally demanding because of the Hessian matrix computation. Thus, quasi-Newton methods are often used to reduce algorithm complexity. These methods are based on an approximation of the Hessian matrix using only current, next position and gradients. The most famous ones are known as Davidon-Fletcher-Powell (DFP), Broyden-Fletcher-Goldfarb-Shanno (BFGS), and Murtaugh-Sargent(MS) algorithms [141].
Once energy is minimized, iterations can be started. From then on, there are several possibilities regarding the thermodynamic ensemble that is required.

2.6.3.2 Periodic boundary conditions

If we consider a protein in a box filled with water molecules, two problems arise. First, water molecules can escape the box, and eventually end up as free particles in vacuum. Second, it might be difficult to deal with non-bonded forces. The problem of particles leaving the box is easily solved by making their image at opposite sides enter the box. For non-bonded forces, the issue comes from the slow convergence of summation to compute pairwise interactions for all atoms.
A typical solution is to use the Ewald sum employing the following identity

$$\frac{1}{r} = \frac{f(r)}{r} + \frac{1 - f(r)}{r} \tag{2.52}$$

In addition, all charges are surrounded by a Gaussian distribution of neutralizing charges in real space, plus another distribution - of opposite charges from the first one - in reciprocal space. The complementary error function is usually used for f(r), and some correction terms are needed, such that the full expression is given by

$$V = \frac{1}{2}\sum_{\mathrm{i,j=1}}^{N} \begin{cases} \sum_{|\mathrm{n}|=0}^{\infty} \frac{q_\mathrm{i} q_\mathrm{j}}{4\pi\epsilon_0} \frac{erfc(\alpha|\boldsymbol{r}_\mathrm{ij}+\boldsymbol{n}|)}{|\boldsymbol{r}_\mathrm{ij}+\boldsymbol{n}|} \\ + \sum_{\mathrm{k}\neq 0} \frac{q_\mathrm{i} q_\mathrm{j}}{4\pi^2 L^3 \epsilon_0} \frac{4\pi^2}{k^2} e^{-k^2/(4\alpha^2)} \cos(\boldsymbol{k}.\boldsymbol{r}_\mathrm{ij}) \\ - \frac{\alpha}{\sqrt{\pi}} \sum_{\mathrm{k}=1}^{N} \frac{q_\mathrm{k}^2}{4\pi\epsilon_0} + \frac{2\pi}{3L^3} \left| \sum_{\mathrm{k}=1}^{N} \frac{q_\mathrm{k}}{4\pi\epsilon_0} r_\mathrm{k} \right|^2 \end{cases} \tag{2.53}$$

Improvements have been proposed, and other methods do exist to efficiently compute non-bonded interactions. Some of them are reviewed by Leach [89].

2.6.4 Complementing the experimental data

In order to run simulations that are comparable with experiments, we need first to correctly set up the system.
In this work, we used the ADC-LIG complex structure which was solved and refined by me whereas the ADC-APO structure was extracted from PDB (PDBID - 1AW8) [43]. Both systems were hydrated to the same level as for neutron scattering measurements (0.4 g H_2O / g protein).
Once simulations are finished, the results can be easily compared with neutron experiment using *n*Moldyn [142] or MDANSE [143] which can be used to compute the dynamic structure factor and the EISF from molecular dynamics simulations. In this work, we have used MDANSE in order to compute the generalized density of states and the EISF for ADC-APO and ADC-LIG samples (figure 6 in appendices)
Neutron scattering data provide thermally averaged information. However, local structure and dynamics might also be of great help in deciphering the aggregation process. This is where molecular dynamics simulations comes to play, by providing full trajectories of each individual atom. Various analysis algorithms are available now, either in the VMD software [144], or in other packages such as MDAnalysis [145].

Chapter 3

Materials and methods

In this chapter, recipes of growth media will be described. This will be followed by description of the optimized purification protocols and a brief explanation of the major instruments which were used for performing the experiments.

3.1 Preparation of media for cell growth, cloning and expression

3.1.1 Luria bertani media

LB media was obtained as a freeze-dried powder. A 25 g/L suspension of LB in H_2O was autoclaved at 120 °C for 20 min.

3.1.2 Luria bertani agar

A suspension of 25 g/L LB powder and 15 g/L of Agar powder in H_2O was autoclaved at 120 °C for 20 min.

3.1.3 Autoinduction media

10 g of Tryptone, 5g of yeast extract, 40 mL of 25x salts solution, 10 mL of 50% v/v aqueous glycerol, 1 mL of 2 M $MgCl_2$ aqueous solution and 200 μL of 1000x metals solution were mixed in a 2 L conical flask. H_2O was added up to 1L and the suspension autoclaved at 120 °C for 20 min. To the cooled sterilised solution was added 1 mL of 50% w/v aqueous glucose(filtered through a 0.22 μm filter) and 10 mL of 20% w/v aqueous lactose(filtered through a 0.22 μm filter).

3.1.4 Terrific broth media

A suspension of 47.6 g/L Terrific Broth powder and 10 g/L of glycerol in H_2O was autoclaved at 120 °C for 20 min.

3.1.5 H or D-minimal media

6.86 g of $(NH_4)_2SO_4$, 1.56 g KH_2PO_4, 6.48 g $Na_2HPO_4{\cdot}2H_2O$, 0.49 g $(NH_4)_2$-H-citrate and 5 g glycerol were mixed in H_2O to make a final volume of 1 L. The medium was then autoclaved and allowed to cool which was followed by the addition of 1 mL of filter sterilized 1M $MgSO_4$ (0.23 g/L) and 1 mL of metal salts. Metal salts solution is prepared by mixing 0.5 g of $CaCl_2{\cdot}2H_2O$, 16.7 g of $FeCl_3.6H_2O$, 0.32 g $CuSO_4{\cdot}5H_2O$, 0.15 $MnSO_4{\cdot}4H_2O$, 0.18$CoCl_2{\cdot}6H_2O$, 20.1 Na-EDTA in 1 L of H_2O.

In case of D-minimal medium, D_2O is used instead of H_2O. [1].

3.1.6 Agarose gel electrophoresis

Standard 1% agarose gel was used for testing the purity of the plasmid samples. 0.4 g of agarose was dissolved in 40 mL of TAE buffer

3.1.7 Transformation of chemically competent *E.coli* cells

The following protocol was used for the transformation of BL21(DE3), XL10 and XL1B *E.coli* cells.

1-3 μL of plasmid was mixed with 23-25 μL of cells in a sterile 1.5 mL eppendorf tube on ice. The cells were incubated on ice for 30 minutes. This was immediately followed by a heat shock at 42 °C for 40 seconds. After this step, the mixture was immediately transferred to the ice and incubated for 2-3 minutes. 250 μL of LB medium was added. The cells were then incubated at 37 °C for 1 hour. 100 μL of cells were then spread onto the agar plate containing the appropriate antibiotic for selection (kanamycin or ampicillin). The plate was incubated at 37 °C for 15 hours.

[1]The conventional recipe of metal salts solution contains 0.18 g $ZnSO_4{\cdot}7H_2O$ and 0.16 g $CuSO_4{\cdot}5H_2O$ in 1 L of H_2O. Here we did not use the Zinc salt and doubled the Copper salt as introduding Zinc into the active site results into inactive protein (section 1.3)

3.1.8 Overexpression of PanD in autoinduction media

E.coli BL21(DE3) cells were used for the overexpression of N-terminal 6X His-tagged wild-type PanD. This was done by transforming the *E.coli* MG1655 ΔpanZ ΔpanD DE3 cells with pRSETA-ADC-WT plasmid [146]. Overexpression in the autoinduction media was based on previously established protocols which employs induction of T7 RNA polymerase [147]. 1L of sterilized autoinduction media with 100 μg/L ampicillin in a 2 litre conical flask was inoculated with 1 mL of overnight grown miniculture. The culture was then incubated at 37°C with shaking (180 rpm) for 20 hours. This was followed by centrifugation at 10000 g for 15 minutes in order to harvest the induced cells. The induced pellet was then incubated at 37°C overnight in order to get activated ADC. The activated pellet was then stored at -80°C prior to purification.

3.2 Purification of ADC

3.2.1 Purification of 6x His-tagged ADC

E.coli cells stored after the overnight 37°C activation were re-suspended in Ni-NTA lysis buffer (50 mM K_2HPO_4, 300 mM NaCl, 10 mM imidazole pH 7.4). Protease inhibitor cocktail tablet (cOmplete EDTA-free, Roche) was added to the mixture at 4°C according to the manufacturer's instructions. The cell slurry was lysed using sonicator (4 to 5 cycles of 1 min each with 25% amplitude with 1 minute gap after every cycle). The lysed cell slurry was then centrifuged at 10000 g for an hour. The prepacked Histrap column from GE healthcare was then attached to AKTA purifier followed by passage of 5 column volumes (CV) of water. The column was then equilibrated with 5 CV of Ni-NTA lysis buffer (50 mM K_2HPO_4, 300 mM NaCl, 10 mM imidazole pH 7.4). The supernatant post centrifugation of the lysate was then loaded onto the column at a flow rate of 2 mL/min to ensure complete binding of the protein to the column. This was followed by the passage of 7 CV of Ni-NTA wash buffer (50 mM K_2HPO_4, 300 mM NaCl, 50 mM imidazole pH 7.4). The protein was then eluted first using Ni-NTA elution buffer 1 (50 mM K_2HPO_4, 300 mM NaCl, 250 mM imidazole pH 7.4). The elution buffer 1 was passed through the column till there was no further increase in the UV absorbance (280 nm). This was followed by the passage of Ni-NTA elution buffer 2 ((50 mM K_2HPO_4, 300 mM NaCl, 500 mM imidazole pH 7.4) to ensure that most of the protein which is bound to the nickel column is eluted. The eluted protein was collected as 2 mL fractions. The protein size, expression level and purity were assessed using SDS-PAGE.

3.2.2 Regeneration of Ni-NTA column

Agarose resin loaded with Ni-NTA is regenerated in order to remove any contaminants post elution. The column was first washed with 6 CV of 0.1 w/v sodium dodecyl sulphate and then with 6 CV of 0.1M NaOH. 6 CV 20 mM EDTA was then used to remove Ni^{+2} by chelation. This was followed by washing of the column with 20 CV of H_2O and regenerated with 5 CV of 500 mM $NiSO_4$. Excess Ni^{+2} was removed by washing the column with 20 CV of H_2O and the column was stored in 20% ethanol.

3.2.3 Size exclusion chromatography

The elution fractions collected after Ni-NTA chromatography were combined and concentrated to 2 mL (Amicon ultra centrifugal concentrator 10000 MWCO) by centrifuging at 4000 g. The concentrated 2 mL protein solution was then transferred to a 2 mL eppendorf tube and then centrifuged again at high speed (10000 g) for 10 minutes to remove any possible precipitate. This was immediately followed by loading of the sample onto a Superdex®200 (HiLoad®16/60 or 26/60, GE Healthcare) and eluted isocratically in SEC buffer (Äkta purifier, GE healthcare) in 2 mL fractions and stored at 4°C. The protein concentration was monitored by absorption at 280 nm.

3.2.4 Determination of protein concentration

Protein concentration was determined from the UV absorbance at 280 nm. 2 μL of protein solution was loaded onto the Nanodrop instrument and the absorbance was measured at 280 nm. Protein concentration was determined by using the appropriate ε value. For PanD, this value is 26470 $M^{-1}cm^{-1}$ as determined by loading the protein sequence on the online platform ProtParam [148].

3.3 Dynamic light scattering

Dynamic light scattering experiments were conducted on an ALV-7004 instrument for both the apo-ADC and D-Serine liganded ADC samples. The recorded measurements covered the scattering angles from 30 to 150°. All measurements were conducted at $T = 298.8\,\mathrm{K}$. The ADC concentrations used covered a range from 5 to 20 mg/mL (16.67 to 83.3 μM) with an interval of 5 mg/mL. The concentration of D-Serine used was 20 times that of the ADC tetramer (in terms of μM) for each of the samples to ensure complete saturation of all the binding sites. The solvent used was D_2O.

3.4 Neutron spectroscopy

Experiments were performed on solutionsof ADC in D_2O using both the IN16B and IN5 spectrometers at the ILL [149, 150]. IN16B [151, 152] has an energy resolution of 0.75 μeV FWHM at 6.27 Å (Si111 crystal analyzer configuration), and IN5 has an energy resolution of approximately 80 μeV FWHM at 5 Å incident wavelength. A cylindrical double walled aluminium sample holder sealed with indium wire was used for the measurements, with the difference in the radius between the two walls being 0.15 mm and the outer diameter 22 mm. The total liquid sample volume was 1.2 mL. The identical samples were used consecutively in both the IN16B and associated IN5 experiments. The temperature was controlled with a standard Orange cryostat.

Independent inelastic neutron spectroscopy experiments were conducted on IN1 and IN5 with ADC powders hydrated with D_2O to 40% in order to capture neutron data over a broad energy range [153, 154]. We also conducted elastic/inelastic fixed window scan (EFWS) experiments on IN16B using ADC powder samples to observe Debye-Waller factor in case of ADC-APO and ADC-LIG [155].

3.5 Sample details for the neutron spectroscopy experiments

3.5.1 Experiment 1 - QENS studies on ADC solution (low concentration)

The concentration of ADC was 50 mg/mL corresponding to a dry protein volume fraction of 0.03 [81] and translating to 6.82×10^{23} hydrogens in total. The concentration of D-Serine was 30 mg/mL which translates to 12.046×10^{23} hydrogens in total.

3.5.2 Experiment 2 - QENS studies on ADC solution (high concentration)

The total quantity of ADC used was 162.3 mg, corresponding to a dry protein volume fraction of 0.09 [81]. To ensure the complete saturation of all the ADC binding sites, 45 mg of D-Serine was added per mL of the sample volume (135 mg/mL of ADC). This means there are a total of 18.29×10^{23} protons from the D-Serine against 18.07×10^{23} protons from the ADC molecule and that there are 47.6 molecules of D-

Serine per ADC monomer. [2] The protein solution and pure D_2O reference samples were measured at 280, 295, and 310 K. For reference, a pure D-Serine solution at 45 mg/mL was also measured on IN5 at one temperature (295 K).

3.5.3 Experiment 3 - Elastic fixed window scans and deep inelastic neutron scattering studies on ADC hydrated powder

ADC concentrated to ≈ 100 mg/mL was used for lyophilization. This concentration was measured using the UV absorbance at 280 nm as described in subsection 3.2.4. For lyophilization, the concentrated ADC solution was dialyzed against 100 mM ammonium acetate. The dialyzed solution was transferred to a 50 mL falcom tube which was then transferred to a glass bottle. The glass bottle was screwed tightly onto the lyophilizer and a vacuum pressure of 0.06 mbar was applied at -47 °C. The resulting dry powder weighed 97 mg for ADC-APO and 138 mg for ADC-LIG. Both the powder samples were then transferred to a plastic boat which was transferred to a clean dessicator containing Phosphorous pentoxide to ensure complete drying and then to a dessicator containing D_2O. The dessicator was then tightly closed by applying grease onto the lid. The samples were incubated till the weight of the samples increased by 40%, corresponding to 40% D_2O hydration. These samples were then transferred to an aluminium foil which was closed from all the sides, folded and transferred to the cylindrical sample holder. It was ensured that the process of transferring the hydrated samples from the dessicator to the sample holder was as quick as possible since there is always an exchange between the atmospheric H_2O and the D_2O hydration layer of the sample. This procedure was used for all the powdered ADC samples [153–155]. The details of all the experiments are listed in appendices (table 3).

3.5.4 Data reduction and fit algorithms

The Mantid software package [156] was used for the initial reduction of the neutron data recorded at IN16B. The Lamp software package [3] provided by the ILL was

[2] We ensured that the amount of D-Serine is 20-30 times higher than that of ADC in all the experiments. This is because the K_d for D-Serine is not known and also because soaking the crystals of ADC-APO in solution containing 20-30 times higher concentration of D-Serine resulted in a D-Serine occupied active site (chapter 4).

[3] For lamp software, please refer to: https://www.ill.eu/users/support-labs-infrastructure/software-scientific-tools/lamp/

used for the IN5 data. All fits were carried out using python3 scripts employing the *scipy.optimize.curve_fit* command. Errors in the fit parameters were calculated from the square root of the diagonal of the covariance matrix. The Voigt profiles used to calculate the scattering functions convoluted with the energy resolution functions were obtained from the real part of the Faddeeva function provided by *scipy.special*.

3.6 Moleular dynamics simulations

MD simulations were performed using Gromacs 2016.3 [157–159] with the Amber99SB-ILDN force field [160]. Parametrization of D-Serine was carried out by our collaborator, Dr. Stephan Niebling. Two different side-chain conformers of D-Serine were used and charges averaged over these two conformations. For the pyruvate residue, existing force-field parameters from acetate and the amide carbonyl were used. For analysis and visualisation of MD trajectories we used self-written Python scripts in combination with the modules MDAnalysis, [161] NumPy [162] and Matplotlib. [163] The ADC-LIG complex determined in this work and the ADC-APO structure (1AW8 [43]) were used as the starting structures for the MD simulations. For each simulation an ADC tetramer (apo and D-Serine bound) was placed in a cubic box with periodic boundary conditions (1 nm initial minimum distance of protein to all boundaries). The box was filled with water (ca. 48000 molecules). To neutralize the negative charge of the protein, a number of water molecules were replaced with sodium and chloride ions to reach a concentration of 100 mM of NaCl. For each system MD simulations were prepared at two different temperatures (285 K and 310 K). After an energy minimisation (50000 steps or maximum force $< 10\,\mathrm{kJ\,mol^{-1}\,nm^{-1}}$), an NVT equilibration with modified Berendsen thermostat and velocity rescaling [164], and a 0.1 ps timestep (separate heat bath for protein and solvent+ions) was run. This was followed by a NPT equilibration using a Parrinello-Rahman pressure coupling [165, 166] at 1 bar with a compressibility of 4.5x10−5 $\mathrm{bar^{-1}}$ and a 2 ps time constant. During both runs, a position restraint potential with a force constant of $1000\,\mathrm{kJ\,mol^{-1}\,nm^{-2}}$ was added to all protein atoms (including the ligand). All bonds to hydrogen atoms were constrained with the Linear Constrained Solver (LINCS) [167] with an order of 4 and one iteration. MD simulations were run with a time step of 1 fs and the leap-frog integrator. Coordinates were saved every 10 ps. A grid-based neighbor list with a threshold of 1 nm was used and updated every 10 fs. For long-range electrostatic interactions above 1 nm the particle-mesh Ewald method [168,169] was used with a fourth order interpolation and a maximum spacing for the FFT grid of 1.6 Å . Lennard-Jones interactions were

cut-off above 1 nm. A long range dispersion correction for energy and pressure was used to compensate for the Lennard-Jones interaction cut-off. [158] For each of the four MD simulations 250 ns were acquired.

3.7 Expression and purification of hydrogenated *E.coli* copper amine oxidase

3.7.1 Expression of *E.coli* copper amine oxidase induced by IPTG

Single colonies of BL21DE3 were picked up from LB agar plates with ampicillin. These were used to inoculate 10 mL of LB broth. The starter culture was then grown at 37 °C for 15 hours with shaking (180 rpm).

The large scale culture was then setup by inoculating 10 mL of starter culture into 1 L of LB broth containing ampicillin. The culture was then allowed to grow at 37 °C with shaking (180 rpm) till the $O.D_{600}$ reaches 0.7 after which the protein expression was induced by adding 1 mL of 1M IPTG stock solution to the culture. After the addition of IPTG, the culture was induced at a particular temperature for 18-20 hours with shaking (150 rpm).

We used three temperatures (25, 30 and 37 °C) to test the expression.

3.7.2 Purification of *E.coli* copper amine oxidase (pKK233-3 vector)

3.7.2.1 Isolation of the periplasmic fraction from *Escherichia coli*

The isolation of the periplasmic fraction from induced *E.coli* BL21DE3 cells was carried out by the hydrolysis of $\beta-1,4-$glycosidic linkages between N-acetylmuramic acid and N-acetylglucosamine in the peptidoglycan layer for which we used Hen Egg White Lysozyme (HEWL) and chelation of metal ions using EDTA. In order to prevent the lysis of the inner membrane, these steps were carried out at 4 °C and Sucrose was used to maintain osmotic pressure which was gradually reduced during the course of the lysis.

Buffers B1 to B5 were used for cell lysis (cf. subsection A.2 in appendices) After proceeding with the induction, the cells were harvested by centrifugation at 10000 g for 15 minutes followed by resuspension in 20 mL of B1. This was followed by the slow and dropwise addition of B2 with gentle stirring. This lysis mixture was

then incubated for an hour with gentle stirring. After this, 6 mL of B3 was added dropwise with gentle stirring. This was followed by the addition of 12 mL of B4 and 27 mL of B5. This was followed by incubation for about 10 minutes with gentle stirring.

This mixture was then centrifuged at 10000 g for 45 minutes to separate the periplasmic fraction (supernatant) and the partially lysed cells. One tablet of EDTA-free complete protease inhibitors (Roche) was added to the supernatant and completely dissolved. The supernatant was then dialysed against 20 mM Tris-HCl pH 7.0 overnight. In order to remove the precipitate, the dialysed periplasmic fraction was centrifuged at 10000 g for 45 minutes.

3.7.2.2 Purification using ion exchange chromatography

A prepacked Q-sepharose fast flow column was attached to the AKTA pure FPLC system and 5 column volumes (CV) of deionised water was passed through the column. This was followed by charging of the resin by passing 10 CV of 20 mM Tris pH 7.0, 1 M NaCl. The supernatant obtained after the centrifugation of the dialysed periplasmic fraction is then slowly loaded onto the Q-sepharose column at a flow rate of 1 mL/min. The column was then washed by passing 8 to 10 CV of 20% buffer B and 80% buffer A through the column which corresponds to 35 mM NaCl concentration. The concentration of buffer B was then increased by 5% in order to elute ECAO. 1.8 mL fractions were collected and the fractions containing the protein as monitored from A_{280} were assessed for the presence of ECAO using SDS-PAGE. The fractions which contained ECAO were pooled and dialysed against 20 mM Tris pH 7.0 overnight at 4°C. After passing 100% buffer B, 5 CV of 20 mM Tris pH 7.0, 1M NaCl was passed through the column in order to elute all the non specifically bound protein molecules. After this step, 5 CV of deionised water was passed through the column followed by 5 CV of 20% ethanol which is used for storing the column.

3.7.2.3 Size exclusion chromatography

The dialysis was followed by concentration of the sample to about 2 mL using Amicon® Ultra 50 kDa centrifugal filter unit (Millipore™). The prepacked HiLoad™ 26/60 Superdex™ 200 column was equilibrated with 1.5 CV of 20 mM Tris HCl pH 7.0 at a flow rate of 1 mL/min. The concentrated protein was then loaded to the column using a 2 mL sample loading loop and syringe. 1 mL protein fractions were collected using the AKTA pure fraction collection robot The fractions which were part of the elution peak as monitored from A_{280} were assessed by SDS-PAGE to

check for the presence of ECAO. The pure fractions were then pooled and dialysed against 4L of 20 mM Tris HCl pH 7.0 supplemented with 0.1 mM $CaCl_2$ and 0.05 mM of $CuSO_4$ for an hour at 4°C. Excess $CaCl_2$ $CuSO_4$ were removed by dialysis against 5 L of 20 mM Tris HCl pH 7.0 overnight at 4°C.

3.7.2.4 Storage of *E.coli* copper amine oxidase protein

The purified ECAO was concentrated to about 15 mg/mL using Amicon® Ultra 50 kDa centrifugal filter unit (Millipore™). The protein sample was then divided into 50 μl aliquotes and then flash- cooled in liquid nitrogen. These aliquotes were then stored at -80°C.

3.8 Dynamic light scattering studies on wild-type and mutant ECAO

Dynamic light scattering experiments were conducted as described in section 3.3. The solvent used was D_2O. The experiments were carried out with wild-type and double mutant (E573Q/I342F) of ECAO in order to compare the diffusion coefficients with those from QENS. The concentration range used was 1 to 8 mg/mL with an interval of 1 mg/mL, the solvent used was D_2O and the data were recorded at 295 K.

3.9 QENS studies on wild-type and mutant ECAO

QENS experiments were carried out with wild-type and double mutant (E573Q/I342F) of ECAO with the IN16B configuration described in section 3.4 at 280 and 295K [150]. Our analysis in this thesis will focus mainly on the spectra recorded at 295 K (section 6.6 in chapter 6).

3.10 Perdeuteration of *E.coli* copper amine oxidase

Ampicillin selection is not suitable for perdeuteration as β-lactamase enzyme which is produced by the bacteria against ampicillin leaks into the growth medium resulting in loss of the selection pressure and a high frequency of plasmid-less cells in a deuterated

high cell density culture. On the other hand, the kanamycin resistant protein is expressed cytosolically and not periplasmically, thereby avoiding the leakage of the protein as in the case of ampicillin.

3.10.1 *In-vitro* transposition

Our constructs encoding wild-type and single mutant (E573Q) of ECAO contained ampicillin resistance marker. In order to change this resistance marker to kanamycin, we used *in-vitro* transposition. We used EZ-Tn5<kan2> insertion kit from Epicentre (Lucigen). In this method, an enzyme transposase randomly inserts the transposon which contains the kanamycin resistance tag into the target plasmid DNA [170]. Since kanamycin tag is essential for fermenter run required for perdeuteration, we use this method to obtain plasmids containing protein construct with kanamycin resistance tag.

3.11 Moleular dynamics simulations

All-atom molecular dynamics (MD) simulations of enzyme copper amine oxidase and its mutants (E573Q and I342F/E573Q) were performed on software GROMACS v5.1.1 and Amber 14SB force field was used to derive all the interaction parameters. Parametrization of the copper center as well as running of the production simulations was carried out by our collaborator Ms. Shilpa Sharma from IITD, India. System preparation and analysis was done by equal contribution from myself and Ms. Shilpa. For the initial structure, the crystal structure of wild-type ECAO (PDBID 1D6Z) was used and its mutants (E573Q, I342F) were prepared with the mutagenesis tool of PyMOL [44]. In the case of E573Q mutant of ECAO, TYR was retained as residue 466 instead of TPQ whereas in case of E573Q/I342F mutant, TPQ was retained in one of the subunits (subunit A) whereas it was replaced with TYR in the other subunit (subunit B). The choice between TPQ and TYR residues for each of the subunits for the three systems was made based on the previously reported experimental studies [77]. In this work, 1.5, 0.2 and 1 TPQ was found to be present per subunit for wild-type, E573Q mutant and E573Q/I342F mutant respectively in solution. The systems were prepared by solvating the three proteins using TIP3P water model in cubical box and 100 mM calcium chloride was added in each of the systems. To remove any bad contacts these solvated electroneutral systems were subjected to energy minimization using steepest descent algorithm with a maximum gradient of 1000 kJ/mol/nm. Equilibration was done in two steps. First under

NVT ensemble, the system was heated to 295 K using velocity-rescale temperature coupling algorithm for 500 ps and then under NPT ensemble for another 500 ps using Parrinello-Rahman barostat [165, 166]. Coupling constants of 0.1 ps and 2 ps were used for NVT and NPT ensembles, respectively. The LINCS constraint algorithm was utilized to fix the bond lengths or bond angles after integration of forces [167]. Convergence of potential energy and density was monitored. Production run was carried out for 100 ns at NPT conditions, with time step of 2 fs. The study utilized periodic boundary conditions (PBC) to eliminate any problem due to finite size and Particle-mesh Ewald (PME) method for electrostatic calculations. Leap-frog algorithm was used to integrate the equations of motion. Coordinates and velocities were saved after every 10 ps.

3.12 SDS-PAGE analysis

3.12.1 Sample preparation

Purified protein solution

20 μL of protein sample was mixed with an equal volume of 2x SDS-PAGE loading buffer. The solutions were heated at 100°C for 10 minutes so as to ensure complete denaturation of the protein.

E.coli cells

1 mL of culture was centrifuged at 13000 rpm. The supernatant was discarded and the cells were resuspended in Ni-NTA lysis buffer. 20 μL of the suspension was mixed with an equal volume of SDS-PAGE loading buffer and heated at 100°C for 10 minutes so as to ensure complete denaturation of the protein.

3.12.2 Tris-glycine PAGE

A 12% running gel was prepared according to the table 3.1. Except ammonium persulphate (APS) and tetramethylethylenediamine (TEMED), the rest of the components were mixed thoroughly by inversion. APS and TEMED were added just before casting the gel. The mixture was poured and a layer of isopropanol or H_2O-saturated n-butanol was added to ensure a levelled surface. A stacking gel was prepared by following the same steps. Once the running gel was set, n-butanol or isopropanol was blotted off which was followed by the addition of stacking gel solution and the insertion of a comb. The gel was loaded with desired protein samples in

Table 3.1: Protocol for preparing Tris-glycine PAGE

Tris-glycine PAGE	Separating gel (12%)	Stacking gel
1.5 M Tris pH 8.8	3.75 mL	-
1.0 M Tris pH 6.8	-	625 μL
30 % Acrylamide (29:1 acrylamide:bis-acrylamide)	6.00 mL	660 μL
10 % SDS	150 μL	50 μL
Water	5.10 mL	3.66 mL
15% APS	200 μL	100 μL
TEMED	10 μL	10 μL

1x SDS-PAGE loading buffer and run in 1x SDS-PAGE Tris glycine running buffer at a constant voltage (180V).

3.12.3 Tris-tricine PAGE

A 10% separating gel was prepared according to table 3.2. All the components except APS and TEMED were mixed by vortexing. The separating gel was poured and allowed to set followed by the stacking gel. The comb was then inserted and allowed to set. The gel was loaded with desired protein samples in 1x SDS-PAGE Tris-tricine cathode and SDS-PAGE Tris-tricine anode buffers were added. The gel was run at 175 V.

Table 3.2: Protocol for preparing Tris-tricine PAGE

Tris-tricine PAGE	Separating gel (10%)	Stacking gel
Glycerol	2.0 mL	-
Tris-tricine gel buffer	5.00 mL	3.10 mL
30 % Acrylamide (29:1 acrylamide:bis-acrylamide)	5.08 mL	1.66 mL
Water	2.92 mL	7.74 mL
15% APS	200 μL	100 μL
TEMED	10 μL	10 μL

3.12.4 Visualization of SDS-PAGE gels

The gel was immersed into 50 mL of Coommassie blue stain. The solution was then warmed in a microwave for 1 minute (full 750W) and the gel was incubated at room temperature for an hour with rocking. The Coomassie blue stain was removed and

Coomassie destain was added to the gel and was incubated overnight with rocking. The gel was then washed with H_2O and photographed using a Gel Doc XR system (Bio-Rad).

3.13 Crystallization of *E.coli* copper amine oxidase

E.coli copper amine oxidase was concentrated to 12 mg/mL and was dialysed against 20 mM Tris pH 7.0 in order to proceed for crystallization. We used 1.4 to 1.8 M sodium citrate as precipitant stock and 1.4 M HEPES at pH ranging from 6.8 to 7.3 as the buffer.stock.

We tested a range of concentrations of sodium citrate in order to crystallize the hydrogenated wild-type ECAO first and then the perdeuterated wild-type ECAO. In addition, we used a stock concentration of 3.8 M Ammonium sulphate in combination with 1.8 M sodium citrate and 1 M Tris pH 8.0 for obtaining bigger crystals of ECAO. The optimization protocols which were tested in order to obtain bigger crystals suitable for neutron diffraction are explained in section 6.1 of chapter 6.

3.14 Diffraction data collection

Full X-ray diffraction datasets were collected for hydrogenated wild-type, H/D exchanged E573Q, E573Q/I342F and perdeuterated wild-type ECAO. A few neutron diffraction images were collected during one test day experiment at LADI-III, ILL, Grenoble.

3.14.1 Single crystal X-ray diffraction

The data for H/D exchanged E573Q, E573Q/I342F mutants of ECAO, hydrogenated wild-type ECAO and perdeuterated wild-type ECAO were collected at ID23.1 beamline at European Synchrotron Radiation Facility (ESRF), Grenoble and at I03 beamline at Diamond Light Source (DLS), UK. The crystals were transferred from H_2O to D_2O buffer and allowed to equilibrate for one to two weeks. This process was carried out thrice to ensure a complete H/D exchange. The crystals were cryoprotected step-wise by transferring the crystals to mother liquor containing 5, 15 and 20 % D-glycerol by soaking for a few seconds in each of the conditions and then flash-cooled in liquid nitrogen. For every crystal, three images were first collected in order

to evaluate the data quality. The X-ray data for H/D exchanged mutant crystals were then collected at 0.9851 Å at ID23.1 beamline and 0.976 Å at I03 beamline, at 100 K. 1000 frames at 0.2 ° oscillation, 0.5 s exposure and 40 % transmission were collected to a maximum resolution of 1.65 Å in case of H/D exchanged crystals and to 1.9 Å in case of perdeuterated wild-type ECAO.

3.14.2 Single crystal neutron diffraction

The largest crystals (>0.06 mm^3) of perdeuterated E573Q mutant of ECAO were used to collect a few test frames LADI-III beamline at Institut Laue-Langevin (ILL), Grenoble. 2 to 3 frames were collected at room temperature with 2 h exposure for each of the crystals. The neutron wavelength range used was 2.8 to 3.8 Å. The crystal was stationary as the Laue method of data collection is used at LADI-III.

3.15 Brief description of the neutron instruments used

The neutron spectroscopic techniques used were inelastic neutron scattering using cold neutrons, employing time-of-flight (IN5) and backscattering (IN16B) spectrometers, and using hot neutrons, employing IN1 vibrational spectrometer (LAGRANGE).

For high-resolution quasi-elastic neutron scattering (QENS) at IN16B, ILL, a polychromatic beam from the reactor is converted into a monochromatic beam of a wavelength of ≈ 6 $\mathring{A}$ using a velocity selector and, importantly, a monochromator crystal in backscattering. As described in section 2.2 of chapter 2, QENS denotes a special case of inlelastic scattering which is centered around the elastic scattering and arises from diffusive motions, i.e. overdamped vibrations. The energy resolution of IN16B is 0.75 μeV FWHM and the dynamic window that is accessible is from the elastic line up to +/- 30 μeV. This energy range and resolution translates to a time scale of picosecond to nanosecond. Hence, QENS can be used to access motions within these time scales which are biologically relevant. The lorentzian line widths Γ and γ obtained from QENS can be used to access apparent center-of-mass and ensemble-averaged internal diffusion coefficients respectively (cf subsection 2.2.3) and will be used extensively in this thesis to attain the objectives mentioned in the Introduction section.

For QENS on a wider energy range and low-energy inelastic neutron scattering (IN5, ILL), a polychromatic neutron flux is converted into a monochromatic beam

by a set of 6 choppers wherein the final wavelength is typically 5 to 10 Å , although the flux is the highest at 5 Å and hence is the optimal wavelength for this technique. The elastic energy resolution for the IN5 beamline is 80 meV FWHM at 5 Å whereby the energy range is theoretically 2.5 meV on the neutron energy loss side and is in principle unlimited on the energy gain side. IN5 was, thus, also employed to record low-energy vibrational spectra. IN5 data can thus be used to access the fast and slow internal diffusive motions which are on a relatively lower time scales, which are typically within picosecond range. Due to the accessibility of IN5 to faster time scales, the data collected is complementary to that from IN16B which can access slower time scales. The IN16B and IN5 data will be complemented further by Dynamic Light Scattering (DLS). DLS can be used to calculate the linewidth Γ_{DLS} (cf. section 2.5 in chapter 2) which is used to access the translational diffusion coefficient and can be correlated with line width of the first lorentzian Γ obtained from IN16B.

In addition, for deep inelastic neutron scattering we used IN1 (ILL) in the LAGRANGE (LArge GRaphite ANalyser for Genuine Excitations) mode. In this mode, the monochromators are built from copper single crystals, the spectra are typically recorded at sub-zero temperatures (10-200K) and the maximum possible energy transfer can be upto 500 meV, recorded on the neutron energy loss side, with the energy resolution being 2-3 meV within the energy transfer range of 50-200 meV. Beyond 200 meV , the energy resolution steadily reduces to about 5 meV. LAGRANGE, thus, allowed to complement the vibrational spectra recorded on IN5 towards higher frequencies. We also used elastic fixed window scan experiments (IN16B), in which elastic intensity is recorded at increasing temperatures and is used to obtain the atomic vibrational mean squared displacements. Mean squared displacements can reveal the temperature at which the onset of localized vibrational motions occurs. Hence, these results are also complementary to those obtained from IN1 which can be used to assign specific frequencies to a particular vibrational mode or modes which are usually seen as distinct peaks in the spectrum.

Finally, neutron Laue diffraction was carried out during one test day on the Laue Diffractometer LADI (LADI-III) at the ILL. This is a quasi-Laue neutron diffractometer which can be used for collecting the data at both room as well as cryo-temperatures and can be used to obtain the positions of deuterium atoms in a molecule. In the present study, the test data were collected at room temperature.

These experimentally obtained parameters from a range of techniques will be used extensively in this thesis to study the structure and dynamics of ADC as a result of addition of a ligand (D-Serine) and of ECAO as a result of the two non-active site mutations which severely affect its activity. [4]

[4]Details of all the instruments can be found at https://www.ill.eu.

Chapter 4

Results and discussion: aspartate α-decarboxylase

In this chapter, the results of ADC are explained. Expression and purification of ADC is followed by structure solution and refinement of ADC-LIG complex. This will be followed by assessment of the effect of binding of D-Serine ligand on the internal and global diffusive dynamics of ADC using a combination of quasi-elastic neutron scattering (QENS), dynamic light scattering (DLS) and MD simulation studies. These results are from Raskar et al.[1] which is currently being prepared for submission to a peer reviewed journal.

[1]Raskar, Tushar, Stephan Niebling, Juliette M. Devos, Briony A. Yorke, Michael Härtlein, Nils Huse, Trevor V. Forsyth, Tilo Seydel, and Arwen R. Pearson. "Diffusive dynamics of Aspartate α-decarboxylase (ADC) liganded with D-serine in aqueous solution." bioRxiv (2020).

4.1 The effect of binding of D-Serine on the structure and dynamics of ADC

4.1.1 Processing and analysis of X-day data for ADC-D-Serine complex crystal

The crystal structure of Aspartate α-decarboxylase in complex with D-Serine was determined to a resolution of 1.9 Å and is deposited in protein data bank with the ID 7A8Y (Table: 4.1). The structure was refined to final crystallographic R_{work} and R_{free} values of 17.9% and 19.4% respectively (Table: 4.1). As in the apo-ADC stucture, the ADC tetramer is formed by a crystallographic two-fold, with two ADC monomers in the asymmetric unit. As in the apo structure, a fraction of mis-processed ADC is present, where the backbone of the zymogen is cleaved, but the β subunit has an N-terminal serine instead of a pyruvate. This mis-processed form is present at 40 % occupancy.

Comparison of the ADC-APO structure (1AW8) with the D-Serine complex shows that the α C-terminal loop opens upon D-Serine binding (figure 4.1).

The D-Serine molecule adopts two conformations in both subunits of the asymmetric unit. One conformer (60% occupancy) forms hydrogen bonds with the mainchain of ALA-75, the sidechains of ARG-54 and THR-57, and the pyruvoyl carbonyl (figure 4.2) adopting a similar binding conformation to the native substrate L-aspartate. The second conformer (40% occupancy) forms hydrogen bonds with the sidechains of LYS-9 and TYR-58 and with the nitrogen atom of the misprocessed β N-terminal SER-25 (figure 4.2). In the apo-ADC structure the α C-terminal loop (residues 21 to 24) adopts two conformations, whereas in the D-Serine complex, the loop adopts a single, open conformation (figure 4.1, bottom). The degree of "openness" is not the same in the two ADC subunits of the asymmetric unit. There is a significant change in the conformation of the α C-terminal loop of chain A (subunit 1) which shows a displacement of 4.3 Å of the C_α of GLU-23 from its position in the apo structure and a clear change in the conformation of its sidechain (figure 4.1, bottom). However, for chain D (subunit 2), the structural change is relatively small with a C_α-C_α distance of only 0.8 Å for the same residue, although D-Serine is bound in both subunits (figure 4.1, top).

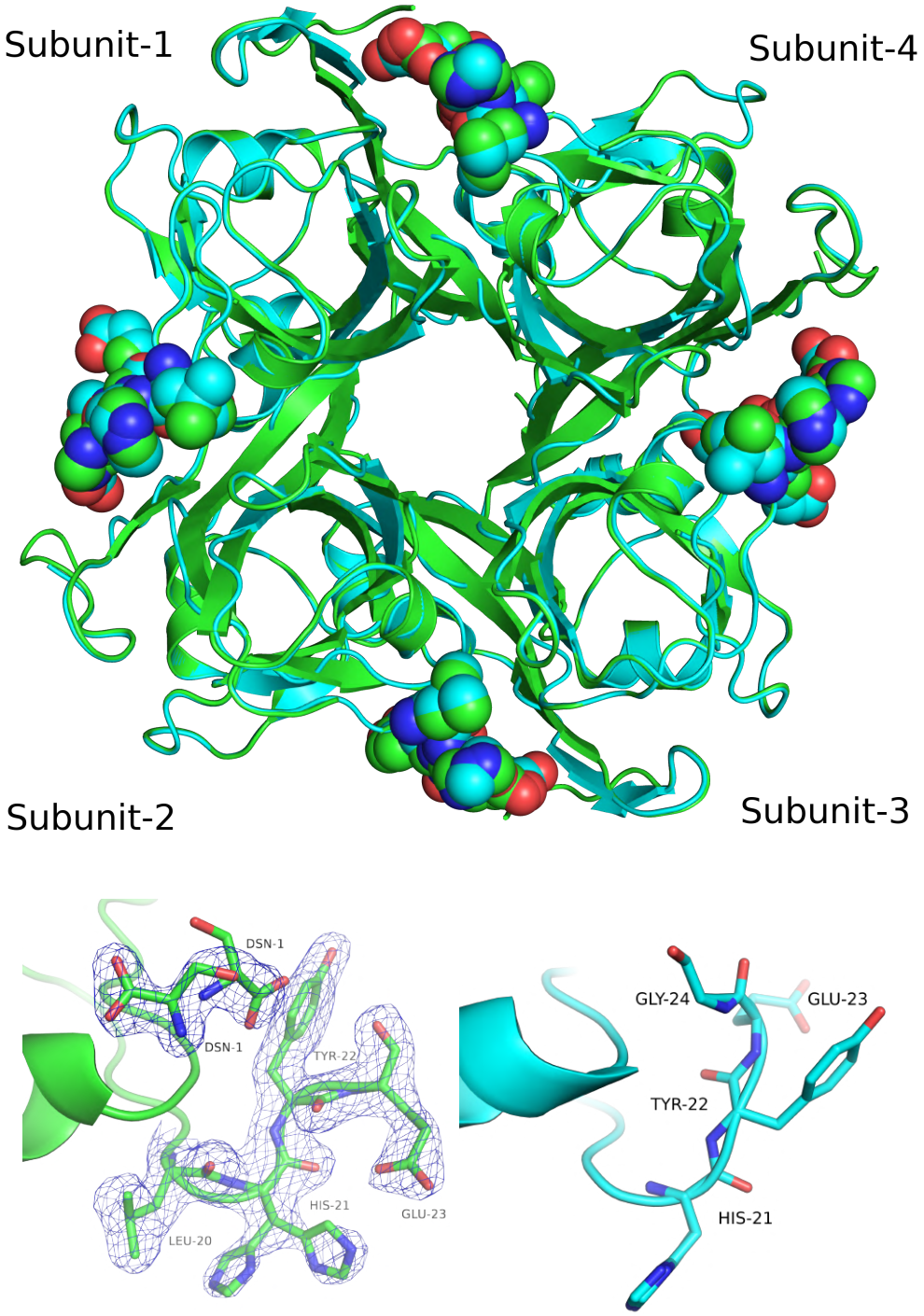

Figure 4.1: Top: superposition of ADC-APO (1AW8) (blue) and ADC-D-Serine (ADC-LIG) complex structure (green). The loop regions in four subunits are shown as spheres. Bottom: Change in the conformations of the α C-terminal loop upon binding of D-Serine in subunit 3 (bottom left) and the corresponding loop conformation in ADC-APO (PDBID 1AW8) (bottom right). $2mF_o$-DF_c electron-density map is displayed at a contour level of 1σ.

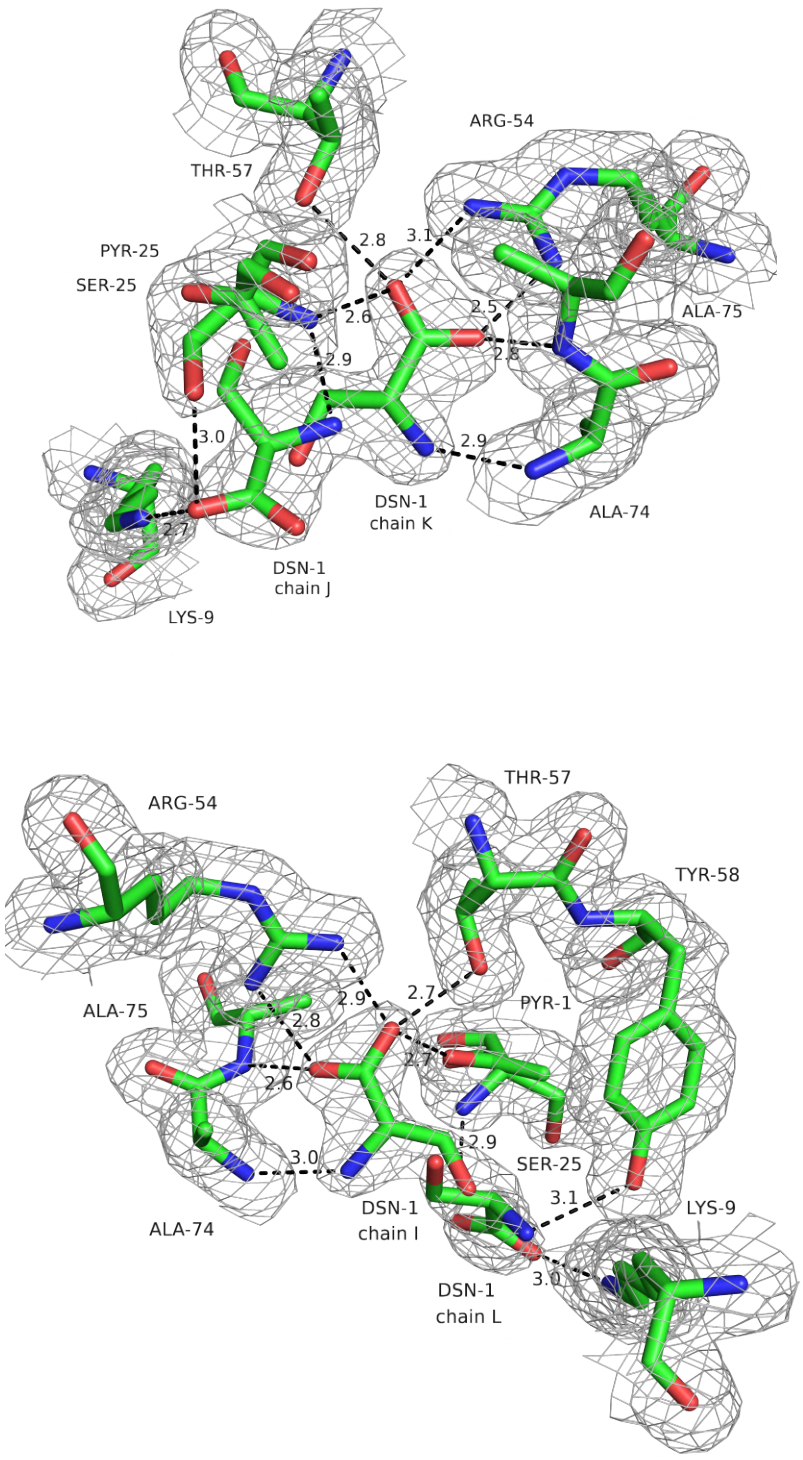

Figure 4.2: Hydrogen bonding network between the alternate conformations in subunit 1 (top) and subunit 3 (bottom) from the crystal structure of ADC-D-Serine (ADC-LIG) complex showing the slightly different interaction patterns.
Chain J and chain L lack the OG atom as we couldn't observe the density for it.

Table 4.1: Crystallographic refinement statistics for ADC-D-Serine (ADC-LIG) complex
The numbers in parenthesis represent the high resolution shell

	ADC-D-Serine complex (ADC-LIG)
Wavelength (Å)	0.9778
Resolution range (Å)	46.93 - 1.75 (1.81 - 1.75)
Space group	P 61 2 2
Unit cell (Å), °	71.3 71.3 216.6, 90 90 120
Total reflections	67840 (6601)
Unique reflections	33921 (3301)
Multiplicity	2.0 (2.0)
Completeness (%)	99.99 (99.97)
Mean I/sigma(I)	21.50 (5.25)
Wilson B-factor (Å^2)	19.07
R-merge	0.024 (0.16)
R-meas	0.034 (0.22)
R-pim	0.024 (0.16)
CC1/2	0.99 (0.93)
CC*	1.00 (0.98)
Reflections used in refinement	33921 (3301)
Reflections used for R-free	1702 (162)
R-work	0.1791 (0.2110)
R-free	0.1937 (0.2266)
CC(work)	0.97 (0.84)
CC(free)	0.96 (0.81)
Number of non-hydrogen atoms	2172
macromolecules	1950
ligands	14
solvent	208
Protein residues	244
RMS(bonds) (Å)	0.014
RMS(angles) (°)	1.67
Ramachandran favored (%)	96.51
Ramachandran allowed (%)	3.06
Ramachandran outliers (%)	0.44
Rotamer outliers (%)	3.5
Clashscore	5.15
Average B-factor (Å^2)	23.17
macromolecules (Å^2)	21.83
ligands (Å^2)	27.57
solvent (Å^2)	35.45

4.1.2 Molecular dynamics simulations

MD simulations were performed in order to analyze how the binding of D-Serine affects the internal motions of ADC. The simulations were validated by plotting the distance between the C_α atoms of a fixed residue in ADC (CYS-26) and that of D-Serine ligand against the simulation time. As seen in figure 4 in appendices, the distance varies between $\approx$ 6 to 8 Å. Hence, it was confirmed that the interaction between D-Serine ligand and protein was stable throughout the production run. As seen in figure 4.1, the conformational change associated with the displacement of the C-terminal loop of subunit 3 mainly involves residues 21 to 24. Hence, we monitored the change in the C_α-C_α distance profile in the production simulation between:

1. HIS 21 and GLY 24
2. TYR 22 and GLY 24.

As seen from figure 4.3, the distance histogram profiles are significantly different for ADC-APO and ADC-LIG. The histogram profile on the top panel shows that there are three peaks in case of ADC-APO at 8.2, 9.7 and 10.5 Å whereas for ADC-LIG, there are two peaks at 8.7 and 9.5 Å for distance 1. Similarly, there is a clear shift in the peak from 7.25 Å in ADC-APO to 5.75 Å in ADC-LIG for distance 2. This clear difference in the probability distribution for distance 1 and distance 2 indicates that the motion of the C-terminal loop is more confined in ADC-LIG than in ADC-APO. This observation is in agreement with the results from IN16B and IN5 which also show that the internal diffusive motions are slower and more confined for ADC-LIG than for ADC-APO (figure 4.10). These distance histogram profiles were also plotted for the individual subunits (figure 5). The changes in the two distances for the individual subunits of ADC-APO and ADC-LIG show that the influence of binding of D-Serine on the dynamics of the C-terminal loop is neither completely symmetrical nor consistent among the four subunits. This result agrees with the observations from the crystal structure of the ADC-LIG complex which shows the distinct structural changes associated with D-Serine binding in subunits 1 and 3 (figure 4.1).

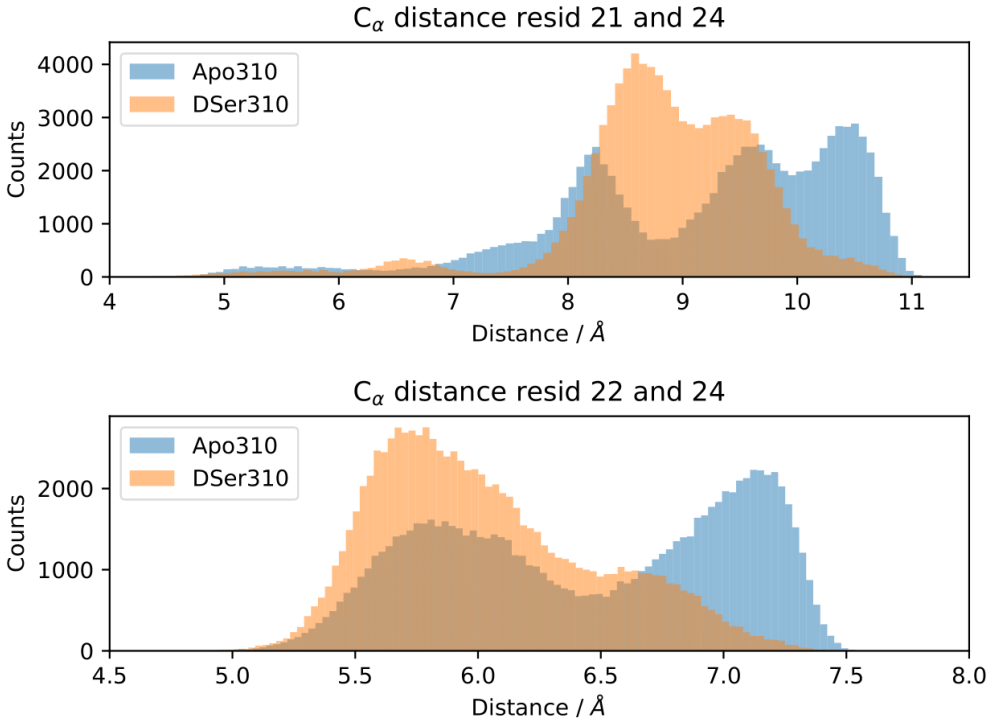

Figure 4.3: Histograms for the average C_α-C_α distances for all the four subunits between HIS 21-GLY 24 (top) and between TYR 22-GLY 24 (bottom) for ADC-APO (blue) and ADC-LIG (orange).

We also extracted $\chi 1$ and $\chi 2$ dihedral distributions for residues 21 and 23 of the C-terminal loop of ADC. Here, $\chi 1$ torsional is defined as $N - C_\alpha - C_\beta - C_\gamma$ angle whereas $\chi 2$ torsional is defined as $C_\alpha - C_\beta - C_\gamma - C_\delta$ angle. As seen in figure 4.4, the $\chi 1$ angle for one of the HIS-21 conformers overlaps with that of ADC-APO. In case of GLU-23, $\chi 1$ and $\chi 2$ angles are different for ADC-LIG when compared with both the conformers of ADC-APO. In this context, we compared the dynamics of HIS-21 and GLU-23 residues for the C-terminal loops of subunit 1 and 3 of ADC-APO and ADC-LIG complex. We calculated the χ dihedral angle distribution from the MD simulation trajectories of ADC-APO and ADC-LIG for these residues in subunit 1 and 3.

As seen in figure 4.5, there is a reasonable agreement between the dihedral angles in the crystal structures of ADC and those extracted from the MD simulation trajectories. The $\chi 1$ distribution for HIS-21 which is in the relatively rigid part of the C-terminal loop, is in agreement with the values from the crystal structure. In the case of $\chi 1$ and $\chi 2$ angles of GLU-23, a significant difference is observed not only between the two simulations but also between the different subunits of the same structure. This is also the case with the crystal structure. However, a reasonable agreement is found for the $\chi 1$ angle of GLU-23 only for subunit 3 but not for subunit 1 unlike the $\chi 1$ of HIS-21. The $\chi 2$ angle of GLU-23 is found to be in a better agreement than $\chi 1$ with the crystal structure. Eventhough there are some differences in the dihedral angles of GLU-23 in MD simulations and crystal structures, it can be clearly seen that in general, the distribution for ADC-LIG is relatively monomodal, unlike that for ADC-APO. Distinct peaks can be seen which correspond to the different conformers of HIS-21 and GLU-23 in ADC-LIG, indicating a lower degree of freedom of these residues compared to ADC-APO. This observation also supports our results which indicate that the C-terminal loop is more disordered in ADC-APO in comparison with ADC-LIG. The distance histogram was calculated to get an overview of the loop dynamics in all the subunits, whereas the dihedral angle distribution was calculated in order to compare the dynamics of the loop residues in specific subunits (subunit 1 and 3). The results from MD simulations agree with those from QENS and crystal structure of ADC-LIG complex, which indicate that binding of D-Serine confines the motion of the C-terminal loop of ADC, making it harder than in its APO state.

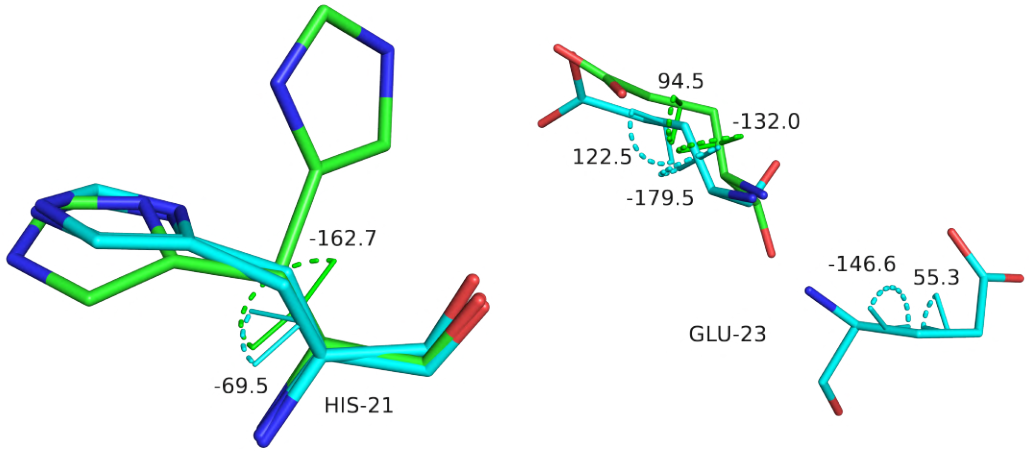

Figure 4.4: $\chi 1$ angle of HIS-21 (left); $\chi 1$ and $\chi 2$ angles of GLU-23 (right) for ADC-APO (cyan) and ADC-LIG (green) structures after superposition.
The conformers of HIS-21 and ASP-23 presented here are obtained by superposing the subunit 3 of ADC-APO (blue) and ADC-LIG (green). The measure of the alternate conformer of HIS-21 of ADC-LIG is not shown for ease of reading. However, we have included this value in figure 4.5.

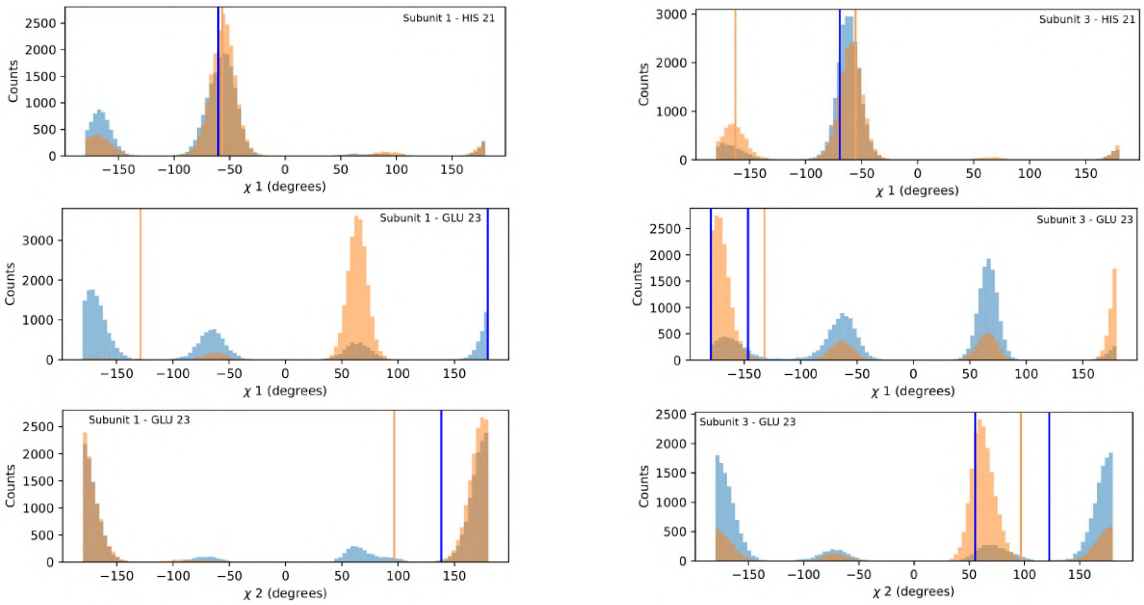

Figure 4.5: Dihedral angle distribution for - χ 1 angle of HIS 21 for subunit 1 (top left) and subunit 3 (top right), χ 1 angle of GLU 23 for subunit 1 (middle left), and subunit 3 (middle right), χ 2 angle of GLU 23 for subunit 1 (bottom left) and subunit 3 (bottom right) for ADC-APO (blue) and ADC-LIG (orange) from MD simulations. The vertical lines are the dihedral angle values extracted from the crystal structures of APO (1AW8) and D-Serine complexed (current structure) ADC.

4.1.3 Picosecond diffusive dynamics

The data from the time-of-flight spectrometer IN5 contain information on both quasielastic scattering arising from diffusive motions and inelastic scattering arising from vibrational motions. Here, we present the QENS part only. The reduced QENS data from IN5 were fitted for each momentum transfer q independently in a two-step procedure as follows. First, the spectra from the buffered D_2O solvent were fitted by [171]

$$\begin{aligned} S(q,\omega) =& \mathcal{R}(q,\omega) \otimes [I_{\mathrm{D_2O},1}\mathcal{L}(\gamma_{\mathrm{D_2O},1},\omega) \\ &+ I_{\mathrm{D_2O},2}\mathcal{L}(\gamma_{\mathrm{D_2O},2},\omega) + I_{\mathrm{D_2O},\delta}\delta(\omega)] \\ &+ s\omega + c. \end{aligned} \tag{4.1}$$

where, $\mathcal{R}(q,\omega)$ represents the apparent energy resolution function of IN5, which also includes effects from the sample container geometry, and $\otimes$ is the convolution in the energy transfer $\hbar\omega$. Technically, the convolution $\mathcal{R}\otimes$ is carried out by modeling $\mathcal{R}$ as a sum of an arbitrary number of Gaussian functions. Therefore, the observable $S(q,\omega)$ can be fitted by a sum of Voigt functions [172–174]. $\mathcal{L}(\sigma,\cdot)$ represents a Lorentzian function with the width σ, and $\delta(\omega)$ is the Dirac function describing the elastic scattering arising from the sample container. $I_{\mathrm{D_2O},1,2,\delta}$ denote scalar q-dependent intensity factors, and s and c are q-dependent scalars describing an apparent background arising from the sample, container, and instrument itself.

Second, the QENS spectra from the protein solution samples and pure D-Serine reference sample are fitted by

$$\begin{aligned} S(q,\omega) &= \mathcal{R}(q,\omega) \otimes \{I_\gamma\mathcal{L}(\gamma\mathrm{IN5},\omega) \\ &+ I_\Gamma\mathcal{L}(\Gamma_{\mathrm{IN5}},\omega) + I_\delta\delta(\omega) \\ &+ (1-\varphi)\,[I_{\mathrm{D_2O},1}\mathcal{L}(\gamma_{\mathrm{D_2O},1},\omega) \\ &+ I_{\mathrm{D_2O},2}\mathcal{L}(\gamma_{\mathrm{D_2O},2},\omega)]\} \\ &+ s\omega + c. \end{aligned} \tag{4.2}$$

where, $I_{\mathrm{D_2O},1,2}$, s, and c are fixed from the fit results of the corresponding pure solvent (equation 4.1), and φ is the known protein volume fraction in the sample. The term $I_\delta\delta(\omega)$ accounts for apparent elastic scattering arising from both, the sample container and the proteins which are quasi-static on the observation scale of IN5. The fit results for the Lorentzian widths assigned to ADC in equation 4.2 are depicted in figures 4.7 and 4.8

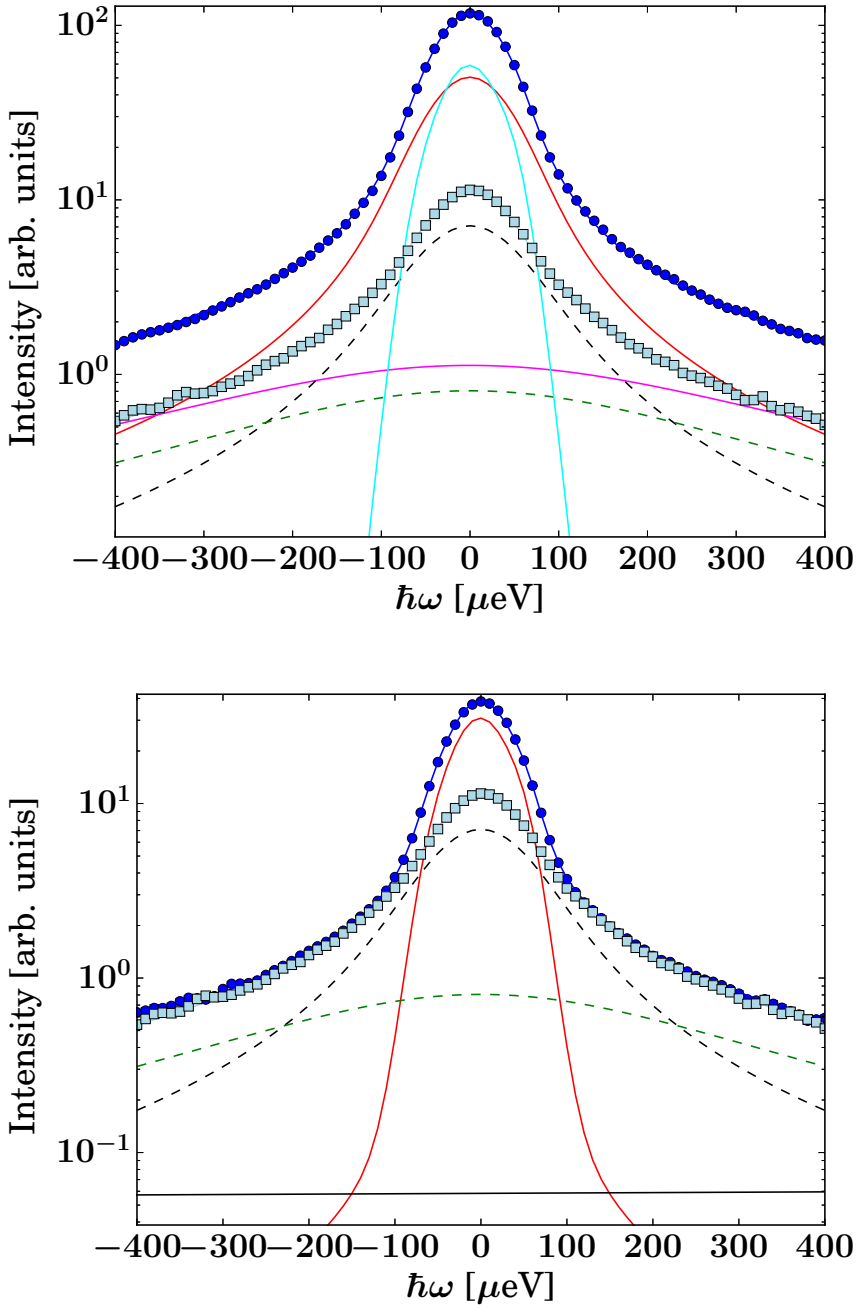

Figure 4.6: Example spectra recorded on IN5 for ADC-LIG. (top) and ADC-APO (bottom), respectively, (symbols) at $q = 0.6\,\text{Å}^{-1}$ at $T = 295\,\text{K}$. The dark blue circle symbols represent the protein solution signal, the light blue square symbols the corresponding solvent signal. The lines superimposed on the protein sample spectra represent fits of equation 4.2. The dashed lines represent the two Lorentzians decribing the buffer, equation 4.1. The other solid lines depict the two Lorentzians associated with the protein and the apparent elastic contribution (equation 4.2). The difference in the spectral intensity between the two samples arises from the D-Serine present in the ADC-LIG sample.

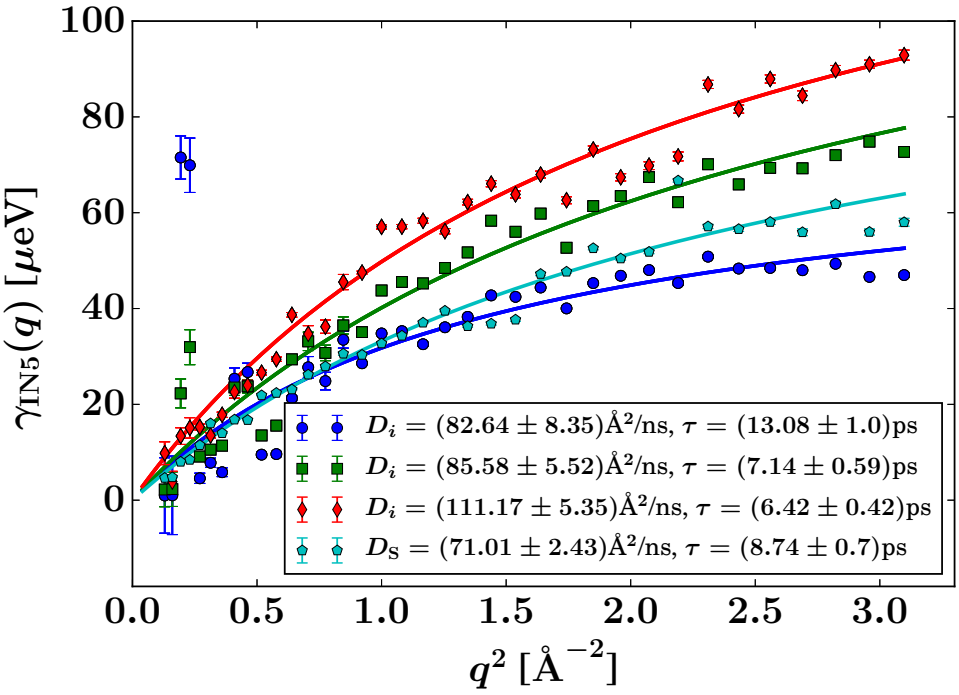

Figure 4.7: Width γ_{IN5} of the Lorentzian accounting for slow internal diffusive motions observed on IN5 (equation 4.2) for ADC-LIG at different temperatures (circles – 280 K, squares – 295 K, and diamonds – 310 K) as well as for the pure D-Serine reference sample (pentagrams – 295 K), and fits using the jump diffusion model (equation 4.5).

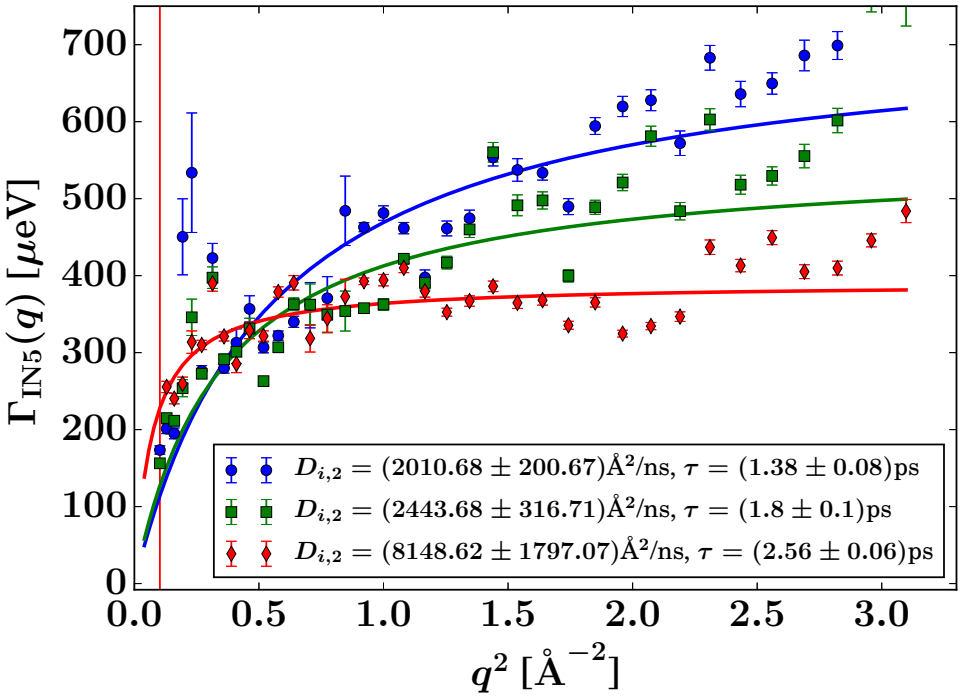

Figure 4.8: Width Γ_{IN5} of the Lorentzian accounting for fast internal diffusive motions observed on IN5 (equation 4.2) for ADC-LIG at different temperatures (symbols) and fits using the jump diffusion model (equation 4.5), the fit parameters being reported in the legend.

In the case of the ADC-APO sample, the amplitude of the Lorentzian associated with the slower part of the internal diffusive dynamics is very small (figure 4.7). In contrast, this contribution is significant in the pure D-Serine reference sample (figure 4.7), although with a different linewidth, at the same measured temperature of 295 K. In the ADC-D-Ser samples, the first Lorentzian can thus be attributed to the dynamic contribution of the D-Serine. It should be noted that this Lorentzian with width γ_{IN5} in equation 4.2 accounts for an average over multiple dynamic contributions that cannot be further discerned with the current accuracy of the data and modeling. In the case of the ADC-D-Serine sample, this Lorentzian likely reflects both bound and unbound D-Serine. The observed diffusion coefficient of pure D-Serine is in reasonable agreement with earlier findings [175–177]. Overall, it appears that few accessible diffusive dynamic contributions on the picosecond time scale are associated with the protein itself, suggesting an overall highly rigid protein consistent with its high content of β-sheet (approximately 40%) as determined using the DSSP server [178, 179].

4.1.4 Nanosecond diffusive dynamics

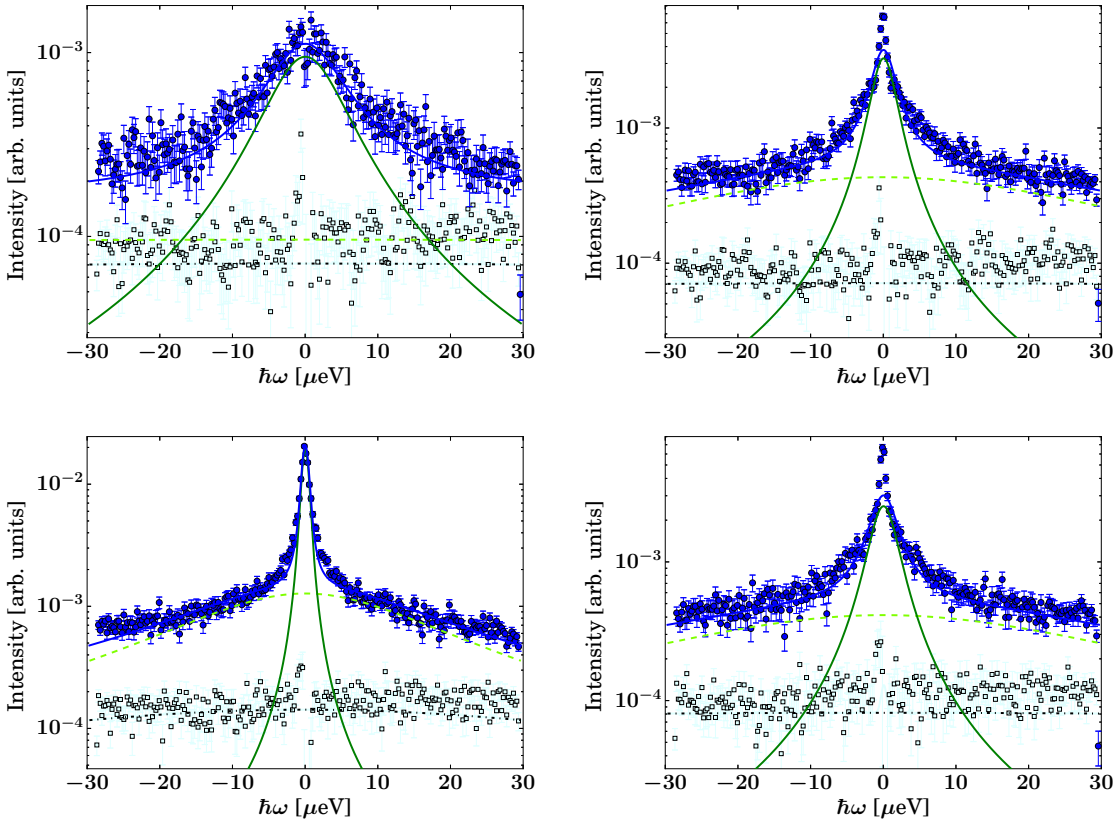

Figure 4.9: Example spectra (filled symbols) of ADC-APO (top left), ADC-LIG (top right) at $q = 1.57\,\text{Å}^{-1}$ and the example spectra of ADC-LIG at $q = 0.69$ (bottom left) and 1.79 (bottom right) Å^{-1} respectively recorded on IN16B at $T = 295\,\text{K}$, and fit using equation 4.3 (solid line superimposed on the symbols consisting of the components (other lines) as given by eq. 4.3). The lower light open symbols represent the pure solvent signal.

With its high energy resolution, the spectrometer IN16B accesses quasi-elastic scattering containing information on superimposed center-of-mass and internal diffusion. Figure 4.9, shows the spectra for ADC-APO and ADC-LIG. It is clear from the figure that the raw data for the two samples are significantly different and that the bottom two plots indicate the q-dependence of the Lorentzian linewidth for ADC-LIG sample. We employ the following model for the scattering function observable

on IN16B [172–174]:

$$\begin{aligned} S(q,\omega) = \mathcal{R}(q,\omega) \otimes \{\beta(q)\,[\,A_0(q)\,\mathcal{L}(\gamma(q),\omega) \\ + (1 - A_0(q))\,\mathcal{L}(\gamma(q) + \Gamma(q),\omega)] \\ + \beta_{\mathrm{D_2O}}(q)\,\mathcal{L}(\gamma_{\mathrm{D_2O}}(q),\omega)\} \end{aligned} \tag{4.3}$$

where, $\mathcal{R} = \mathcal{R}(q,\omega)$ denotes the spectrometer resolution function, and $\mathcal{L}(\mathring{\mathrm{u}},\sigma)$ is a Lorentzian function with the HWHM σ. $\beta(q)$, $A_0(q)$, $\gamma(q)$, and $\Gamma(q)$ are scalar fit parameters. The scalar parameters for the solvent water contribution $\beta_{\mathrm{D_2O}}(q)$ and $\gamma_{\mathrm{D_2O}}(q)$ are fixed based on a pure solvent measurement using established protocols [180].

4.1.4.1 Nanosecond apparent center-of-mass diffusion

In a global fit approach of the spectra for all q simultaneously using equation 4.3, we further assume a Fickian center-of-mass diffusion of the proteins described by the observable apparent diffusion coefficient D obeying the relation which was confirmed for other proteins [173, 174, 180],

$$\gamma(q) = Dq^2. \tag{4.4}$$

The resulting fit parameters for D according to equations 4.3 and 4.4 for all ADC solution samples measured on IN16B are summarized in the figure 4.11 and will be discussed in section 4.1.4.3.

4.1.4.2 Sub-nanosecond internal dynamics

To examine the internal diffusion, a similar global fitting approach was taken, using equation 4.3, and assuming a jump diffusion model for the internal diffusion within ADC [181],

$$\Gamma(q) = \frac{D_i q^2}{1 + D_i q^2 \tau}. \tag{4.5}$$

where, D_i is the jump diffusion coefficient and τ is the so-called residence time between diffusive jumps. In earlier work [182], this jump diffusion model has been shown to be a reasonable approximation. Importantly, in the fits of the IN16B spectra, the values of D_i and τ are fixed based on the results from IN5 for D-Serine

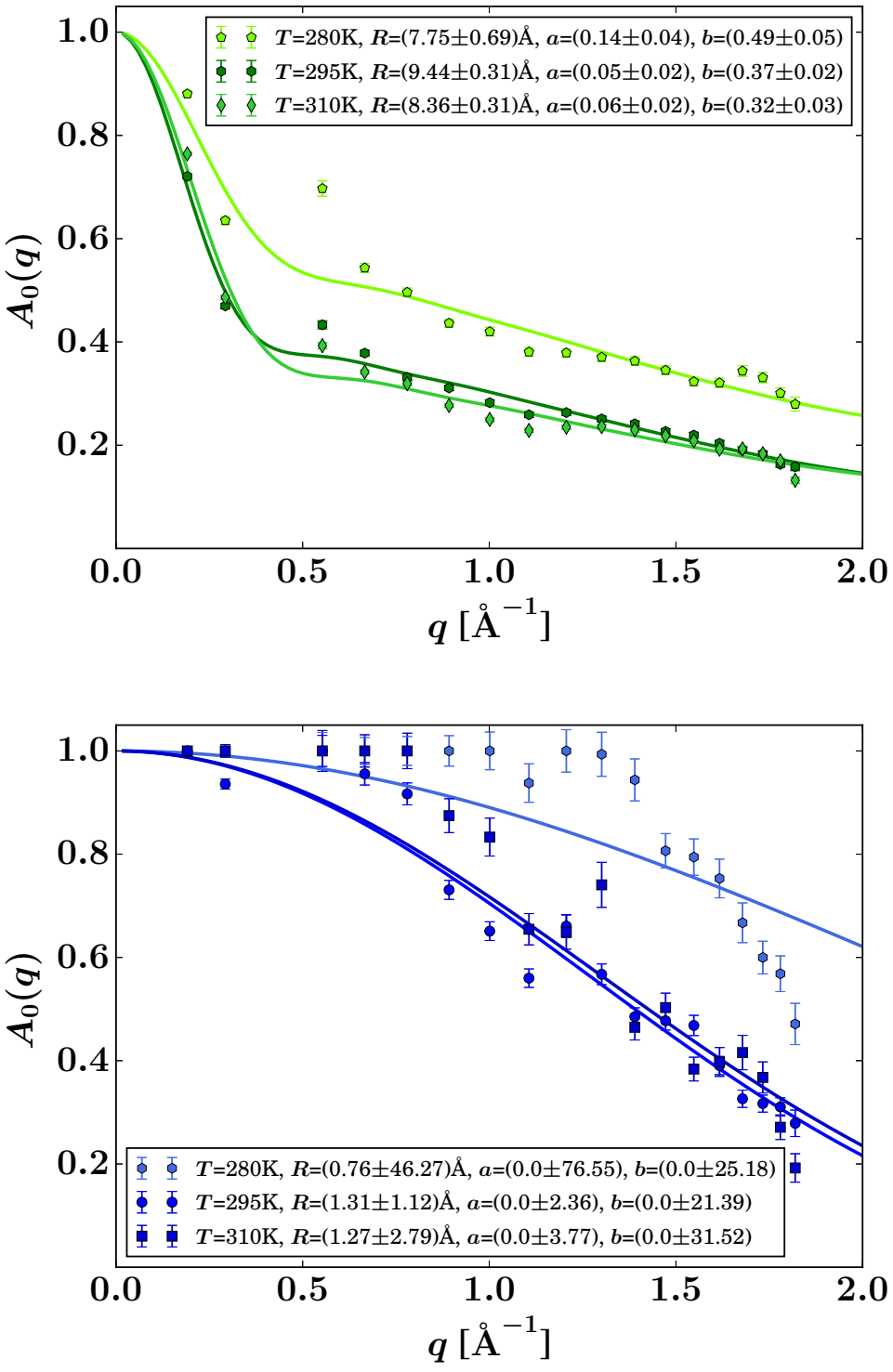

Figure 4.10: EISF $A_0(q)$ (symbols) for ADC-LIG (top) and ADC-APO (bottom) obtained from equation 4.3 when fixing Γ by using the IN5 result, and fits by equation 4.6 (solid lines). The resulting fit parameters are given in the legends.

liganded ADC (figure 4.7). An example spectrum and fit using equation 4.3 is shown in figure 4.9.

In equation 4.3, $A_0(q)$ can be identified with the elastic incoherent structure factor EISF [183] (figure 4.10). This EISF is modeled as follows:

$$\begin{aligned} A_0(q) = a + (1-a)\,(&bA_{3-\mathrm{jump}}(q,d) \\ &+ (1-b)A_{\mathrm{sphere}}(q)\ ) \end{aligned} \tag{4.6}$$

where, a represents the fraction of scatterers within the protein that are immobile (apart from the protein center-of-mass diffusion) on the observation time or coherence time of the measurement (on the order of a nanosecond for IN16B). $A_{3-\mathrm{jump}}(q,d)$ describes atoms undergoing reorientational jumps between three equivalent sites, which are associated with the methyl groups,

$$A_{3-\mathrm{jump}}(q,d) = \frac{1}{3}\left[1 + 2j_0(qd)\right]. \tag{4.7}$$

The jump length $d = 1.715\,\text{Å}$ is fixed for these methyl groups [174, 183, 184]. We further assume that the protein side chains diffuse on the surface of a sphere with the average radius R, such that

$$A_{\mathrm{sphere}} = \left|\frac{3j_1(qR)}{qR}\right|^2. \tag{4.8}$$

where, j_0 and j_1 are the spherical Bessel functions of zeroth and first order, respectively. b denotes the relative weight of the contributions from $A_{3-\mathrm{jump}}(q)$ and A_{sphere}. Results for the EISF $A_0(q)$ and fits using equation 4.6 are reported in figure 4.10 for both samples at the higher protein concentration $\varphi = 0.09$.

We note that the fit result for the EISF is strongly dependent on the fixing of the linewidth Γ in equation 4.3 using the fit results from IN5. In contrast, when not fixing the internal dynamics (not shown), a finite internal linewidth $\Gamma(q)$ in equation 4.3 can only be seen for the samples with the ligands. For the Apo-samples, $\Gamma(q) \to \infty$ in the fits of the IN16B spectra. At the same time, the errors on the internal diffusion fit parameters diverge. This could suggest very fast internal motions beyond the dynamic window accessible by IN16B.

Remarkably, the internal diffusive dynamics seem to change substantially depending on whether or not the ligand is present, although this it should be noted that this change may simply reflect the dynamics of the D-Serine itself. However,

independently from the assumptions on the internal dynamics, the fit results for the center-of-mass diffusion coefficient D appear rather robust. The systematic error due to assumptions on the internal dynamics is on the order of $\pm 0.5 \text{Å}^2/ns$, thus, larger than the fit parameter confidence bounds, but significantly smaller than the difference in the diffusion between the too samples (table 4.3).

4.1.4.3 Summary of the center-of-mass diffusion fit results

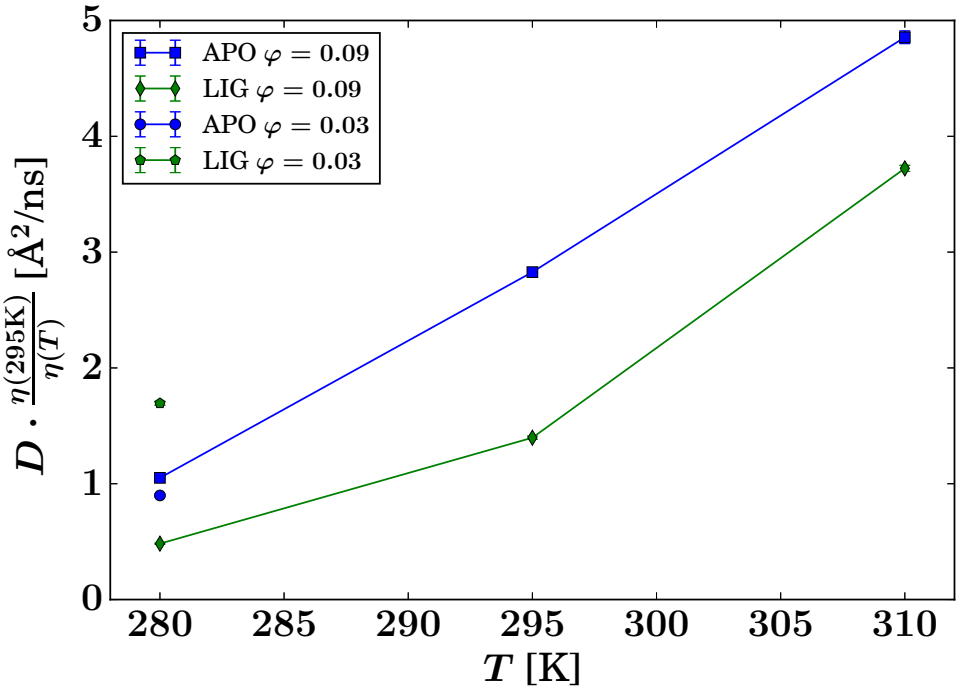

Figure 4.11: Observable apparent center-of-mass diffusion coefficients D (symbols) obtained from the IN16B spectra for all samples and protein volume fractions φ, as given in the legend, rescaled by the temperature-dependent solvent viscosity $\eta(T)$, versus sample temperature T. The lines are guides to the eye.

The global apparent center-of-mass diffusion D obtained from fitting equation 4.3 is expected to follow a Stokes-Einstein temperature dependence

$$D(T, \varphi) = \frac{k_b T}{6\pi\eta(T)R(T)} f(\varphi), \tag{4.9}$$

where η is the solvent viscosity and R the effective protein hydrodynamic radius. $f(\varphi)$ is a scalar function of the protein volume fraction φ and does not depend on T or R [81, 173, 185]. Thus, D directly reflects the average hydrodynamic size of the diffusing particle, which can be a protein monomer, oligomer, or cluster. Interestingly, the results in figure 4.11 indicate that the hydrodynamic size of the ADC tetramer or aggregate depends strongly on whether or not ligand is present. For ADC-APO, a perfect Stokes-Einstein dependence, equation 4.9 can be observed, as illustrated by the linear dependence on T, suggesting that the size of the ADC-APO assembly is constant within the observed temperature range. In contrast, for D-Serine liganded ADC, a larger assembly seems to be present, although there is evidence of partial dissociation at higher temperatures. Since the crystal structure indicates that ligand binding does not significantly alter the hydrodynamic size of the ADC tetramer (table 4.2), the difference between apo-ADC and liganded ADC samples is best explained by the formation of a higher order protein oliogmer in solution in the presence of D-Serine [81, 186].

4.1.5 Dynamic Light Scattering

In order to further investigate this higher order oligmer formation in the presence of ligand, and to ensure that it is not simply a concentration artifact, dynamic light scattering (DLS) measurements were carried out. DLS accesses the collective diffusion of relatively dilute samples as opposed to the self-diffusion in concentrated samples accessed by spatially incoherent QENS. It also observes substantially longer diffusive time scales on the order of milliseconds as opposed to the nanosecond diffusive short-time regime explored by QENS. Due to the low momentum transfers, DLS in general only accesses the translational part D_t of the diffusion coefficient in the case of proteins. Using DLS, a time autocorrelation function was measured over the angular range 30-150 ° (figure 4.12). The correlation function for a monodisperse sample is given as

$$g^1(q,t) = a\exp(-2\Gamma t). \tag{4.10}$$

In case of more than one populations of clusters with different diffusion coefficients, the following equation is used as the fit function:

$$\begin{aligned} g^1(q,t) = {} & a_1\exp(-2\Gamma_1 t) + a_2\exp(-2\Gamma_2 t) \\ & + a_3\exp(-2\Gamma_3 t) + ..., \end{aligned} \tag{4.11}$$

where Γ is the decay rate and t the time. The first order autocorrelation function was treated as a monoexponential decay (equation 4.10) in order to extract the decay rate.

Γ can then be plotted versus q^2 which follows Fickian diffusion (cf. equation 4.4),

$$2\Gamma = D_t q^2, \tag{4.12}$$

where D_t is the diffusion coefficient and q is the momentum transfer defined as

$$q = \frac{4\pi n_0}{\lambda} \sin\left(\frac{\theta}{2}\right). \tag{4.13}$$

where θ is the scattering angle, n_0 the refractive index of the sample, and λ the wavelength of the incident beam. Due to the linear relation of Γ and q^2 (equation 4.12) a linear fit of the q-dependence gives the long-time translational diffusion coefficient D_t (figure 4.13).

By fitting the diffusion coefficients for all five measured concentrations for both the samples apo-ADC and D-Serine liganded ADC, average diffusion coefficients were obtained for for both samples at the low-concentration limit (figure 4.14, symbols at $\varphi_t \approx 0$). These diffusion coefficients of 4.05 ± 0.02 and 4.30 ± 0.03 Å^2/ns for liganded ADC and apo-ADC, respectively indicate that the two samples behave significantly differently and that the nearly dilute-limit ($\varphi_t \approx 0$) diffusion coefficient is higher for apo-ADC than that for liganded ADC. This result follows the same trend observed in the QENS results (symbols at $\varphi_t >> 0$ in figure 4.14), indicating that the presence of ligand does indeed drive formation of higher order oligomers.

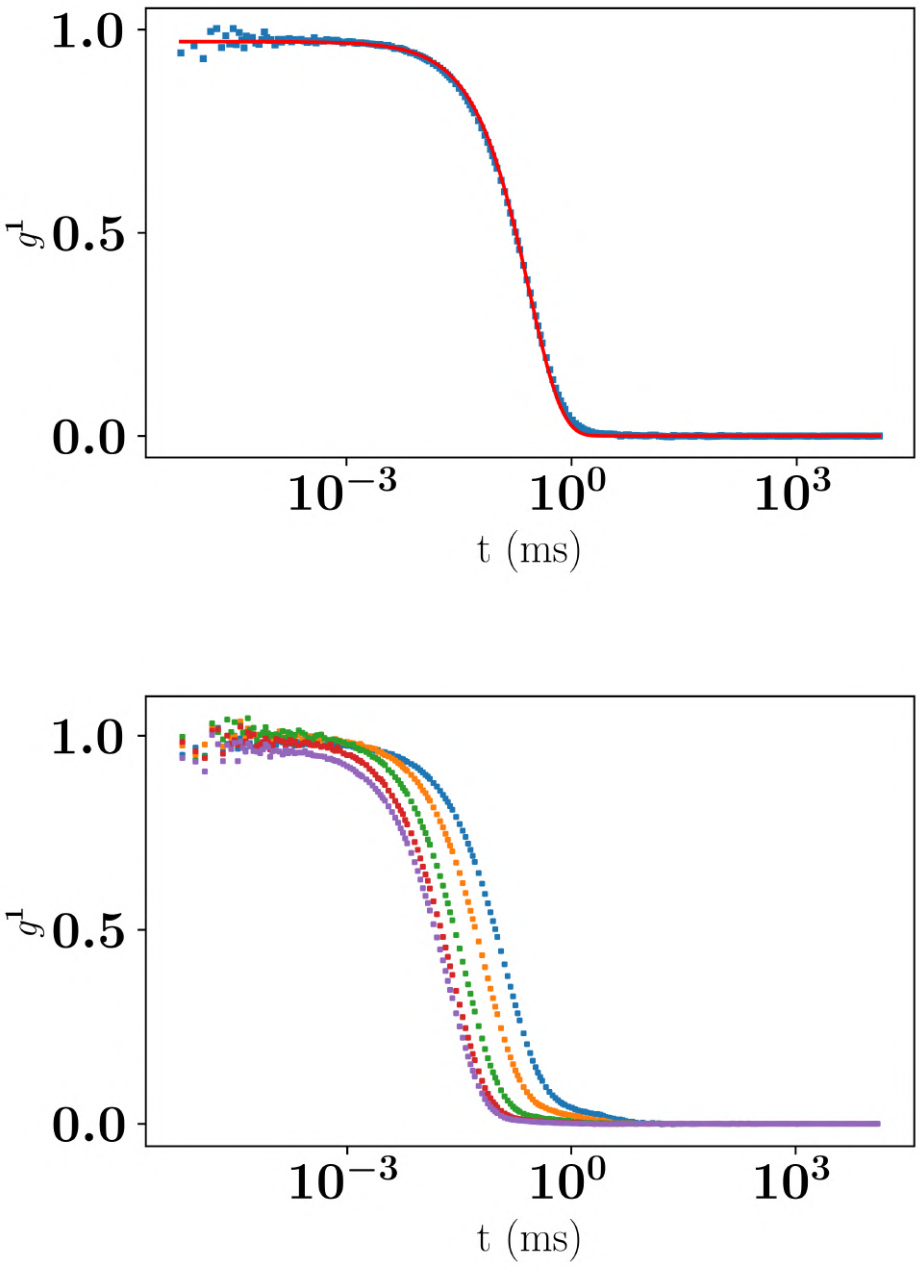

Figure 4.12: Top: DLS data recorded at $\Theta = 120°$ and single-exponential fit (equation 4.10) of the autocorrelation function g^{-1} for D-Serine liganded ADC at $c_p = 15\,\mathrm{mg/mL}$ and $T = 298\,\mathrm{K}$ versus correlation time t. Bottom: DLS data recorded at 70 (violet), 80 (red), 100 (green), 120 (orange), 140 (blue) ° for the same sample. Symbols represent the data whereas the solid line represents the fit.

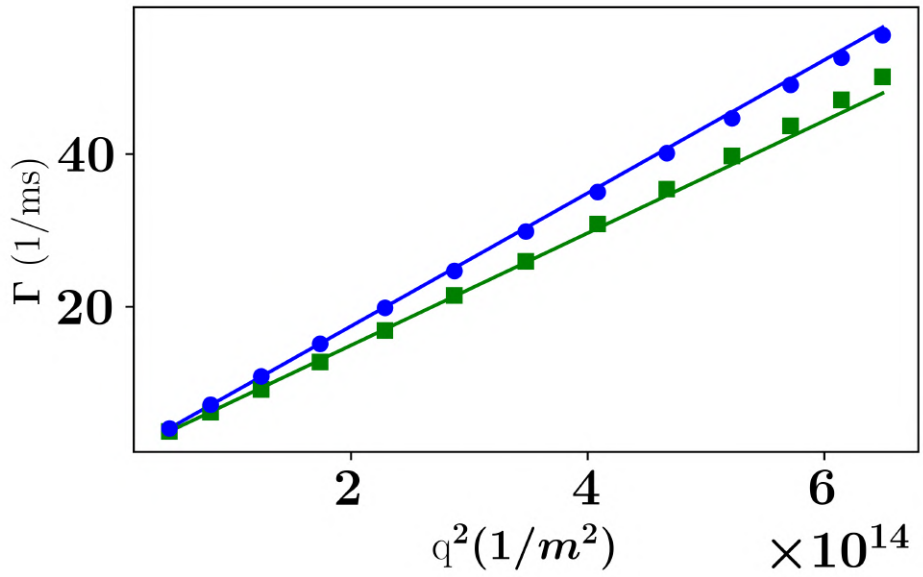

Figure 4.13: Linear fit of the decay rate obtained from the DLS data for D-Serine liganded ADC (ADC-LIG) (squares) and ADC-APO (circles) at 15 mg/mL

4.1.6 Comparison of experimentally observed and theoretical diffusion

4.1.6.1 Colloid models and separation of rotational and translational diffusion coefficient

The observable apparent diffusion coefficient $D = D(D_r, D_t)$ obtained by fitting equations 4.3 and 4.4 is an implicit function of the rotational $D_r = D_r(\varphi)$ and translational $D_t = D_t(\varphi)$ diffusion coefficients of the diffusing protein assemblies [81, 172–174, 186]. The dependence of $D_t = D_t(\varphi)$ and $D_r = D_r(\varphi)$ on the effective hydrodynamic volume fraction φ_t of hard spheres can be calculated using models from colloid physics for the short-time self-diffusion [185]. The validity of the approximation of proteins as hard spheres on the short-time scale, on which proteins diffusive by less than their radius and on which protein-protein collisions are negligible, has been shown for other proteins [81,172–174,186]. Moreover, $D = D(D_r, D_t)$ observable with QENS can be calculated using established frameworks [172,173]. In this way, $D_t(\varphi_t) = D_{r,0} f_t(\varphi_t)$ and $D(\varphi_t) = D_0 f(\varphi_t)$ of colloidal hard spheres can be calculated for the two studied systems, namely ADC-APO and ADC-LIG, where f and f_t are scalar functions being determined by the colloid models [185] and solely depending on the theoretical effective hydrodynamic volume fraction φ_t. The translational and rotational diffusion coefficients at infinite dilution, $D_{t,0} = D_t(\varphi_t = 0)$ and $D_{r,0} = D_r(\varphi_t = 0)$, and, thus, $D_0 = (D_{t,0}, D_{r,0})$ can be obtained using the software HYDROPRO as explained in section 4.1.7. Taking the above steps together, both $D(\varphi_t)$ and $D_t(\varphi_t)$ may be obtained for both ADC-APO and ADC-LIG (lines in figure 4.14).

4.1.6.2 Experimental results in the context of theory

For the experimental results for the diffusion coefficients from QENS (section 4.1.4), the effective hydrodynamic volume fraction can be calculated as [173]

$$\varphi_t = \varphi \left(\frac{R_h}{R} \right)^3 . \tag{4.14}$$

The theoretical hydrodynamic volume fraction φ_t thus obtained is used as the X-axis in figure 4.14.

4.1.7 Theoretical diffusion coefficient at infinite dilution

For comparison with the experimental data, we calculated the theoretical diffusion coefficients at infinite dilution from the structures of ADC-APO (PDBID:1AW8) and

Table 4.2: HYDROPRO output for ADC-APO (PDBID 1AW8) and ADC-LIG (ref. section: 4.1.1) structures for 280, 295 and 310K temperatures

		ADC-APO			ADC-LIG	
PDB file (tetramer)	1AW8	1AW8	1AW8	Ref. section: 4.1.1	Ref. section: 4.1.1	Ref. section: 4.1.1
Radius of elements in primary model (Å)	2.90	2.90	2.90	2.90	2.90	2.9
Temperature (K)	280	295	310	280	295	310
Solvent viscosity (pure D2O) (poise)	0.01830	0.01175	0.00830	0.01976	0.01270	0.00896
Molecular weight (Da)	50391	50391	50391	50811	50811	50811
Solute partial specific volume (cmˆ3/g)	0.702	0.702	0.702	0.702	0.702	0.702
Solution density (g/cmˆ3)	1.110	1.110	1.110	1.110	1.110	1.110
Translational diffusion coefficient ($D_t(\varphi=0)$) (cmˆ2/s)	3.784E-07	6.209E-07	9.236E-07	3.163E-07	5.188E-07	7.721E-07
Radius of gyration (cm)	2.257E-07	2.257E-07	2.257E-07	2.446E-07	2.446E-07	2.446E-07
Volume (cmˆ3)	8.235E-20	8.235E-20	8.235E-20	9.476E-20	9.476E-20	9.476E-20
Rotational diffusion coefficient ($D_r(\varphi=0)$) (sˆ-1)	3.174E+06	5.208E+06	7.747E+06	2.124E+06	3.484E+06	5.185E+06

of the ADC-LIG complex (section 4.1.1) using HYDROPRO [187]. Table 4.2 summarizes the output from HYDROPRO based on the input of the structures explained in section 4.1.1. To account for the actual viscosities of the deuterated solvent as input to HYDROPRO, the following assumptions were made: we considered the contribution from the addition of D-Serine to the viscosity and, hence, to the overall calculation of the diffusion coefficient. In general, the viscosity depends on both the ionic concentration as well as on the temperature [188]. The change in the viscosity upon addition of L-Serine has been studied previously [188–191]. The temperature strongly influences the viscosity [188, 190, 191]. In contrast, the change in viscosity with an increase in the concentration of D-Serine is approximately 5-8% when the concentration is increased from 0 to 3.6 M [191]. To account for this increase in viscosity in the calculation of the theoretical diffusion coefficient calculated from HYDROPRO, we increased the original viscosity for deuterium oxide (D_2O) by 8% and subsequently re-calculated the theoretical diffusion coefficient (Table 4.2). It has been found in other studies that due to the presence of a polar -OH group and the absence of a neutral alkyl group, D-Serine has a higher hydration number [191] and hence has a lower viscosity compared to the relatively hydrophobic amino acids such as Threonine or Alanine which are similar in terms of bulkiness of the side chain functional groups [188, 191].

4.1.8 Comparison of experimental and calculated diffusion coefficient

Figure 4.14 displays the comparison between the experimental and calculated diffusion coefficients. Both the translational diffusion coefficients D_t obtained from DLS at $\varphi_t \approx 0$ and the apparent diffusion coefficients $D = D(D_r, D_t)$ from QENS at $\varphi_t >> 0$ are depicted as symbols in figure 4.14. The different values for φ_t for the QENS results follow directly from equation 4.14 due to the different R_h resulting from the different $D_t(\varphi = 0)$ according to table 4.2 and equation 4.9. The translational diffusion coefficients calculated from HYDROPRO for ADC-APO and ADC-LIG are 3.78 and 3.42Å^2/ns at 280K and 6.20 and 5.62Å^2/ns at 295K, cf. table 4.2. These diffusion coefficients from HYDROPRO appear as intersections of the dashed lines with $\varphi_t = 0$ in figure 4.14. They differ from the DLS results but follow the same trend. The dashed lines in figure 4.14 represent the theoretical translational self-diffusion coefficients in the diffusive short-time limit – assuming ADC-tetramers – and are calculated using the HYDROPRO results (table 4.2) and PDB structures, section 4.1.1, employing established protocols based on colloid models [81, 173, 185]. (Note that for $\varphi = 0$, the short-time and long-time limits as well as collective and self-diffusion are identical, since the distinctions arise solely from multiple-particle effects.)

The solid lines in figure 4.14 represent the theoretical apparent short-time self diffusion coefficients $D = D(D_t, D_r)$ – assuming ADC-tetramers - observable in QENS, which are calculated from D_t using further established protocols [81, 172, 173, 185], involving notably the radial hydrogen density distribution function determined from the pdb structure. We note that we have assumed spherical proteins for our calculations, neglecting possible effects of a more complicated shape [192].

The lower experimental values from both DLS and QENS compared to the respective theoretical expectations for ADC tetramers indicates that the hydrodynamic size of the diffusing objects is larger than tetramers. This observation suggests a formation of ADC clusters in solution both with and without D-Serine. We emphasize that these clusters might be transient. We note that inaccuracies may arise from uncertainties in the determination of the protein volume fraction φ in all samples.

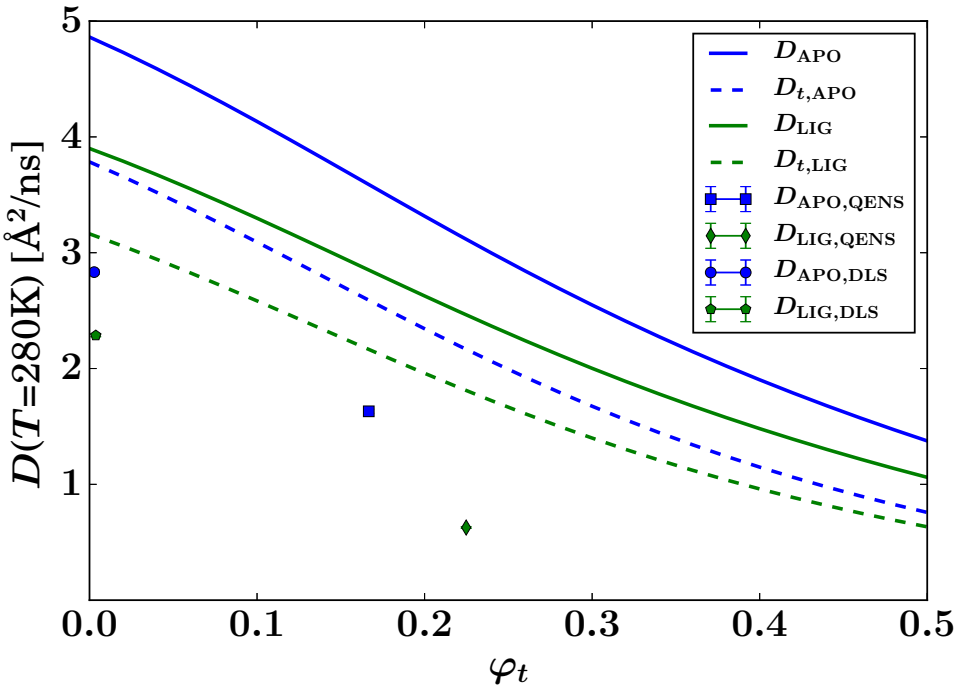

Figure 4.14: Summary of the QENS and DLS results for the global center-of-mass diffusion coefficient D of the proteins versus the theoretical volume fraction φ_t of effective hard spheres, at $T = 280\,\mathrm{K}$. For DLS, the symbols (at $\varphi \approx 0$) represent the translational diffusion D_t probed by this method. For QENS, the symbols represent the apparent short-time self-diffusion consisting of contributions from both rotations and translations. The solid lines report the theoretical apparent diffusion of effective spheres with the hydrodynamic size of ADC-APO and ADC-LIG tetramers, respectively. The dashed lines represent the corresponding translational diffusion (cf. legend).

Table 4.3: Diffusion coefficients obtained from QENS, DLS and HYDROPRO at 295K in Å^2/ns. The errors denote 67% confidence bounds on the fits and do not account for systematic errors arising from the choice of the model.

	QENS $D(\varphi$=0.09)	HYDROPRO $D_t(\varphi$=0)	DLS $D_t(\varphi \approx 0)$
ADC-APO	2.82±0.03	6.21	4.30±0.03
ADC-LIG	1.39±0.02	5.19	4.05±0.02

4.2 Discussion

Previous work on diffusive dynamics in protein solution has focused on standard samples that are commercially available. Here, for the first time, the effect of ligand binding on the diffusive dynamics of a recombinantly expressed non-standard protein sample is studied, combining results from X-ray crystallography, quasi-elastic neutron scattering, dynamic light scattering, and MD simulations. A key challenge for this type of study is discerning the protein internal diffusive dynamics from those of the ligand, as in order to ensure the saturation of the protein binding sites a large excess of the ligand must be present in solution. Therefore, although the elastic incoherent structure factor seems to undergo a qualitative change upon ligand binding, this apparent change still has to be determined with higher accuracy.

We will conduct isothermal titration calorimetry (ITC) experiments in order to calculate the dissociation constant for the D-Serine ligand which will reveal the affinity of D-Serine for ADC. We will then know the amount of free D-Serine after all the binding sites of ADC are saturated. This quantity will be useful in discerning the signal from ADC and that from free D-Serine, enabling the modeling of the internal diffusive motions of ADC in ligand bound form.

Nevertheless, by combining the QENS data with information from the protein structures, HYDROPRO calculations based on these structures, calculations of the radial hydrogen density distribution functions based on the structures, and theoretical predictions of the short-time self-diffusion of colloidal hard spheres an interpretation of the measured center-of-mass diffusion of the ADC protein in aqueous suspension in the presence and absence of the D-Serine ligand is possible.

The diffusion coefficients obtained from QENS and DLS follow the same trend as the those predicted from calculations, but differ quantitatively. The calculations were made based on the assumption of that the ADC tetramer is the protein assembly present in solution. However, the observed deviation provides strong evidence that larger objects than the ADC tetramers determine the diffusion. Further work will be needed to determine how stable these larger assemblies are, as the time-scales accessed in this study are milli-second (DLS) and shorter (QENS), and how relevant the oligomers are to the biological function of this enzyme. Gel filtration studies show that the dominant species in solution is the ADC tetramer [52], suggesting that the higher order species formed here are transient.

It is clear that ligand binding affects both the nature of these transient oligomers, as well as the internal protein dynamics of ADC. In general, ADC appears rather stiff on the pico- to nanosecond time scale, consistent with the crystallogaphic data (resolution and Wilson B factor). Binding of D-Serine to ADC causes a change in

the conformation of the C-terminal loop (figure 4.1), which is observed in the crystal structure, MD simulations and, indirectly via a change in the dynamics, in the QENS data. Specifically, the results from MD simulations suggest that the conformational dynamics of individual residues of the loop are different for ADC-LIG in comparison with ADC-APO forms (figure 4.5). We also observe that the main chain conformation is different in APO and D-Serine complexed states (figure 4.3).

Other studies have suggested that the more dynamical a region is in the protein, the more influence it has on the propensity of the protein to aggregate as a result of unfavorable entropic terms [193, 194]. Our studies therefore are important to re-emphasize that that small structural changes in highly dynamic regions of the protein, such as the C-terminal loop in ADC, can potentially cause larger changes in the aggregation dynamics of the protein which can be biologically important.

Chapter 5

Inelastic neutron scattering to understand the dynamics of aspartate α-decarboxylase with and without the ligand in amorphous state

We attempted to cover a broad range of dynamical modes of ADC in presence and absence of D-Serine. For this purpose, we used inelastic time-of-flight (IN5) spectrometer and neutron vibrational spectrometer IN1-Lagrange at Institut Laue-Langevin (ILL), Grenoble. These were complemented with elastic incoherent neutron scattering experiments in order to understand the difference between the dynamics of these two systems in powdered form. These experiments were also important because in case of classical solution QENS studies, we could clearly see the difference in the global dynamics of these systems (ADC-APO and ADC-LIG).

5.1 Deep inelastic neutron scattering experiments on apo and D-Serine complexed aspartate α-decarboxylase

We attempted to capture the vibrational modes of ADC in presence and absence of the ligand (D-Serine) so as to understand whether the binding of D-Serine also induces changes in the global vibrational dynamics of different groups. For this purpose, we used hydrated powders of ADC-APO, ADC-LIG and D-Serine alone. The powders were hydrated to $\approx$ 40 % as this level of hydration is reported to have been essential to retain the biological activity of the protein [195]. The spectra were measured at different temperatures (50-300K) on the neutron vibrational spectrometer IN1-Lagrange and IN5 time-of-flight spectrometer so as to assess if and how the vibrational dynamics changes after dynamical transition. The major part of the signal comes from the vibrations associated with hydrogens. Figure 5.1, shows the example IN1 spectrum for ADC-APO and ADC-LIG samples at 50K. We have attempted to assign the vibrational modes to the observed peaks based on those for the IR spectra.

It can be clearly seen that overall, the peaks are sharper for ADC-LIG than those for ADC-APO sample. A peak can be seen at 3000 cm^{-1} which is usually attributed to the C-H stretching mode. In case of ADC-LIG the vibrational peak for C-H bending is shifted to a slightly lower frequency, indicating a possible slowing down of the C-H bending motion as a result of binding of D-Serine. We also observe additional peaks close to the C-H bending mode for the ADC-LIG sample. The additional peaks may have originated due to the vibrations of ADC-LIG being more ordered than those of ADC-APO sample which also agrees with the results from solution QENS studies. Figure 5.2 shows the generalized density of states [196] from IN5 (left) and IN1 vibrational spectra (right) of ADC-APO (blue) and ADC-LIG (green) at 100, 200 and 300 K. The spectra indicate a gradual loss of the sharpness of the peaks indicating onset of vibrational dynamics, especially at 300K. In general, the spectral features at $\approx$ 200 cm^{-1} which consistently differ for ADC-APO and ADC-LIG, maybe due to intra-sheet hydrogen bonding [197]. The region between 1200 and 1700 cm^{-1} likely contains information on a combination of bending, twisting and wagging motions of -CH, -CH_2 and -CH_3 functional groups [198] along with amide I, II and III bands [199]. Here, amide I band mainly denotes the stretching vibrations of C=O ($\approx$80 %) and C-N groups ($\approx$10-20 %) and is typically found between the frequency range of 1600 and 1700 $cm^{-}1$ [200].

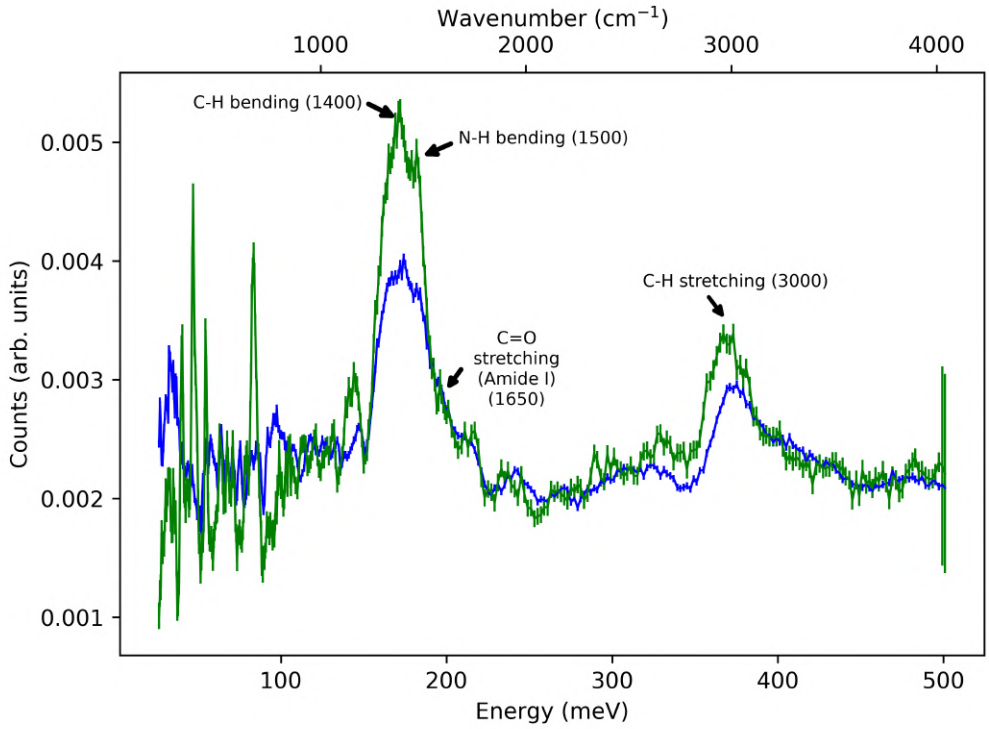

Figure 5.1: Assignment of frequencies to various vibrational modes of ADC-APO (blue) and ADC-LIG (green) samples at 50K for the data measured on IN1 instrument. The sample used is hydrogenated ADC hydrated with D_2O to $\approx$ 40 %. The signal was normalised to monitor counts and then the empty can signal was substracted. The normalization for the two samples was carried out by dividing the spectrum of each of the samples by the total spectral intensity.
Numbers in the parenthesis are vibrational frequencies of the specific functional groups as shown in the figure

As seen in figure 5.1, the Amide I band is seen as a small bump around 1650 cm^{-1}. The low intensity for this mode is probably because a majority of the hydrogens in N-H groups on the surface of the protein are exchanged with the deuterium from D_2O solvent. The shoulder maybe because of 'a riding mode' involving hydrogen atoms which are riding on another atom. Amide II region on the other hand denotes the N-H bending and is seen as an intense peak around 1500 cm^{-1} in figure 5.1.

It can also be seen that the fine splitting of the peak which corresponds to C-H bending (1400 cm^{-1}) and N-H bending (1500 cm^{-1}) is distinguishable at 100 K (figure 5.2). At 200 K, it is hardly distinguishable whereas at 300K, the two peaks merge into one making these two vibrational frequencies indiscernible (figure 5.2). This observation mainly applies to ADC-LIG, indicating that the binding of D-Serine influences the C-H bending modes. It should be noted that it is difficult to understand the reasons for finer changes in the peaks of these spectra since they originate from superposition of all the functional groups, vibrating at the corresponding frequencies.

We also recorded spectra at 100, 200 and 300 K on IN5 and at 50 K on IN1 on D-Serine powder (figure 5.3). In case of IN5, generalized density of states is calculated using the following equation [196]:

$$g_{exp}(Q,\omega) = \frac{\omega}{Q^2} S_{\exp}(Q,\omega) \left[1 - \exp\left(-\frac{\omega}{k_B T}\right)\right] \tag{5.1}$$

where, ω is the frequency, Q is the momentum transfer, k_B is the Boltzmann constant, T is the temperature in Kelvin and $S_{exp}(Q,\omega)$ is the powder average of the product of the scattering function and the Debye-Waller factor. For a detailed explanation on GDOS, please refer to reference [196] and for the fundamental concepts, please refer to the book [87].

The left panel shows the GDOS obtained for D-Serine at 100-300 K. We observed sharp vibrational signatures arising from crystalline order in the GDOS calculated from IN5 spectra of D-Serine. This observation is consistent even in the case of GDOS for ADC-LIG samples (figure 5.2). These signatures become weaker as the temperature is increased. These are consistently present at all the temperatures including 300 K which means that D-Serine crystallites are present at least upto room temperature. In the case of ADC-LIG sample, there is always excess D-Serine present and this results in the formation of crystallites which is possibly the cause of these sharp features of the spectrum. This interpretation is also consistent with the presence of sharp features in the case of spectra recorded on pure D-Serine (without ADC) (figure 5.3) and with their absence in the case of ADC-APO sample.

Our vibrational data recorded on D-Serine can be compared with the previous reports [201]. Our vibrational spectra were found to be very similar to those reported

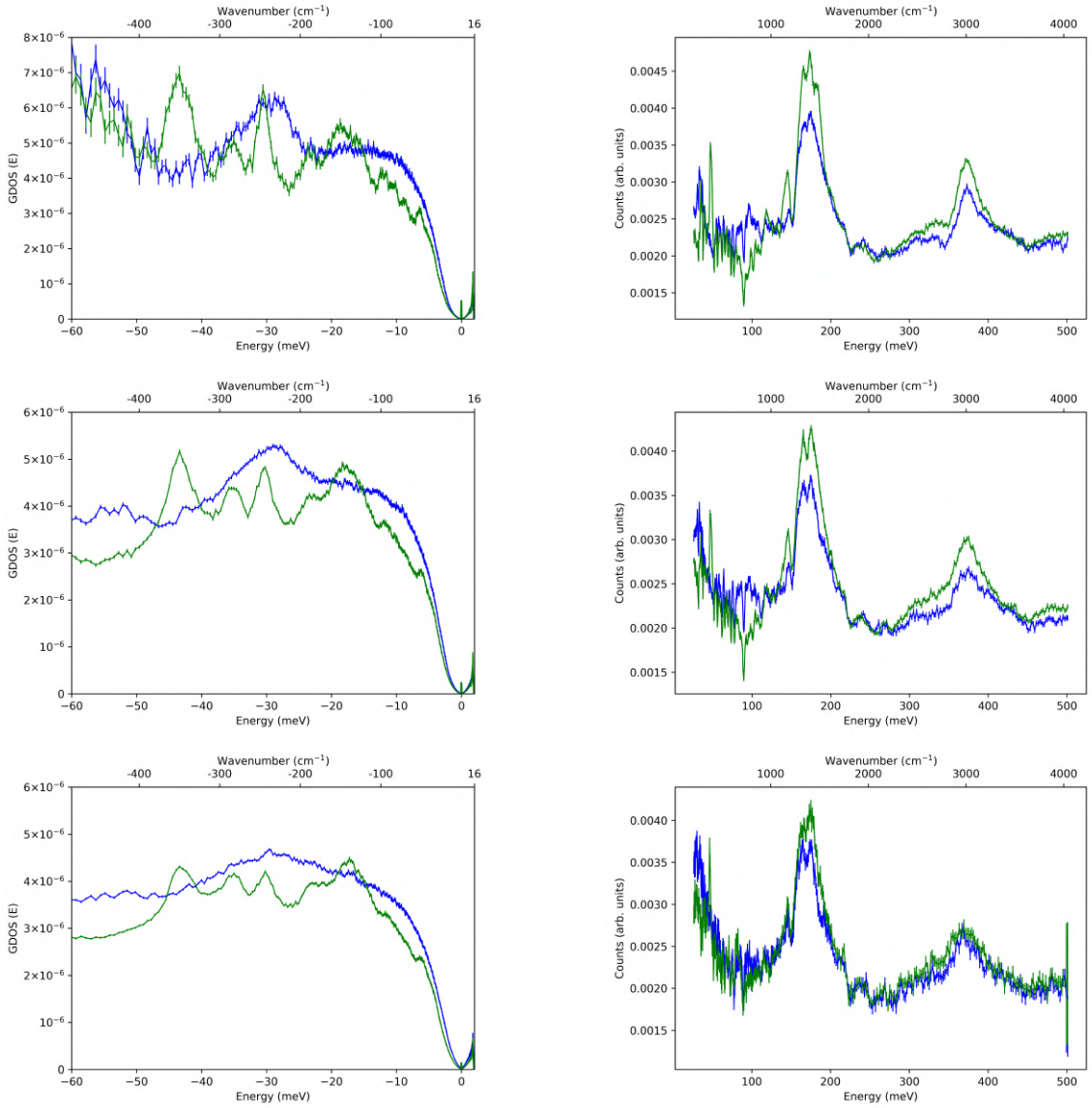

Figure 5.2: Generalized density of states (GDOS) (left) calculated from 5 Å IN5 data and IN1 vibrational spectra (right) for ADC-APO (blue) and ADC-LIG (green) at 100 (top), 200 (middle), 280 (bottom right) and 300 K (bottom left). The negative energies in case of IN5 GDOS signify the neutron energy gain side i.e. the X-axis is in terms of the energy of the scattering nuclei. The signal is empty can subtracted and normalized. The normalization was carried out by dividing the spectrum of a sample by the total spectral intensity of that sample.

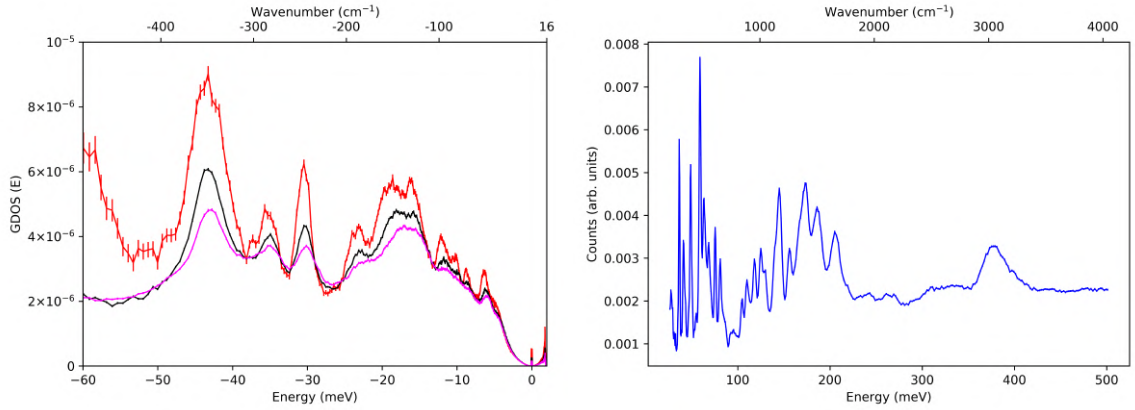

Figure 5.3: IN5 generalized density of states (left) at 100 (red), 200 (black) and 300 (magenta) K; IN1 vibrational spectrum (right) for D-Serine. The signal is empty can subtracted and normalized. The normalization was carried out by dividing the spectrum for the corresponding sample by the total spectral intensity.

for the 'wet' state of the D-Serine as we could see a sharp and intense peak at 80 meV which is also reported in the aforementioned study, hence indicating D_2O hydrated D-Serine in our case as expected. The peak positions for pure D-Serine were also similar to those observed in this work. [201]

5.2 Elastic/inelastic fixed window scans (E/IFWS) of apo and D-Serine bound ADC in powdered state

In order to access the changes in the dynamics of ADC after binding of D-Serine, we performed elastic and inelastic incoherent fixed window neutron scattering with an energy resolution higher than 1 μeV using IN16B. We used hydrated ADC powders in these experiments and the samples used were identical to those used for IN1 vibrational spectroscopy measurements. The prevailing incoherent signal in these experiments is dominated by the hydrogen atoms of the protein, thus allowing the dynamics of ADC in D-Serine bound and APO forms to be monitored. Unlike inelastic neutron scattering, we observe the changes in terms of internal motions of ADC wherein all the modes (backbone, side chain and relaxation motions) are superimposed. EFWS accesses an apparent thermal Debye-Waller or Lamb-Moessbauer factor through the

q- and temperature dependent loss of elastic scattering intensity arising from both diffusive and vibrational motions [84, 107, 202].

Figure 5.4 shows the elastic incoherent neutron scattering intensity (EINS) as a function of temperature and momentum transfer (q). It can be seen that there is a difference in the change in the intensity with respect to both q as well as temperature. In case of ADC-APO, the change in intensity is gradual whereas in case of the ADC-LIG complex, a drastic change in intensity is observed between 200 and 300 K. A similar trend can be seen for the change in intensity with respect to momentum transfer (q) with a sudden increase in intensity observed within the q-range of 1.0-0.8. This indicates a change in protein dynamics on a relatively lower lengthscales. This trend indicates that in its apo state, ADC shows higher plasticity than in D-Serine complexed state. The drastic reduction in the intensity at $\approx$ 250 K in case of ADC-LIG complex indicates relatively rigid or ordered internal motions below 250 K.

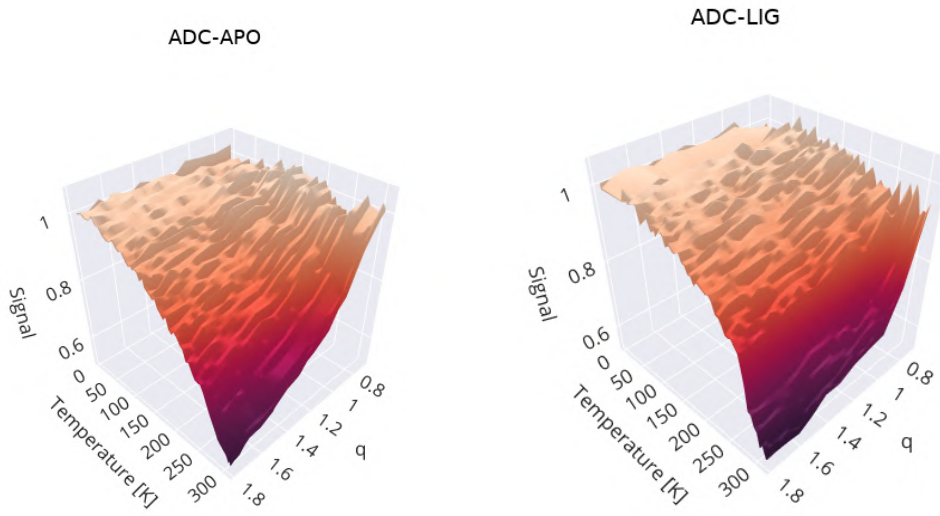

Figure 5.4: Intensity vs. momentum transfer (q) vs. temperature (K) for ADC-APO (left) and ADC-LIG (right) obtained from EFWS recorded on IN16B.

In order to further assess the change in softness of ADC upon binding of D-Serine, we obtained vibrational mean squared displacement from EFWS. The mean squared displacements were obtained by fitting the angular dependence of the elastic incoherent neutron intensity in the Gaussian approximation at each of the temperatures (subsubsection 2.2.4 in chapter 3) [202]. In general, there are two dynamical

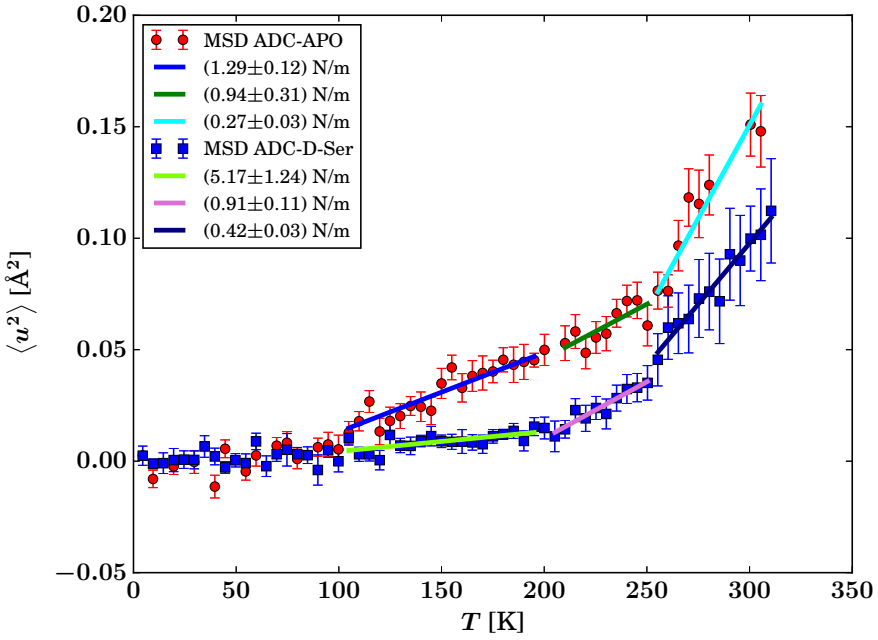

Figure 5.5: Apparent vibrational mean squared displacement (Å^2) vs. temperature (K) derived from the elastic fixed window scans for ADC-APO (red) and ADC-LIG (blue) samples

transitions which can be seen at 200 and 250 K as seen from the deviation with respect to the straight line. At 250 K, the mean squared displacements are approximately twice for ADC-APO than that for ADC-LIG. This difference increases with increasing temperature. We obtained the effective force constants from the slopes of different parts of the plot [202] and is lower for ADC-APO between 50-200 K and between 250-330 K. Hence, we conclude that the binding of D-Serine hardens ADC, causing the overall dynamics of ADC to be more confined for the ADC-LIG complex.

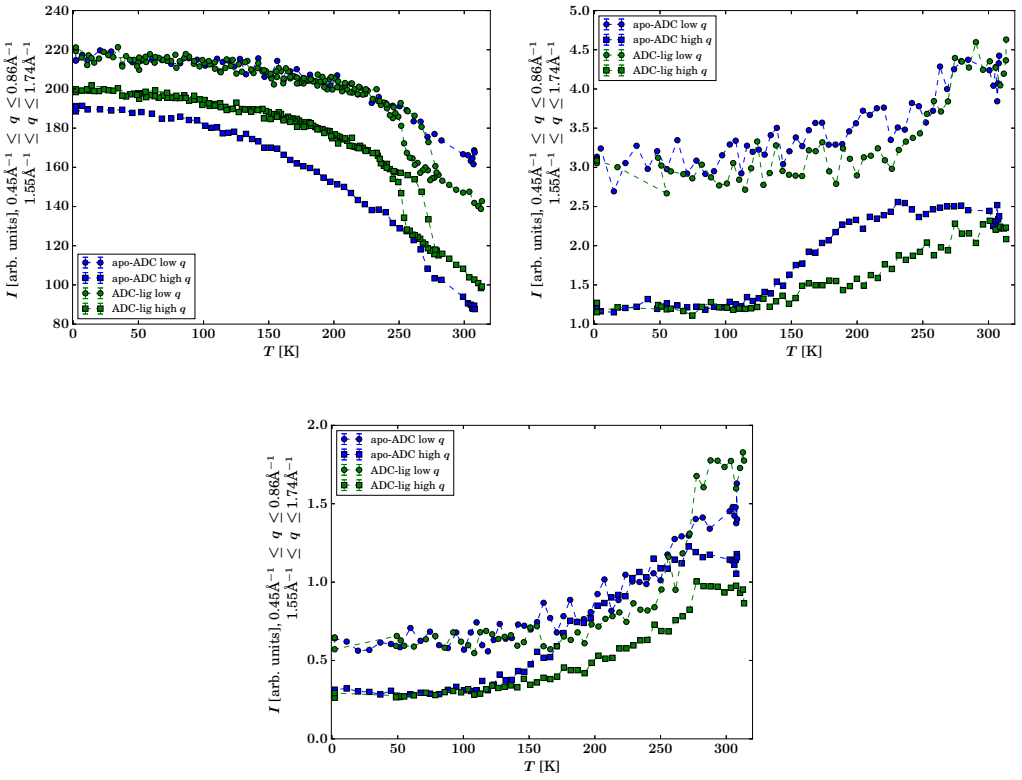

Figure 5.6: Intensity vs. temperature (K) from elastic (top left) and inelastic (1.3 μeV offset, top right and 2.6 μeV offset, bottom) fixed window scans for ADC-APO (blue) and ADC-LIG (green) samples. In each plot, integrated intensities for two different q-ranges, as specified by the y-axis label, are given (square and circle symbols, respectively).

We also collected the inelastic fixed window scans at 1.3 (top right) and 2.6 μeV (bottom) offsets (figure 5.6). A clear difference between ADC-APO and ADC-LIG

complex samples can be seen especially above 150 K. The detailed analysis of IFWS will be followed up in future. The THz time domain spectroscopy experiments are pending which will be used to probe the complementary solvent dynamics of ordered and bulk solvent.

5.3 Discussion

Vibrational spectroscopy has been applied to protein samples to a relatively less extent in the past. This is mainly because of the challenges associated with evaluating the spectral contribution from various secondary structural elements [203]. Hence, we are limited to assessing changes in the global properties of the system under consideration. However, the vibrational data can yield some very important aspects when complemented by other techniques such as THz-TDS spectroscopy.

In the preliminary results presented here, we have attempted to assess the global changes in vibrational spectrum of ADC associated with binding of the ligand (D-Serine) using a combination of inelastic neutron spectroscopy and elastic/inelastic fixed window scan experiments. We were able to assign the peaks from neutron vibrational spectra to the specific vibrational modes, based on the infrared spectroscopy database by the commercial company Merck and another study similar to the present work [197]. We also collected IN5 time-of-flight data on infinite neutron energy gain side and finite ($\approx$ 1 meV) neutron energy loss side. We could observe vibrational signatures corresponding to crystalline order of excess D-Serine crystallites. These vibrational signatures were observed in the generalized density of states (GDOS) of pure D-Serine and ADC-LIG complex but not in ADC-APO. Although the intensities of these peaks reduced with increasing temperature, it was surprising that the peaks were present at all of these temperatures as this corresponds to a high stability of these crystallites. This observation implies that the free D-Serine which is left after all the binding sites of ADC are saturated, probably tends to form these crystallites which are stable even at temperatures as high as 300K (figure 5.2).

In order to corroborate our findings and to further assess the temperature dependence of atomic mean squared displacements, we carried out elastic (EFWS) and inelastic fixed window scan (IFWS) experiments which enabled us to access the onset of motions faster than $\approx$ 4 ns corresponding to an instrumental resolution of 0.75-1.00 μeV. mean squared displacements from both the samples followed a similar trend. In general, ADC was found to be more rigid than the other proteins which have been studied using EFWS. Both the dynamical transitions for ADC are above 180 K whereas for hydrated apoferritin which is also a rigid protein, the first

dynamical transition is observed at $\approx$ 120 K [204]. First dynamical transition for Ribonuclease A from Bovine is also observed at a similar temperature [205]. In spite these common features for the two ADC samples, clear differences were observed in terms of the 'softness' of the two systems.

We observe that the change in intensity with both temparature and momentum transfer indicate a more rigid ADC-LIG sample. Slopes of different parts of mean squared displacements indicate the 'force constants' which quantify the 'molecular resilience' of a protein molecule [202]. In this context, we observe higher force constants for ADC-LIG than those for ADC-APO which is indicative of a harder ADC-LIG sample in comparison with ADC-APO. These observations are consistent with a more confined motion of the C-terminal loop of ADC-LIG than that of ADC-APO as seen from the distance between HIS-21 and GLY-24 obtained from MD simulations (figure 4.3). The change in vibrational neutron intensities of the two samples observed at different temperatures indicating the change in softness being an effect of the binding of D-Serine. In order to observe the change in dynamics of hydration and bulk water upon binding of D-Serine, we will perform THz time domain spectroscopy experiments in future as THz radiation is extremely sensitive to changes in the dynamics of bulk and ordered water [206, 207]. These experiments in combination with the current results from inelastic neutron scattering spectroscopy will enable us to observe consorted protein and hydration shell dynamics by combining neutron and photon spectroscopic techniques.

Chapter 6

Results and discussion: E.coli copper amine oxidase

In this chapter, the results for ECAO are described. Optimization of expression, purification and crystallization of hydrogenated ECAO will be followed by perdeuteration of wild-type and E573Q mutant of ECAO. Crystallization of perdeuterated wild-type and mutant ECAO will then be attempted with the aim of obtaining crystals > 0.1 mm^3 in volume. This will be followed by QENS experiments on wild-type and E573Q/I342F mutants of ECAO to access the possible changes in the internal and global diffusive dynamics. These experiments will be complemented by DLS studies. MD simulations will be used analyze changes in the dynamics of specific residues. These results will be discussed in context of previously reported crystallographic studies on E573Q and E573Q/I342F mutants of ECAO.

6.1 Expression, purification and crystallization of hydrogenated wild-type E.coli copper amine oxidase

Prior attempts to obtain higher yields of ECAO resulted in expression of inactive *E.coli* copper amine oxidase (ECAO) due to divalent Zinc ions entering the active site instead of copper ions. Hence, we decided to add copper sulphate to the growth medium. Since different cell types have varying levels of tolerance of metal salts like copper sulphate, we had to optimize the expression of *E.coli* copper amine oxidase for *E.coli* BL21(DE3) cells as these were used for protein expression. The optimization was carried out by varying three parameters: the growth medium, the temperature

and the concentration of copper sulphate. We used terrific broth (TB) which was previously used to successfully express ECAO [72,77]. In addition, we used luria betrani (LB) broth which is a conventional and widely used choice of growth medium for expression in *E.coli* BL21(DE3) cells. Protein expression in cultures grown in these media was induced by adding 1 mM IPTG (subsection 3.7.1) at three temperatures: 25, 30 and 37°C (figure 8). As it can be seen in figure 8, the protein yield was almost the same in both terrific and luria bertani broth. However, there were more impurities in case of terrific broth. We therefore decided to proceed with LB broth to standardize the next set of conditions for optimal expression. After selection of the growth medium and the temperature, we induced the expression of wild-type ECAO at two concentrations of copper sulphate (figure 9).

We also varied the temperature for these expression studies in order to obtain optimal conditions (growth medium, concentration of copper sulphate and the temperature) for the expression for which the impurities would be minimal whilst maintaining a high yield of wild-type ECAO. As seen in figure 9, the protein yield was found to be higher for 2 mM copper sulphate. Hence, we decided to use this concentration for the large scale bacterial growth and protein purification. The purity level for 25 and 30 °C was found to be very similar. However, we decided to induce the expression at 30 °C since the cells would grow better at this temperature than 25 °C resulting in a higher total protein yield. Induced *E.coli* BL21DE3 cells were centrifuged at 6500 rpm at room temperature. The wet weight of the pellet obtained was usually $\approx$ 4 gm. The pellet obtained was then partially lysed at 4 °C as described in subsubsection 3.7.2.1 of chapter 3. The lysate was then centrifuged at 12000 rpm at 4 °C. The pellet obtained was then discarded and the supernatant dialyzed against 20 mM Tris-Cl, pH 7.0. After preparing the Q-sepharose column as described in subsubsection 3.7.2.2 of chapter 3, the supernatant was loaded and the purification was carried out. Figure 6.1 (top) shows the fractions obtained after ion exchange purification of wild-type ECAO.

It can be seen that wild-type ECAO eluted slowly as the concentration of NaCl is increased from 14 % to 40 %. However, the majority of the protein starts to elute at 28 % NaCl and then continues to elute upto 44 % NaCl (fraction no. 21). The obtained fractions containing ECAO were then pooled, dialyzed against 20 mM Tris, pH 7.0 and concentrated to $\approx$ 2 mL. The dialyzed protein sample and the supernatant obtained after the partial lysis of the induced cell pellet were both compared. The purity of dialyzed ECAO was found to have improved significantly as seen in figure 6.1 (bottom). The dialyzed and concentrated sample was then loaded onto superdex-200 column. ECAO typically eluted at the 60 mL mark as seen in figure 11 in appendices. Figure 6.2 shows the SDS-PAGE gel for the fractions obtained after gel filtration. It can be seen that the protein obtained was pure and also high in yield.

The purest fractions obtained after gel filtration were then dialyzed against 5 L of 20 mM Tris-Cl, pH 7.0 supplemented with 0.05 mM $CuSO_4$ and 0.1 mM $CaCl_2$ for an hour at 4 °C so as to ensure that the surface and the buried Calcium binding sites are saturated and that the active centre contains Cu(II) ion. The protein was then concentrated until its concentration reaches 12-15 mg/mL. This sample was then used to setup crystallization trials. The same protocol was followed for the single (E573Q) and the double (E573Q/I342F) mutants of ECAO. The only difference with respect to wild-type ECAO was that the elution concentration of NaCl (buffer B) during ion exchange chromatography was different for the mutants. As seen in figure 10 in appendices, the E573Q/I342F double mutant was found to elute between 30 and

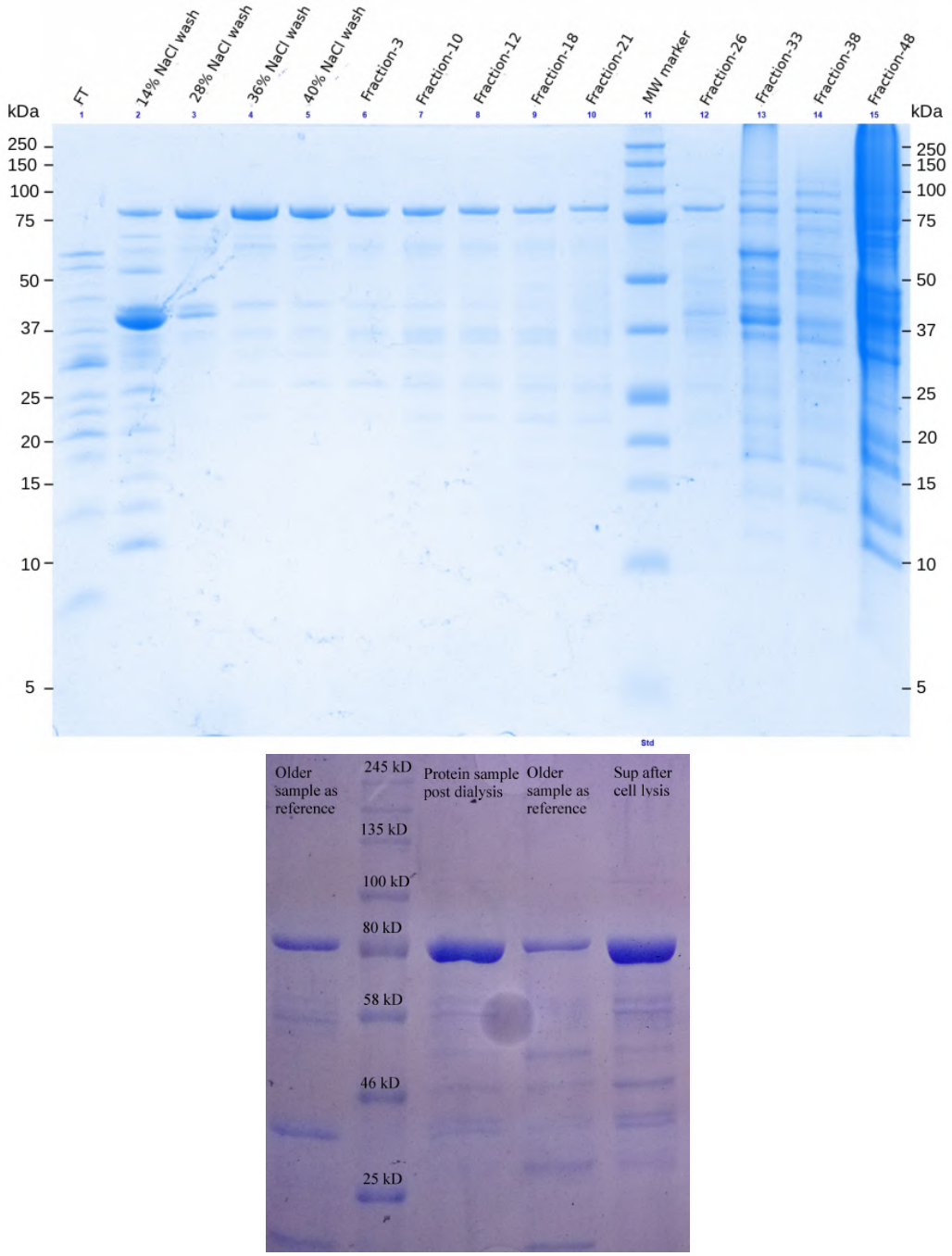

Figure 6.1: Fractions of wild-type ECAO collected after ion-exchange chromatography (top) and the pooled fractions after dialysis against 20 mM Tris-Cl, pH 7.0 (bottom).
Labels (top panel): FT - flowthrough, 14, 28, 36 and 40 % washes - fractions of ECAO obtained after the passage of 14, 28, 36 and 40 % NaCl solution through the column. Fractions 3-26 - continuation of 40 % NaCl wash, MW marker - molecular weight marker, fraction-33 - 60 % NaCl wash, fraction-38 - 100 % NaCl wash, fraction-48 - 1M NaCl wash

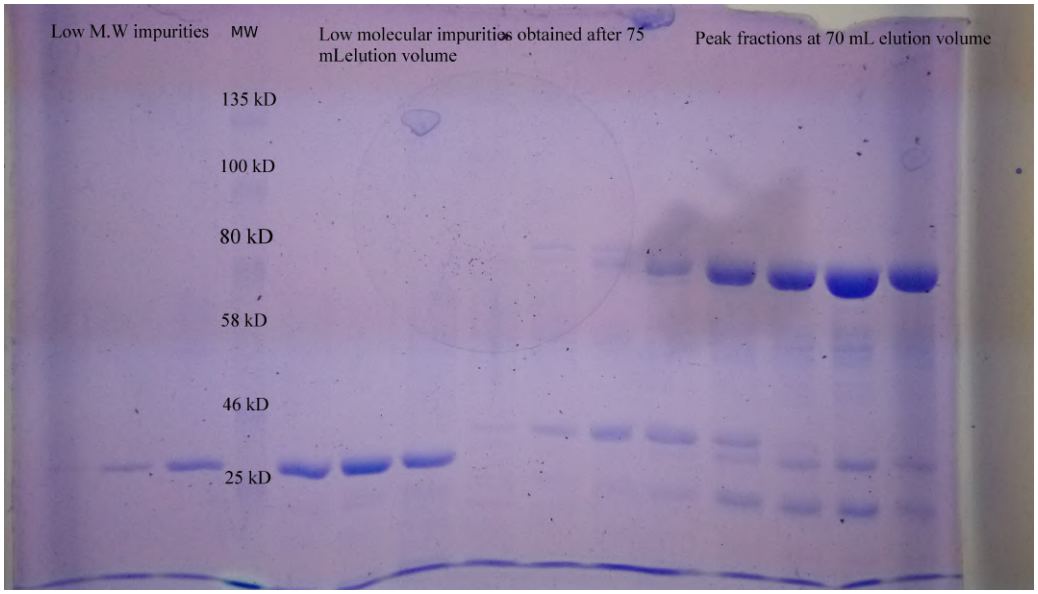

Figure 6.2: Fractions of hydrogenated wild-type ECAO collected after gel filtration run through Superdex-200 column in 20 mM Tris-Cl, pH 7.0 (bottom).

40 % NaCl. Hence, the NaCl concentration range for elution of double mutant was found to be narrower than that for wild-type ECAO. For hydrogenated *E.coli* copper amine oxidase, the following crystallization conditions were tested.

Table 6.1: Crystallization scheme 1 for *E.coli* copper amine oxidase

	pH						
Concentration of sodium citrate (M)		6.8	6.9	7.0	7.1	7.2	7.3
	1.0 M			*			
	1.1 M					+	
	1.2 M				+		
	1.3 M				*		

The boxes marked with '+' sign are the conditions in which well diffracting crystals of wildtype *ECAO* were obtained while those indicated with a '*' are the conditions in which well diffracting crystals of the double mutant (E573Q/I342F) of ECAO were obtained.

Table 6.1 summarizes the very first conditions that were tested for crystal growth. The best crystals of hydrogenated wild-type and double mutant (E573Q/I342F) were obtained at 1.2 and 1.3 M sodium citrate respectively (Table 6.1). Based on these results, we tried to optimize the crystallization conditions further by increasing the concentration of the protein to 15-20 mg/mL. The best crystals were obtained at 15 mg/mL (figure 6.3) and were pinkish orange in color, as expected when TPQ is in the off-copper conformation [77].

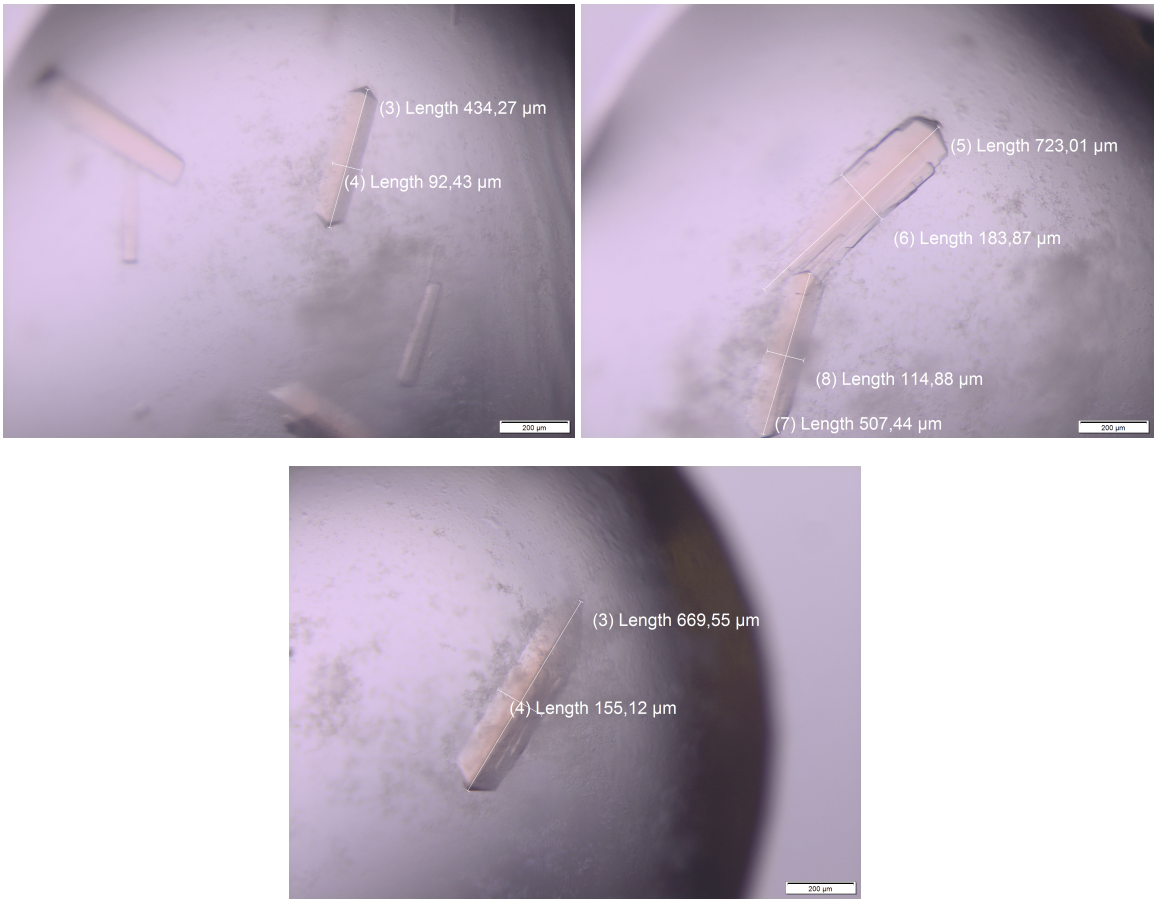

Figure 6.3: The best crystals obtained from hydrogenated wild-type ECAO

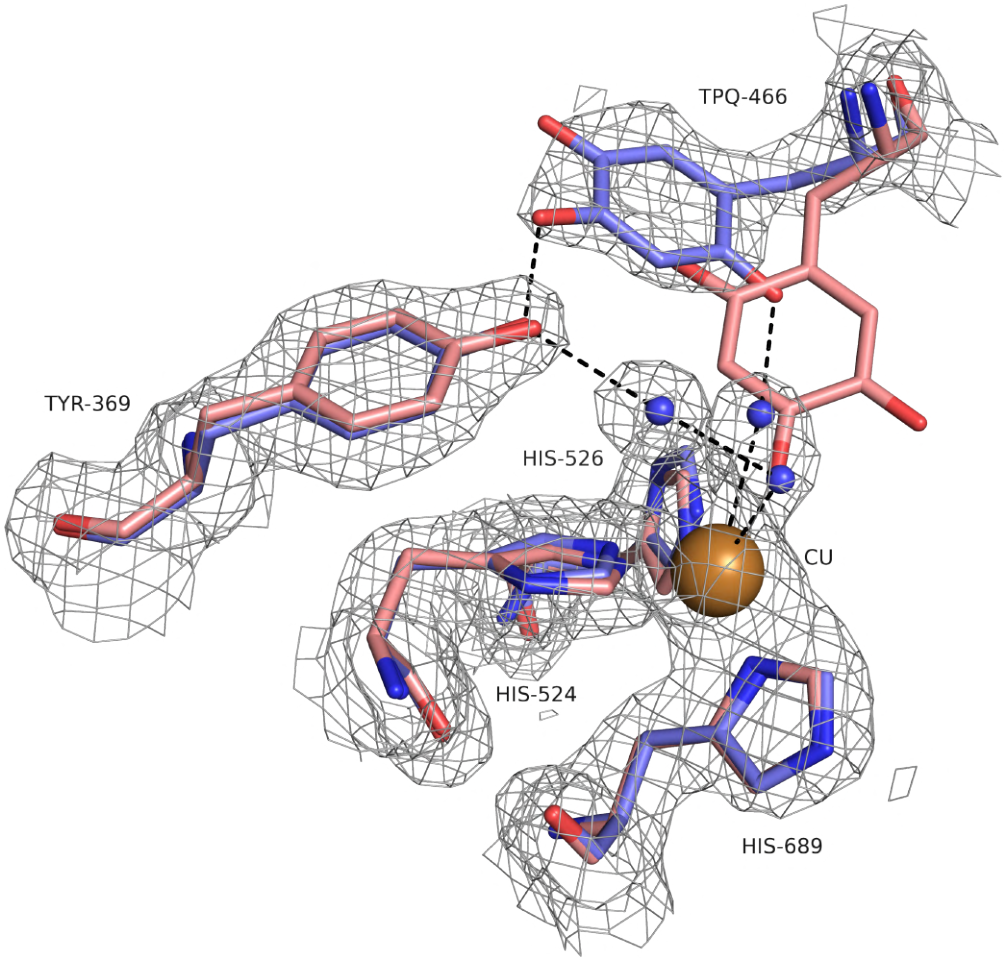

Figure 6.4: Superposition of the active site of the refined structure of wild-type *E.coli* copper amine oxidase (blue) and its inactive form (PDBID - 1OAC) (pink). The labeled components are - the cofactor 1,4,5-topaquinone (TPQ) 466, the divalent copper ion coordinated with the four Histidines and the hydrogen bonded network of water molecules which are catalytically important. $2F_o$-F_c map of the wild-type ECAO solved in the present work is displayed at a contour level of 1 σ.

These crystals were used for X-ray diffraction experiments at ID23.1 beamline. The structure was solved using PDB 16DZ as the search model and refined to a final R_{work} of 20.5 % and the R_{free} of 23.4 % (table 1). The superposition of the active site of the solved hydrogenated wild-type ECAO and that of the inactive ECAO (PDBID 1OAC) is shown in figure 6.4. It can be clearly seen that the TPQ in our structure is in the off-copper conformation unlike the inactive ECAO. In addition, the catalytic water molecules can also be seen. It was therefore confirmed that the wild-type ECAO expressed and purified using the current protocol was active.

6.2 Perdeuteration and crystallization of perdeuterated wild-type E.coli copper amine oxidase

6.2.1 *In-vitro* transposition in order to change the antibiotic marker

In-vitro transposition was carried out as described in chapter 3. About 30 % of the total streaked colonies were found to be ampicillin -ve and kanamycin +ve. Some of these are indicated by white circles in figure 6.5. Four of the colonies were randomly selected. Plasmids were purified from the overnight grown cultures of these colonies. Commercially available purification kits were used for the plasmid purification. The purity of the plasmids was checked using agarose gel electrophoresis. As seen in figure 12 in appendices, the plasmids were found to be pure.

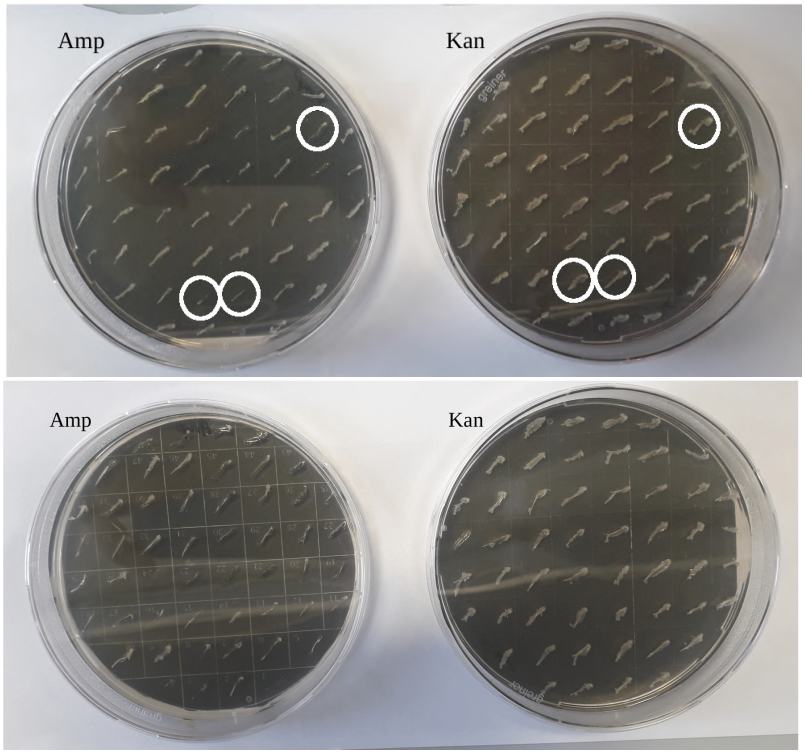

Figure 6.5: Growth obtained after streaking cells transformed with *in−vitro* transposed pKKECAO plasmid on Luria Agar with Ampicillin (left) and Kanamycin (right). The white circles are some of the randomly selected colonies.

6.2.2 Expression test for the *in-vitro* transposed ECAO wild-type construct in pKK 233-3 vector

BL21(DE3) cells were transformed with the purified plasmids containing the Kanamycin resistance marker as described in chapter 3. The transformed cells were used for testing the expression of plasmids containing the wild-type ECAO construct and kanamycin resistance marker in LB medium. Since the optimum temperature for expression was found to be 30 °C for the cells transformed with pKKECAO plasmid, this was the only temperature which was tested for the expression of *in−vitro* transposed pKKECAO plasmid. An additional fifth plasmid was also selected from the colonies obtained in in-vitro transposition. As seen in figure 6.5, the expression of ECAO was seen for all the selected colonies. After the expression check in LB medium, we proceeded to check the expression of wild-type ECAO in H and D-minimal media in order to obtain perdeuterated wild-type ECAO.

6.2.3 Preculture, adaptation and expression check in D-enfors media

We selected the F2 and E8 plasmids (figure 13) from those which were found to express ECAO in LB medium. Overnight grown LB culture of BL21(DE3) cells harboring the F2 and E8 plasmids were used to inoculate the H-minimal medium (the recipe is given in subsection 3.1.5. The cultures grown in H-minimal media were induced at 20, 25 and 30 °C. 37 °C temperature was excluded as the amount of impurities was found to be higher at this temperature when the expression of original pKKECAO plasmid was tested (figure 9).

The expression of wild-type ECAO in case of both F2 and E8 plasmids was found to be reasonably good at all the three temperatures. However, the purity of ECAO was higher in case of the E8 plasmid (figures 6.6 and 6.7). Hence, E8 plasmid was selected for testing the expression in D-minimal medium. In the case of the expression profile for E8 plasmid, the yield was found to be the highest at 30 °C (figure 6.7).

The cultures grown in H-minimal media were inoculated into D-minimal media. In order to adapt the cells to the deuterated components of the D-minimal medium, the cultures were transferred to fresh 10 mL medium after overnight growth. This process was repeated four times so as to ensure complete adaptation of the cells to the D-minimal medium. This step was followed by induction of protein expression in cultures of cells harboring E8 plasmid. Expression was induced at 20, 25 and 30 °C. As seen in figure 6.8, the yield of the perdeuterated wild-type ECAO was the lowest at 20 °C. Hence, 30 °C was chosen as the induction temeprature for the large scale fermenter run.

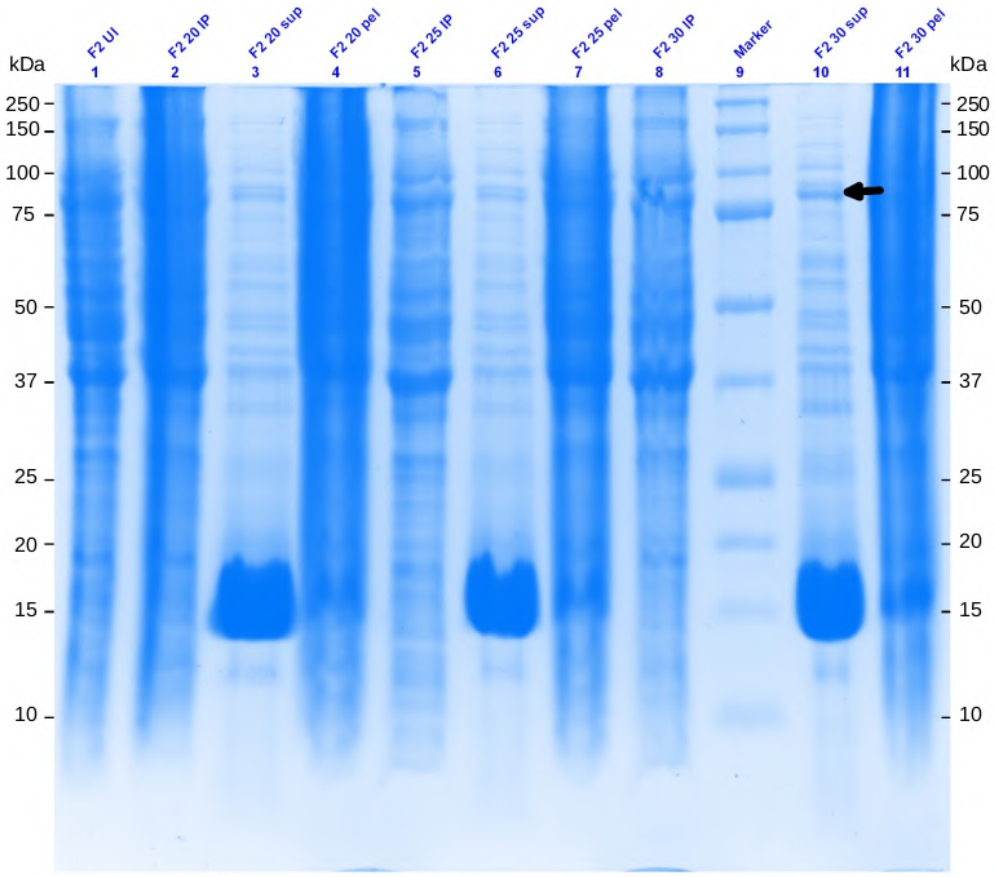

Figure 6.6: Expression test of *in-vitro* transposed clone F2 from wild-type *E.coli* copper amine oxidase in H−minimal medium. ECAO band is marked.
Labels: UI - uninduced pellet, IP - the induced pellet (whole cells), pel - pellet obtained post enzymatic lysis, sup - supernatant obtained post enzymatic lysis at the respective temperatures (20, 25 and 30 °C) for F2 clone.

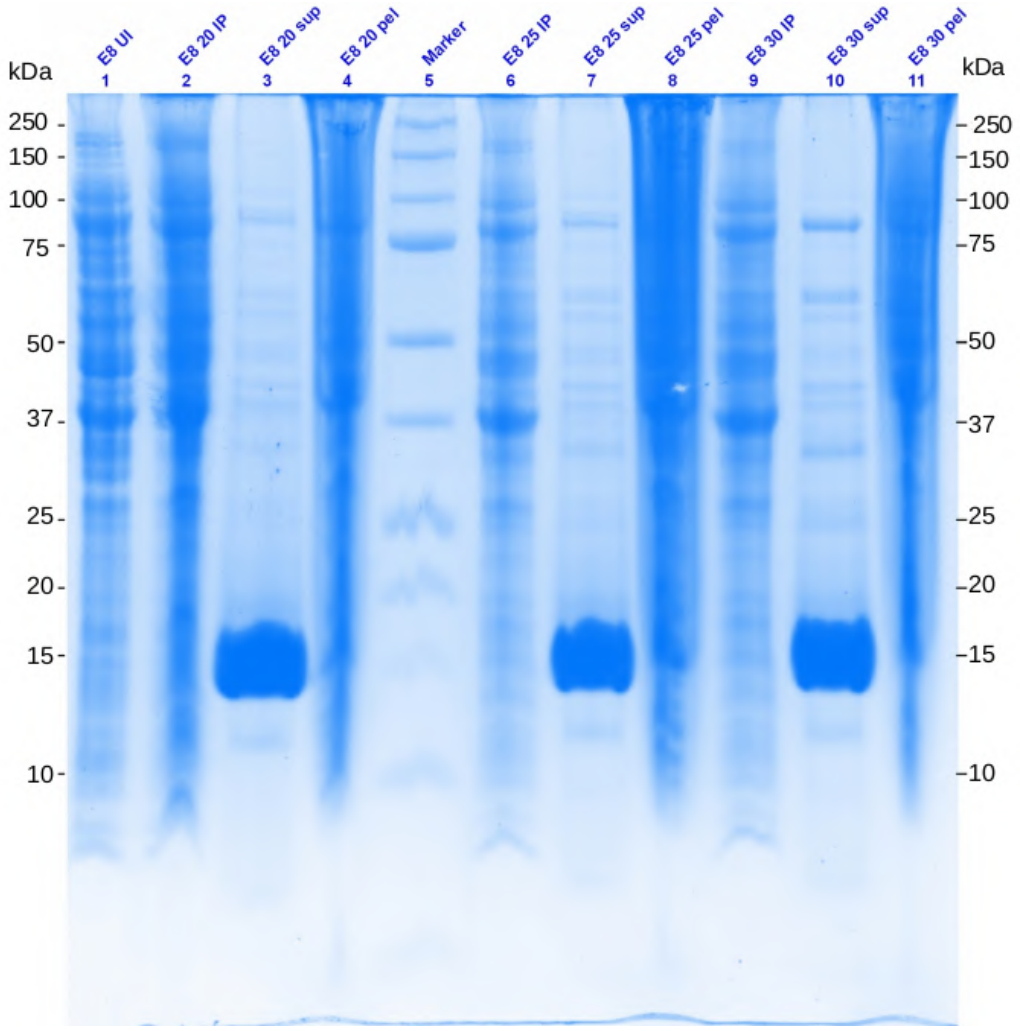

Figure 6.7: Expression test of *in−vitro* transposed clone E8 of wild−type *E.coli* copper amine oxidase in H−minimal medium.
Labels: UI - uninduced pellet, IP - the induced pellet (whole cells), pel - pellet obtained post enzymatic lysis, sup - supernatant obtained post enzymatic lysis at the respective temperatures (20, 25 and 30 °C) for E8 clone.

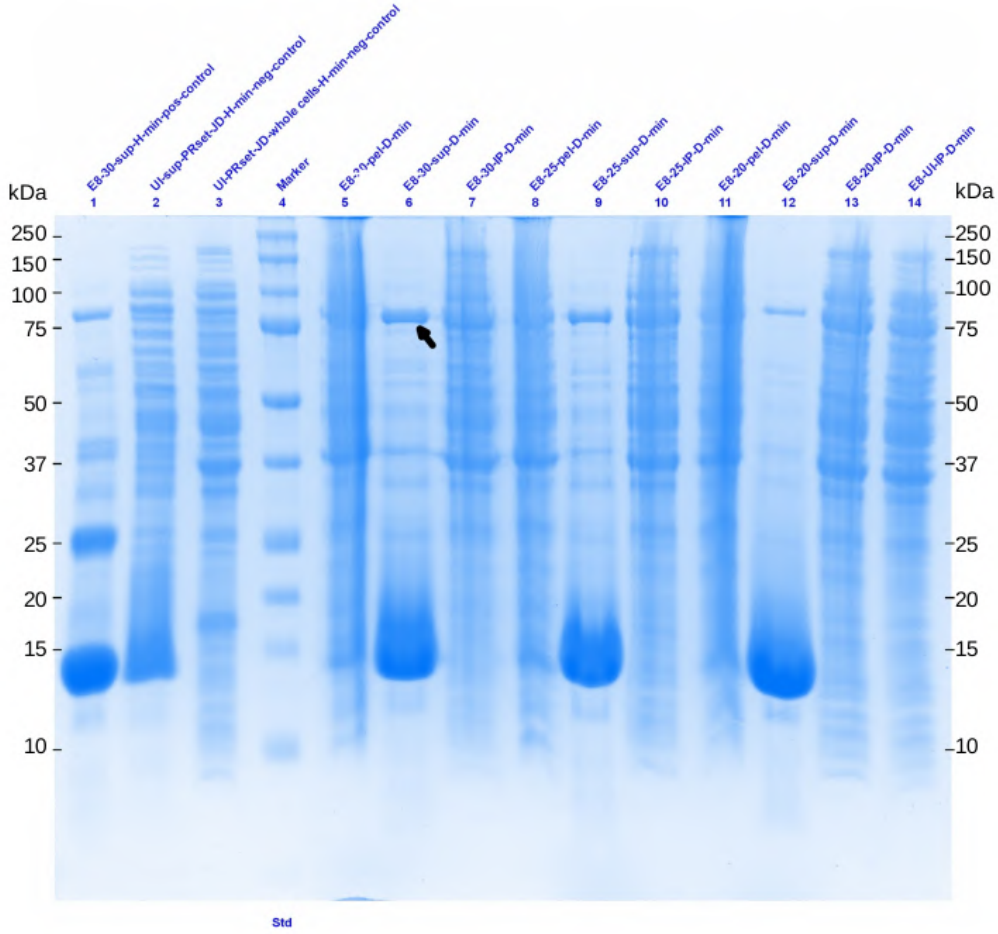

Figure 6.8: Expression test of *in−vitro* transposed clone E8 of wild-type *E.coli* copper amine oxidase in D−minimal medium. The ECAO band is marked with an arrow.

Labels: Marker - molecular weight marker, E8-30-sup-H-min-pos-control - supernatant of the E8, clone obtained after induction at 30 °C (positive control for hydrogenated ECAO), UI-sup-PRset-JD-H-min-neg-control - supernatant of empty cells (negative control), UI-PRset-JD-whole cells-H-min-neg-control - whole empty, cells (negative control), Marker - molecular weight marker, E8-30-pel-D-min - post-lysis pellet of E8 clone after induction at 30°C, E8-30-sup-D-min - induced supernatant of E8 clone at 30°C, E8-30-IP-D-min - whole cells induced at 30°C, E8-25-pel-D-min - post-lysis pellet of E8 clone after induction at 25°C, E8-25-sup-D-min - induced supernatant of E8 clone at 25°C, E8-20-pel-D-min - post-lysis pellet of E8 clone after induction at 20°C, E8-20-sup-D-min - post-lysis supernatant of E8 clone after induction at 20°C, E8-20-sup-D-min - induced supernatant of E8 clone at 20°C, E8-UI-IP-D-min - uninduced whole cells of E8 clone.

6.2.4 Expression test after the final scale up

The expression of perdeuterated wild-type ECAO was checked after the final fermenter run as shown in figure 6.9. The expression was found to be reasonably good and hence we proceeded with the purification of perdeuterated wild-type ECAO.

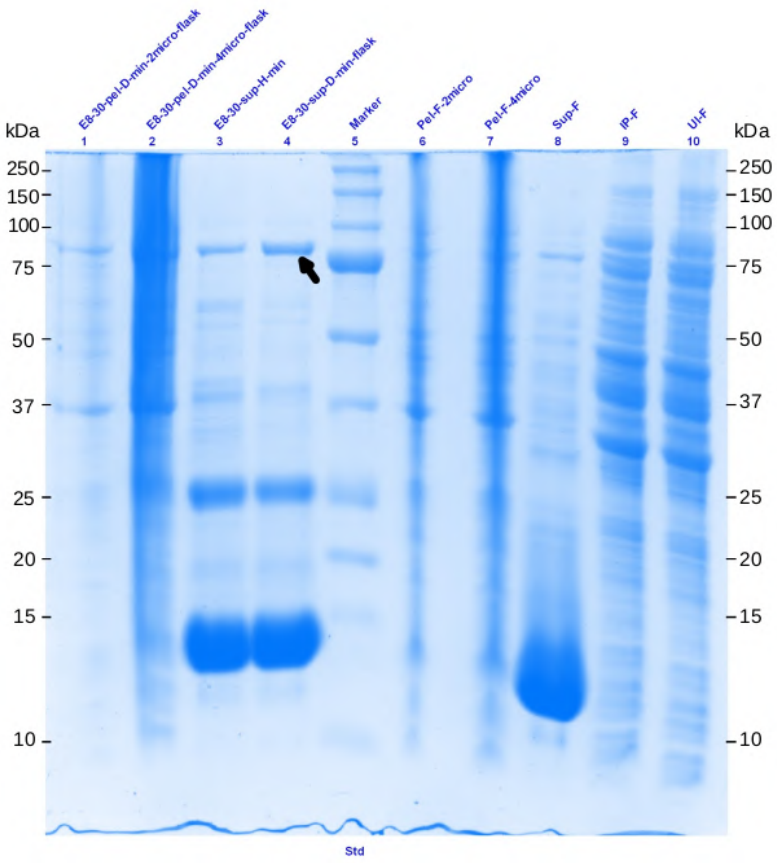

Figure 6.9: Expression test of *in−vitro* transposed clone E8 of wild-type *E.coli* copper amine oxidase from fermenter run.
Labels: E8-30-pel-D-min-2micro-flask - E8 clone induced at 30 °C in D−minimal medium in flask (2 μl sample), E8-30-pel-D-min-2micro-flask - same as the prior sample but 4 μl in volume, E8-30-sup-H-min - supernatant post enzymatic lysis after induction at 30 °C in H−minimal medium, E8-30-sup-D-min-flask - supernatant obtained from the culture of E8 clone grown in D-minimal medium in flask, Marker - molecular weight marker, Pel-F-2micro - pellet post enzymatic lysis from the cells obtained after induction carried out in fermenter (2 μl sample), Pel-F-4micro same as prior sample but 4 μl in volume, Sup-F - supernatant after enzymatic lysis of cells induced in fermenter run, IP-F - induced whole cells obtained after fermenter run, UI-F - uninduced whole cells obtained after fermenter run.

6.2.5 Confirmation of perdeuteration of wild-type ECAO

Perdeuteration of ECAO leads to an increase in the molecular weight of the protein since deuterium (^{2}H) atom has an atomic mass number twice as that of protium (^{1}H) which forms hydrogen. We performed electrospray ionization mass spectrometry (ESI−MS) on purified perdeuterated wild-type ECAO sample in order to confirm the deuteration of wild-type ECAO and to calculate the percentage of deuteration.

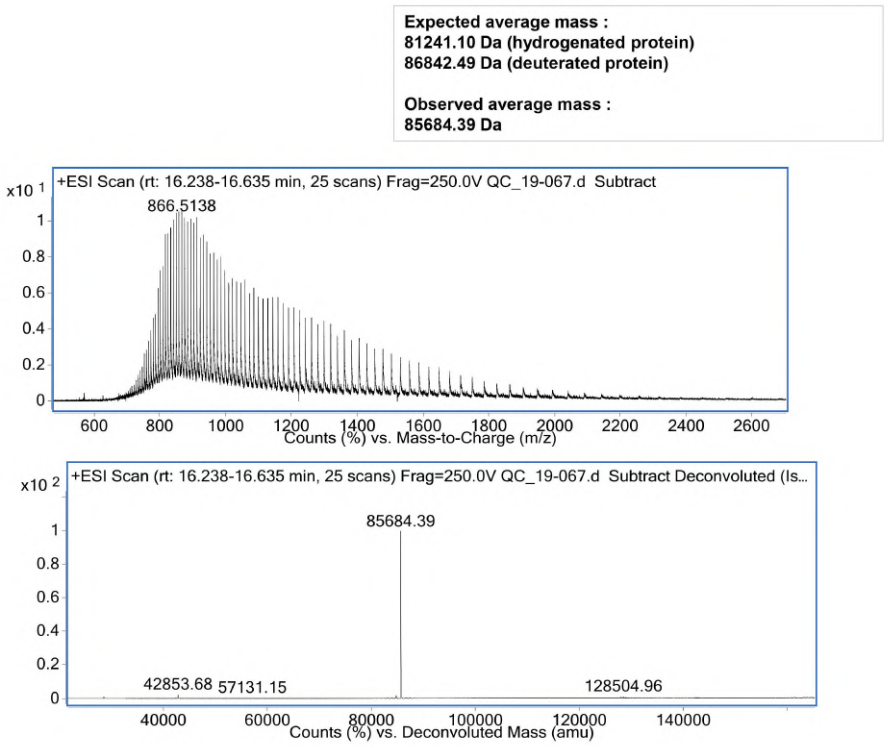

Figure 6.10: Mass spectrum recorded for perdeuterated wild−type ECAO

As seen in the top panel of figure 6.10, the mass of hydrogenated ECAO is 81241.1 Da and that of perdeuterated ECAO is expected to be 86842.5 Da. However, the expected mass is calculated based on the assumption that all the hydrogens are replaced with deuterium atoms. It should be noted that there are labile deuterium atoms on the protein surface which get exchanged with the hydrogen atoms of the solvent i.e. H_2O. These mainly include the hydrogens in main chain peptide -NH groups and the hydrogens bound to N, O and S groups of polar amino acid groups [208]. Since ESI-MS experiments are carried out in light water, this exchange takes

place, resulting in a lower molecular weight than expected. In the case of wild-type ECAO, the expected mass after 100% deuteration is 86842.5 Da (figure 6.10). The number of exchangeable hydrogens was found to be 1223 based on the Biomolecular Scattering Length Density Calculator [209] which translates to 85619.5 Da. It can be seen in figure 6.10 that the observed average mass of perdeuterated wild-type ECAO is 85864.4 Da which is higher than the expected 85619.5 Da. Hence, it can be inferred that there are caged deuterium atoms present in perdeuterated wild-type ECAO.

6.2.6 Purification and crystallization of perdeuterated wild-type ECAO

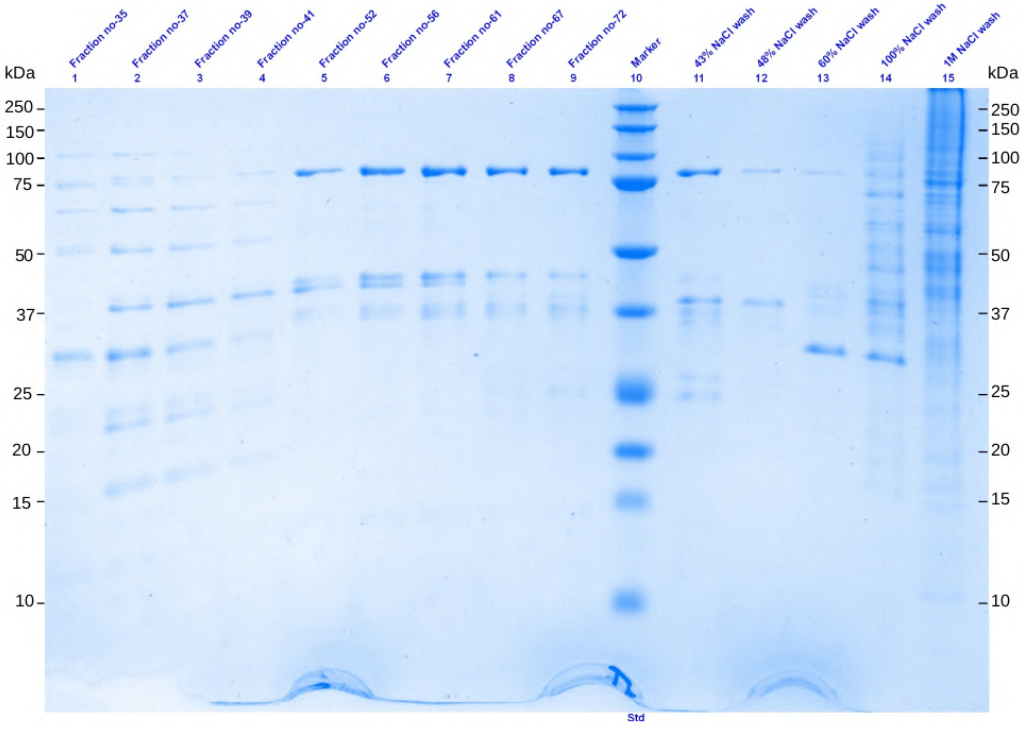

Figure 6.11: Fractions of perdeuterated wild-type ECAO obtained after ion exchange chromatography.
Labels: fractions 35 to 41 - 0 to 20 % NaCl washes, fractions 52 to 72 - 23 to 40 % NaCl.

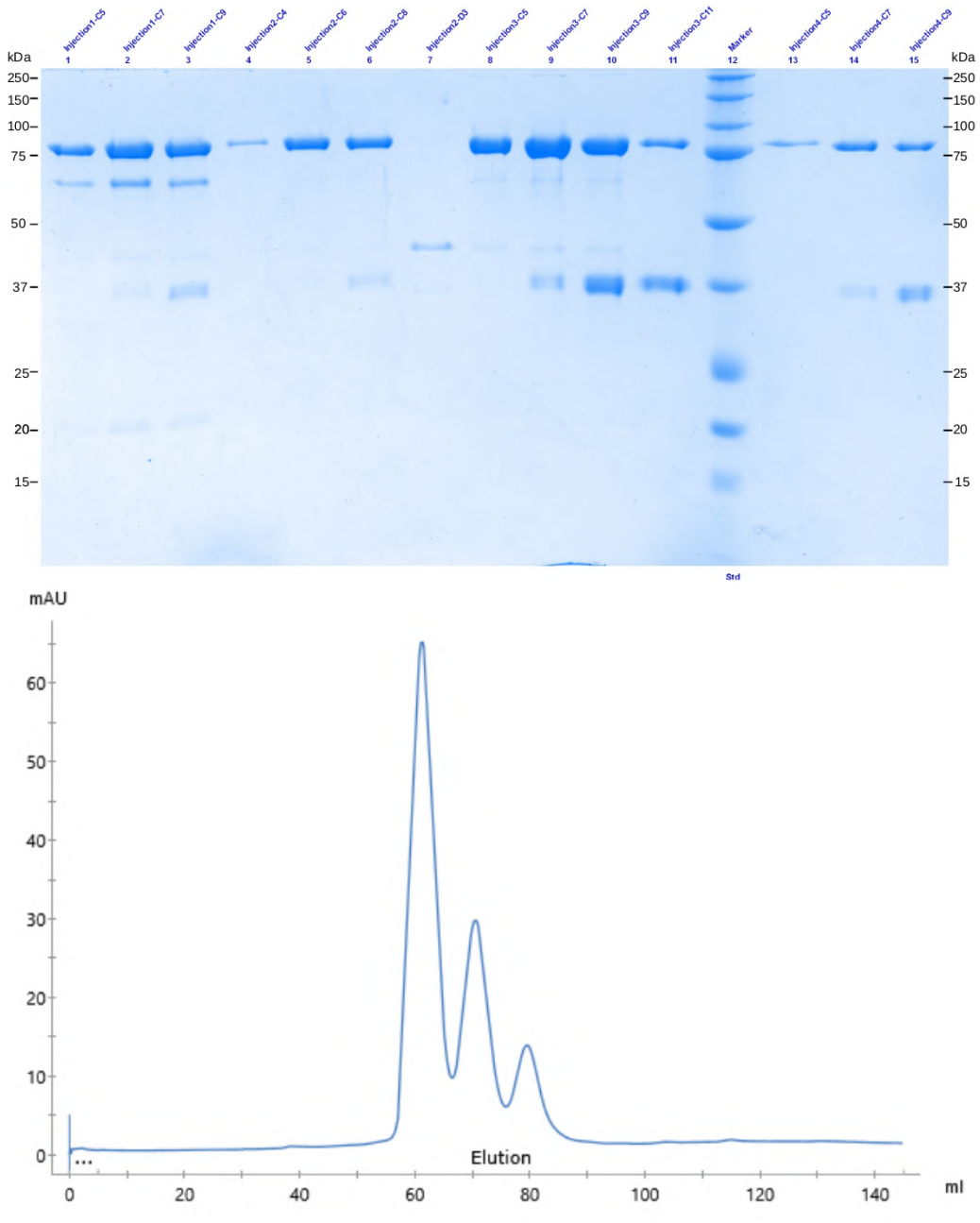

Figure 6.12: Fractions of perdeuterated wild-type ECAO obtained after gel filtration using superdex-200 column (top) and the corresponding UV_{280} trace (bottom).

Ion exchange chromatography was carried out as described in the methods section of chapter 3. As seen in figure 6.11, the protein eluted between 35 and 43 % NaCl. As compared to the elution profile for hydrogenated wild-type ECAO described in the previous section, the perdeuterated wild-type ECAO is found to elute faster (35-43 % vis-a-vis 28-45 % NaCl). This maybe due to the altered dynamics of the entire molecule following the replacement of hydrogen atoms by the heavier deuterium atoms. It should be noted that the elution was carried out by increasing the concentration of NaCl by 5 %. The UV_{280} was allowed to stabilize after each of the increments. The fractions obtained in ion exchange chromatography i.e. fractions 61 to 72 and the 43% NaCl wash in figure 6.11 were pooled, concentrated to a final volume of $\approx$ 2 mL and loaded onto superdex-200 gel filtration column.

The fractions obtained from the major peak at 60 mL after gel filtration were pooled and concentrated to 15 mg/mL as optimized in section 6.1. The pooled and concentrated D-ECAOWT was dialyzed overnight against 0.1 mM $CaCl_2$ and 0.05 mM $CuSO_4$ at 4 °C. The dialyzed protein solution was then exchanged with the same buffer as described in section 3.13 but prepared with heavy water (D_2O) instead of light water (H_2O). The pD of the buffer was adjusted to 7.0 in the same way as the pH 7.0 of the H_2O buffer was adjusted. The pH corresponding to a particular pD is obtained by subtracting 0.4 from the pD value [210]. So, in order to attain a pD of 7.0, the pH of the D_2O buffer was adjusted to 6.6. The D_2O buffer was used for dilution after every cycle. A total of three cycles of dilution and concentration were carried out to ensure a complete exchange. We initially used the same conditions as described in section 6.1. However, we obtained smaller crystals (longest dimension of $\approx$ 0.14-0.21 mm) than was the case for hydrogenated wild-type ECAO, as seen in figure 6.13. We increased the highest concentration of sodium citrate to 1.4 M to determine whether we can get larger crystals of perdeuterated *E.coli* copper amine oxidase. We also changed the pH range slightly in order to cover a wider range of conditions. For the next optimization, we used the pH range from 6.5 to 7.5 with intervals of 0.2 instead of 0.1 (Table 6.2) The new conditions used for optimization 1 were as follows.

Table 6.2: Crystallization scheme 1 - optimization 1 for *E.coli* copper amine oxidase

	pH						
Concentration of sodium citrate (M)		6.5	6.7	6.9	7.1	7.3	7.5
	1.3 M			*		+/*	+
	1.4 M					+	+

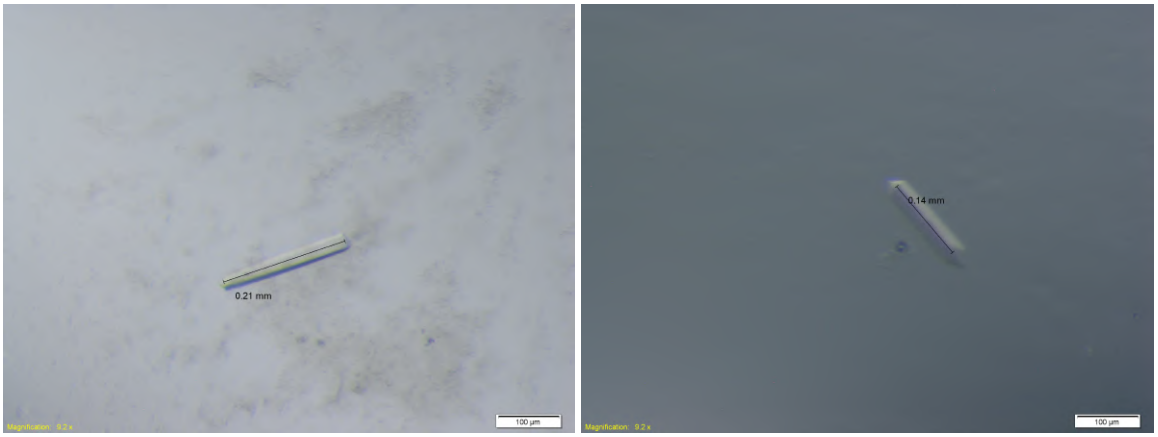

Figure 6.13: Crystals of perdeuterated wild-type ECAO obtained in 1.2 M sodium citrate pH 6.9 (left) and pH 7.3 (right) as described in the scheme described in table 6.1

As seen in figure 6.14, the crystals obtained in 1.3-1.4 M sodium citrate were viable in size but still lacked the desired crystal volume of at least 0.1 mm^3 for the neutron diffraction. Besides the smaller volume, these crystals were also defective as seen in the left panel of figure 6.14.

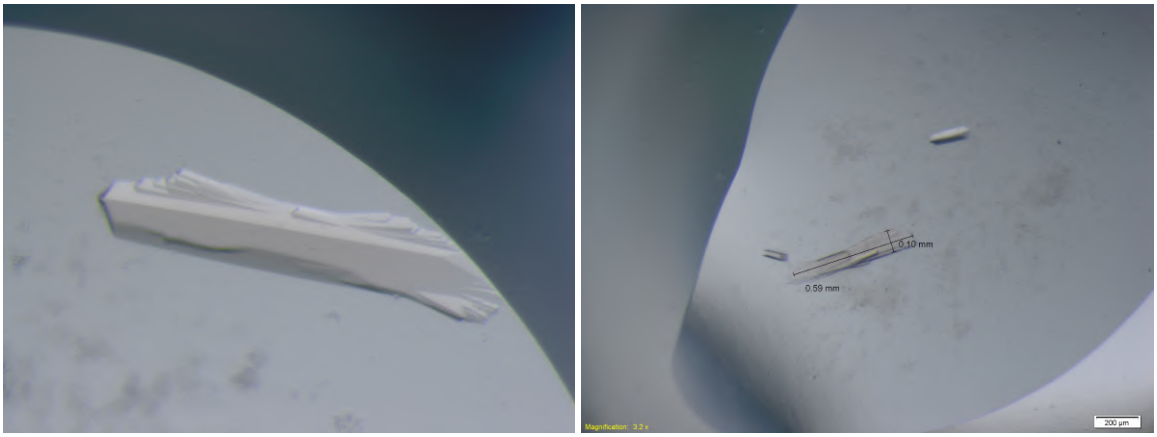

Figure 6.14: Crystals of perdeuterated wild-type ECAO obtained in 1.4 M sodium citrate pH 7.3 (left) and in 1.3 M sodium citrate pH 6.9 (right) as described in the scheme described in table 6.2

Figure 6.15: Crystals of perdeuterated wild-type ECAO obtained in 1.30 (top left), 1.35 (top right) and 1.4 (bottom) M sodium citrate at pH 7.3 as indicated in table 6.2

We increased the drop size from 2 μL of protein and 2 μL of mother liquor to 10 + 10 μL of the same. As seen in figure 6.15, we could obtain crystals of much larger volume (longest dimension of $\approx$ 0.4-0.5 mm). However, these crystals also contained defects as seen in figure 6.15. The reason for these defects can be excessive nucleation which is usually due to the specific precipitant.

Table 6.3: Crystallization scheme 1 - optimization 2 for *E.coli* copper amine oxidase

	AS			
		1.6	1.7	1.8
Concentration of sodium citrate (M)	0.15			
	0.20	+		
	0.25	+		

In order to avoid this excessive nucleation and hence the crystal defects, we used a different crystallization condition. We used a combination of ammonium sulphate and sodium citrate as precipitant and Tris-Cl (pH 8.0) as the buffer in the mother liquor (Table 6.3).

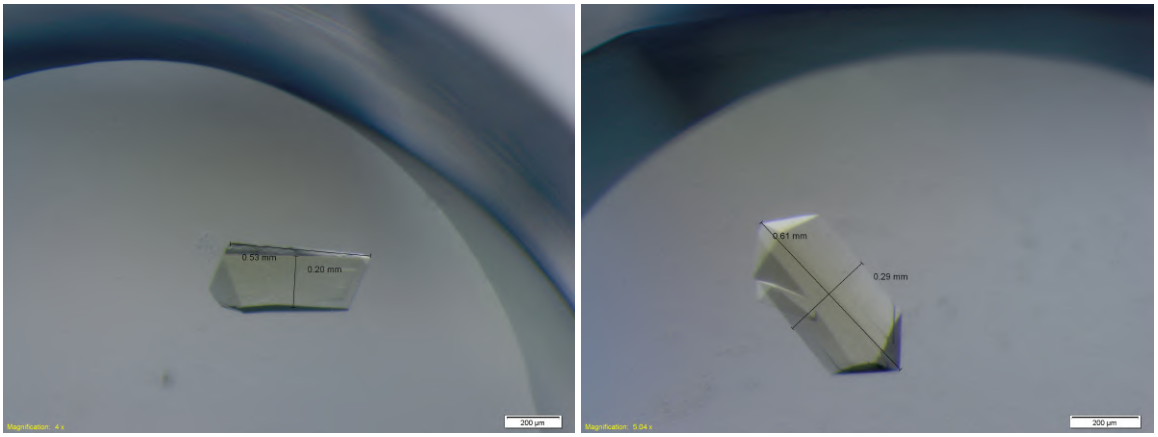

Figure 6.16: Crystals of perdeuterated wild-type ECAO obtained in 0.20 and 0.25M sodium citrate with 1.6M ammonum sulphate at pH 8.0 as indicated in table 6.3

We could obtain crystals with a longest crystal dimension of $\approx$ 0.6 mm and a shortest dimension of $\approx$ 0.2-0.3 mm (figure 6.16). Although the length of the longest dimension was the same as that obtained in previous conditions, the short dimension was almost twice longer resulting in an increase in the overall crystal volume.

6.3 Perdeuteration and crystallization of perdeuterated single mutant (E573Q) E.coli copper amine oxidase

We carried out the *in−vitro* transposition of the E573Q mutant plasmid of pKKE-CAO in the same way as described in subsection 6.2.1. We randomly selected five colonies from those which were ampicillim -ve and kanamycin +ve. Plasmids were purified from these colonies and were used for transforming BL21(DE3) cells.

6.3.1 Expression test for the *in−vitro* transposed ECAO single mutant (E573Q) construct in pKK 233-3 vector

BL21(DE3) cells transformed with the randomly selected colonies after *in−vitro* transposition were used for testing the protein expression in LB medium. We used two temperatures: 25 and 30 °C for expression studies. 20 °C was not included as the induction of wild-type ECAO after *in−vitro* transposition indicated a lower protein yield. As seen in figure 14, we could see a significant level of expression of E573Q mutant of ECAO at 30 °C in all the clones (colonies) except B5. The expression at 20 °C was very low which is consistent with the expression studies with wild-type ECAO post in-vitro transposition. We selected the clones A8 and B6 for the expression check in D-minimal medium. The cultures were adapted as described in the previous section. This step was followed by expression tests at 25 and 30 °C. The 20 °C temparature was excluded as it consistently resulted in very low yields for both wild-type and E573Q mutant of ECAO after in-vitro transposition. As seen in figure 15, the yield was found to be the same for all the samples. However, the expression profile was slightly cleaner at 30 °C for both the selected clones (figure 15). Hence, we used the plasmid from A8 and selected 30 °C as the induction temperature. The fermenter run was carried out and the expression of perdeuterated E573Q mutant of ECAO was found to be significantly good.

6.3.2 Purification and crystallization of perdeuterated E573Q mutant of ECAO

Lysis of the induced pellet and the subsequent ion exchange chromatography were carried out as described in the methods section of chapter 3. As seen in figure 6.17, the protein eluted at a slightly lower concentration of NaCl (between 28 and 38 %) compared to perdeuterated wild-type ECAO. The fractions 7 to 24 were pooled and concentrated using a 50 kDa cutoff centrifugal filter and loaded onto superdex-200 gel filtration column. The fractions obtained after gel filtration (figure 6.18) were pooled, dialyzed and exchanged with D_2O buffer as explained in the previous section and used to setup crystallization trials at a concentration of 15 mg/mL.

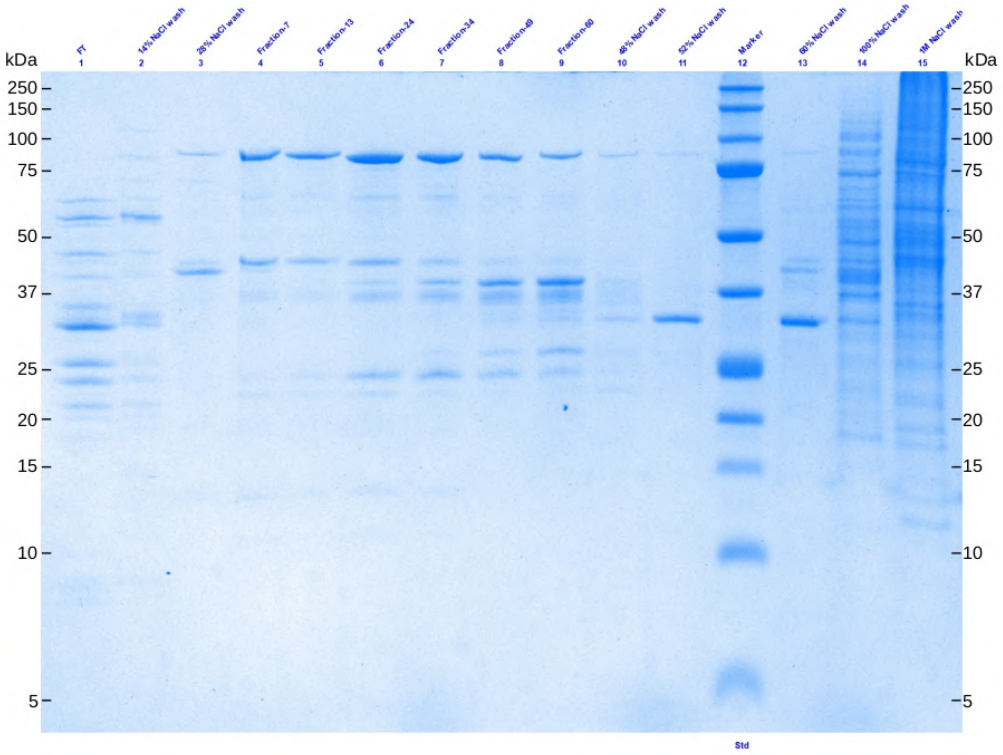

Figure 6.17: Fractions of perdeuterated E573Q mutant of ECAO obtained after ion exchange chromatography
Apart from the NaCl washes labeled in the figure, fractions 7 to 60 are fractions obtained after the passage of 30 to 45 % NaCl through the column.

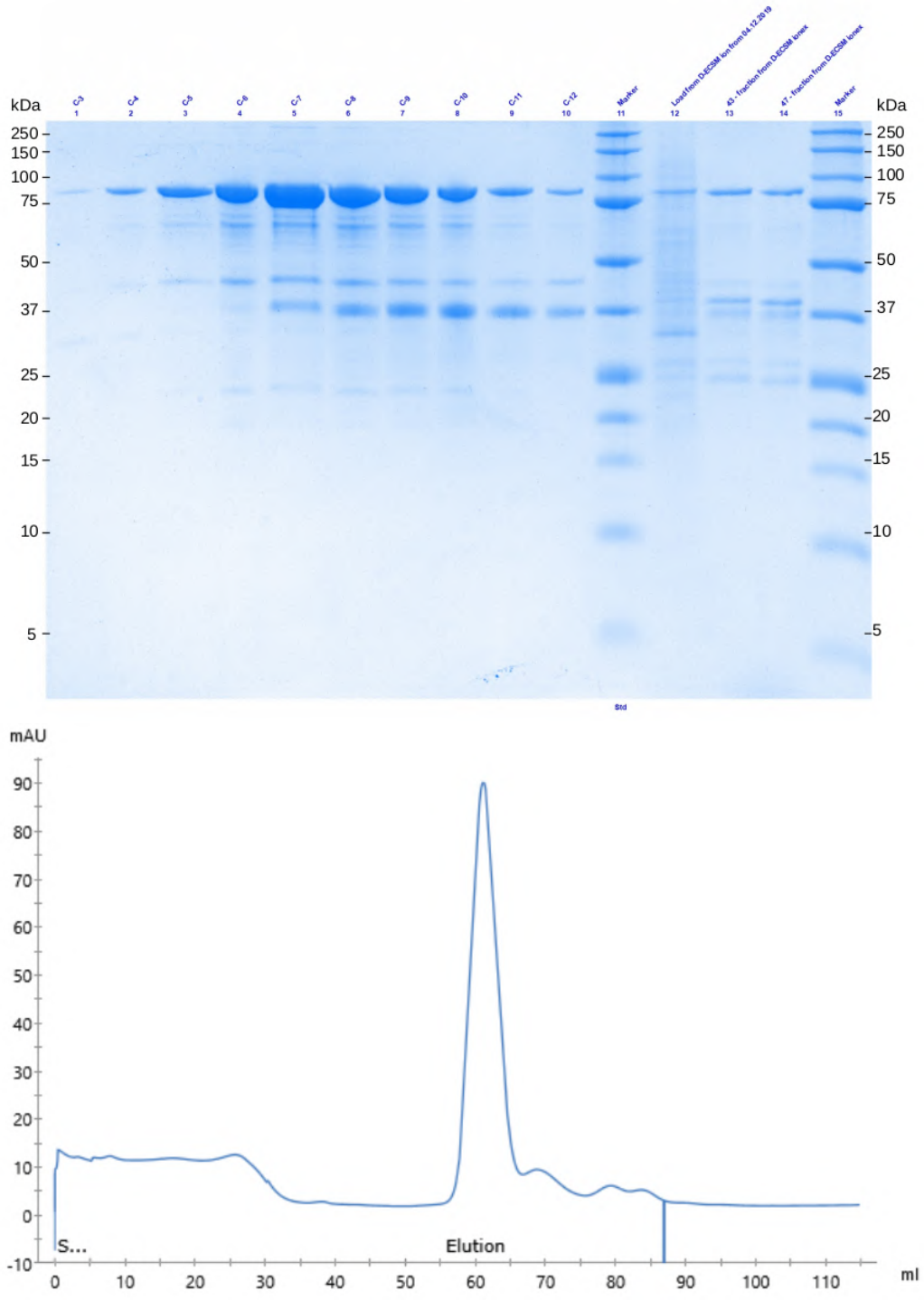

Figure 6.18: Fractions of perdeuterated E573Q mutant of ECAO obtained after gel filtration using superdex-200 column (top) and the corresponding UV_{280} trace (bottom).
Wells 12 to 15 (top) contain some of the fractions obtained during ion exchange chromatography of E573Q mutant of ECAO.

Table 6.4: Crystallization scheme 1 - optimization 1 for perdeuterated E573Q mutant of *E.coli* copper amine oxidase

	AS				
		1.4	1.6	1.8	2.0
Concentration of sodium citrate (M)	0.10		+	+	+
	0.20		+	+	
	0.25	+	+	+	

As seen in table 6.4, the number of hits were higher for perdeuterated E573Q mutant of ECAO than for perdeuterated wild-type ECAO. However, the best crystals were obtained in:

- 1.8 M ammonium sulphate and 0.10 M sodium citrate
- 1.8 M ammonium sulphate and 0.20 M sodium citrate
- 1.6 M ammonium sulphate and 0.20 M sodium citrate
- 1.6 M ammonium sulphate and 0.25 M sodium citrate
- 1.8 M ammonium sulphate and 0.25 M sodium citrate

In general, it was observed that the increase in sodium citrate was found to promote crystal growth whereas increase in the concentration of ammonium sulphate caused a reduction in the crystal volume, possibly due to excessive nucleation.

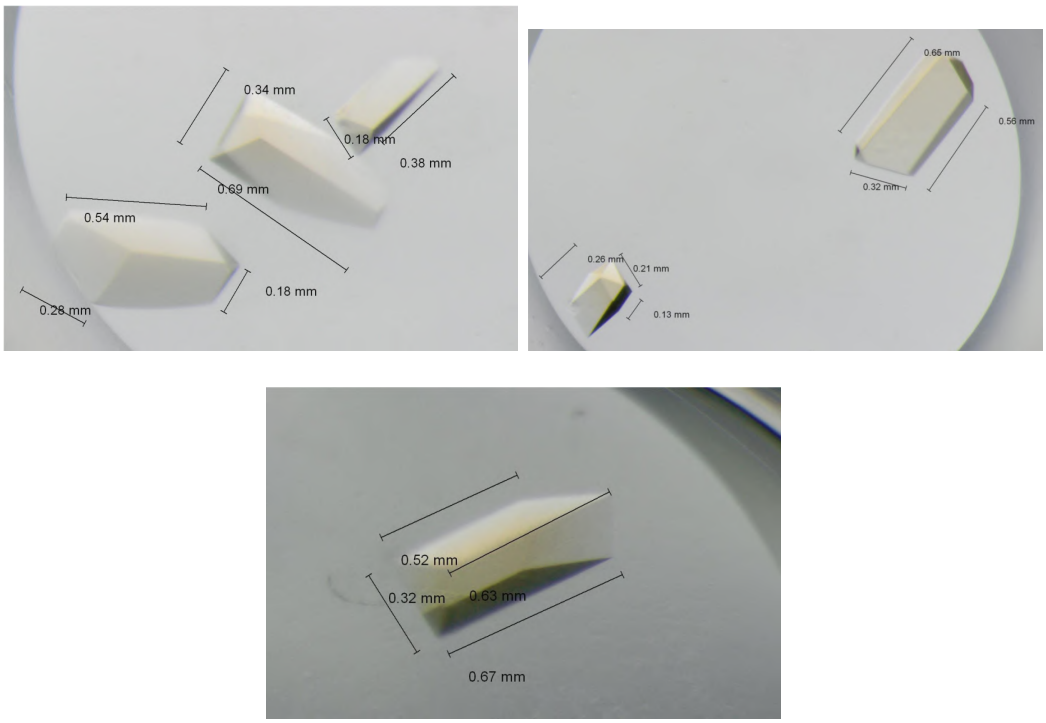

Figure 6.19: Crystals of perdeuterated E573Q mutant of ECAO obtained in 1.8 M ammonium sulphate and 0.1 M (top left), 0.2 M (top right) sodium citrate. Crystal obtained in 1.6 M ammonium sulphate and 0.2 M sodium citrate (bottom).

6.4 X-ray data collection and analysis of wild-type D-ECAO and H/D exchanged single mutant (E573Q) and double mutant (E573Q/I342F) H-ECAO crystals

6.4.1 X-ray structure of perdeuterated wild-type ECAO

Crystals of perdeuterated wild-type ECAO were used for the X-ray diffraction experiment at I03 beamline at the Diamond Light Source. We used the X-ray ray structure 1D6Z as the search model. The model was refined to R_{work} of 18.7 and R_{free} of 20.4 % (table 6.5). Assessment of the influence of perdeuteration on the structure of ECAO is important in order to differentiate the structural changes due to mutations from those due to perdeuteration. In order to assess the influence of perdeuteration, we superposed the hydrogenated (1D6Z) and the perdeuterated wild-type ECAO. The C_α RMSD averaged over all the residues was found to be 0.19 Å upon rigid superposition and 0.15 Å upon iteratively excluding the residues with highest RMSD. We also plotted the main chain and the side chain RMSD to study the influence of perdeuteration on the structure of ECAO. It can be seen in figure 6.20 that the atomic positions in case of perdeuterated ECAO are significantly different from those in case of hydrogenated ECAO for both chain A (blue) and chain B (green). This is particularly observed for residues close to the N and the C-terminal loops of the protein as expected. The RMSD in case of the two chains is also not strictly symmetric. This is especially observed for residues 50-200 and 400-470. This also includes two of the active site residues: TPQ-466 (2,4,5-trihydroxyphenylalanine quinone) which plays a crucial role in catalysis (cf. section 1.1 of chapter 1) and TYR-369 which stabilizes TPQ-466 by forming a short hydrogen bond with the O4 position of TPQ [211]. These structural alterations that result from perdeuteration must be taken into consideration when the datasets for the crystals of perdeuterated mutant ECAO are collected and analyzed.

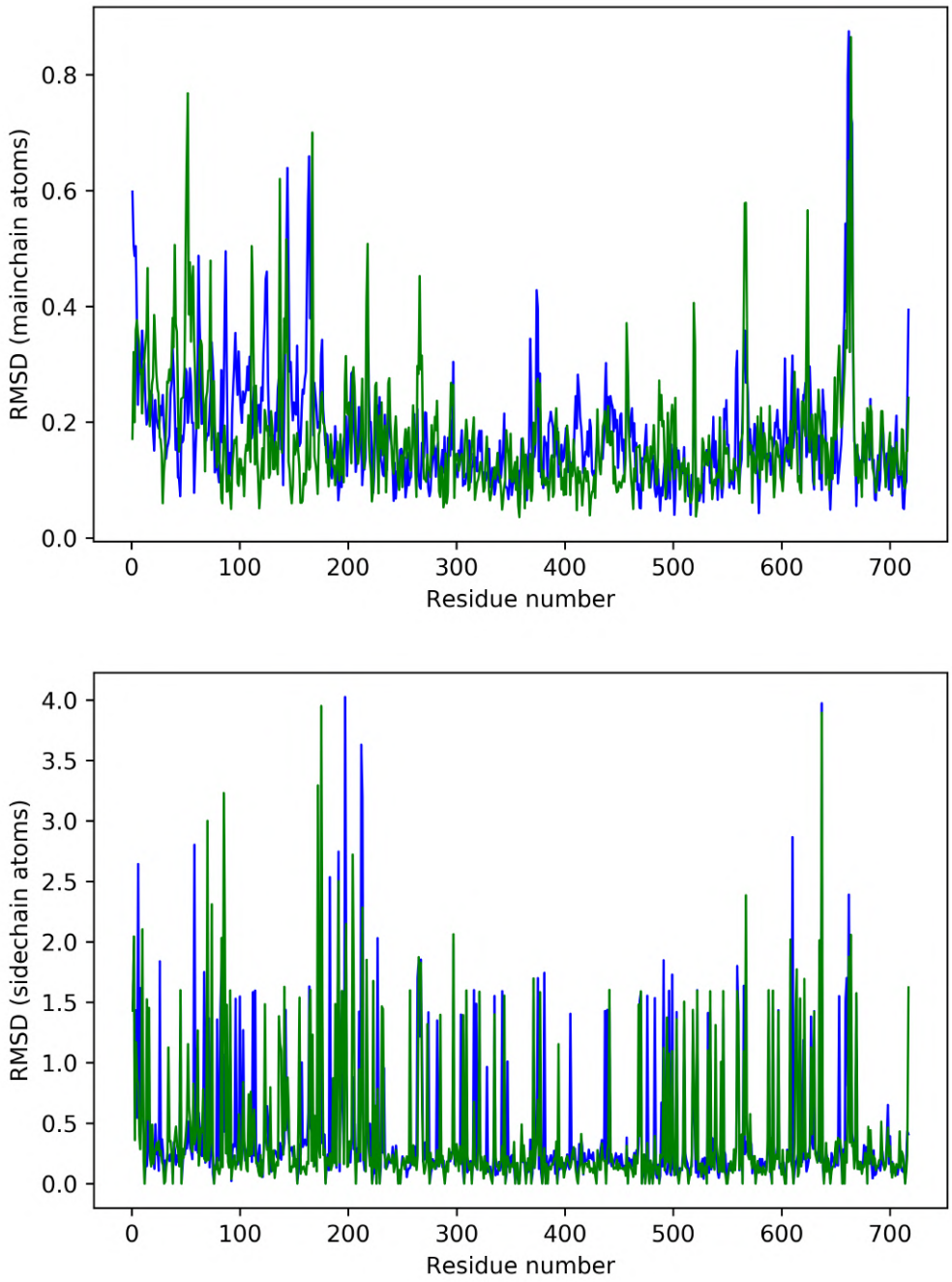

Figure 6.20: Main chain (top) and side chain (bottom) root mean square deviation for perdeuterated wild-type ECAO after superposition with hydrogenated wild-type ECAO.
The B-factors were calculated using *Baverage* module in CCP4 program [123]

Table 6.5: Crystallographic refinement statistics for perdeuterated wild-type ECAO. The numbers in parenthesis represent the high resolution shell

	D-ECAO (wild-type)
Wavelength (Å)	0.976
Resolution range (Å)	70.88 - 1.91 (1.978 - 1.91)
Space group	P 21 21 21
Unit cell	79.5257 134.713 166.69 90 90 90
Total reflections	1052516 (105952)
Unique reflections	139150 (13788)
Multiplicity	7.6 (7.7)
Completeness (%)	99.88 (99.77)
Mean I/sigma(I)	7.77 (1.16)
Wilson B-factor (Å^2)	31.25
R-merge	0.1305 (1.229)
R-meas	0.1401 (1.319)
R-pim	0.05066 (0.4743)
CC1/2	0.998 (0.751)
CC*	0.999 (0.926)
Reflections used in refinement	138986 (13760)
Reflections used for R-free	6921 (705)
R-work	0.1871 (0.2742)
R-free	0.2039 (0.3319)
CC(work)	0.949 (0.819)
CC(free)	0.938 (0.670)
Number of non-hydrogen atoms	11361
macromolecules	11347
ligands	14
Protein residues	1438
RMS(bonds) (Å)	0.008
RMS(angles) (°)	1.55
Ramachandran favored (%)	96.99
Ramachandran allowed (%)	2.87
Ramachandran outliers (%)	0.14
Rotamer outliers (%)	2.5
Clashscore	5.97
Average B-factor (Å^2)	46.13
macromolecules (Å^2)	37.13
ligands (Å^2)	55.19

6.4.2 X-ray structures of H/D exchanged crystals of E573Q and E573Q/ I342F mutants of ECAO

The crystals obtained in hydrogenated buffers (table 6.1) were used for the H/D exchange experiments. These crystals were equilibrated using identical buffer conditions in which these crystals were obtained. However, the buffers used were prepared with D_2O instead of H_2O. The crystals were transferred to freshly prepared buffers at least thrice in order to ensure the maximum possible exchange of H by D. The best crystals obtained were used for these experiments since these were also to be used for neutron diffraction experiments in future which typically require crystals having volumes in excess of 0.1 mm^3. The crystals used for H/D exchange protocol were between 0.1 and 0.2 mm^3 and will be used in future for the neutron diffraction experiments. Table 6.6, represents the refinement statistics for the H/D exchanged X-ray crystal structure of the E573Q/I342F mutant. The structure was refined to a final R_{work} of 21.0 and the R_{free} of 24.1 % (table 6.6). In case of E573Q mutant, these values are: 20.0 and 23.1 % respectively (table 2 in appendices).

Figure 6.21 shows the $2F_o$-F_c density for the two mutant structures of ECAO (E573Q and E573Q/I342F). It can be seen that the PHE-342 residue is well defined at a contour level of 1 σ. However, it is difficult to confirm the E573Q mutation since there is a difference of only an atom between Glutamate (E) and Glutamine (Q). Here, it should be noted that E573 is a Calcium binding site. This Calcium ion coordinates with other residues as well (GLU-647, TYR-667, ASP-670). However, Calcium is absent in case of Q573 [77], probably due to the net zero charge on Q573 as against a net -1 charge on E573. We superposed the wild-type crystal structure (1D6Z) with our structure in order to check the conformers of these key residues which coordinate with the Calcium ion in wild-type ECAO structure and to assess if Calcium ion is present in our structure.

As seen in figure 6.22, the side chain conformation sampled by GLU 672 in wild-type ECAO is such that its oxygen is one of the atoms which the calcium ion coordinates with. In the case of H/D exchanged E573Q mutant (current structure), this conformation is not sampled, which in combination with the $2F_o$-F_c peak confirms the absence of calcium ion. The reason for the absence of Calcium ion is that the mutated residue GLN-573 has a net zero charge at the pH used in the crystallization condition (pH 7.0). On the other hand, GLU-573 in wild-type structure has a net -1 charge and hence requires the presence of calcium ion to neutralize the charge.

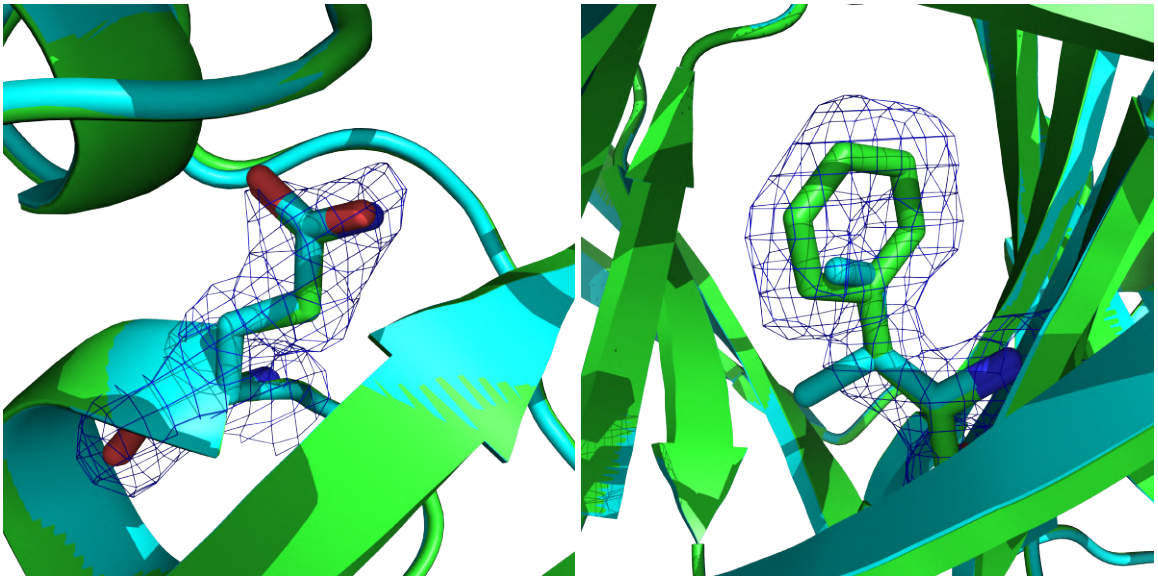

Figure 6.21: Superposition of the mutated Q573 (left) and F342 (right) residues of the refined E573Q mutant (left) and E573Q/I342F mutant (right) structures of *E.coli* copper amine oxidase (green) and the same residues from the active wild-type ECAO (1D6Z) (cyan). $2F_o$-F_c density for the mutated residues is shown at a contour level of 1 σ.

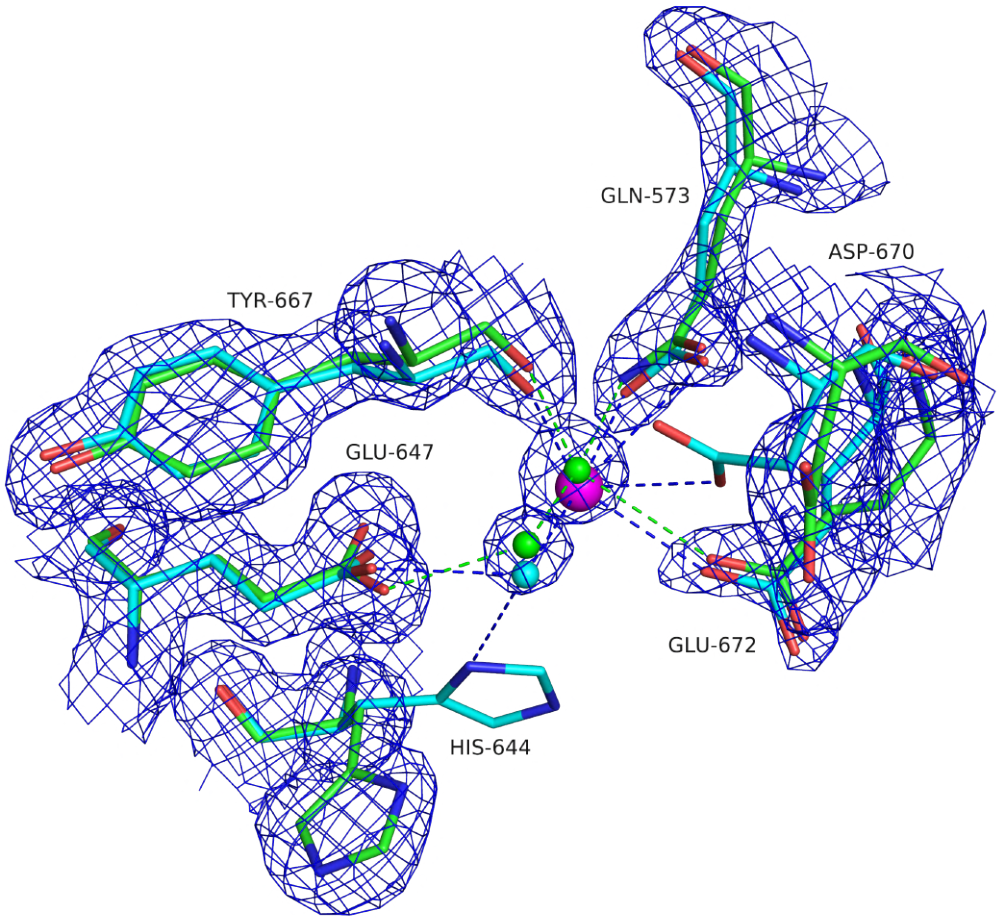

Figure 6.22: Superposition of wild-type (PDBID 1D6Z, cyan) and H/D exchanged E573Q mutant (current structure, green) ECAO. The contacts for the mutant are shown in black and those for the wild-type ECAO are shown in blue color. Calcium ion is shown as sphere (magenta). $2F_o$-F_c density for the mutant structure is shown at a contour level of 1 σ.

Table 6.6: Crystallographic refinement statistics for H/D exchanged E573Q/I342F mutant of ECAO
The numbers in parenthesis represent the high resolution shell

	ECAO (E573Q/I342F mutant)
Wavelength (Å)	0.985
Resolution range (Å)	83.56 - 1.57 (1.65 - 1.57)
Space group	P212121
Unit cell	79.85 134.9 167.1 90 90 90
Total reflections	1229538 (247030)
Unique reflections	247030 (36078)
Multiplicity	5.0 (5.0)
Completeness (%)	97.9 (97.0)
Mean I/sigma(I)	20.8 (5.5)
Wilson B-factor (Å^2)	20.21
R-merge	0.052 (0.176)
R-meas	0.058 (0.196)
R-pim	0.025 (0.804)
CC1/2	0.997 (0.935)
CC*	1 (0.99)
Reflections used in refinement	247030 (36078)
Reflections used for R-free	12283 (1801)
R-work	0.2100 (0.3205)
R-free	0.2406 (0.3720)
CC(work)	0.988 (0.811)
CC(free)	0.975 (0.799)
Number of non-hydrogen atoms	13076
macromolecules	11376
ligands	4
solvent	1696
Protein residues	1440
RMS(bonds) (Å)	0.006
RMS(angles) (°)	1.33
Ramachandran favored (%)	95.04
Ramachandran allowed (%)	4.54
Ramachandran outliers (%)	0.42
Rotamer outliers (%)	4.45
Clashscore	3.5
Average B-factor (Å^2)	22.6
macromolecules (Å^2)	20.4
ligands (Å^2)	26.7
solvent (Å^2)	37.6

6.5 Neutron diffraction experiments with the crystals of perdeuterated single mutant (E573Q) of ECAO

Some of the crystals of perdeuterated E573Q mutant of ECAO were tested for suitability on LADI-III at Institut-Laue Langevin (ILL), Grenoble. Figure 16 in appendices shows the diffraction obtained for two of the largest crystals of mutant ECAO which we could obtain. The figure shows that there are only a handful of diffraction spots at low resolution. However, the spots are well ordered and do not show any visible defects like tailing. These observations show that although the crystals are well ordered, they are not sufficiently large to obtain a full neutron dataset. The fact that there are two long axes in these crystals (table 6.6) adds to the problem since the total number of unit cells is proportionately smaller. Diffraction experiments at LADI-III typically require a minimum of 10^{14} unit cells. In order to fulfill these requirements, attempts to grow larger crystals of perdeuterated wild-type and E573Q mutant of ECAO will be continued.

6.6 QENS studies on ECAO to understand the effect of E573Q and I342F mutations on the dynamics in solution

QENS datasets were collected at using the IN16B neutron backscattering spectrometer (ILL, Grenoble) at 295K for wild-type and double mutant (E573Q/I342F) of ECAO in order to assess the effect of mutations on the dynamics of ECAO. The top panel of figure 6.23 shows the comparison of the Lorentzian fits of the raw data obtained for the two samples. It can be seen that there is a significant difference between the two spectra.

The two figures in the bottom panel show a comparison between the QENS spectra for double mutant (E573Q/I342F) of ECAO at 295 K at two different q values. The q-dependence of the Lorentzian linewidth can be clearly seen from the increase in the linewidth at 1.79 compared to 0.69 Å^{-1}. We proceeded to calculate the center-of-mass diffusion coefficient. This was done in the same was as described in chapter 4. The center-of-mass diffusion coefficient was found to be $2.45\pm0.05\,\text{Å}^2$/nm for wild-type and $1.65\pm0.03\,\text{Å}^2$/nm for E573Q/I342F mutant of ECAO. Hence, the behavior of the two systems varied significantly in solution in picosecond-to-nanosecond time scale.

We proceeded to the assessment of the internal diffusion of the two systems. For this purpose, we used equation 4.5 and equation 4.3 as described in chapter 4. When the linewidths γ in equation 4.3 were not fixed, the fits were unstable for both the samples as in the case of ADC-APO. Ideally, we would need the IN5 linewidths for ECAO at this stage if we are to fix the linewidths as was done in the case of ADC-APO. However, we are yet to perform IN5 experiment on ECAO samples. Hence, we needed a reasonable guess of the linewidths in order to model the internal diffusion of ECAO. For this purpose, we used the IN5 linewidths obtained from ADC-APO sample in order to define reasonable bounds for the linewidth in equation 4.3. As seen in figure 6.24, we obtained a reasonable fit with a goodness of fit of ≈ 1.43 for both wild-type and double mutant of ECAO. The q-dependent reduction in $A_0(q)$ is observed for both the systems. However, this loss of information is faster in the case of the double mutant (E573Q/I342F) than in the case of wild-type ECAO which may indicate a possible induction of plasticity or disorder in the double mutant. As opposed to this observation, the center-of-mass diffusion was found to be faster for wild-type ECAO than that for the E573Q/I342F double mutant. It should be noted that these are preliminary trial fits and that the interpretation may change after we obtain the IN5 linewidths for the two samples.

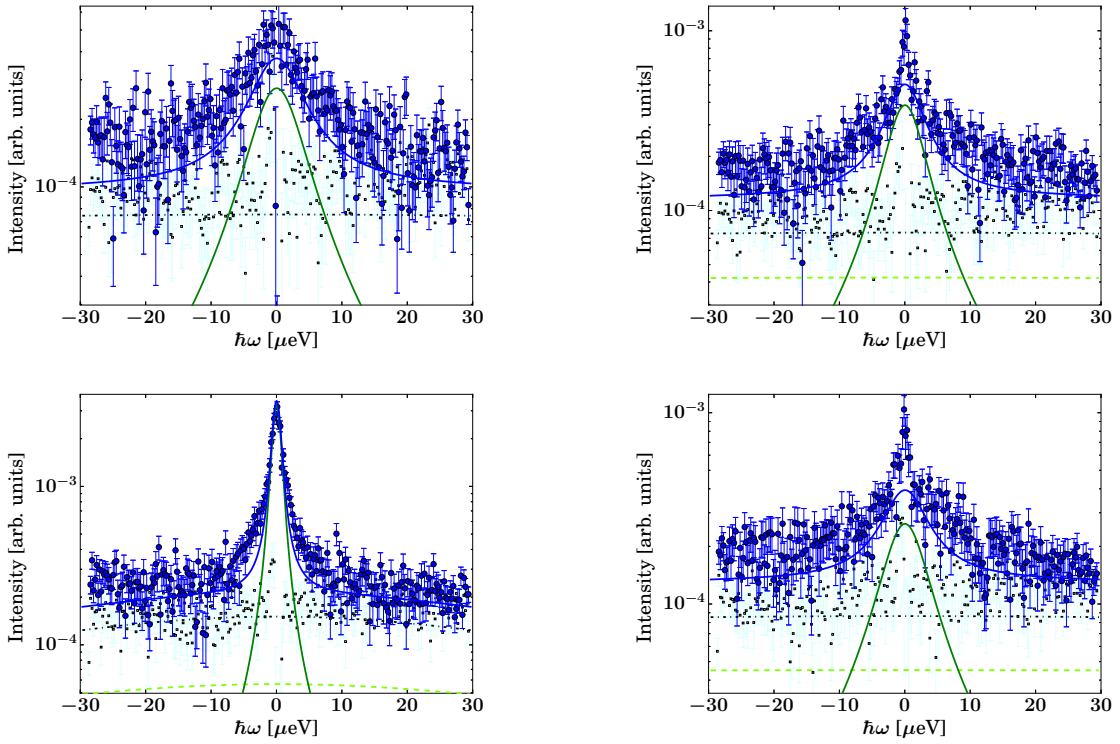

Figure 6.23: Example spectra (filled symbols) of wild-type (top left) and E573Q/I342F mutant (top right) of ECAO at $q = 1.57\,\text{Å}^{-1}$ and the example spectra of the E573Q/I342F mutant of ECAO at $q = 0.69$ (bottom left) and 1.79 (bottom right) Å^{-1} respectively recorded on IN16B at $T = 295\,\text{K}$, and fit using equation 4.3 (solid line superimposed on the symbols consisting of the components (other lines) as given by eq. 4.3). The lower light open symbols represent the pure solvent signal.

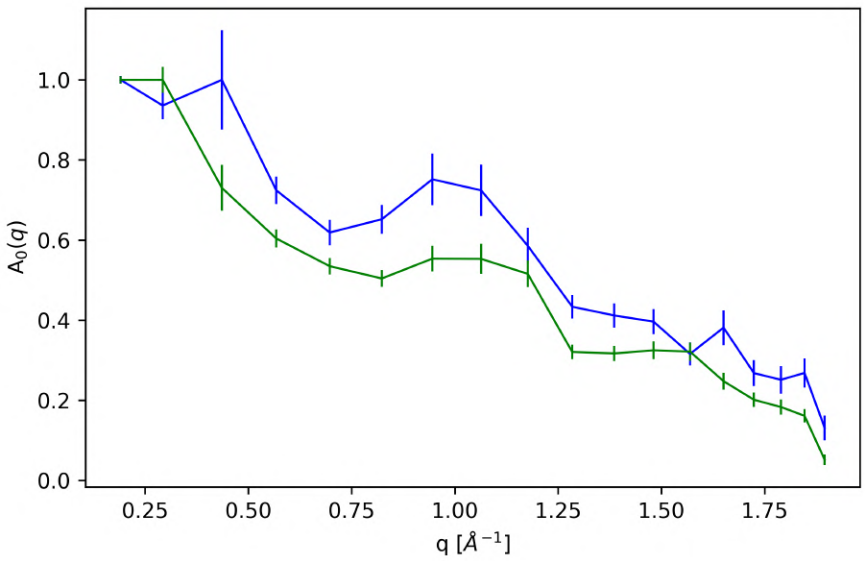

Figure 6.24: EISF $A_0(q)$ for wild-type (blue) and E573Q/I342F double mutant (green) of ECAO (top).

6.7 Dynamic light scattering studies on wild-type and E573Q/I342F double mutant of ECAO

We performed dynamic light scattering on wild-type and double mutant (E573Q/I342F) of ECAO in order to compare the translational diffusion coefficients of these two samples with those obtained from QENS (ref. previous section). The fitting of the data was carried out in the same way as that explained in chapter 4. It is clear from figure 6.26 that there is a small difference between the translational diffusion coefficients of the two samples. By fitting the diffusion coefficients for all five measured concentrations for both the wild-type and the double mutant of ECAO, the average diffusion coefficients were obtained for both samples at the low-concentration limit.

These diffusion coefficients are 3.85 ± 0.02 and 3.78 ± 0.01 Å^2/ns for wild-type and E573Q/I342F mutant of ECAO respectively indicating that the nearly dilute limit ($\varphi_t \approx 0$) diffusion coefficient is slightly higher for wild-type than that for E573Q/I342F mutant of ECAO. This trend is the same as that for center-of-mass diffusion coefficient obtained for both these systems from QENS (section 6.6). Hence, we conclude that the diffusive dynamics for the two systems differ significantly in the nanosecond regime for both internal and global dynamics whereas in the millisecond regime, there is a relatively smaller difference in the translational diffusion of the two systems.

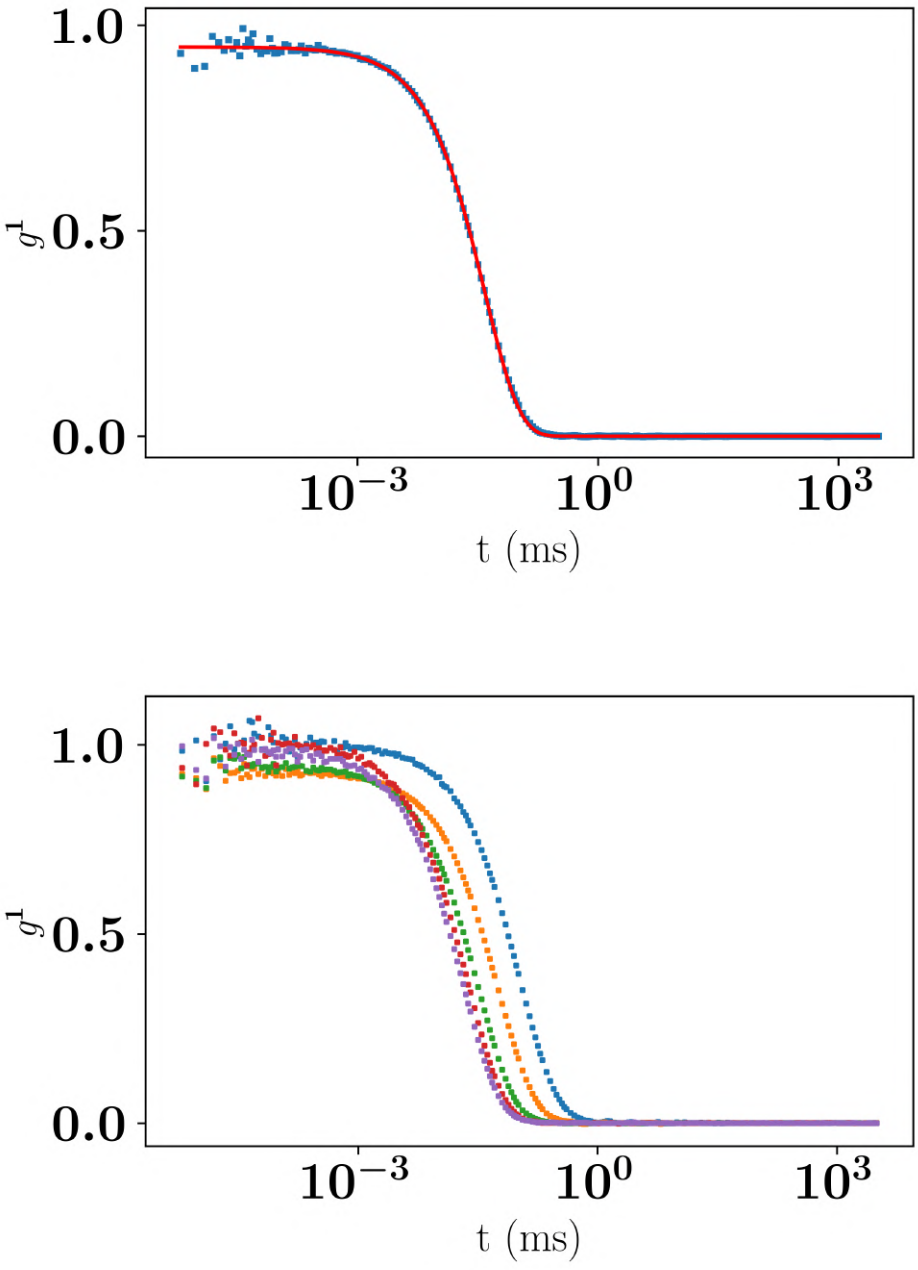

Figure 6.25: Top: DLS data recorded at $\Theta = 120°$ and single-exponential fit (equation 4.10) of the autocorrelation function g^{-1} for wild-type ECAO at $c_p = 8\,\mathrm{mg/mL}$ and $T = 298\,\mathrm{K}$ versus correlation time t. Bottom: DLS data recorded at 70 (violet), 80 (red), 100 (green), 120 (orange), 140 (blue) ° for the same sample. Symbols represent the data whereas the solid line represents the fit.

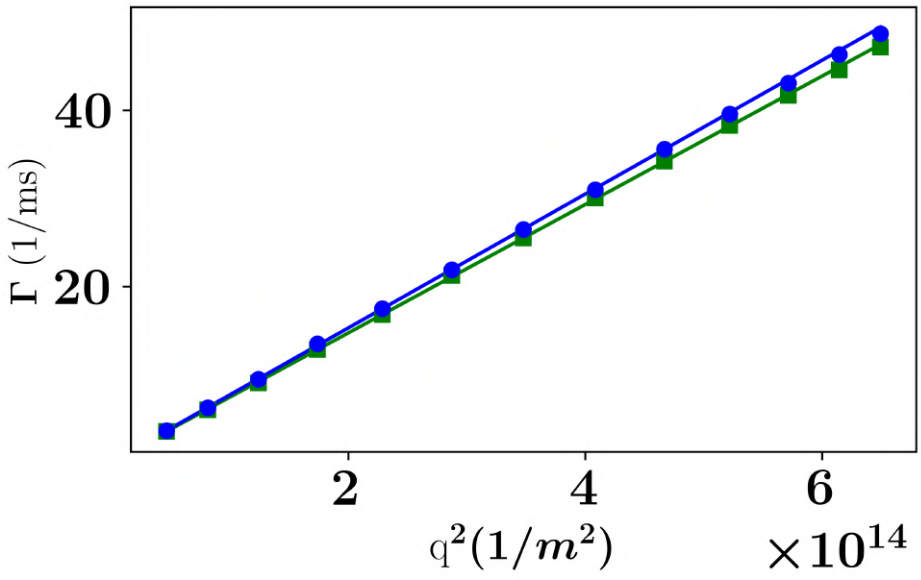

Figure 6.26: Linear fit of the decay rate obtained from the DLS data for E573Q/I342F mutant (squares) and wild-type (circles) ECAO at 8 mg/mL

6.8 Molecular dynamics simulations of wild-type, E573Q and E573Q/I342F mutants of ECAO

We performed molecular dynamics simulations with wild-type, single (E573Q) and double (E573Q/I342F) mutants of ECAO in order to analyze the possible changes in dynamics of ECAO as a result of the two mutations. For the preliminary analysis, we used the simulations which were carried out at 295K.

Figure 6.27 shows the mutation sites, the active site (dotted circle) with the three Histidines coordinating the Cu(II) ion and the key residues which influence the oxygen entry channel number 2 (cf. section 1.3 in chapter 1). Of these residues, TPQ-466 and TYR-381 (blue) are the most important in context of Oxygen entry. Residues ASP-383 and MET-699 (magenta) are found to have altered rotamer conformations and show an increased flexibility in the both the mutant structures [77]. It has been proposed that these residues may also influence the oxygen trafficking as they are in close proximity to the other residues lining the oxygen entry channel [77]. Hence, our preliminary analysis of MD simulation trajectories will focus on some of the key observables based on the proposed effects on dynamics. Figure 6.27 shows these key observables. As seen in the figure, the residues MET-699 and ASP-383 are in the vicinity of both Oxygen and substrate entry sites. In our analysis here, we will mainly focus on these residues along with TPQ-466 and TYR-381 which also play an important role in 'gating' of the substrate [211].

We first validated the copper parametrization in three systems by plotting the distances between the Nitrogens of the three Histidines (HIS-524, HIS-526, HIS-689) at the active site and the Cu(II) ion for both the subunit A of ECAO to ensure a stable active center. As seen in figure 6.28, the distance stays approximately at 1.9 Å throughout the production simulation, and hence agrees with our the observations from our crystal structure of wild-type ECAO (figure 6.4). The bottom-right figure shows the structural model of the copper center containing the three Histidines and the Cu(II) ion which coordinates the Nitrogens of these Histidines [57]. Wild-type ECAO structure solved and refined by me was used for producing this model.

In order to understand the dynamics of TPQ and whether it stays in on or off-copper conformation, we analyzed the distance between Cu(II) and the O4 atom of TPQ. The conformation of TPQ was found to be off-copper for most of the production simulation run for wild-type ECAO whereas for the E573Q and E573Q/I342F mutants, it was found to be switching between on and off-copper conformations (figure 18) as opposed to the on-copper conformation seen in the X-ray crystal structure [77]. We analyzed the global changes in the dynamics of all three systems by

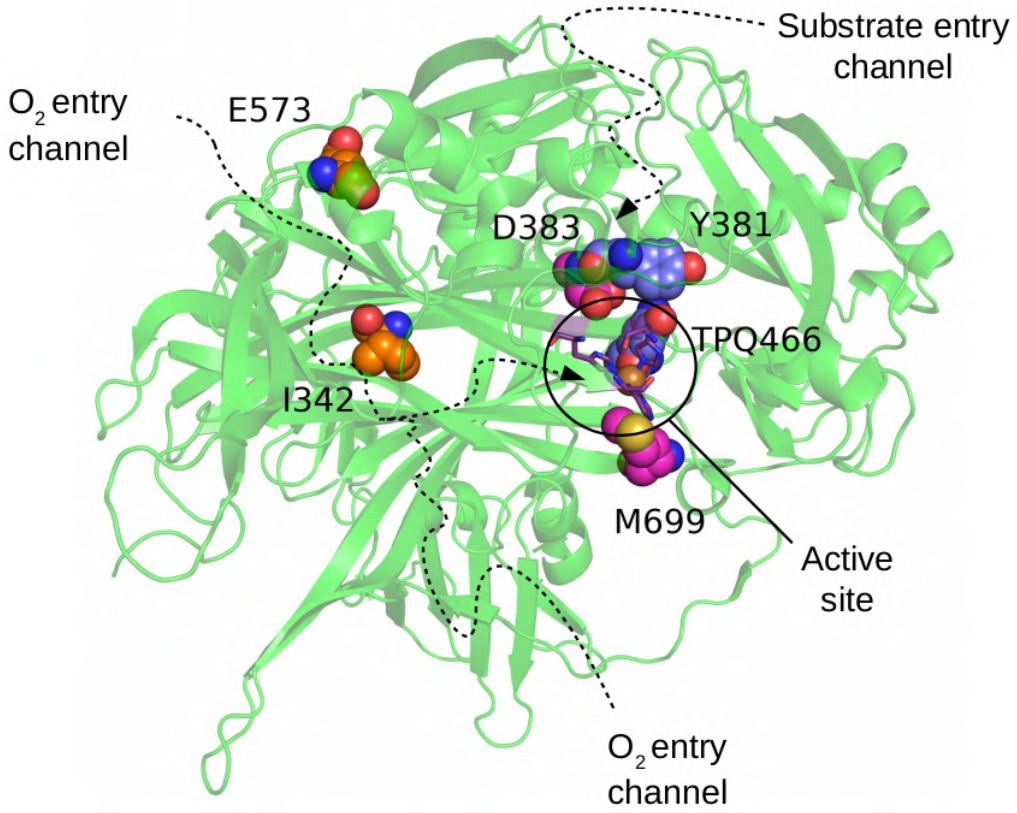

Figure 6.27: Residues shown in orange (I342) and green (E573) are the mutation sites. The dotted circle shows the active site. Residues Y381 (blue), TPQ-466 (blue) are the key residues important for the entry of molecular oxygen via the pathway which is shown as dotted arrow whereas residues D383 and M699 are those which have been reported to have altered rotamer conformation in the crystal structures of E573Q and E573Q/I342F mutants.
The model is generated using PyMol [44].

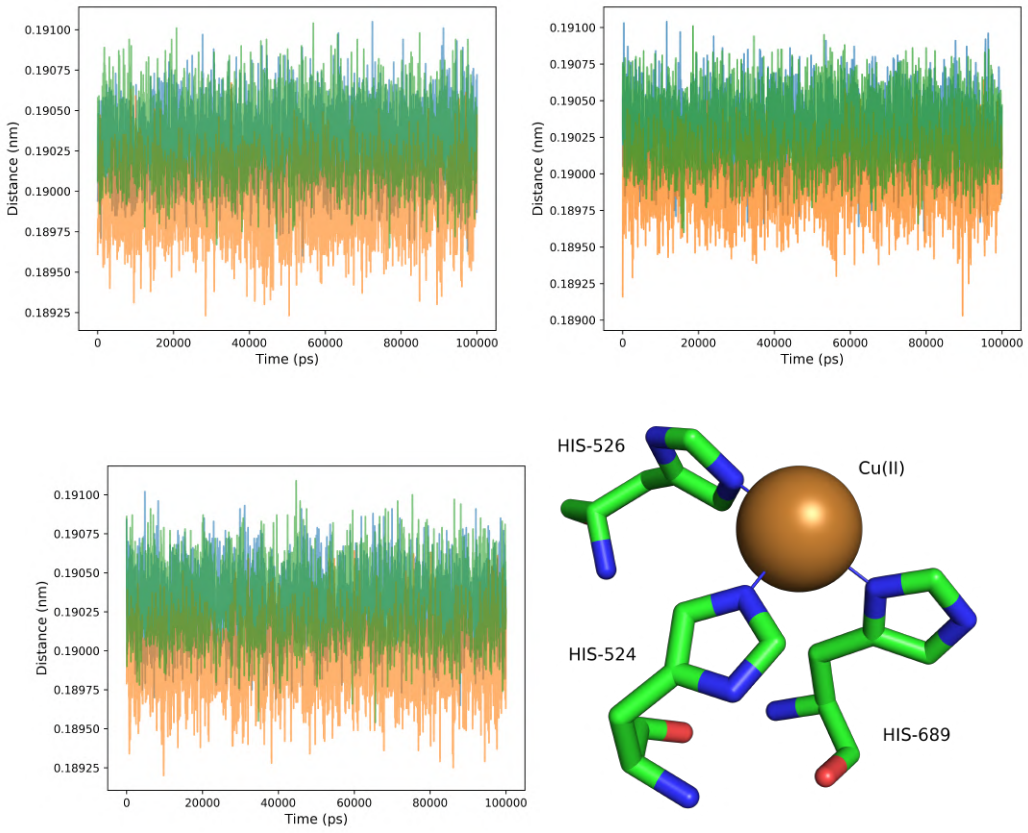

Figure 6.28: Distance between Cu(II) and the Nitrogens from HIS-524 (blue), HIS-526 (orange), HIS-689 (green) for wild-type (top-left), E573Q (top-right), E573Q/I342F (bottom-left) mutants of ECAO and the example figure for copper-center in ECAO (PDBID 1D6Z) (bottom-right)
The model on bottom-right is rendered using PyMol software [44]

monitoring the 'radius of gyration' against the simulation time. As seen in figure 17 in appendices, the fluctuations in radius of gyration are similar for the three systems although the mutant enzyme systems undergo a slight compaction (≈ 0.2Å) with time.

6.8.1 Conformational flexibility of ASP-383 and MET-699

We analyzed the conformational dynamics of specific residues which were found to have altered conformation in the X-ray crystal structures of mutant ECAO [77]. In these crystal structures, the residues which were mainly found to be affected in the mutant structures were ASP-383 and MET-699. The C3 from MET-699 was found to be oriented towards C3 position of TYR-466.

In order to get a deeper insight into the orientational preference of MET-699, we calculated the χ1-χ2 and ϕ-ψ ramachandran plots for MET-699 (figure 6.29). Here, the ϕ dihedral is defined as $C_{(N-1)} - N - C_\alpha - C$ angle whereas the ψ dihedral is defined as $N - C - C_\alpha - C - N_{(N+1)}$ angle. It can be seen that the main chain conformational dynamics represented by ϕ and ψ ramachandran plot of MET-699 was found to be very similar for the three systems. In case of side chain dynamics reflected by χ1 and χ2 dihedrals, MET-699 in the double (E573Q/I342F) mutant of ECAO was found to be more flexible than that in wild-type ECAO (figure 6.29).

We then analyzed the dynamics of ASP-383 as this residue is present in close proximity to the active site. In the crystal structures of the mutant ECAO, ASP-383 was found to adopt a conformation similar to that found in the case of on-copper conformation of TPQ/TYR [77]. As seen in figure 6.30, a change in both main chain (ϕ-ψ) and side chain (χ1-χ2) conformational dynamics of ASP-383 was observed. The rotamers were found to show orientational preference in case of E573Q and E573Q/I342F mutants.

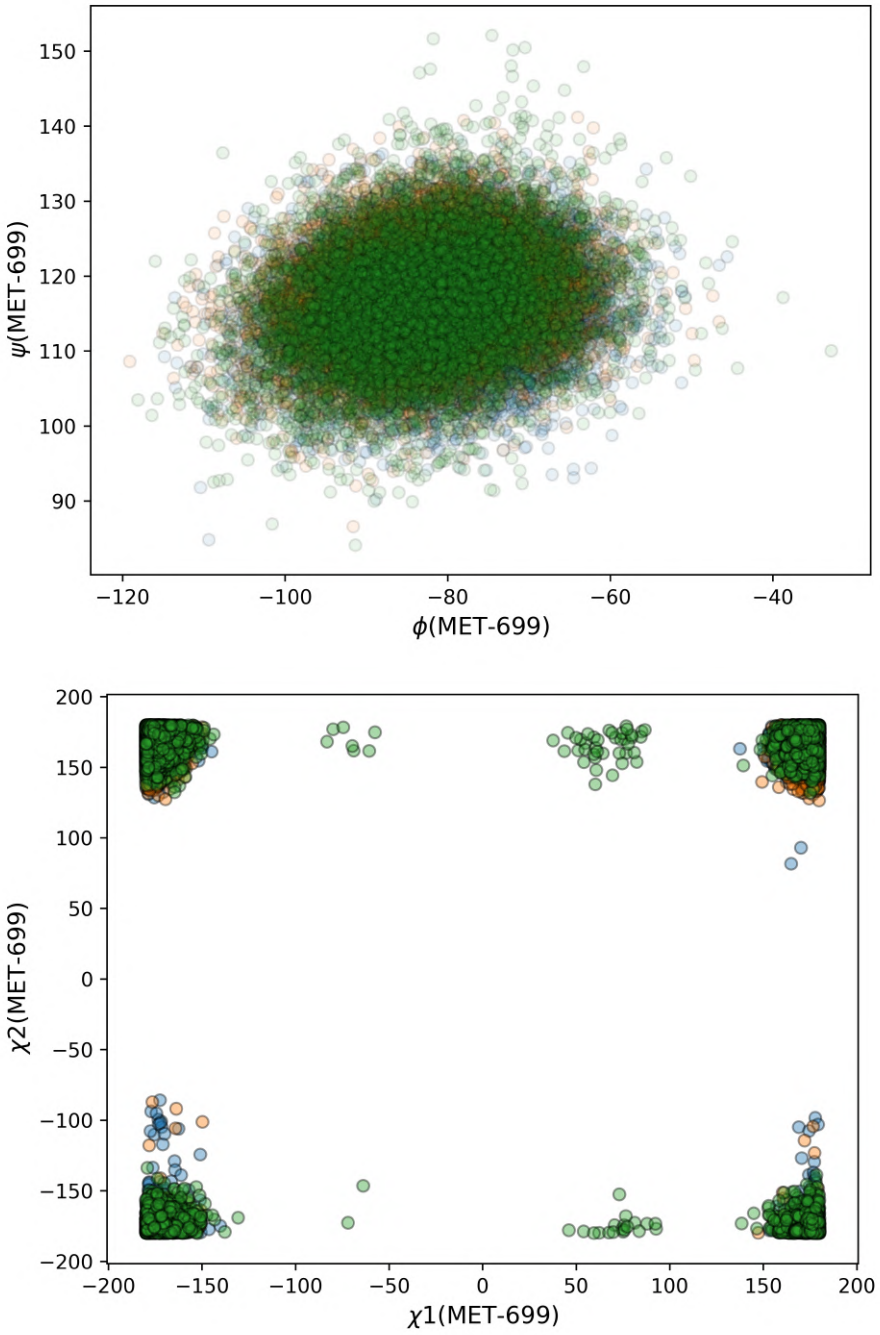

Figure 6.29: ϕ-ψ (top) and χ1-χ2 ramachandran plots for MET-699 from wild-type (blue), E573Q (orange) and E573Q/I342F (green) mutants of ECAO
Subunit A was used for plotting.

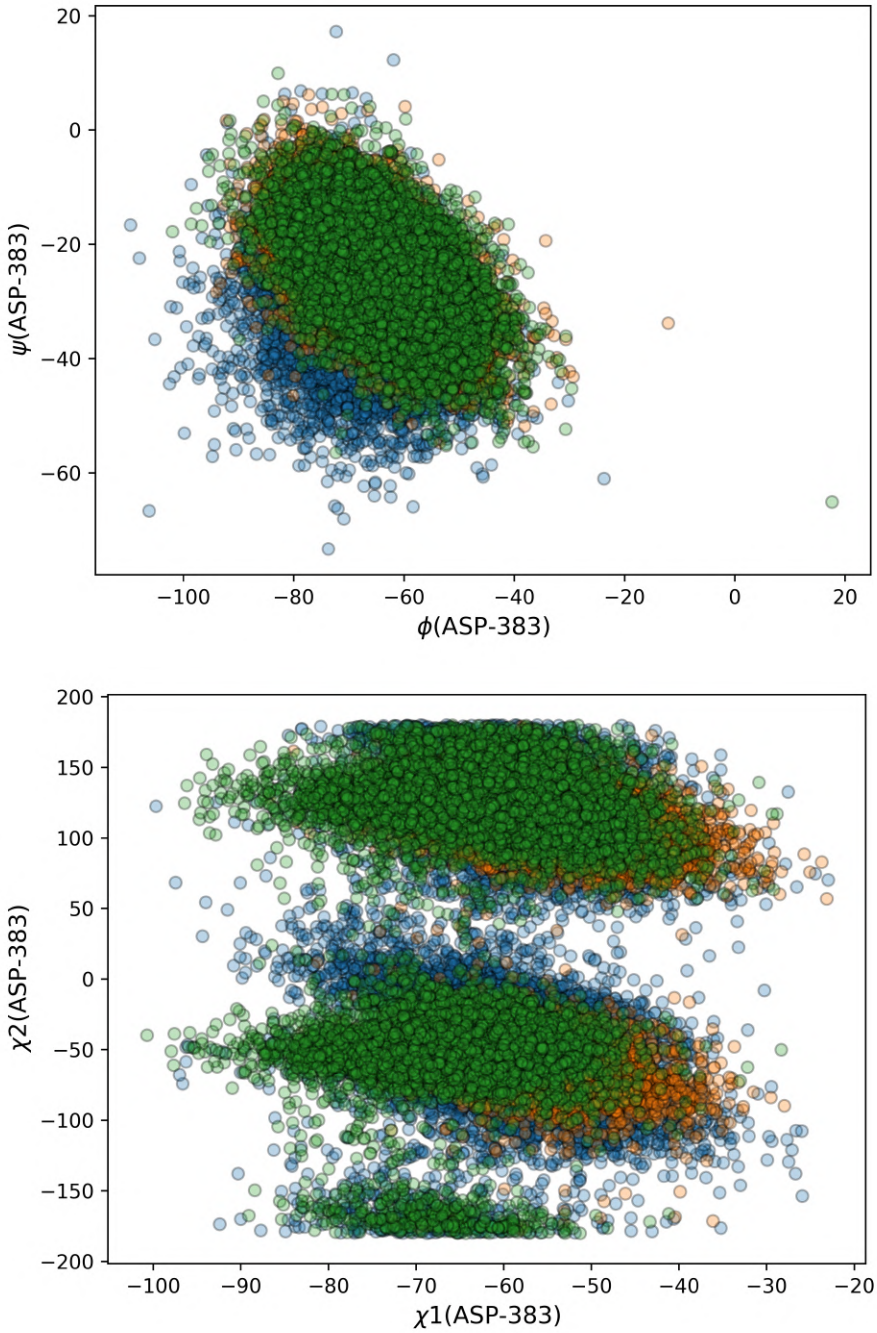

Figure 6.30: ϕ-ψ (top) and χ1-χ2 ramachandran plots ASP-383 from wild-type (blue), E573Q (orange) and E573Q/I342F (green) mutants of ECAO
Subunit A was used for plotting.

6.8.2 Effect of E573Q and I342F mutations on the dynamics and hydration of TPQ/TYR-466

It has been proposed that the E573Q mutation also affects the dynamics of TPQ/TYR [77]. In order to explore the dynamics of TPQ, we analyzed the MD trajectories further. In case of the crystal structures of E573Q and E573Q/I342F mutants, the conformation of TPQ/TYR-466 was found to be on-copper (inactive). Distance between O4 of TPQ/TYR-466 and Cu(II) was analyzed and it was found that wild-type ECAO was in the off-copper conformation during all of the production run in case of subunit A and most of the production run in case of subunit B (figure 18 in appendices). In case of mutants, the distances were found to be different for the two subunits indicating that TPQ/TYR-466 switches between 'on' and 'off-copper' conformations which is contrary to the results from X-ray crystal structures of the two mutants [77]. O4 of TPQ/TYR-466 was found to be closer to Cu(II) of subunit A of the mutants than that of wild-type ECAO. In case of subunit A of E573Q/I342F mutant, TPQ/TYR was present in an on-copper conformation throughout the production run whereas in case of E573Q, the average distance was ≈ 4.5 Å. In case of subunit B, the average distance was similar for wild-type and E573Q/I342F mutant whereas for E573Q mutant, it was slightly lower (≈ 4.0 Å). This observation shows that O4 of TPQ being in close proximity to Cu(II) is only a qualitative indicator of reduction in the activity. The ϕ-ψ distributions for TPQ/TYR-466 residues for all the three systems were plotted and were found to change significantly in case of the mutant enzymes. In case of wild-type and E573Q/I342F mutant of ECAO, the main chain corresponding to TPQ/TYR was found to be more confined than that in case of the E573Q mutant (figure 6.31). In general, this trend shows that TPQ/TYR in case of E573Q mutant, explores a larger the conformational subspace in comparison with wild-type and E573Q/I342F mutant of ECAO. The similarity in the dynamical behavior of wild-type and E573Q/I342F mutant of ECAO is in direct agreement with the previously reported THz-TDS studies in which an increase in the THz absorption coefficient of the double mutant and wild-type ECAO is observed in comparison to E573Q mutant of ECAO [212]. Hence, in the case of both the mutants, the conformational subspace explored by TPQ/TYR-466 is different from that in the case of wild-type ECAO. This shows altered conformational dynamics of TPQ/TYR-466 as a result of E573Q and I342F mutations.

The flexibility of TPQ/TYR is directly related with the conformational dynamics of TYR-381 which is the so called 'gating residue' as its conformation changes depending on the presence of the absence of the substrate [211]. Thus, we explored the correlated dynamics between TPQ/TYR-466 and 'the gating residue' TYR-381.

We analyzed the C_α-C_α distance between TPQ/TYR-466 and TYR-381 figure 6.32. It can be seen that this distance is lower for the two mutants when compared with the wild-type ECAO.

Since TYR-381 and TPQ/TYR-466 residues are a part of the oxygen and amine substrate entry channels, the change in conformational flexibility of TPQ/TYR-466 can also affect the oxygen and the substrate entry which maybe directly correlated with the inactivation of ECAO [72, 77]. We analyzed the hydration dynamics of TPQ/TYR residues for all the systems by calculating the number of water molecules within 3.5 Å of TPQ/TYR-466. The hydration number of TPQ/TYR-466 was found to change significantly in case of wild-type, single and double mutant of ECAO. In case of both TPQ and TYR-466 residues from the mutants, around 4 fold reduction in the number of water molecules was observed (figure 6.32). Hence, we could observe that the E573Q and I342F mutations are found to affect both, the dynamics of TPQ/TYR-466 as well as the conformational flexibility of MET-699 and ASP-383. The former results in a reduction in the C_α-C_α distance between TPQ/TYR-466 and TYR-381 and thus results in reduced solvent access to the active site. This observation is important in context of the reduced activity of E573Q mutant as the reduced solvent access can also be an indicator of the reduced access to molecular oxygen and to the substrate as a result of the changes in the correlated dynamics of TYR-381 and TPQ/TYR-466. The latter changes such that MET-699 main chain dynamics remain unchanged whereas its side chain is more flexible in case of the mutants indicating the possible effects on the Oxygen entry pathways (figure 6.27). As opposed to MET-699, the side chain dynamics of ASP-383 are found to be significantly affected as a result of the mutations such that there is an orientational preference in case of mutants resulting in relatively rigid side chain dynamics when compared with wild-type ECAO. Overall, we observe an increase in the flexibility of some of the key residues which are essential for the activity of ECAO and oxygen entry (MET-699, TPQ/TYR-466) whereas in case of ASP-383, we observe a reduction in the flexibility. The increase in flexibility is in agreement with interpretation of the EISF for wild-type and E573Q/I342F mutant of ECAO (figure 6.24).

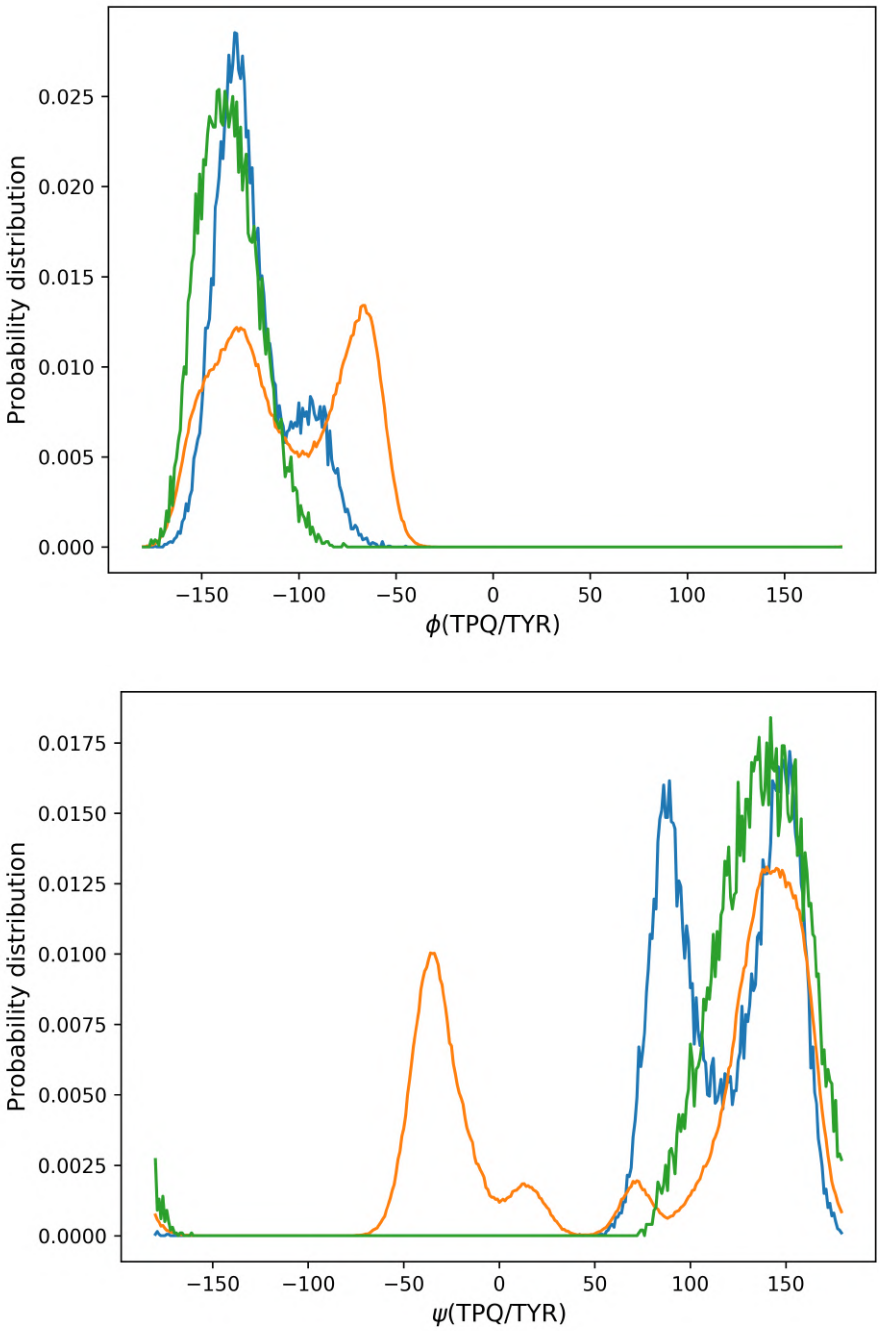

Figure 6.31: ϕ (top) and ψ (bottom) dihedral angle distribution for TPQ/TYR-466 in wild-type (blue), E573Q (orange) and E573Q/I342F (green) mutants of ECAO. Subunit A was used for plotting.

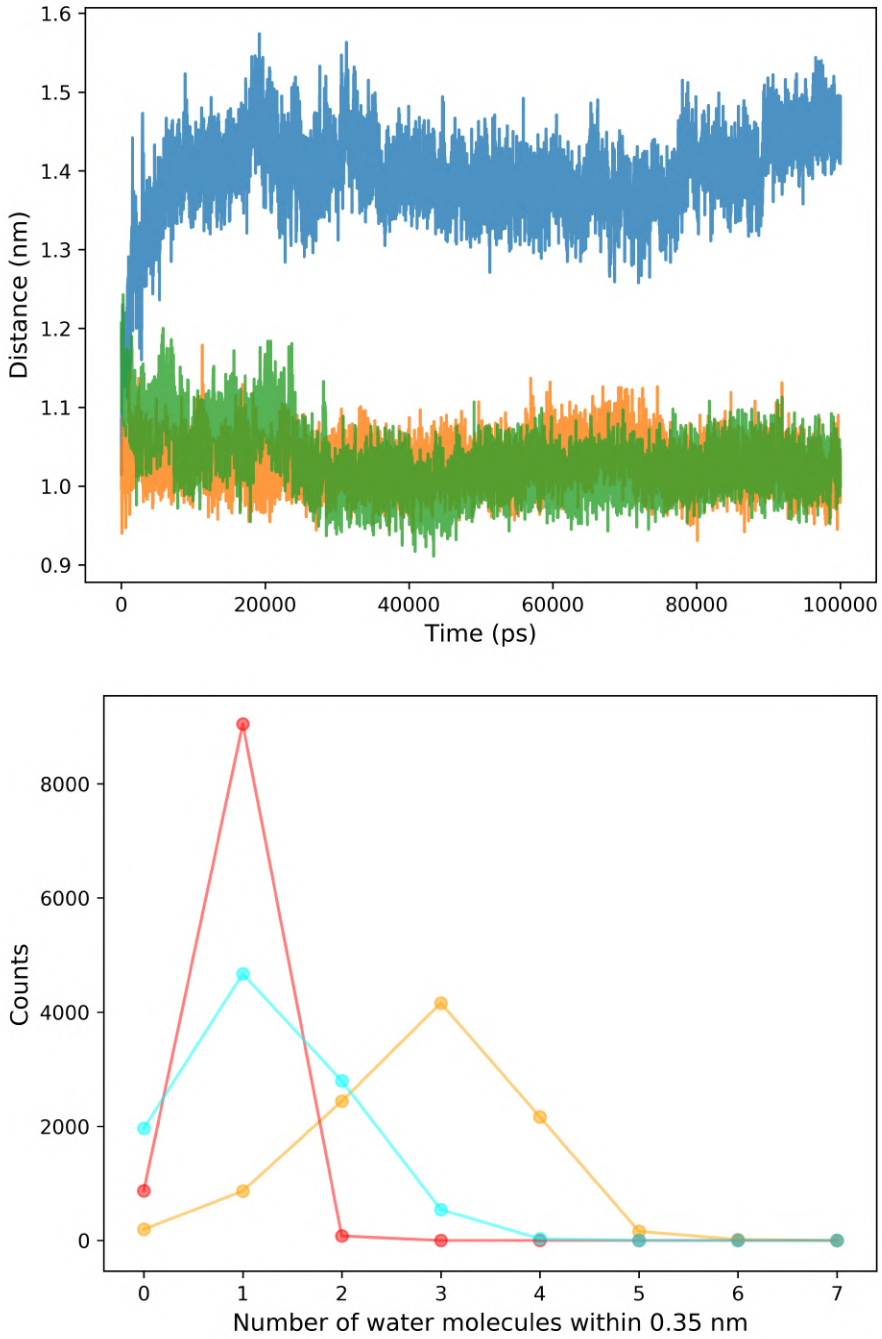

Figure 6.32: C_α-C_α distance between TPQ/TYR-466 and TYR-381 (top) and probability distribution for the number of water molecules within 3.5 Å from TPQ/TYR-466 in case of wild-type (blue), E573Q (orange) and E573Q/I342F (green) mutants of ECAO throughout the production trajectory.
Subunit A was used for the analysis.

6.9 Discussion

We have investigated the effect of the two mutations (E573Q, I342F) on the structure and dynamics of ECAO using a combination of X-ray and neutron diffraction to understand the structural changes and using QENS, dynamic light scattering (DLS) and molecular dynamics simulations to probe the changes associated with the dynamics of ECAO. We also carried out the perdeuteration of wild-type and single mutant (E573Q) of ECAO. Perdeuteration was confirmed using electrospray ionization mass spectrometry (ESI-MS) which indicated the presence of caged deuterium in the perdeuterated wild-type ECAO. The mutation I342F in the H/D exchanged X-ray crystal structure of the double mutant was confirmed by calculating the $F_o - F_c$ omit map for the mutation I342F. The mutation E573Q was confirmed indirectly as a water molecule was found to interact with Q573 rather than a calcium ion as a result of the zero net charge on Q573. Neutron diffraction tests resulted in very few but discrete diffraction spots indicating that the crystals are well ordered but have to be $\approx$ 2 to 5 times larger in volume for a viable study. We will continue to attempt to grow larger crystals in order to collect a full neutron diffraction data on perdeuterated wild-type and single mutant (E573Q) of ECAO.

We used QENS in combination with DLS to compare the global diffusive dynamics of the two systems. QENS results also gave an insight into the internal diffusive dynamics of the E573/I342F mutant and wild-type ECAO. The global center-of-mass diffusion coefficients obtained from QENS and DLS follow the same trend but differ quantitatively. The nanosecond center-of-mass diffusion coefficients differ significantly for the two samples whereas the millisecond diffusion differs to a lesser extent. The trends indicate a slower apparent center-of-mass diffusion for the E573Q/I342F double mutant in comparison with the wild-type ECAO. On the other hand, the preliminary estimations of elastic incoherent structure factor (EISF) indicate that the internal dynamics of the two systems also differ with the internal motions in (E573Q/I342F) being more flexible than wild-type ECAO. This observation is in a general agreement with the crystallographic results which indicate that ASP-383, MET-699 and HIS-526 residues which are associated with the oxygen entry channel, display positional flexibility in E573Q and E573Q/I342F mutant structures of ECAO [77].

MD simulations were performed in order to explore if these changes observed in the X-ray crystal structures of the two mutants can be reproduced. We could also correlate the QENS data with the specific structural changes in both the mutants ECAO. Comparison of the results from MD simulations with the previously reported findings indicates that the conformational flexibility of both ASP-383 and MET-699 is signi-

ficantly affected which is in agreement with the crystallographic studies [77]. On the other hand, the on-copper conformation of TPQ/TYR-466 found in the mutant X-ray structures is not retained during the production run and rather switches between the on and the off-copper conformation in case of both the mutants. This observation re-emphasizes the fact that the conformational subspace which is sampled in a crystal structure can be significantly different from that sampled in solution. The TPQ-466 in case of wild-type was found to be in the off-copper conformation during most of the production simulation. The analysis of ϕ and ψ dihedral distribution of TPQ/TYR-466 in wild-type, E573Q and E573Q/I342F mutants indicate a similar conformational subspace explored by TPQ/TYR-466 in case of wild-type and double mutant. In case of E573Q mutant, this residue is found to be more flexible and explores a larger conformational subspace. This result is in a general agreement with the THz-TDS measurements on the crystals of ECAO which indicate an increase in the absorption coefficient of the single (E573Q) in comparison with the double (E573Q/I342F) mutant and wild-type ECAO, indicating an increase in the dynamics of E573Q mutant of ECAO on picosecond-to-nanosecond time scale [212]. We could also observe a change in the correlated dynamics of TPQ/TYR-466 and TYR-381 such that the C_α-C_α distance between these residues reduces significantly in the case of the two mutants leading to a reduction in the solvent accessibility to the active site. This can be directly correlated with a ≈ 3 fold reduction in hydration number of TPQ/TYR-466 in the case of the mutant enzyme systens. This result can be one of the explanations for the drastic reduction in the activity of the two mutants of ECAO. These observations indicate that the local, internal and global diffusive dynamics of ECAO in case of E573Q and E573Q/I342F mutants are significantly affected.

These results agree well with the X-ray crystallographic studies [77] in case of MET-699 but not in the case of ASP-383 and in addition, give a deeper insight into the possible mechanism by which the activity of E573Q and E573Q/I342F mutant enzymes reduces.

Chapter 7

Conclusion and summary

The aim of this project was to combine complementary techniques in such a way that the understanding of the structure and dynamics of enzyme systems is more comprehensive. This was studied in context of assessing the smaller structural changes leading to larger local and global changes in the dynamics of enzymes. To address this larger goal, we selected two well studied enzyme systems: ADC and ECAO, each with a different biological problem to address. We have used a range of spectroscopic and crystallographic techniques to probe the structure and dynamics of ADC and ECAO.

We have probed the changes in the structure and dynamics of ADC upon binding of D-Serine by combining the results from X-ray diffraction, quasi-elastic neutron scattering (QENS), dynamic light scattering (DLS) and molecular dynamics simulations. In the crystal structure of ADC-LIG complex, the C-terminal loop was found to undergo a conformational change and was found to open up in the ADC-LIG complex crystal structure solved by us. We also observed that D-Serine forms distinct hydrogen bonds with the active site residues. We then studied how this change in structure influences the internal and global dynamics on the picosecond to nanosecond times cale using QENS and on the millisecond time scale using DLS. The binding of the ligand was found to significantly influence the global diffusive dynamics of ADC such that ADC-LIG complex was found to form higher order oligomers larger than those observed for ADC-APO on the nanosecond and the millisecond time scales. This is reflected in our results in terms of a lower center-of-mass diffusion coefficient for ADC-LIG in comparison to that for ADC-APO. The elastic incoherent structure factor which contains information about the internal confined motions of the enzyme showed that the internal motions in ADC-APO are faster than the time scales probed by QENS whereas those of ADC-LIG complex lie within

the dynamic range of QENS. Thus, the investigation of dynamics of ADC in these two states showed that the binding of the ligand causes a major change in both the structure and aggregation dynamics upon binding of the ligand. Molecular dynamics simulations showed that the binding of D-Serine causes a change in both main chain and side chain conformational dynamics such that the the motions in ADC-LIG are more restricted than those in ADC-APO. These results corroborate the interpretation from elastic incoherent structure factor in QENS that the binding of the ligand changes the internal dynamics of ADC such that the internal motions are more confined in presence of D-Serine. Hence, we observed that ADC-APO covers a larger conformational space than ADC-LIG which is consistent with the distribution of $\chi 1$ dihedral for HIS-21 and both $\chi 1$, $\chi 2$ dihedrals for GLU-23 of C-terminal loop of ADC extracted from molecular dynamics simulations. Despite these observations regarding the global and internal dynamics, this estimation has to be assessed with a higher accuracy and needs the development of more suitable models which can describe internal motions which are more complex than brownian jumps between multiple states. Such large scale changes in aggregation dynamics due to relatively minor structural alterations in flexible parts of the protein have been reported previously [193, 194].

Neutron vibrational spectroscopy of hydrated ADC-APO and ADC-LIG powders shows a clear distinction in terms of specific vibrational peaks and in general, indicate that the binding of the ligand induces stabilization of ADC. The difference between the onset of the first and the second dynamical transitions of ADC-APO and ADC-LIG corroborates this observation and also indicates that ADC is overall, a rigid protein. However, the binding of D-Serine is found to stiffen ADC even more as seen by reduction in the mean squared displacements and increase in force constants of ADC-LIG complex obtained from Elastic Fixed Window Scans (EFWS). This observation agrees well with the solution QENS results which indicate slower internal motions in ADC-LIG than in ADC-APO. Change in softness of Hexokinase as a result of binding of Glucose has been studied using inelastic neutron scattering and was found to stiffen the protein [213]. Another study involving human acetyl cholinesterase and the ligand Huperzine A shows a result which deviates to a certain extent from our results [214]. The authors observe that the binding of Huperzine causes an increase in the average amplitude of atomic motions whereas the internal diffusive motions are slightly slowed down. While the former observation deviates from our result, the latter agrees with it. Lenton and others studied the effect of covalently bound ligand on Human-like Osteopontin (rOPN) Peptide [215]. They report that phosphorylation of the protein results in blocking of some of the side chain motions on the nanosecond time scale which is in line with our results which indic-

ate a significant influence on the internal dynamics of ADC in sub-nanosecond time regime. On the other hand, there was no influence of phosphorylation on overall vibrational dynamics observed through mean squared displacements from EFWS [215]. However, none of the structural changes directly influenced the overall size or shape of rOPN as reflected in 'radius of gyration' [215]. The strong agreement of our results with the other works, especially in context of the overall stiffening of the internal motions indicates that the slowing down of internal diffusive modes as a result of binding of the ligand is a common phenomenon [216].

In case of ECAO, we used QENS to study the effect of two mutations (E573Q, I342F) on the structure and dynamics of the enzyme as these mutations have been shown to severely affect the activity [77]. The mutations were found to significantly affect the global diffusive dynamics of ECAO on nanosecond timescale whereas the dynamics on millisecond timescale follow a similar trend but differ quantitatively. Preliminary analysis of the elastic incoherent structure factor shows that the internal diffusive dynamics of ECAO is affected by the two mutations in a such way that the flexibility of E573/I342F mutant ECAO is in general, higher than that of the wild-type ECAO. This result can be correlated with the a clear change in the backbone dynamics of TPQ/TYR-466, in both backbone and side chain dynamics of ASP-383. TPQ/TYR-466 and MET-699 are found to be more dynamic in case of the mutant structures of ECAO when compared with the wild-type ECAO whereas ASP-383 is found to be more rigid. While the former observation agrees with the crystallographic results [77], the latter deviates from it. These changes may be related with the Oxygen trafficking as MET-699 and ASP-383 as these residues are close to the residues that form one of these channels. On the other hand, the correlation between change in the distance between TYR-381-TPQ/TYR-466 C_α atoms and the reduced hydration in case of both the mutants may imply that the mutated residues result in inhibition of the entry of amine substrate and a reduced solvent accessibility. Hence, the combined results from QENS and MD simulations indicate the possible changes in substrate and Oxygen entry as well as the changes in the internal dynamics of the protein as a result of the two mutations (E573Q, I342F) A study similar to our work was conducted on Human Myelin Protein P2 to understand the effect of P38G mutation on the structure and dynamics of the protein [217]. The authors report that the P38G mutant is more dynamic than the wild-type protein as the contribution of atoms participating in local displacements was found to be higher for the mutant. Another study on photosynthetic reaction center of the purple bacterium *Rhodobacter sphaeroides* indicates that two non-functional mutations in the protein result in an increased local flexibility of the enzyme when compared to wild-type protein [85]. These observations are generally, in agreement with our studies

which also indicate the induction of flexibility due to the mutated residues which then affects the biological function. As these changes are biologically relevant, the results obtained are of general interest to both biological and pharmaceutical fields.

In the case of calculation of the center-of-mass diffusion coefficient which is directly related with the dynamics of proteins in their native environment, we have used models from colloid physics. Hence, understanding how the control parameters affect the physical properties of the biomolecules is also of interest to the field of colloid physics and chemical engineering aside from the biological significance of these studies.

This work, in combination with other related works presents some of the important results. These are of relevance to understanding how relatively small structural alterations, particularly in the disordered regions of the protein govern their internal and global dynamics. Like in the case of ADC, a small structural change can result in a change in aggregation dynamics on picosecond-to-millisecond timescale. These results can therefore potentially be of interest to the field of biology as aggregation plays a major role in several neurodegenerative diseases but its origin is still not completely known.

This is one of the few studies in which the dynamics of recombinantly expressed proteins have been investigated using a range of neutron spectroscopic, crystallographic and MD simulation methods rather than using commercially available protein samples. We could access the dynamics of ADC and ECAO at various time and lengthscales using a unique combination of techniques. We could cross-correlate these results between various spectroscopic techniques which were used to probe picosecond-to-millisecond dynamics. This revealed how the dynamics of these enzymes are correlated with their structure and also how the smaller structural changes which can be seen using crystallographic methods can cause larger changes in the dynamics in solution. In case of ADC, we could corroborate the crystallographically observed structural changes with MD simulations. The dynamics of ADC in aqueous solution and in hydrated powder was also studied using neutron spectroscopy methods which correlated well with the crystallographic and MD studies and helped disentangle the changes in (local and global) diffusive and vibrational modes of ADC as a result of binding of D-Serine. In the case of ECAO, we could also identify the changes in the dynamics of specific residues which agreed with the observations from QENS and DLS studies. Hence, we could successfully cross-correlate the results from a diverse range of spectroscopic and crystallographic techniques which helped us understand the structure and dynamics of the aforementioned enzyme systems in depth. The comprehensive understanding gained from these studies is ultimately important in context of understanding the structure-dynamics relationship in these

enzyme systems.

Bibliography

[1] CHRISTINE DEBOUCK. The hiv-1 protease as a therapeutic target for aids. *AIDS research and human retroviruses*, 8(2):153–164, 1992.

[2] Laurent Maveyraud and Lionel Mourey. Protein x-ray crystallography and drug discovery. *Molecules*, 25(5):1030, 2020.

[3] Bernhard Rupp. *Biomolecular crystallography: principles, practice, and application to structural biology*. Garland Science, 2009.

[4] Bertil Halle. Flexibility and packing in proteins. *Proceedings of the National Academy of Sciences*, 99(3):1274–1279, 2002.

[5] Dennis E Danley. Crystallization to obtain protein–ligand complexes for structure-aided drug design. *Acta Crystallographica Section D: Biological Crystallography*, 62(6):569–575, 2006.

[6] Ute Krengel and Anne Imberty. Crystallography and lectin structure database. In *Lectins*, pages 15–50. Elsevier, 2007.

[7] Amit Das, Smita Mahale, Vishal Prashar, Subhash Bihani, J-L Ferrer, and MV Hosur. X-ray snapshot of hiv-1 protease in action: observation of tetrahedral intermediate and short ionic hydrogen bond sihb with catalytic aspartate. *Journal of the American Chemical Society*, 132(18):6366–6373, 2010.

[8] Oksana Gerlits, Jianhui Tian, Amit Das, Paul Langan, William T Heller, and Andrey Kovalevsky. Phosphoryl transfer reaction snapshots in crystals insights into the mechanism of protein kinase a catalytic subunit. *Journal of Biological Chemistry*, 290(25):15538–15548, 2015.

[9] Ai Woon Yee, Matteo Aldeghi, Matthew P Blakeley, Andreas Ostermann, Philippe J Mas, Martine Moulin, Daniele de Sanctis, Matthew W Bowler, Chris-

toph Mueller-Dieckmann, Edward P Mitchell, et al. A molecular mechanism for transthyretin amyloidogenesis. *Nature communications*, 10(1):1–10, 2019.

[10] Julian C-H Chen, B Leif Hanson, S Zoë Fisher, Paul Langan, and Andrey Y Kovalevsky. Direct observation of hydrogen atom dynamics and interactions by ultrahigh resolution neutron protein crystallography. *Proceedings of the National Academy of Sciences*, 109(38):15301–15306, 2012.

[11] Fong Shu, Venki Ramakrishnan, and Benno P Schoenborn. Enhanced visibility of hydrogen atoms by neutron crystallography on fully deuterated myoglobin. *Proceedings of the National Academy of Sciences*, 97(8):3872–3877, 2000.

[12] Tim Gruene, Hinrich W Hahn, Anna V Luebben, Flora Meilleur, and George M Sheldrick. Refinement of macromolecular structures against neutron data with shelxl2013. *Journal of applied crystallography*, 47(1):462–466, 2014.

[13] William T Heller. Small-angle neutron scattering and contrast variation: a powerful combination for studying biological structures. *Acta Crystallographica Section D: Biological Crystallography*, 66(11):1213–1217, 2010.

[14] Takeshi Murakawa, Seiki Baba, Yoshiaki Kawano, Hideyuki Hayashi, Takato Yano, Takashi Kumasaka, Masaki Yamamoto, Katsuyuki Tanizawa, and Toshihide Okajima. In crystallo thermodynamic analysis of conformational change of the topaquinone cofactor in bacterial copper amine oxidase. *Proceedings of the National Academy of Sciences*, 116(1):135–140, 2019.

[15] Zygmunt S Derewenda and Peter G Vekilov. Entropy and surface engineering in protein crystallization. *Acta Crystallographica Section D: Biological Crystallography*, 62(1):116–124, 2006.

[16] Marcus D Collins, Chae Un Kim, and Sol M Gruner. High-pressure protein crystallography and nmr to explore protein conformations. *Annual review of biophysics*, 40:81–98, 2011.

[17] Peter J Steinbach, Richard J Loncharich, and Bernard R Brooks. The effects of environment and hydration on protein dynamics: a simulation study of myoglobin. *Chemical physics*, 158(2-3):383–394, 1991.

[18] Javier Pérez, Jean-Marc Zanotti, and Dominique Durand. Evolution of the internal dynamics of two globular proteins from dry powder to solution. *Biophysical journal*, 77(1):454–469, 1999.

[19] Frank Gabel. Protein dynamics in solution and powder measured by incoherent elastic neutron scattering: the influence of q-range and energy resolution. *European Biophysics Journal*, 34(1):1–12, 2005.

[20] Clare L Armstrong, Erik Sandqvist, and Maikel C Rheinstadter. Protein-protein interactions in membranes. *Protein and Peptide Letters*, 18(4):344–353, 2011.

[21] Quentin Berrod, Karine Lagrené, Jacques Ollivier, and Jean-Marc Zanotti. Inelastic and quasi-elastic neutron scattering. application to soft-matter. In *EPJ Web of Conferences*, volume 188, page 05001. EDP Sciences, 2018.

[22] S König, TM Bayerl, G Coddens, D Richter, and E Sackmann. Hydration dependence of chain dynamics and local diffusion in l-alpha-dipalmitoylphosphtidylcholine multilayers studied by incoherent quasi-elastic neutron scattering. *Biophysical journal*, 68(5):1871–1880, 1995.

[23] Takeshi Yamada and Hideki Seto. Quasi-elastic neutron scattering studies on hydration water in phospholipid membranes. *Frontiers in Chemistry*, 8:8, 2020.

[24] S Magazù. Iqens—dynamic light scattering complementarity on hydrogenous systems. *Physica B: Condensed Matter*, 226(1-3):92–106, 1996.

[25] PS Dubey, H Srinivasan, VK Sharma, S Mitra, V Garcia Sakai, and R Mukhopadhyay. Dynamical transitions and diffusion mechanism in dodab bilayer. *Scientific reports*, 8(1):1–16, 2018.

[26] Jan Peter Embs, Fanni Juranyi, and Rolf Hempelmann. Introduction to quasielastic neutron scattering. *Zeitschrift für Physikalische Chemie*, 224(1):5–32, 2010.

[27] H. Hartmann, F. Parak, W. Steigemann, G. A. Petsko, D. R. Ponzi, and H. Frauenfelder. Conformational substates in a protein: Structure and dynamics of metmyoglobin at 80 K. *Proceedings of the National Academy of Sciences*, 79(16):4967–4971, August 1982.

[28] W Doster, S Cusack, and W Petry. Dynamical transition of myoglobin revealed by inelastic neutron scattering. *Nature*, 337:754–756, 1989.

[29] Derya Vural, Liang Hong, Jeremy C Smith, and Henry R Glyde. Long-time mean-square displacements in proteins. *Physical Review E*, 88(5):052706, 2013.

[30] J. H. Roh, V. N. Novikov, R. B. Gregory, J. E. Curtis, Z. Chowdhuri, and A. P. Sokolov. Onsets of Anharmonicity in Protein Dynamics. *Physical Review Letters*, 95(3):038101, July 2005.

[31] Kathleen Wood, Douglas J. Tobias, Brigitte Kessler, Frank Gabel, Dieter Oesterhelt, Frans A A Mulder, Giuseppe Zaccai, and Martin Weik. The low-temperature inflection observed in neutron scattering measurements of proteins is due to methyl rotation: Direct evidence using isotope labeling and molecular dynamics simulations. *Journal of the American Chemical Society*, 132(14):4990–4991, 2010.

[32] Giorgio Schiró, Chiara Caronna, Francesca Natali, and Antonio Cupane. Direct evidence of the amino acid side chain and backbone contributions to protein anharmonicity. *Journal of the American Chemical Society*, 132(4):1371–1376, 2010.

[33] Liang Hong, Dennis C. Glass, Jonathan D. Nickels, Stefania Perticaroli, Zheng Yi, Madhusudan Tyagi, Hugh O'Neill, Qiu Zhang, Alexei P. Sokolov, and Jeremy C. Smith. Elastic and Conformational Softness of a Globular Protein. *Physical Review Letters*, 110(2):028104, January 2013.

[34] Giorgio Schirò, Chiara Caronna, Francesca Natali, M. Marek Koza, and Antonio Cupane. The "protein dynamical transition" does not require the protein polypeptide chain. *Journal of Physical Chemistry Letters*, 2(18):2275–2279, 2011.

[35] Shiou-Ru Tzeng and Charalampos G Kalodimos. The role of slow and fast protein motions in allosteric interactions. *Biophysical Reviews*, 7(2):251–255, 2015.

[36] Ivet Bahar, Ali Rana Atilgan, Melik C. Demirel, and Burak Erman. Vibrational Dynamics of Folded Proteins: Significance of Slow and Fast Motions in Relation to Function and Stability. *Physical Review Letters*, 80(12):2733–2736, March 1998.

[37] Joanne M Williamson and Gene M Brown. Purification and properties of l-aspartate-alpha-decarboxylase, an enzyme that catalyzes the formation of beta-alanine in escherichia coli. *Journal of Biological Chemistry*, 254(16):8074–8082, 1979.

[38] Diana CF Monteiro, Vijay Patel, Christopher P Bartlett, Shingo Nozaki, Thomas D Grant, James A Gowdy, Gary S Thompson, Arnout P Kalverda, Edward H Snell, Hironori Niki, et al. The structure of the pand/panz protein complex reveals negative feedback regulation of pantothenate biosynthesis by coenzyme a. *Chemistry & biology*, 22(4):492–503, 2015.

[39] Bret D Wallace, Jonathan S Edwards, Jamie R Wallen, Wessel JA Moolman, Renier van der Westhuyzen, Erick Strauss, Matthew R Redinbo, and Al Claiborne. Turnover-dependent covalent inactivation of staphylococcus aureus coenzyme a-disulfide reductase by coenzyme a-mimetics: mechanistic and structural insights. *Biochemistry*, 51(39):7699–7711, 2012.

[40] Malin Darnell and Lars Weidolf. Metabolism of xenobiotic carboxylic acids: focus on coenzyme a conjugation, reactivity, and interference with lipid metabolism. *Chemical research in toxicology*, 26(8):1139–1155, 2013.

[41] Mary C Hunt, Veronika Tillander, and Stefan EH Alexson. Regulation of peroxisomal lipid metabolism: the role of acyl-coa and coenzyme a metabolizing enzymes. *Biochimie*, 98:45–55, 2014.

[42] JE Cronan, KJ Littel, and S Jackowski. Genetic and biochemical analyses of pantothenate biosynthesis in escherichia coli and salmonella typhimurium. *Journal of Bacteriology*, 149(3):916–922, 1982.

[43] Armando Albert, Venugopal Dhanaraj, Ulrich Genschel, Ghalib Khan, Manoj K Ramjee, Rosaline Pulido, B Lynn Sibanda, Frank von Delft, Michael Witty, Tom L Blundell, et al. Crystal structure of aspartate decarboxylase at 2.2 å resolution provides evidence for an ester in protein self–processing. *Nature structural biology*, 5(4):289, 1998.

[44] Warren L DeLano et al. Pymol: An open-source molecular graphics tool. *CCP4 Newsletter on protein crystallography*, 40(1):82–92, 2002.

[45] Paul D van Poelje and Esmond E Snell. Pyruvoyl-dependent enzymes. *Annual review of biochemistry*, 59(1):29–59, 1990.

[46] Heather M Wallace and Anthony E Pegg. S-adenosylmethionine decarboxylase. *Essays in biochemistry*, 46:25–46, 2009.

[47] Joanne E Dove, Benjamin Schwartz, Neal K Williams, and Judith P Klinman. Investigation of spectroscopic intermediates during copper-binding and tpq

formation in wild-type and active-site mutants of a copper-containing amine oxidase from yeast. *Biochemistry*, 39(13):3690–3698, 2000.

[48] Tara N Stuecker, Shanti Bramhacharya, Kelsey M Hodge-Hanson, Garret Suen, and Jorge C Escalante-Semerena. Phylogenetic and amino acid conservation analyses of bacterial l-aspartate-α-decarboxylase and of its zymogen-maturation protein reveal a putative interaction domain. *BMC research notes*, 8(1):1–11, 2015.

[49] Diana Canelhas Freitas Monteiro. *New photochemical tools for time-resolved structural studies*. PhD thesis, University of Leeds, 2015.

[50] Florian Schmitzberger, Mairi L Kilkenny, Carina MC Lobley, Michael E Webb, Mladen Vinkovic, Dijana Matak-Vinkovic, Michael Witty, Dimitri Y Chirgadze, Alison G Smith, Chris Abell, et al. Structural constraints on protein self-processing in l-aspartate-α-decarboxylase. *The EMBO Journal*, 22(23):6193–6204, 2003.

[51] Michael E Webb, Briony A Yorke, Tom Kershaw, Sarah Lovelock, Carina MC Lobley, Mairi L Kilkenny, Alison G Smith, Tom L Blundell, Arwen R Pearson, and Chris Abell. Threonine 57 is required for the post-translational activation of escherichia coli aspartate α-decarboxylase. *Acta Crystallographica Section D: Biological Crystallography*, 70(4):1166–1172, 2014.

[52] Shingo Nozaki, Michael E Webb, and Hironori Niki. An activator for pyruvoyl-dependent l-aspartate α-decarboxylase is conserved in a small group of the γ-proteobacteria including escherichia coli. *MicrobiologyOpen*, 1(3):298–310, 2012.

[53] Simon G Møller and Michael J McPherson. Developmental expression and biochemical analysis of the arabidopsis atao1 gene encoding an h2o2-generating diamine oxidase. *The Plant Journal*, 13(6):781–791, 1998.

[54] William S McIntire. Newly discovered redox cofactors: possible nutritional, medical, and pharmacological relevance to higher animals. *Annual review of nutrition*, 18(1):145–177, 1998.

[55] Petri Bono, Marko Salmi, David J Smith, and Sirpa Jalkanen. Cloning and characterization of mouse vascular adhesion protein-1 reveals a novel molecule with enzymatic activity. *The Journal of Immunology*, 160(11):5563–5571, 1998.

[56] Heli Elovaara, Teija Huusko, Mikael Maksimow, Kati Elima, Gennady G Yegutkin, Mikael Skurnik, Ulrich Dobrindt, Anja Siitonen, Michael J McPherson, Marko Salmi, et al. Primary amine oxidase of escherichia coli is a metabolic enzyme that can use a human leukocyte molecule as a substrate. *PloS one*, 10(11):e0142367, 2015.

[57] CM Wilmot, J Hajdu, MJ McPherson, PF Knowles, and SEV Phillips. Visualization of dioxygen bound to copper during enzyme catalysis. *Science*, 286(5445):1724–1728, 1999.

[58] Valerie J Klema and Carrie M Wilmot. The role of protein crystallography in defining the mechanisms of biogenesis and catalysis in copper amine oxidase. *International journal of molecular sciences*, 13(5):5375–5405, 2012.

[59] Brian J Brazeau, Bryan J Johnson, and Carrie M Wilmot. Copper-containing amine oxidases. biogenesis and catalysis; a structural perspective. *Archives of Biochemistry and biophysics*, 428(1):22–31, 2004.

[60] Christa Hartmann and Judith P Klinman. Structure-function studies of substrate oxidation by bovine serum amine oxidase: relationship to cofactor structure and mechanism. *Biochemistry*, 30(18):4605–4611, 1991.

[61] Rongbao Li, Judith P Klinman, and F Scott Mathews. Copper amine oxidase from hansenula polymorpha: the crystal structure determined at 2.4 å resolution reveals the active conformation. *Structure*, 6(3):293–307, 1998.

[62] Danying Cai and Judith P Klinman. Evidence of a self-catalytic mechanism of 2, 4, 5-trihydroxyphenylalanine quinone biogenesis in yeast copper amine oxidase. *Journal of Biological Chemistry*, 269(51):32039–32042, 1994.

[63] Ryuichi Matsuzaki, Toshio Fukui, Hidetoshi Sato, Yukihiro Ozaki, and Katsuyuki Tanizawa. Generation of the topa quinone cofactor in bacterial monoamine oxidase by cupric ion-dependent autooxidation of a specific tyrosyl residue. *FEBS letters*, 351(3):360–364, 1994.

[64] Christy E Ruggiero and David M Dooley. Stoichiometry of the topa quinone biogenesis reaction in copper amine oxidases. *Biochemistry*, 38(10):2892–2898, 1999.

[65] David Mu, SM Janes, AJ Smith, DE Brown, DM Dooley, and JP Klinman. Tyrosine codon corresponds to topa quinone at the active site of copper amine oxidases. *Journal of Biological Chemistry*, 267(12):7979–7982, 1992.

[66] Matthew CJ Wilce, David M Dooley, Hans C Freeman, J Mitchell Guss, Hideyuki Matsunami, William S McIntire, Christy E Ruggiero, Katsuyuki Tanizawa, and Hiroshi Yamaguchi. Crystal structures of the copper-containing amine oxidase from arthrobacter globiformis in the holo and apo forms: implications for the biogenesis of topaquinone. *Biochemistry*, 36(51):16116–16133, 1997.

[67] Misa Kim, Toshihide Okajima, Seiichiro Kishishita, Megumi Yoshimura, Asako Kawamori, Katsuyuki Tanizawa, and Hiroshi Yamaguchi. X-ray snapshots of quinone cofactor biogenesis in bacterial copper amine oxidase. *Nature structural biology*, 9(8):591–596, 2002.

[68] Zhi-wei Chen, Benjamin Schwartz, Neal K Williams, Rongbao Li, Judith P Klinman, and F Scott Mathews. Crystal structure at 2.5 å resolution of zinc-substituted copper amine oxidase of hansenula polymorpha expressed in escherichia coli. *Biochemistry*, 39(32):9709–9717, 2000.

[69] Aaron P McGrath, Tom Caradoc-Davies, Charles A Collyer, and J Mitchell Guss. Correlation of active site metal content in human diamine oxidase with trihydroxyphenylalanine quinone cofactor biogenesis. *Biochemistry*, 49(38):8316–8324, 2010.

[70] MR Parsons, MA Convery, CM Wilmot, KDS Yadav, V Blakeley, AS Corner, SEV Phillips, MJ McPherson, and PF Knowles. Crystal structure of a quinoenzyme: copper amine oxidase of escherichia coli at 2 å resolution. *Structure*, 3(11):1171–1184, 1995.

[71] Mark A Smith, Pascale Pirrat, Arwen R Pearson, Christian RP Kurtis, Chi H Trinh, Thembaninkosi G Gaule, Peter F Knowles, Simon EV Phillips, and Michael J McPherson. Exploring the roles of the metal ions in escherichia coli copper amine oxidase. *Biochemistry*, 49(6):1268–1280, 2010.

[72] Pascale Pirrat, Mark A Smith, Arwen R Pearson, Michael J McPherson, and Simon EV Phillips. Structure of a xenon derivative of escherichia coli copper amine oxidase: confirmation of the proposed oxygen-entry pathway. *Acta Crystallographica Section F: Structural Biology and Crystallization Communications*, 64(12):1105–1109, 2008.

[73] Takeshi Murakawa, Hideyuki Hayashi, Tomoko Sunami, Kazuo Kurihara, Taro Tamada, Ryota Kuroki, Mamoru Suzuki, Katsuyuki Tanizawa, and Toshihide Okajima. High-resolution crystal structure of copper amine oxidase from arthrobacter globiformis: Assignment of bound diatomic molecules as

o_2. *Acta Crystallographica Section D: Biological Crystallography*, 69(12):2483–2494, 2013.

[74] Emily E Scott, Quentin H Gibson, and John S Olson. Mapping the pathways for o_2 entry into and exit from myoglobin. *Journal of Biological Chemistry*, 276(7):5177–5188, 2001.

[75] Anthony P Duff, Daniel M Trambaiolo, Aina E Cohen, Paul J Ellis, Gregory A Juda, Eric M Shepard, David B Langley, David M Dooley, Hans C Freeman, and J Mitchell Guss. Using xenon as a probe for dioxygen-binding sites in copper amine oxidases. *Journal of molecular biology*, 344(3):599–607, 2004.

[76] Bryan J Johnson, Jordi Cohen, Richard W Welford, Arwen R Pearson, Klaus Schulten, Judith P Klinman, and Carrie M Wilmot. Exploring molecular oxygen pathways in hansenula polymorpha copper-containing amine oxidase. *Journal of Biological Chemistry*, 282(24):17767–17776, 2007.

[77] Thembaninkosi G Gaule, Mark A Smith, Katarzyna M Tych, Pascale Pirrat, Chi H Trinh, Arwen R Pearson, Peter F Knowles, and Michael J McPherson. Oxygen activation switch in the copper amine oxidase of escherichia coli. *Biochemistry*, 57(36):5301–5314, 2018.

[78] Stefano Piana and Paolo Carloni. Conformational flexibility of the catalytic asp dyad in hiv-1 protease: an ab initio study on the free enzyme. *Proteins: Structure, Function, and Bioinformatics*, 39(1):26–36, 2000.

[79] Motoyasu Adachi, Takashi Ohhara, Kazuo Kurihara, Taro Tamada, Eijiro Honjo, Nobuo Okazaki, Shigeki Arai, Yoshinari Shoyama, Kaname Kimura, Hiroyoshi Matsumura, et al. Structure of hiv-1 protease in complex with potent inhibitor kni-272 determined by high-resolution x-ray and neutron crystallography. *Proceedings of the National Academy of Sciences*, 106(12):4641–4646, 2009.

[80] Yohta Fukuda, Yu Hirano, Katsuhiro Kusaka, Tsuyoshi Inoue, and Taro Tamada. High-resolution neutron crystallography visualizes an oh-bound resting state of a copper-containing nitrite reductase. *Proceedings of the National Academy of Sciences*, 117(8):4071–4077, 2020.

[81] Christian Beck, Marco Grimaldo, Felix Roosen-Runge, Michal K Braun, Fajun Zhang, Frank Schreiber, and Tilo Seydel. Nanosecond tracer diffusion as a probe of the solution structure and molecular mobility of protein assemblies:

The case of ovalbumin. *The Journal of Physical Chemistry B*, 122(35):8343–8350, 2018.

[82] Utsab R Shrestha, Debsindhu Bhowmik, John RD Copley, Madhusudan Tyagi, Juscelino B Leão, and Xiang-qiang Chu. Effects of pressure on the dynamics of an oligomeric protein from deep-sea hyperthermophile. *Proceedings of the National Academy of Sciences*, 112(45):13886–13891, 2015.

[83] Marco Grimaldo, Hender Lopez, Christian Beck, Felix Roosen-Runge, Martine Moulin, Juliette M Devos, Valerie Laux, Michael Härtlein, Stefano Da Vela, Ralf Schweins, et al. Protein short-time diffusion in a naturally crowded environment. *The journal of physical chemistry letters*, 10(8):1709–1715, 2019.

[84] Kevin Pounot, Hussein Chaaban, Vito Foderà, Giorgio Schirò, Martin Weik, and Tilo Seydel. Tracking internal and global diffusive dynamics during protein aggregation by high-resolution neutron spectroscopy. *The Journal of Physical Chemistry Letters*, 11(15):6299–6304, 2020.

[85] Sophie Sacquin-Mora, Pierre Sebban, Valérie Derrien, Bernhard Frick, Richard Lavery, and Christiane Alba-Simionesco. Probing the flexibility of the bacterial reaction center: the wild-type protein is more rigid than two site-specific mutants. *Biochemistry*, 46(51):14960–14968, 2007.

[86] Tilo Seydel. Dynamics of biological systems. In *Experimental Methods in the Physical Sciences*, volume 49, pages 77–134. Elsevier, 2017.

[87] Gordon Leslie Squires. *Introduction to the Theory of Thermal Neutron Scattering*. N.Y. : Dover publ., Mineola, 1996.

[88] Marc Bée. Quasielastic neutron scattering. 1988.

[89] Andrew R. Leach. *Molecular Modelling: Principles and Applications, 2nd Edition*. Pearson, 2001.

[90] Lionel I Rebhun. Polarized intracellular particle transport: saltatory movements and cytoplasmic streaming. In *International review of cytology*, volume 32, pages 93–137. Elsevier, 1972.

[91] Albert Einstein. Über die von der molekularkinetischen theorie der wärme geforderte bewegung von in ruhenden flüssigkeiten suspendierten teilchen. *Annalen der physik*, 4, 1905.

[92] Igor M Sokolov. Models of anomalous diffusion in crowded environments. *Soft Matter*, 8(35):9043–9052, 2012.

[93] Howard C Berg and Edward M Purcell. Physics of chemoreception. *Biophysical journal*, 20(2):193–219, 1977.

[94] Marco Grimaldo. *Global and Internal Diffusive Dynamics of Proteins in Solution Studied by Neutron Spectroscopy*. PhD thesis, Eberhard Karls Universität Tübingen, 2016.

[95] Gerhard Nägele. *The physics of colloidal soft matter*. Centre of Excellence for Advanced Materials and Structures Warsaw, 2004.

[96] Thomas Franosch, Matthias Grimm, Maxim Belushkin, Flavio M Mor, Giuseppe Foffi, László Forró, and Sylvia Jeney. Resonances arising from hydrodynamic memory in brownian motion. *Nature*, 478(7367):85–88, 2011.

[97] Adolph Fick. V. on liquid diffusion. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 10(63):30–39, 1855.

[98] Felix Fernandez-Alonso and David L Price. *Neutron Scattering–Applications in Biology, Chemistry, and Materials Science*. Elsevier, 2017.

[99] Michael Kovermann, Christin Grundström, A Elisabeth Sauer-Eriksson, Uwe H Sauer, and Magnus Wolf-Watz. Structural basis for ligand binding to an enzyme by a conformational selection pathway. *Proceedings of the National Academy of Sciences*, 114(24):6298–6303, 2017.

[100] Sayoko Oiki, Reiko Kamochi, Bunzo Mikami, Kousaku Murata, and Wataru Hashimoto. Alternative substrate-bound conformation of bacterial solute-binding protein involved in the import of mammalian host glycosaminoglycans. *Scientific reports*, 7(1):1–13, 2017.

[101] Marijn de Boer, Giorgos Gouridis, Ruslan Vietrov, Stephanie L Begg, Gea K Schuurman-Wolters, Florence Husada, Nikolaos Eleftheriadis, Bert Poolman, Christopher A McDevitt, and Thorben Cordes. Conformational and dynamic plasticity in substrate-binding proteins underlies selective transport in abc importers. *Elife*, 8:e44652, 2019.

[102] Julie A Reisz, Nidhi Bansal, Jiang Qian, Weiling Zhao, and Cristina M Furdui. Effects of ionizing radiation on biological molecules—mechanisms of damage

and emerging methods of detection. *Antioxidants & redox signaling*, 21(2):260–292, 2014.

[103] Elspeth F Garman and Martin Weik. X-ray radiation damage to biological samples: recent progress, 2019.

[104] Eugenio de la Mora, Nicolas Coquelle, Charles S Bury, Martin Rosenthal, James M Holton, Ian Carmichael, Elspeth F Garman, Manfred Burghammer, Jacques-Philippe Colletier, and Martin Weik. Radiation damage and dose limits in serial synchrotron crystallography at cryo-and room temperatures. *Proceedings of the National Academy of Sciences*, 117(8):4142–4151, 2020.

[105] Léon Van Hove. Correlations in space and time and born approximation scattering in systems of interacting particles. *Physical Review*, 95(1):249, 1954.

[106] Marcus Hennig, Felix Roosen-Runge, Fajun Zhang, Stefan Zorn, Maximilian WA Skoda, Robert MJ Jacobs, Tilo Seydel, and Frank Schreiber. Dynamics of highly concentrated protein solutions around the denaturing transition. *Soft Matter*, 8(5):1628–1633, 2012.

[107] Felix Roosen-Runge and Tilo Seydel. A generalized mean-squared displacement from inelastic fixed window scans of incoherent neutron scattering as a model-free indicator of anomalous diffusion confinement. In *EPJ Web of Conferences*, volume 83, page 02015. EDP Sciences, 2015.

[108] Zheng Yi, Yinglong Miao, Jerome Baudry, Nitin Jain, and Jeremy C. Smith. Derivation of mean-square displacements for protein dynamics from elastic incoherent neutron scattering. *Journal of Physical Chemistry B*, 116(16):5028–5036, 2012.

[109] Gerald R. Kneller and Konrad Hinsen. Quantitative model for the heterogeneity of atomic position fluctuations in proteins: A simulation study. *Journal of Chemical Physics*, 131(4):1–6, 2009.

[110] D. Zeller, M. T. F. Telling, M. Zamponi, V. García Sakai, and J. Peters. Analysis of elastic incoherent neutron scattering data beyond the Gaussian approximation. *The Journal of Chemical Physics*, 149(23):234908, December 2018.

[111] Judith Peters and Gerald R. Kneller. Motional heterogeneity in human acetylcholinesterase revealed by a non-Gaussian model for elastic incoherent neutron scattering. *The Journal of Chemical Physics*, 139(16):165102, October 2013.

[112] Dominik Zeller, Pan Tan, Liang Hong, Daniele Di Bari, Victoria Garcia Sakai, and Judith Peters. Differences between calcium rich and depleted alpha-lactalbumin investigated by molecular dynamics simulations and incoherent neutron scattering. *Physical Review E*, 101(3):032415, March 2020.

[113] Sabine Petry, Ditlev E Brodersen, Frank V Murphy IV, Christine M Dunham, Maria Selmer, Michael J Tarry, Ann C Kelley, and V Ramakrishnan. Crystal structures of the ribosome in complex with release factors rf1 and rf2 bound to a cognate stop codon. *Cell*, 123(7):1255–1266, 2005.

[114] Helen M Berman, John Westbrook, Zukang Feng, Gary Gilliland, Talapady N Bhat, Helge Weissig, Ilya N Shindyalov, and Philip E Bourne. The protein data bank. *Nucleic acids research*, 28(1):235–242, 2000.

[115] Alexi Assmus. Early history of x rays. *Beam Line*, 25(2):10–24, 1995.

[116] Camelia V Stan, Christine M Beavers, Martin Kunz, and Nobumichi Tamura. X-ray diffraction under extreme conditions at the advanced light source. *Quantum Beam Science*, 2(1):4, 2018.

[117] Garry Taylor. The phase problem. *Acta Crystallographica Section D: Biological Crystallography*, 59(11):1881–1890, 2003.

[118] Garry L Taylor. Introduction to phasing. *Acta Crystallographica Section D: Biological Crystallography*, 66(4):325–338, 2010.

[119] Steven E Ealick. Advances in multiple wavelength anomalous diffraction crystallography. *Current opinion in chemical biology*, 4(5):495–499, 2000.

[120] Luke M Rice, T Earnest, and Axel T Brunger. Single-wavelength anomalous diffraction phasing revisited. *Acta Crystallographica Section D: Biological Crystallography*, 56(11):1413–1420, 2000.

[121] Axel T Brünger. Free r value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature*, 355(6359):472–475, 1992.

[122] Wolfgang Kabsch. Xds. *Acta Crystallographica Section D: Biological Crystallography*, 66(2):125–132, 2010.

[123] Martyn D Winn, Charles C Ballard, Kevin D Cowtan, Eleanor J Dodson, Paul Emsley, Phil R Evans, Ronan M Keegan, Eugene B Krissinel, Andrew GW Leslie, Airlie McCoy, et al. Overview of the ccp4 suite and current developments. *Acta Crystallographica Section D: Biological Crystallography*, 67(4):235–242, 2011.

[124] Alexei A Vagin, Roberto A Steiner, Andrey A Lebedev, Liz Potterton, Stuart McNicholas, Fei Long, and Garib N Murshudov. Refmac5 dictionary: organization of prior chemical knowledge and guidelines for its use. *Acta Crystallographica Section D: Biological Crystallography*, 60(12):2184–2195, 2004.

[125] Paul Emsley and Kevin Cowtan. Coot: model-building tools for molecular graphics. *Acta Crystallographica Section D: Biological Crystallography*, 60(12):2126–2132, 2004.

[126] Thomas F Koetzle and Garry J McIntyre. Single-crystal neutron diffraction. *Characterization of Materials*, pages 1–14, 2002.

[127] Mats HM Olsson, Per EM Siegbahn, and Arieh Warshel. Simulating large nuclear quantum mechanical corrections in hydrogen atom transfer reactions in metalloenzymes. *JBIC Journal of Biological Inorganic Chemistry*, 9(1):96–99, 2004.

[128] Kylie A Vincent, James A Cracknell, Alison Parkin, and Fraser A Armstrong. Hydrogen cycling by enzymes: electrocatalysis and implications for future energy technology. *Dalton Transactions*, (21):3397–3403, 2005.

[129] Joanna Trylska, Paweł Grochowski, and J Andrew McCammon. The role of hydrogen bonding in the enzymatic reaction catalyzed by hiv-1 protease. *Protein science*, 13(2):513–528, 2004.

[130] Zhigang Liu, Yong Wang, Ravikiran S Yedidi, Tamaria G Dewdney, Samuel J Reiter, Joseph S Brunzelle, Iulia A Kovari, and Ladislau C Kovari. Conserved hydrogen bonds and water molecules in mdr hiv-1 protease substrate complexes. *Biochemical and biophysical research communications*, 430(3):1022–1027, 2013.

[131] Oksana Gerlits, Troy Wymore, Amit Das, Chen-Hsiang Shen, Jerry M Parks, Jeremy C Smith, Kevin L Weiss, David A Keen, Matthew P Blakeley, John M Louis, et al. Long-range electrostatics-induced two-proton transfer captured by neutron crystallography in an enzyme catalytic site. *Angewandte Chemie*, 128(16):5008–5011, 2016.

[132] CC Wilson and DA Myles. Single crystal neutron diffraction and protein crystallography. In *Neutron Scattering in Biology*, pages 21–42. Springer, 2006.

[133] Thomas Gutberlet, Udo Heinemann, and Michael Steiner. Protein crystallography with neutrons–status and perspectives. *Acta Crystallographica Section D: Biological Crystallography*, 57(2):349–354, 2001.

[134] B Leif Hanson. Getting protein solvent structures down cold. *Proceedings of the National Academy of Sciences*, 101(47):16393–16394, 2004.

[135] Matthew P Blakeley, Paul Langan, Nobuo Niimura, and Alberto Podjarny. Neutron crystallography: opportunities, challenges, and limitations. *Current opinion in structural biology*, 18(5):593–600, 2008.

[136] Pavel V Afonine, Marat Mustyakimov, Ralf W Grosse-Kunstleve, Nigel W Moriarty, Paul Langan, and Paul D Adams. Joint x-ray and neutron refinement with phenix. refine. *Acta Crystallographica Section D: Biological Crystallography*, 66(11):1153–1163, 2010.

[137] Paul D Adams, Navraj S Pannu, Randy J Read, and Axel T Brünger. Cross-validated maximum likelihood enhances crystallographic simulated annealing refinement. *Proceedings of the National Academy of Sciences*, 94(10):5018–5023, 1997.

[138] Jörg Stetefeld, Sean A McKenna, and Trushar R Patel. Dynamic light scattering: a practical guide and applications in biomedical sciences. *Biophysical reviews*, 8(4):409–427, 2016.

[139] L Verlet. Computer "experiments" on classical fluids. i. *Thermodynamical properties of Lennard-Jones molecules. Phys. Rev., 159: 98—103*, 1967.

[140] MBBJM Tuckerman, Bruce J Berne, and Glenn J Martyna. Reversible multiple time scale molecular dynamics. *The Journal of chemical physics*, 97(3):1990–2001, 1992.

[141] William H Press, Saul A Teukolsky, William T Vetterling, and Brian P Flannery. *Numerical recipes 3rd edition: The art of scientific computing*. Cambridge university press, 2007.

[142] Gerald R Kneller, Volker Keiner, Meinhard Kneller, and Matthias Schiller. nmoldyn: a program package for a neutron scattering oriented analysis of molecular dynamics simulations. *Computer physics communications*, 91(1-3):191–214, 1995.

[143] G Goret, B Aoun, and Eric Pellegrini. Mdanse: An interactive analysis environment for molecular dynamics simulations. *Journal of chemical information and modeling*, 57(1):1–5, 2017.

[144] William Humphrey, Andrew Dalke, Klaus Schulten, et al. Vmd: visual molecular dynamics. *Journal of molecular graphics*, 14(1):33–38, 1996.

[145] Richard J Gowers, Max Linke, Jonathan Barnoud, Tyler John Edward Reddy, Manuel N Melo, Sean L Seyler, Jan Domanski, David L Dotson, Sébastien Buchoux, Ian M Kenney, et al. Mdanalysis: a python package for the rapid analysis of molecular dynamics simulations. Technical report, Los Alamos National Lab.(LANL), Los Alamos, NM (United States), 2019.

[146] S Adrian Saldanha, Louise M Birch, Michael E Webb, Brent K Nabbs, Frank von Delft, Alison G Smith, and Chris Abell. Identification of tyr58 as the proton donor in the aspartate-α-decarboxylase reaction. *Chemical Communications*, (18):1760–1761, 2001.

[147] F William Studier. Protein production by auto-induction in high-density shaking cultures. *Protein expression and purification*, 41(1):207–234, 2005.

[148] Elisabeth Gasteiger, Christine Hoogland, Alexandre Gattiker, Marc R Wilkins, Ron D Appel, Amos Bairoch, et al. Protein identification and analysis tools on the expasy server. In *The proteomics protocols handbook*, pages 571–607. Springer, 2005.

[149] Tushar Raskar, Stephan Niebling, Nicolas Coquelle, Matthew P. Blakeley, Trevor Forsyth, Nils Huse, Tilo Seydel, and Arwen Pearson. Neutron spectroscopy & crystallography combined with THz spectroscopy: complementary methods for the study of protein dynamics, 2017. publisher: Institut Laue-Langevin, https://doi.ill.fr/10.5291/ILL-DATA.8-05-428.

[150] Tushar Raskar, Stephan Niebling, Trevor Forsyth, Nils Huse, Tilo Seydel, and Arwen Pearson. Neutron spectroscopy & crystallography combined with THz spectroscopy: complementary methods for the study of protein dynamics, 2018. publisher: Institut Laue-Langevin, https://doi.ill.fr/10.5291/ILL-DATA.8-05-431.

[151] Bernhard Frick, Eugene Mamontov, Lambert van Eijck, and Tilo Seydel. Recent backscattering instrument developments at the ILL and SNS. *Zeitschrift für Physikalische Chemie*, 224(1-2):33–60, 2010.

[152] Marcus Hennig, Bernhard Frick, and Tilo Seydel. Optimum velocity of a phase-space transformer for cold-neutron backscattering spectroscopy. *J. Appl. Cryst.*, 44(3):467–472, 2011.

[153] Tushar Raskar, Stephan Niebling, Trevor Forsyth, Nils Huse, Tilo Seydel, and Arwen Pearson. Vibrational spectroscopy on aspartate alpha-decarboxylase enzyme with and without ligand to compare with infrared/thz spectroscopy, 2018. publisher: Institut Laue-Langevin, https://doi.ill.fr/10.5291/ILL-DATA.8-05-434.

[154] Tushar Raskar, Stephan Niebling, Trevor Forsyth, Nils Huse, Tilo Seydel, and Arwen Pearson. Vibrational spectroscopy on aspartate alpha-decarboxylase enzyme with and without ligand to compare with infrared/thz spectroscopy, 2019. publisher: Institut Laue-Langevin, https://doi.ill.fr/10.5291/ILL-DATA.8-05-447.

[155] Tushar Raskar, Stephan Niebling, Trevor Forsyth, Nils Huse, Tilo Seydel, and Arwen Pearson. Vibrational spectroscopy on aspartate alpha-decarboxylase enzyme with and without ligand to compare with infrared/thz spectroscopy, 2020. publisher: Institut Laue-Langevin, https://doi.ill.fr/10.5291/ILL-DATA.8-05-459.

[156] Owen Arnold, Jean-Christophe Bilheux, JM Borreguero, Alex Buts, Stuart I Campbell, L Chapon, Mathieu Doucet, N Draper, R Ferraz Leal, MA Gigg, et al. Mantid—data analysis and visualization package for neutron scattering and μ sr experiments. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 764:156–166, 2014.

[157] Sander Pronk, Szilárd Páll, Roland Schulz, Per Larsson, Pär Bjelkmar, Rossen Apostolov, Michael R Shirts, Jeremy C Smith, Peter M Kasson, David van der

Spoel, Berk Hess, and Erik Lindahl. Gromacs 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics*, 29(7):845–854, 2013.

[158] Berk Hess, Carsten Kutzner, David van der Spoel, and Erik Lindahl. Gromacs 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theo. Comput.*, 4(3):435–447, 2008.

[159] David Van Der Spoel, Erik Lindahl, Berk Hess, Gerrit Groenhof, Alan E. Mark, and Herman J. C. Berendsen. GROMACS: Fast, flexible, and free. *J. Comput. Chem.*, 26(16):1701–1718, 2005.

[160] Kresten Lindorff-Larsen, Stefano Piana, Kim Palmo, Paul Maragakis, John L Klepeis, Ron O Dror, and David E Shaw. Improved side-chain torsion potentials for the amber ff99sb protein force field. *Proteins: Structure, Function, and Bioinformatics*, 78(8):1950–1958, 2010.

[161] Naveen Michaud-Agrawal, Elizabeth J Denning, Thomas B Woolf, and Oliver Beckstein. Mdanalysis: a toolkit for the analysis of molecular dynamics simulations. *J. Comput. Chem.*, 32(10):2319–2327, 2011.

[162] Travis E. Oliphant. *Guide to NumPy*. CreateSpace Independent Publishing Platform, USA, 2nd edition, 2015.

[163] John D. Hunter. Matplotlib: A 2d graphics environment. *Computing in Science & Engineering*, 9(3):90–95, 2007.

[164] Giovanni Bussi, Davide Donadio, and Michele Parrinello. Canonical sampling through velocity rescaling. *J. Chem. Phys.*, 126(1):014101, 2007.

[165] Shuichi Nosé and ML Klein. Constant pressure molecular dynamics for molecular systems. *Mol. Phys.*, 50(5):1055–1076, 1983.

[166] Michele Parrinello and Aneesur Rahman. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.*, 52(12):7182–7190, 1981.

[167] Berk Hess, Henk Bekker, Herman JC Berendsen, Johannes GEM Fraaije, et al. Lincs: a linear constraint solver for molecular simulations. *J. Comput. Chem.*, 18(12):1463–1472, 1997.

[168] Tom Darden, Darrin York, and Lee Pedersen. Particle mesh ewald: An n log (n) method for ewald sums in large systems. *J. Chem. Phys.*, 98(12):10089–10092, 1993.

[169] Ulrich Essmann, Lalith Perera, Max L Berkowitz, Tom Darden, Hsing Lee, and Lee G Pedersen. A smooth particle mesh ewald method. *J. Chem. Phys.*, 103(19):8577–8593, 1995.

[170] Thomas J Griffin IV, Andres E Leschziner, Nigel DF Grindley, Linda Parsons, Justine DeVost, and Keith M Derbyshire. In vitro transposition of tn 552: a tool for dna sequencing and mutagenesis. *Nucleic acids research*, 27(19):3859–3865, 1999.

[171] Johan Qvist, Helmut Schober, and Bertil Halle. Structural dynamics of supercooled water from quasielastic neutron scattering and molecular simulations. *The Journal of chemical physics*, 134(14):144508, 2011.

[172] Marco Grimaldo, Felix Roosen-Runge, Niina Jalarvo, Michaela Zamponi, Fabio Zanini, Marcus Hennig, Fajun Zhang, Frank Schreiber, and Tilo Seydel. High-resolution neutron spectroscopy on protein solution samples. In *EPJ Web of Conferences*, volume 83, page 02005. EDP Sciences, 2015.

[173] Felix Roosen-Runge, Marcus Hennig, Fajun Zhang, Robert MJ Jacobs, Michael Sztucki, Helmut Schober, Tilo Seydel, and Frank Schreiber. Protein self-diffusion in crowded solutions. *Proceedings of the National Academy of Sciences*, 108(29):11815–11820, 2011.

[174] Marco Grimaldo, Felix Roosen-Runge, Fajun Zhang, Tilo Seydel, and Frank Schreiber. Diffusion and dynamics of γ-globulin in crowded aqueous solutions. *The Journal of Physical Chemistry B*, 118(25):7203–7209, 2014.

[175] Youguang Ma, Chunying Zhu, Peisheng Ma, and KT Yu. Studies on the diffusion coefficients of amino acids in aqueous solutions. *Journal of Chemical & Engineering Data*, 50(4):1192–1196, 2005.

[176] LG Longsworth. Diffusion measurements, at 25, of aqueous solutions of amino acids, peptides and sugars. *Journal of the American Chemical Society*, 75(22):5705–5709, 1953.

[177] Frederick J Gutter and Gerson Kegeles. The diffusion of α-alanine in water at 25. *Journal of the American Chemical Society*, 75(16):3893–3896, 1953.

[178] Wolfgang Kabsch and Christian Sander. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers: Original Research on Biomolecules*, 22(12):2577–2637, 1983.

[179] Robbie P Joosten, Tim AH Te Beek, Elmar Krieger, Maarten L Hekkelman, Rob WW Hooft, Reinhard Schneider, Chris Sander, and Gert Vriend. A series of pdb related databases for everyday needs. *Nucleic acids research*, 39(suppl_1):D411–D419, 2010.

[180] Marco Grimaldo, Felix Roosen-Runge, Marcus Hennig, Fabio Zanini, Fajun Zhang, Niina Jalarvo, Michaela Zamponi, Frank Schreiber, and Tilo Seydel. Hierarchical molecular dynamics of bovine serum albumin in concentrated aqueous solution below and above thermal denaturation. *Physical Chemistry Chemical Physics*, 17(6):4645–4655, 2015.

[181] KS Singwi and Alf Sjölander. Diffusive motions in water and cold neutron scattering. *Physical Review*, 119(3):863, 1960.

[182] Marco Grimaldo, Felix Roosen-Runge, Fajun Zhang, Frank Schreiber, and Tilo Seydel. Dynamics of proteins in solution. *Quart. Rev. Biophys.*, 52, 2019.

[183] M Bée. A physical insight into the elastic incoherent structure factor. *Physica B: Condensed Matter*, 182(4):323–336, 1992.

[184] J Fitter. Confined molecular motions of globular proteins studied in powder samples and in solution. *Le Journal de Physique IV*, 10(PR7):Pr7–265, 2000.

[185] Adolfo J Banchio and Gerhard Nägele. Short-time transport properties in dense suspensions: from neutral to charge-stabilized colloidal spheres. *The Journal of chemical physics*, 128(10):104903, 2008.

[186] Michal K Braun, Marco Grimaldo, Felix Roosen-Runge, Ingo Hoffmann, Orsolya Czakkel, Michael Sztucki, Fajun Zhang, Frank Schreiber, and Tilo Seydel. Crowding-controlled cluster size in concentrated aqueous protein solutions: structure, self-and collective diffusion. *The Journal of Physical Chemistry Letters*, 8(12):2590–2596, 2017.

[187] Alvaro Ortega, D Amorós, and J García De La Torre. Prediction of hydrodynamic and other solution properties of rigid proteins from atomic-and residue-level models. *Biophysical journal*, 101(4):892–898, 2011.

[188] Rajeev Jindal, Meenu Singla, and Harsh Kumar. Transport behavior of aliphatic amino acids glycine/l-alanine/l-valine and hydroxyl amino acids l-serine/l-threonine in aqueous trilithium citrate solutions at different temperatures. *Journal of Molecular Liquids*, 206:343–349, 2015.

[189] Kirtanjot Kaur and Harsh Kumar. Viscometric measurements of l-serine with antibacterial drugs ampicillin and amoxicillin at different temperatures:(305.15 to 315.15) k. *Journal of Molecular Liquids*, 177:49–53, 2013.

[190] Ashwani Kumar, Ruby Rani, Tanu Sharma, Rajinder Kumar Bamezai, et al. Effect of concentration and temperature variations on interactions in (l-serine/l-valine+ aqueous glucose/sucrose/lactose) systems: Viscometric and activation parametric study. *Indian Journal of Pure & Applied Physics (IJPAP)*, 57(4):225–235, 2019.

[191] Zhenning Yan, Jianji Wang, Wei Kong, and Jinsuo Lu. Effect of temperature on volumetric and viscosity properties of some α-amino acids in aqueous calcium chloride solutions. *Fluid Phase Equilibria*, 215(2):143–150, 2004.

[192] Jin Suk Myung, Felix Roosen-Runge, Roland G Winkler, Gerhard Gompper, Peter Schurtenberger, and Anna Stradner. Weak shape anisotropy leads to a nonmonotonic contribution to crowding, impacting protein dynamics under physiologically relevant conditions. *The Journal of Physical Chemistry B*, 122(51):12396–12402, 2018.

[193] Mohamad Zahid Kamal, Virender Kumar, Kundarapu Satyamurthi, Kushal Kumar Das, and Nalam Madhusudhana Rao. Mutational probing of protein aggregates to design aggregation-resistant proteins. *FEBS open bio*, 6(2):126–134, 2016.

[194] Alfonso De Simone, Craig Kitchen, Ann H Kwan, Margaret Sunde, Christopher M Dobson, and Daan Frenkel. Intrinsic disorder modulates protein self-assembly and aggregation. *Proceedings of the National Academy of Sciences*, 109(18):6951–6956, 2012.

[195] Vandana Kurkal, Roy M Daniel, John L Finney, Moeava Tehei, Rachel V Dunn, and Jeremy C Smith. Enzyme activity and flexibility at very low hydration. *Biophysical journal*, 89(2):1282–1287, 2005.

[196] Hannu Mutka, Michael Marek Koza, Mark Robert Johnson, Zenji Hiroi, Jun-Ichi Yamaura, and Yohei Nagao. Generalized density-of-states and anharmonicity of the low-energy phonon bands from coherent inelastic neutron scattering

response in the pyrochlore osmates a os 2 o 6 (a= k, rb, cs). *Physical Review B*, 78(10):104307, 2008.

[197] Christopher Alan Crain, Nicholas A Strange, and JZ Larese. Inelastic neutron scattering studies of natural silkworm proteins. *MRS Online Proceedings Library Archive*, 1793:41–46, 2015.

[198] Anne V Goupil-Lamy, Jeremy C Smith, Junko Yunoki, Stewart F Parker, and Mikio Kataoka. High-resolution vibrational inelastic neutron scattering: a new spectroscopic tool for globular proteins. *Journal of the American Chemical Society*, 119(39):9268–9273, 1997.

[199] HD Middendorf, RL Hayward, SF Parker, J Bradshaw, and A Miller. Vibrational neutron spectroscopy of collagen and model polypeptides. *Biophysical journal*, 69(2):660–673, 1995.

[200] S Yu Venyaminov and NN Kalnin. Quantitative ir spectrophotometry of peptide compounds in water (h2o) solutions. i. spectral parameters of amino acid residue absorption bands. *Biopolymers: Original Research on Biomolecules*, 30(13-14):1243–1257, 1990.

[201] Ying Zhang, Peng Zhang, Robert C Ford, Shenghao Han, and Jichen Li. Inelastic neutron scattering studies of the interaction between water and some amino acids. *The Journal of Physical Chemistry B*, 109(38):17784–17786, 2005.

[202] Giuseppe Zaccai. How soft is a protein? a protein dynamics force constant measured by neutron scattering. *Science*, 288(5471):1604–1607, 2000.

[203] Ángela I López-Lorente and Boris Mizaikoff. Mid-infrared spectroscopy for protein analysis: potential and challenges. *Analytical and bioanalytical chemistry*, 408(11):2875–2889, 2016.

[204] Mark TF Telling, Cameron Neylon, Luke Clifton, Spencer Howells, Lambert van Eijck, and Victoria García Sakai. Thermal motion in the multi-subunit protein, apoferritin, as probed by high energy resolution neutron spectroscopy. *Soft Matter*, 7(15):6934–6941, 2011.

[205] Kathleen Wood, Chiara Caronna, Peter Fouquet, Wolfgang Haussler, Francesca Natali, Jacques Ollivier, Andrea Orecchini, Marie Plazanet, and Giuseppe Zaccai. A benchmark for protein dynamics: ribonuclease a measured by neutron scattering in a large wavevector-energy transfer range. *Chemical physics*, 345(2-3):305–314, 2008.

[206] U Heugen, G Schwaab, E Bründermann, M Heyden, X Yu, DM Leitner, and Martina Havenith. Solute-induced retardation of water dynamics probed directly by terahertz spectroscopy. *Proceedings of the National Academy of Sciences*, 103(33):12301–12306, 2006.

[207] Ellen M Adams, Oliver Lampret, Benedikt König, Thomas Happe, and Martina Havenith. Solvent dynamics play a decisive role in the complex formation of biologically relevant redox proteins. *Physical Chemistry Chemical Physics*, 22(14):7451–7459, 2020.

[208] S Walter Englander, Tobin R Sosnick, Joan J Englander, and Leland Mayne. Mechanisms and uses of hydrogen exchange. *Current opinion in structural biology*, 6(1):18–23, 1996.

[209] Varley F Sears. Neutron scattering lengths and cross sections. *Neutron news*, 3(3):26–37, 1992.

[210] Kenneth A Rubinson. Practical corrections for p (h, d) measurements in mixed h 2 o/d 2 o biological buffers. *Analytical Methods*, 9(18):2744–2750, 2017.

[211] Christian RP Kurtis, Peter F Knowles, Mark R Parsons, Thembaninkosi G Gaule, Simon EV Phillips, and Michael J McPherson. Tyrosine 381 in e. coli copper amine oxidase influences substrate specificity. *Journal of neural transmission*, 118(7):1043–1053, 2011.

[212] Katarzyna Maria Tych. *Terahertz time-domain spectroscopy of biological macromolecules*. PhD thesis, University of Leeds, 2011.

[213] Bernard Jacrot, Stephen Cusack, Albert J Dianoux, and Donald M Engelman. Inelastic neutron scattering analysis of hexokinase dynamics and its modification on binding of glucose. *Nature*, 300(5887):84–86, 1982.

[214] Melek Saouessi, Judith Peters, and Gerald R Kneller. Asymptotic analysis of quasielastic neutron scattering data from human acetylcholinesterase reveals subtle dynamical changes upon ligand binding. *The Journal of chemical physics*, 150(16):161104, 2019.

[215] Samuel Lenton, Marco Grimaldo, Felix Roosen-Runge, Frank Schreiber, Tommy Nylander, Roger Clegg, Carl Holt, Michael Härtlein, Victoria García Sakai, Tilo Seydel, and Susana C. Marujo Teixeira. Effect of phosphorylation on a human-like osteopontin peptide. *Biophys. J.*, 112(8):1586–1596, 2017.

[216] Utsab R Shrestha, Suchithranga MDC Perera, Debsindhu Bhowmik, Udeep Chawla, Eugene Mamontov, Michael F Brown, and Xiang-Qiang Chu. Quasielastic neutron scattering reveals ligand-induced protein dynamics of a g-protein-coupled receptor. *J. Phys. Chem. Lett.*, 7(20):4130–4136, 2016.

[217] Saara Laulumaa, Michael Marek Koza, Tilo Seydel, Petri Kursula, and Francesca Natali. A quasielastic neutron scattering investigation on the molecular self-dynamics of human myelin protein p2. *The Journal of Physical Chemistry B*, 123(39):8178–8185, 2019.

Appendices

A Supporting information for chapter 3

A.1 Stock solution recipes: ADC

The constituents for the stock solutions were dissolved in 15 MΩ filtered H_2O. All the stock solutions except 1000X metals and 25X salts were sterilized by filtration through 0.22μm membrane. 25X salts (1L) - 88.7 g Na_2HPO_4, 85 g KH_2PO_4, 67 g

NH_4Cl and 17.8 g Na_2SO_4.
1000X metal salts (100 mL) - 1.35 g $FeCl_3.6H_2O$, 438 mg $CaCl_2$, 198 mg $MnCl_2.4H_2O$, 288 mg $ZnSO_4.7H_2O$, 47.6 mg $CoCl_2.6H_2O$, 34.1 mg $CuCl_2.2H_2O$, 48.4 mg $Na_2MoO_4.2H_2O$, 34.6 mg Na_2SeO_3, 12.4 mg H_3BO_3, 52.6 mg $NiSO_4.6H_2O$, 500 μL HCl (37 % in H_2O).
10% Ampicillin - 1.0 g ampicilin sodium salt in 10 mL of H_2O split into 1 mL aliquotes and stored at -20 °C.
5% Kanamycin - 0.5 g kanamycin sulphate in 10 mL H_2O split into 1 mL aliquotes and stored at -20 °C.
10X TAE buffer (1L) - 4.84 g Tris, 5.71 mL glacial acetic acid and 292 mg EDTA. 6x DNA loading buffer - 5 mM Tris-Cl pH 8.0, 0.15 % w/v Orange G, 2.5% w/v Ficoll 400, 10 mM EDTA

A.1.1 Buffer preparation recipes: ADC

All the buffers were filter sterilized by using 0.22 μm membrane filter under vaccume pressure.

Ni-NTA stock buffer: 500 mM K_2HPO_4, 3000 mM NaCl pH 7.4 (15.0 g KH_2PO_4, 67.8 g K_2HPO_4, 175.3 g NaCl).
Ni-NTA lysis buffer: 50 mM K_2HPO_4, 300 mM NaCl, 10 mM imidazole pH 7.4 (100 mL Ni-NTA stock buffer, 681 mg imidazole).
Ni-NTA wash buffer: 50 mM K_2HPO_4, 300 mM NaCl, 50 mM imidazole pH 7.4 (100 mL Ni-NTA stock buffer, 3.40 g imidazole).
Ni-NTA elution buffer 1: 50 mM K_2HPO_4, 300 mM NaCl, 250 mM imidazole pH 7.4 (100 mL Ni-NTA stock buffer, 17.02 g imidazole).
Ni-NTA elution buffer 2: 50 mM K_2HPO_4,300 mM NaCl, 500 mM imidazole pH 7.4 (100 mL Ni-NTA stock buffer, 34.04 g imidazole).
Buffer for size exclusion chromatography: 50 mM K_2HPO_4,100 mM NaCl, 0.1 mM DTT pH 7.5 (4.93 g Tris, 2.92 g Tris-HCl, 11.70 g NaCl, 30.9 mg DTT).

A.2 Buffer preparation recipes: ECAO

All the buffers were filter sterilized using 0.22 μm.

Buffers for partial cell lysis:
Buffer 1 (B1) - 50 mM Tris-HCl pH 8.0, 20% (w/v) sucrose
Buffer 2 (B2) - 50 mM Tris-HCl pH 8.0, 10 mg/mL Hen Egg White Lysozyme (HEWL)
Buffer 3 (B3) - 42 mM EDTA pH 8.0
Buffer 4 (B4) - 50 mM Tris-HCl pH 8.0, 9% (w/v) sucrose
Buffer 5 (B5) - 50 mM Tris-HCl pH 8.0, 9% (w/v) sucrose, 15 mM $MgCl_2$

Buffers for ion exchange chromatography:
Ion exchange buffer A: 20 mM Tris pH 7.0
Ion exchange buffer B: 20 mM Tris pH 7.0, 175 mM NaCl

Buffer for gel filtration: 20 mM Tris pH 7.0

A.3 Buffer recipes: SDS-PAGE

5X SDS-PAGE Tris-glycine running buffer - 15.1 g Tris base, 94 g glycine, 5 g SDS.
SDS-PAGE Tris-tricine running buffer - 3 M Tris, 0.3% SDS, pH 8.45 (the pH was adjusted by addition of 2M NaOH).
SDS-PAGE Tris-tricine cathode buffer - 0.1 M Tris, 0.1 M tricine, 0.1% SDS, pH 8.23 (the pH was adjusted by addition of 2M NaOH).
SDS-PAGE Tris-tricine anode buffer - 0.2 M Tris pH 8.90 (the pH was adjusted by addition of 2M NaOH).
2x SDS-PAGE loading buffer - 50 mM Tris-HCl pH 6.8 (the pH was adjusted by addition of 2M NaOH), 100 mM dithiothreitol (DTT), 2% w/v SDS, 0.1% bromophenol blue, 10% w/v glycerol.
Commassie blue stain - 1.0 g Commassie blue R-250, 400 mL MeOH, 500 mL H_2O, 100 mL glacial acetic acid.

B Supporting information for chapter 4

B.1 Expression and purification of aspartate α-decarboxylase

BL21(DE3) cells were transformed with the pRSET-A plasmid containing the ADC construct. The expression of ADC was checked in both LB and autoinduction media. In case of LB medium, the expression was induced by adding IPTG to a final concentration of 1 mM at an $O.D_{600}$ of 0.6. The expression was checked for 30 and 37 °C. In case of the former, the cells were grown overnight after the addition of IPTG whereas in case of latter, they were grown for 4 hours in order to compensate for a higher growth rate at 37 °C. The induced pellet was then incubated overnight at 37 °C in case of all the samples except those in the first three lanes in figure 1. This is done so that the zymogen form of ADC is converted to an active form. It should be noted that this conversion is never 100%.

As seen in figure 1, the expression of ADC is observed in case for all the samples except those in which the overnight incubation was not performed. The purity level of ADC expressed at 37 °C was found to be lower than that for 30 °C and in case of autoinduction medium. Hence, we selected the autoinduction medium for cell growth, expression and purification. The induced cell pellet was lysed using a sonicator as described in chapter 3. The supernatant obtained was used for Ni-NTA purification. Figure 2 shows the fractions obtained in Ni-NTA affinity chromatography. It can be seen that the protein is high in yield. Fractions 2 to 20 indicate the protein eluted with 500 mM imidazole. These fractions were concentrated using a 10 kDa cutoff centrifugal filter to a final volume of 2 mL. The concentrated protein solution was then loaded onto superdex-200 column. As seen in figure 3, the fractions obtained after gel filtration were $\approx$ 95% pure and suitable to use in QENS experiments.

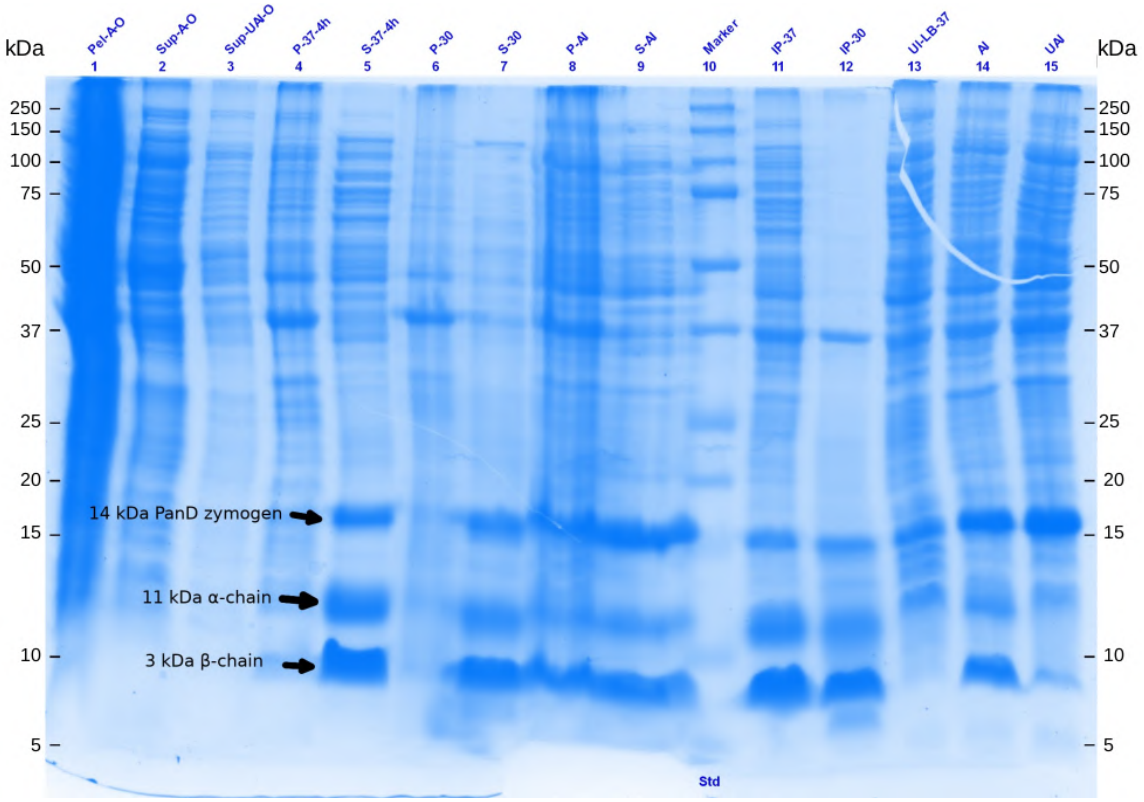

Figure 1: Expression check for ADC in BL21(DE3) cells transformed with pRSET-A plasmid containing the ADC construct. The expression is tested for the cells grown in LB and autoinduction media. ADC bands (PanD, α and β-chains) are indicated with three arrows

Labels: Pel-A-O - induced BL21DE3 cells with empty vector, Sup-A-O - supernatant of the BL21DE3 cells with empty vector, Sup-UAI-O - uninduced BL21DE3 cells with empty vector, P-37-4h - post-lysis pellet from the cells induced at 37 °C in LB medium for 4 hours, S-37-4h - post-lysis supernatant from the cells induced at 37 °C for 4 hours in LB medium, P-30 - post-lysis pellet from cells induced at 30 °C, S-30 - supernatant obtained from cells induced at 30 °C, P-AI - pellet post-lysis after autoinduction, S-AI - supernatant post-lysis after autoinduction, Marker - molecular weight marker, IP-37 - whole cells induced at 37 °C in LB, IP-30 - whole cells induced at 30 °C, UI-LB-37 - uninduced whole cells grown at 37 °C, AI - autoinduced whole cells, UAI - uninduced whole cells

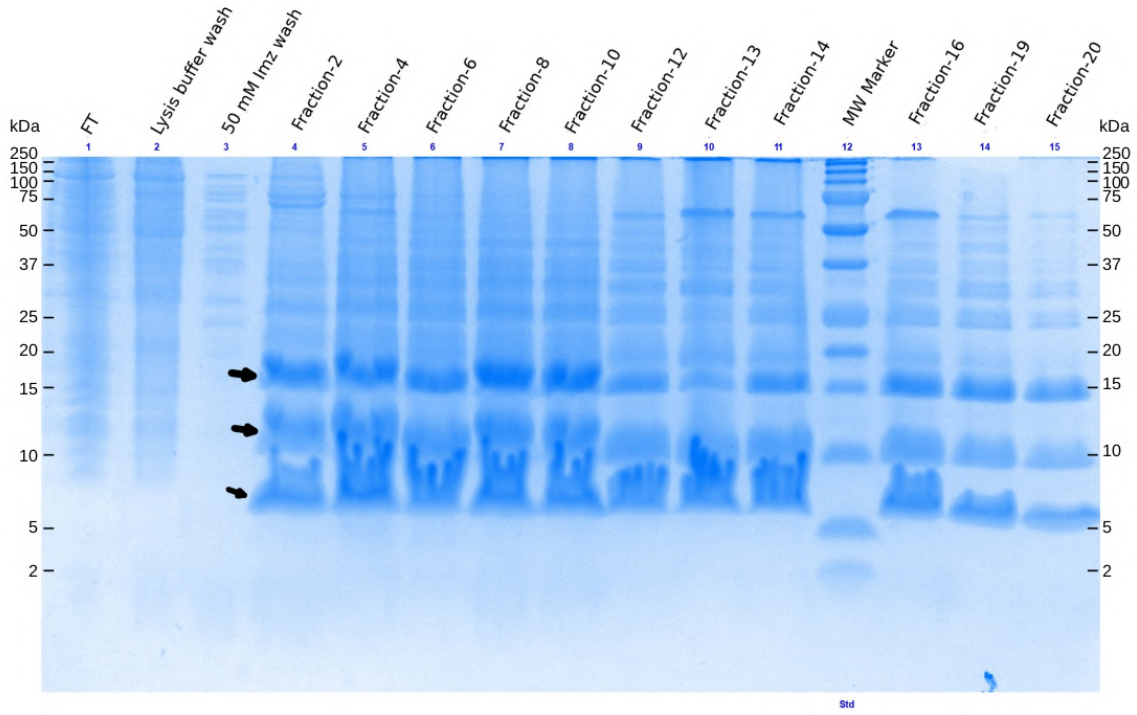

Figure 2: Fractions of ADC obtained after Ni-NTA affinity chromatography. The three arrows indicate the protein bands, same as for those seen in figure 1.
Labels: FT - flowthrough, Lysis buffer wash - impurities obtained after buffer wash, 50 mM Imz wash - impurities obtained after washing the Ni-NTA column with 50 mM imidazole, Fractions-2 to 20 - fractions obtained after elution with 500 mM imidazole.

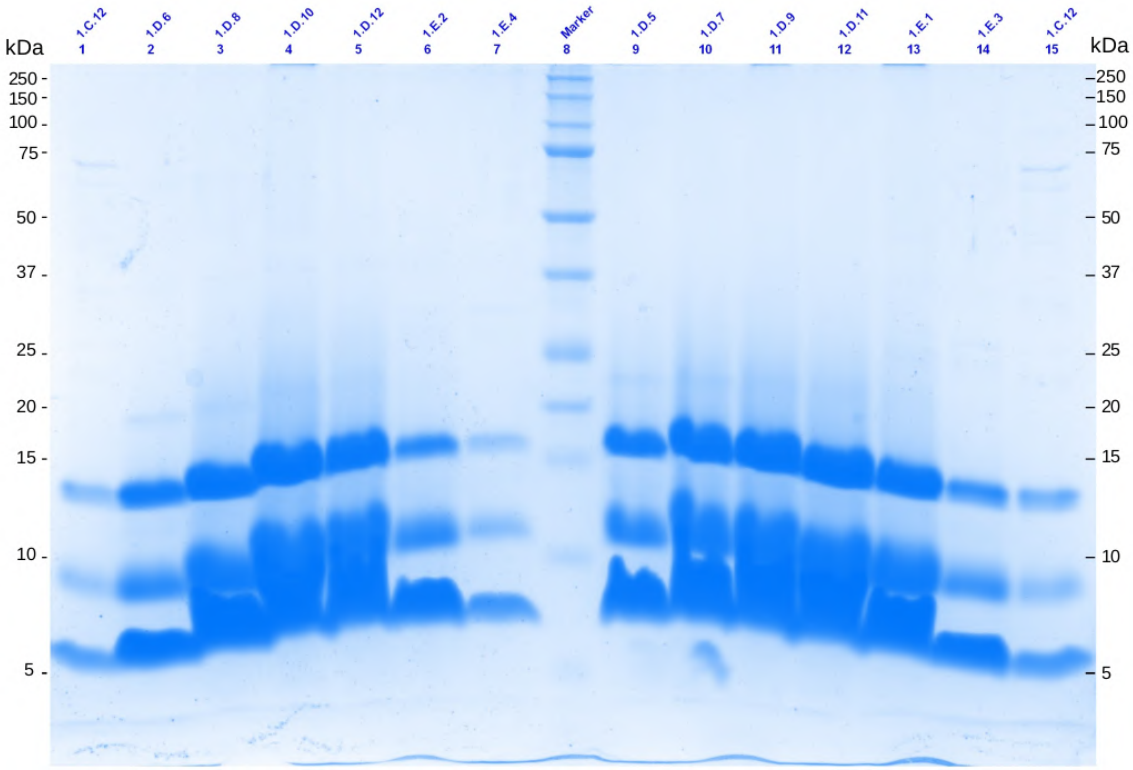

Figure 3: Fractions of ADC obtained after gel filtration using superdex-200 column

B.2 Molecular dynamics simulations: ADC

In order to ensure the stable ADC-LIG interaction, we plotted the distance between C_αs of CYS-26 and D-Serine ligand. The results are shown in figure 4

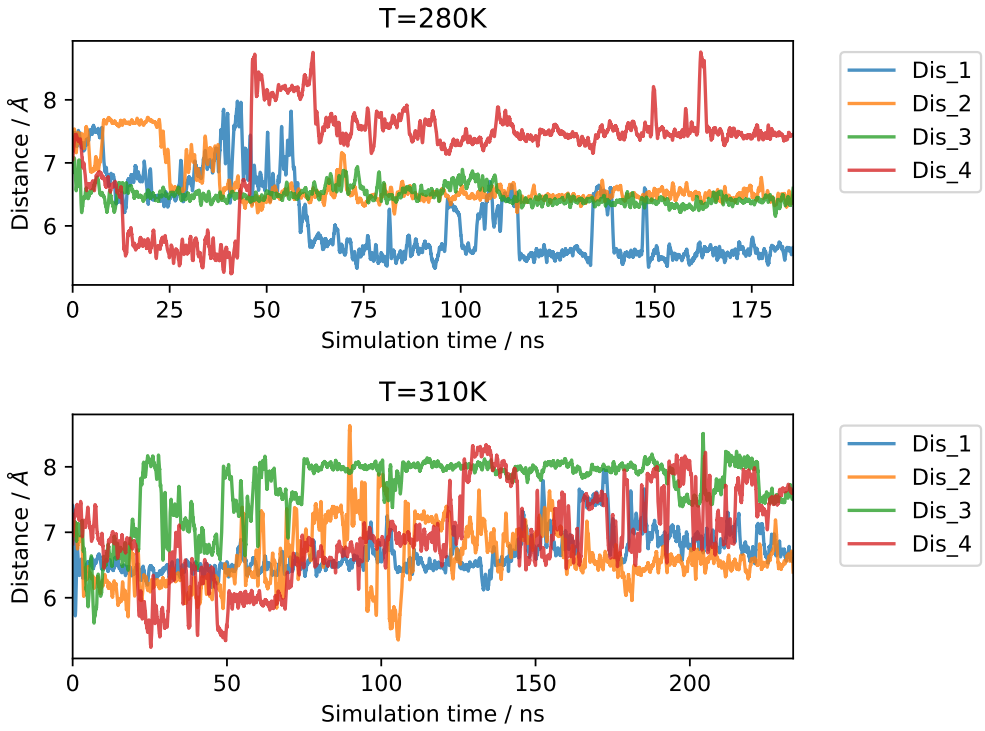

Figure 4: Distance between C_αs of CYS-26 and the D-Serine ligand in subunit-1 (blue), subunit-2 (orange), subunit-3 (green) and subunit-4 (red) of ADC-LIG system for the simulations at 280 (top) and 310 K (bottom).

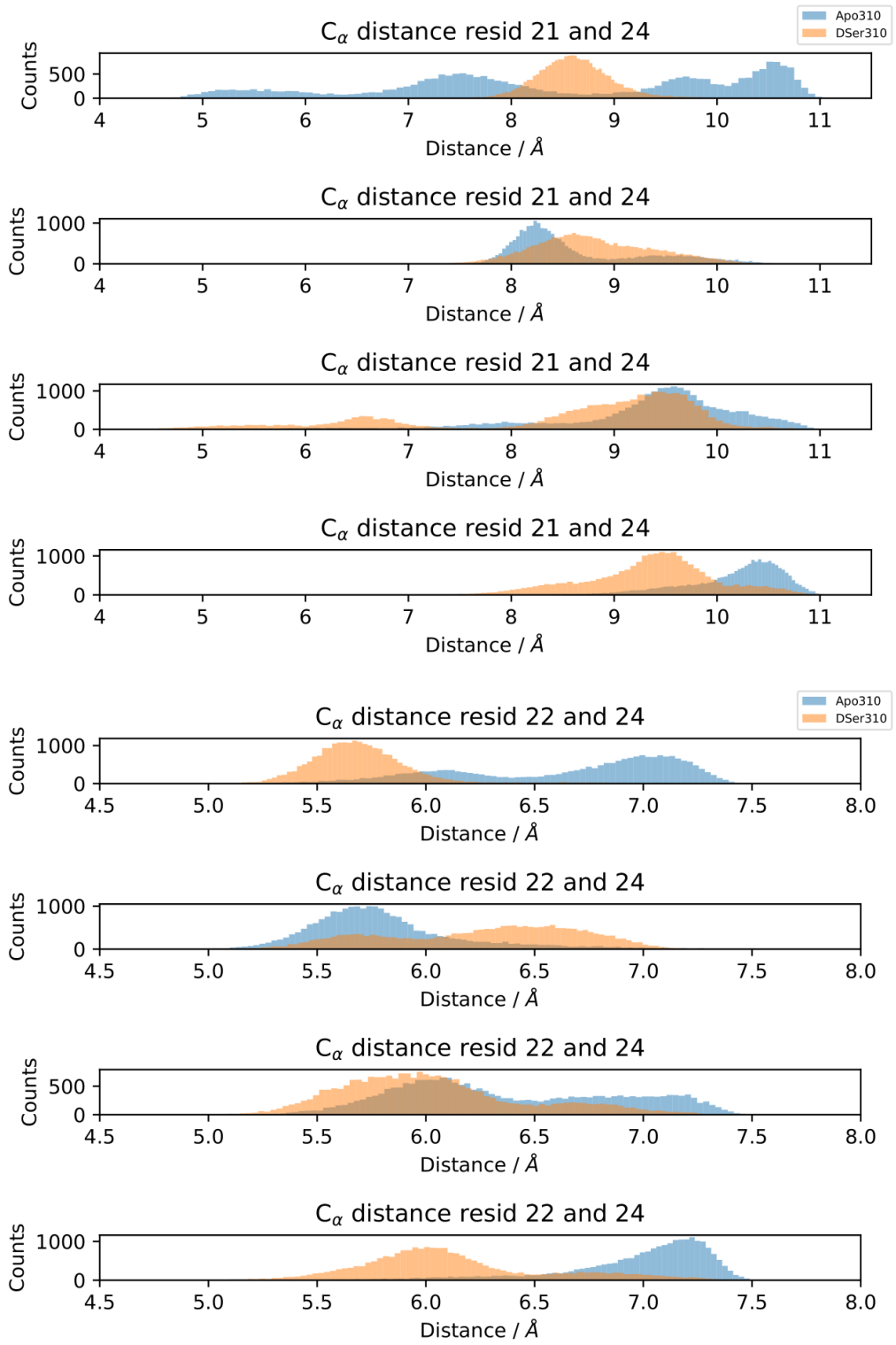

Figure 5: Histograms for the average C_α-C_α distances between HIS 21-GLY 24 (top) and between TYR 22-GLY 24 (bottom) for the individual subunits of ADC-APO (blue) and ADC-LIG (orange).

MDANSE software was used to calculate the EISF for ADC-APO and ADC-LIG systems from MD simulations. The EISF plots are generated at increasing timesteps and an increasing time of production simulation.

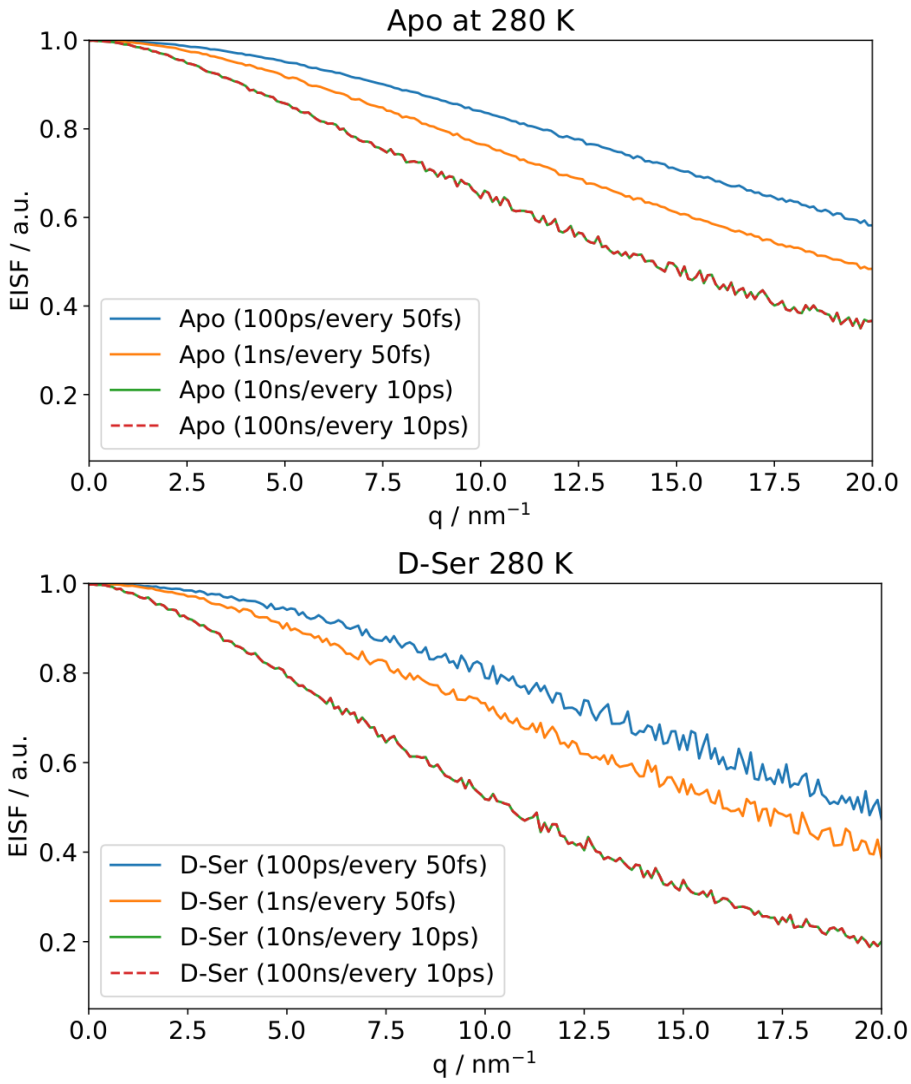

Figure 6: Elastic incoherent structure factor calculated from the molecular dynamics simulation of ADC-APO and ADC-LIG at different timesteps in the production run.

C Supporting information for chapter 5

C.1 Inelastic neutron scattering: ADC

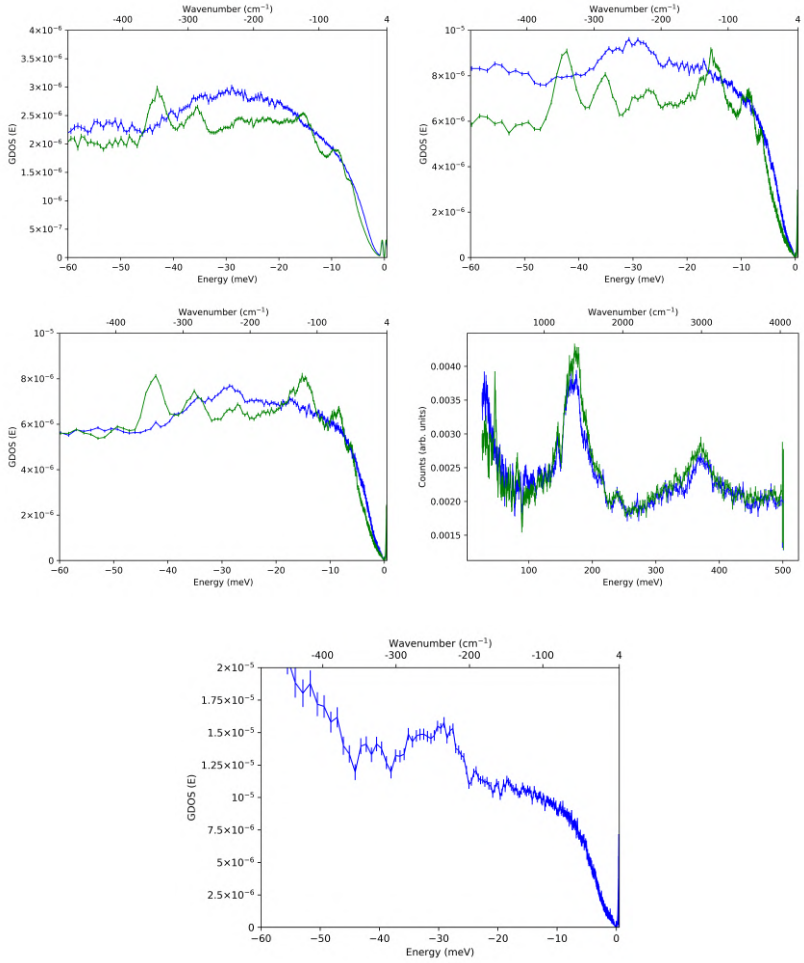

Figure 7: GDOS from IN5 data at 2.5 Å collected at 200K (top left), at 10 Å collected at 200K (top right), at 10 Å collected at 250K (middle left), at 10 Å collected at 100K and IN1 vibrational spectra collected at 250K (bottom). The signal is empty can subtracted and normalized. The normalization was carried out by dividing the spectrum of a sample by the total spectral intensity of that sample.

D Supporting information for chapter 6

D.1 Optimization of the growth medium and temperature for expression of wild-type ECAO

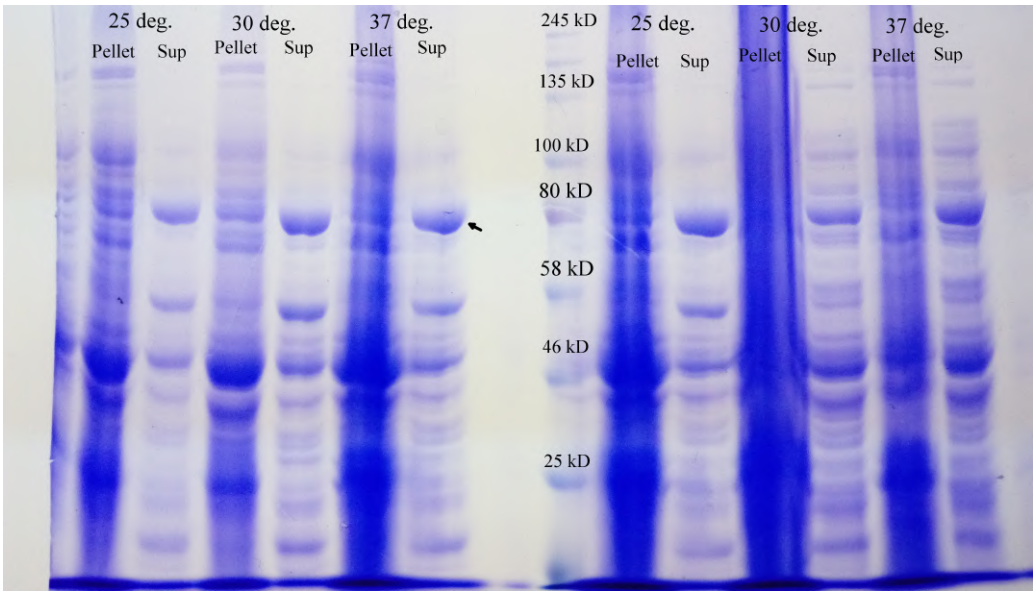

Figure 8: Expression test for the wild-type *E.coli* copper amine oxidase at different temperatures and in different media. Starting from the left, first six lanes denote expression in Luria Bertani (LB) broth and the last six in Terrific broth (TB). The lane in the center denotes the molecular weight marker. 'Sup' denotes supernatant and 'Pel' denotes pellet obtained after the enzymatic lysis of the cell pellet. Bands in the 8th lane represent the molecular wright marker.

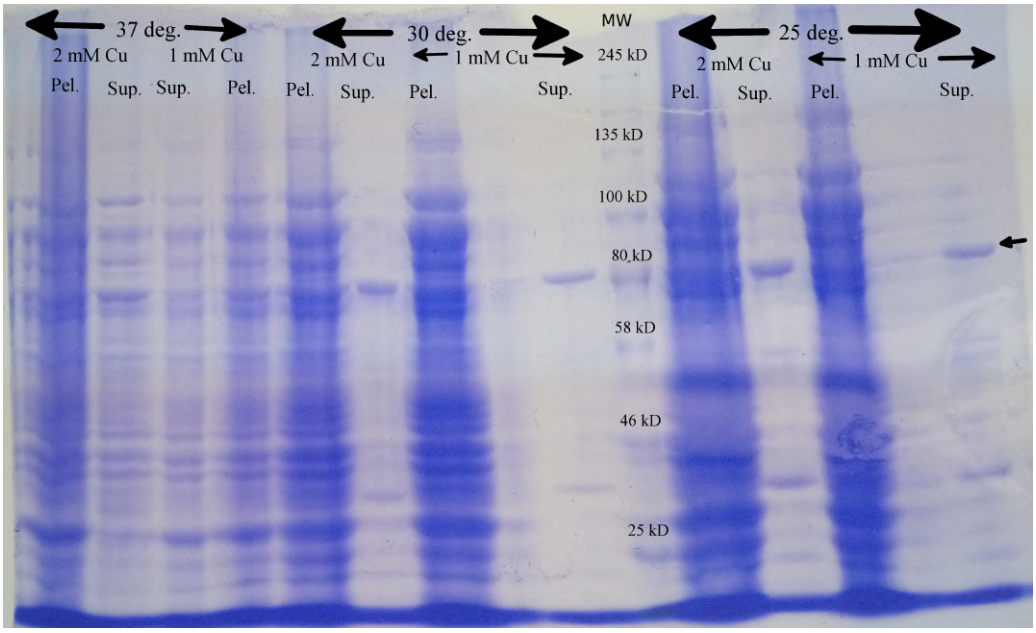

Figure 9: Expression of wild-type *E.coli* copper amine oxidase at different temperatures and at different concentrations of $CuSO_4$. 'Sup' denotes supernatant and 'Pel' denotes pellet obtained after the enzymatic lysis of the cell pellet. The ECAO band is marked with an arrow. MW - molecular weight marker with the calibration bands indicated in kilodaltons (kDa)

D.2 Purification of wild-type and E573Q/I342F mutant of ECAO

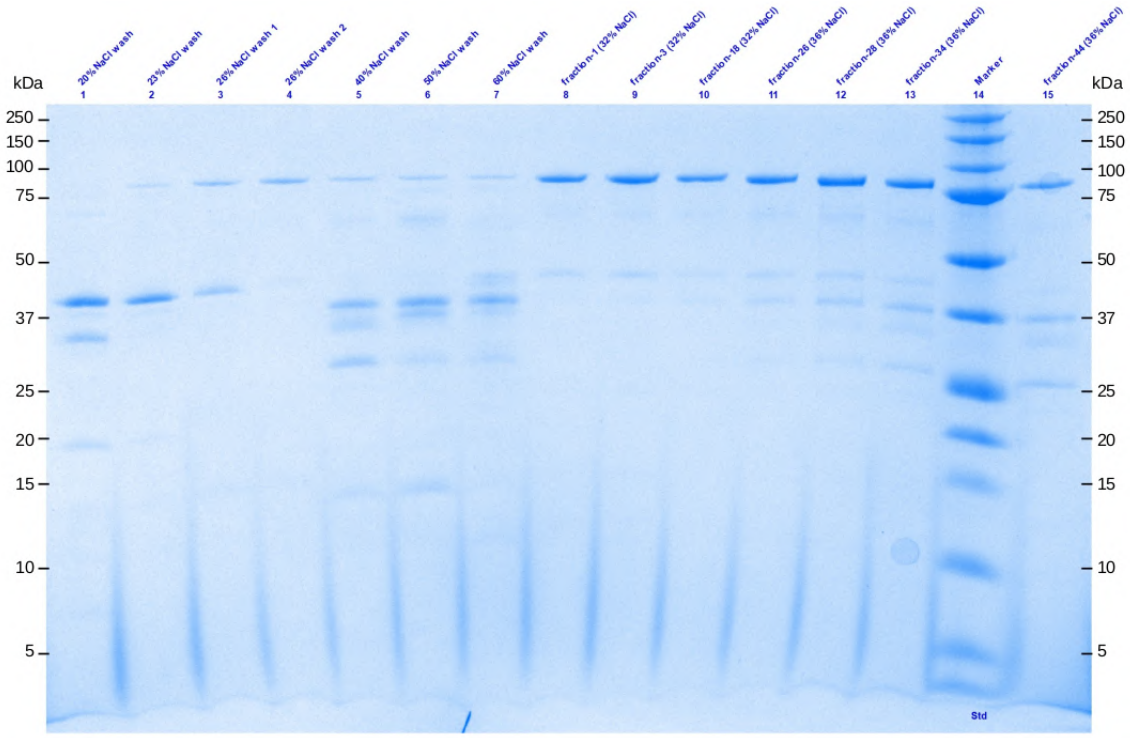

Figure 10: Fractions obtained after ion exchange chromatography of E573Q/I342F double mutant of ECAO.
Labels: Apart from the NaCl washes in the wells 1 to 7, the rest of the wells show the eluted fractions of E573/I342F double mutant of ECAO. It can be seen that NaCl elution concentration lies between 30 and 40 %.

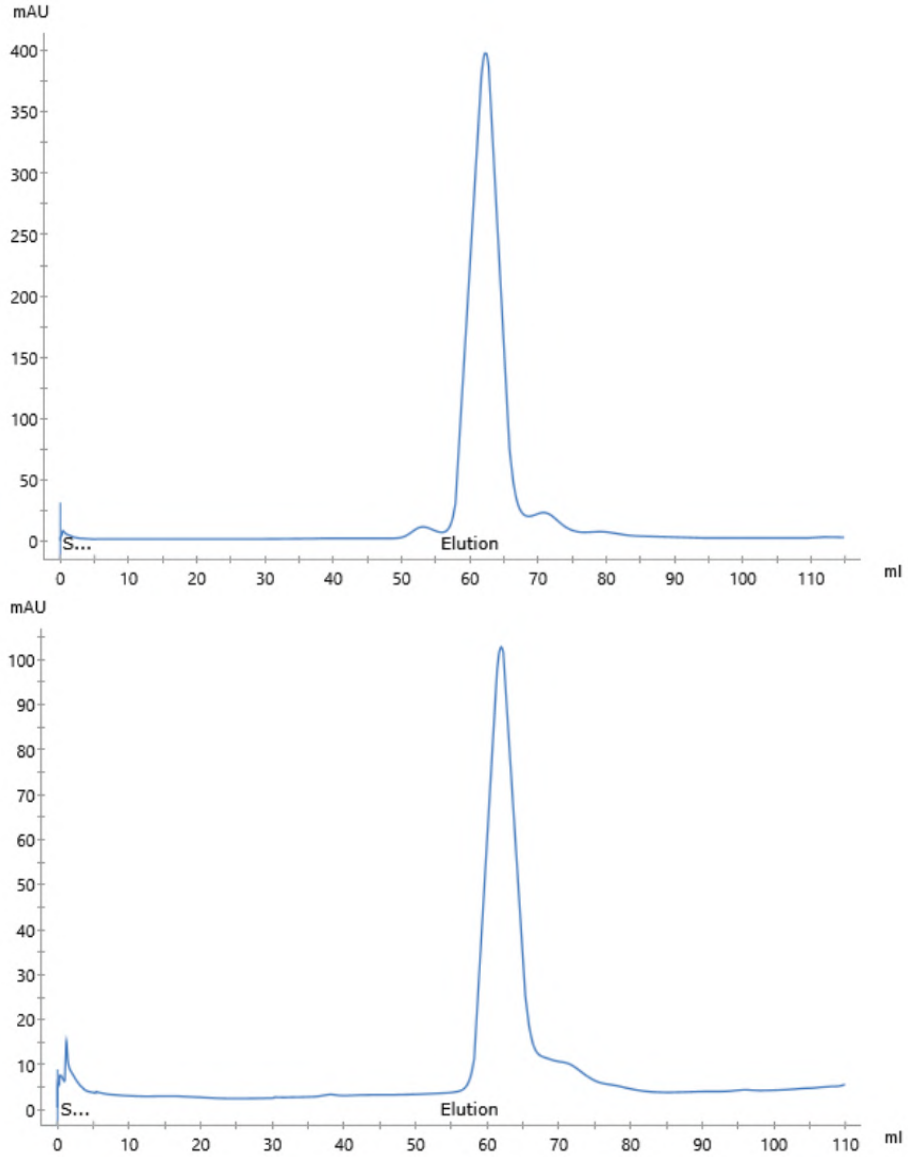

Figure 11: UV_{280} trace for hydrogenated wild-type (top) and double mutant (E573Q/I342F) (bottom) *E.coli* copper amine oxidase during S200 gel filtration.

D.3 X-ray refinement statistics of hydrogenated wild-type ECAO

Table 1: Crystallographic refinement statistics for hydrogenated wild-type ECAO
The numbers in parenthesis represent the high resolution shell

	Hydrogenated wild-type ECAO
Wavelength (Å)	0.985
Resolution range (Å)	83.56 - 1.52 (1.60 - 1.51)
Space group	P212121
Unit cell	79.85 134.9 167.1 90 90 90
Total reflections	1215398 (243080)
Unique reflections	235032 (35827)
Multiplicity	5.0 (5.0)
Completeness (%)	98.9 (98.0)
Mean I/sigma(I)	22.8 (4.5)
Wilson B-factor (Å^2)	22.21
R-merge	0.051 (0.168)
R-meas	0.054 (0.176)
R-pim	0.027 (0.604)
CC1/2	0.997 (0.935)
CC*	1 (0.99)
Reflections used in refinement	245021 (34168)
Reflections used for R-free	11788 (1731)
R-work	0.2052 (0.3136)
R-free	0.2336 (0.3476)
CC(work)	0.998 (0.740)
CC(free)	0.965 (0.728)
Number of non-hydrogen atoms	13280
macromolecules	11376
ligands	6
solvent	1768
Protein residues	1440
RMS(bonds) (Å)	0.009
RMS(angles) (°)	1.43
Ramachandran favored (%)	94.04
Ramachandran allowed (%)	5.54
Ramachandran outliers (%)	0.42
Rotamer outliers (%)	4.40
Clashscore	2.5
Average B-factor (Å^2)	21.6
macromolecules (Å^2)	20.4
ligands (Å^2)	24.7
solvent (Å^2)	34.4

D.4 *In-vitro* transposition of pKKECAO plasmid, expression and test neutron diffraction experiment on perdeuterated wild-type ECAO and X-ray data collection statistics of hydrogenated H/D exchanged E573Q mutant of ECAO

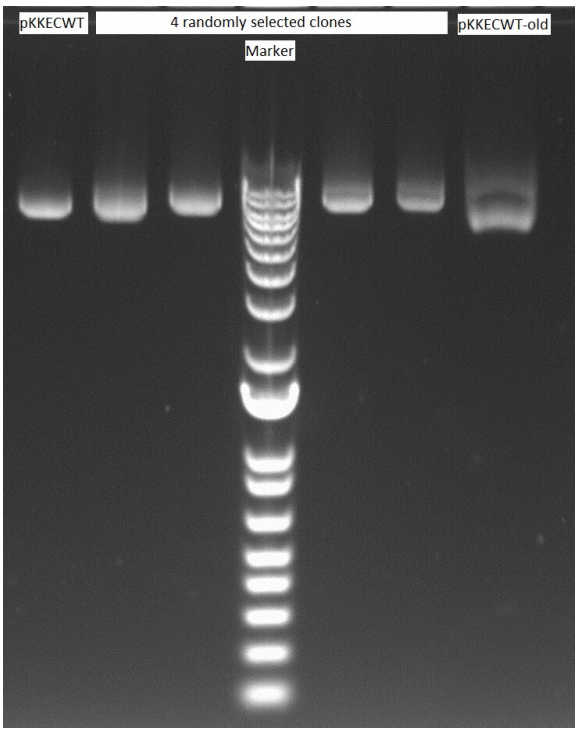

Figure 12: Agarose gel electrophoresis of the four plasmids of randomly selected in−vitro transposed clones of wild-type *E.coli* copper amine oxidase.

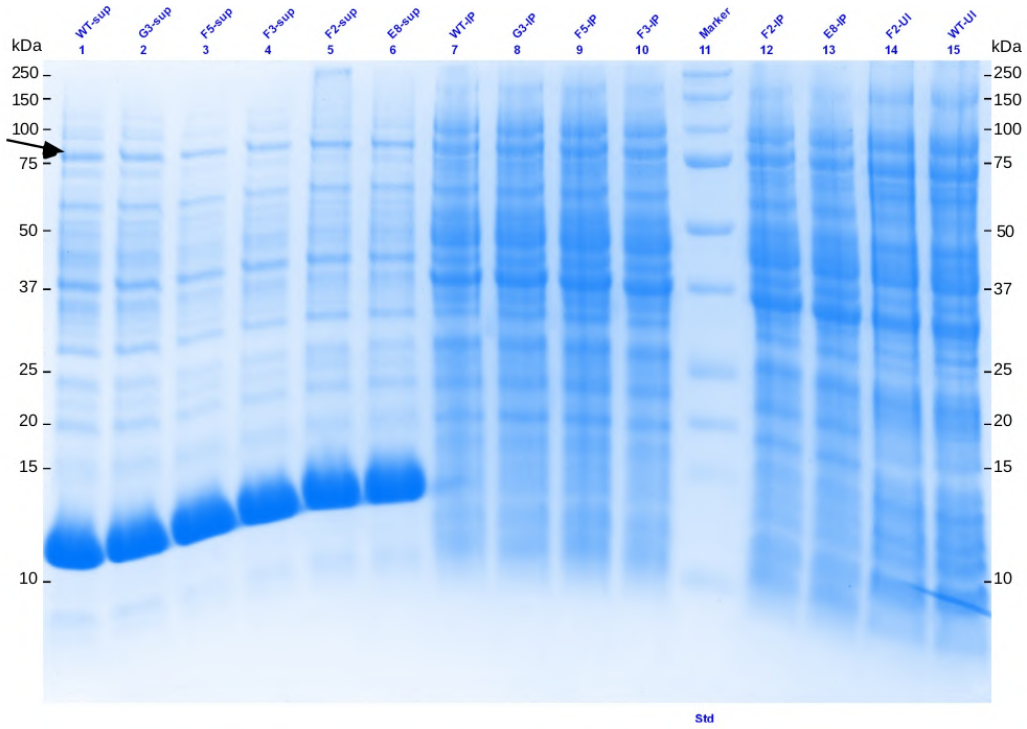

Figure 13: Expression test of five randomly selected *in−vitro* transposed clones of wild-type *E.coli* copper amine oxidase. The black arrow indicates the ECAO band. The prominent band around 14 kDa is the Lysozyme which was used for partial lysis of cells.

Labels: Labels with the suffix 'sup' are the supernatants obtained after enzymatic lysis of the respective induced pellets, the suffix 'IP' represent whole cells from the respective clones obtained after *in−vitro* transposition. The samples with suffix 'UI' are the whole uninduced cells. Marker: the molecular weight marker.

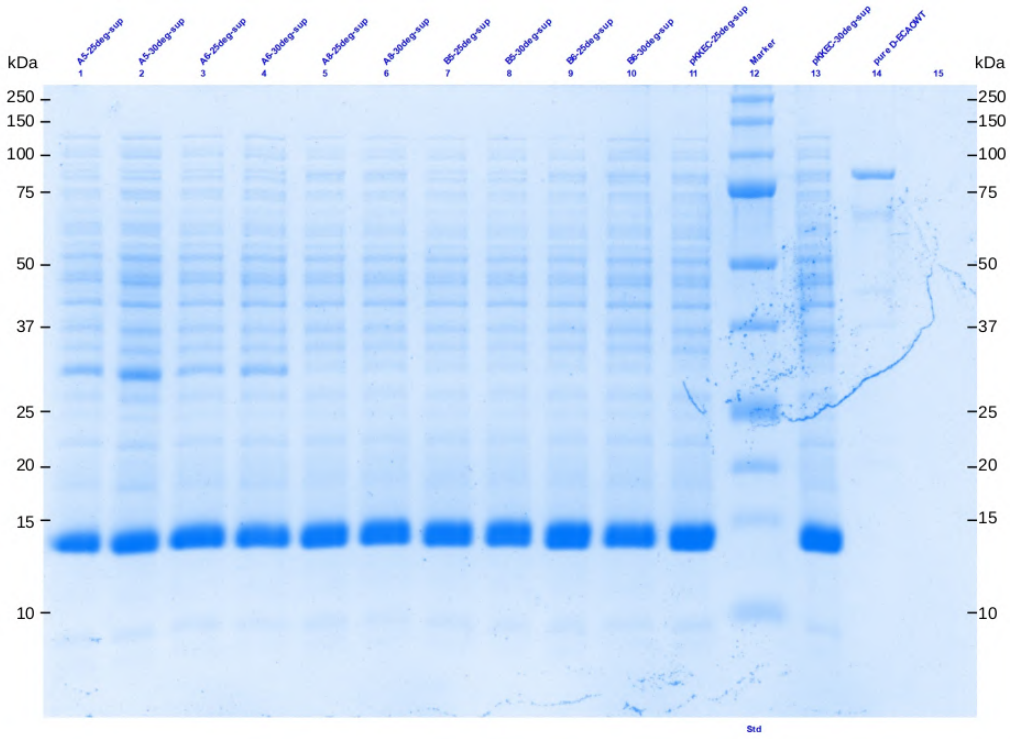

Figure 14: Expression test of five randomly selected *in−vitro* transposed clones of E573Q mutant of *E.coli* copper amine oxidase in LB medium.
Supernatants for clones A5, A6, A8, B5 and B6 induced at 25 and 30 °C. Pure D-ECAOWT - perdeuterated wild-type ECAO as a reference.

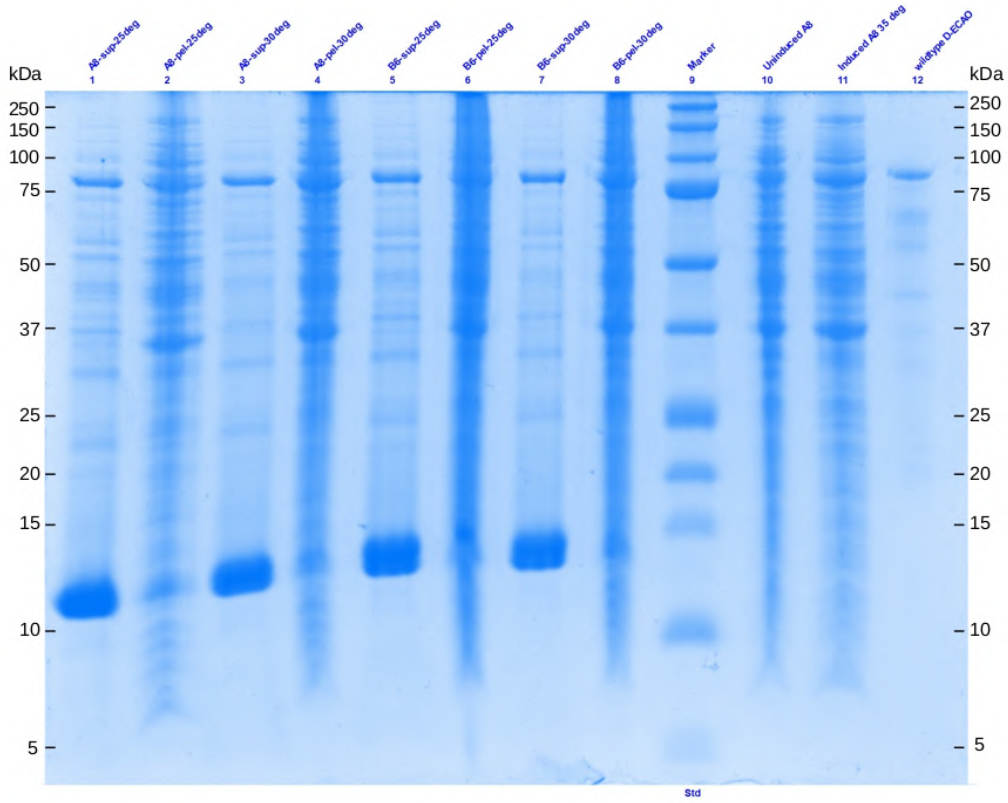

Figure 15: Expression test for plasmids from A8 and B6 clones in D-minimal medium. Labels: Labels with the suffix 'sup' are the supernatants obtained after enzymatic lysis of the respective induced pellets, the suffix 'IP' represent whole cells from the respective clones obtained after *in−vitro* transposition. The samples with suffix 'UI' are the whole uninduced cells. Marker: the molecular weight marker..

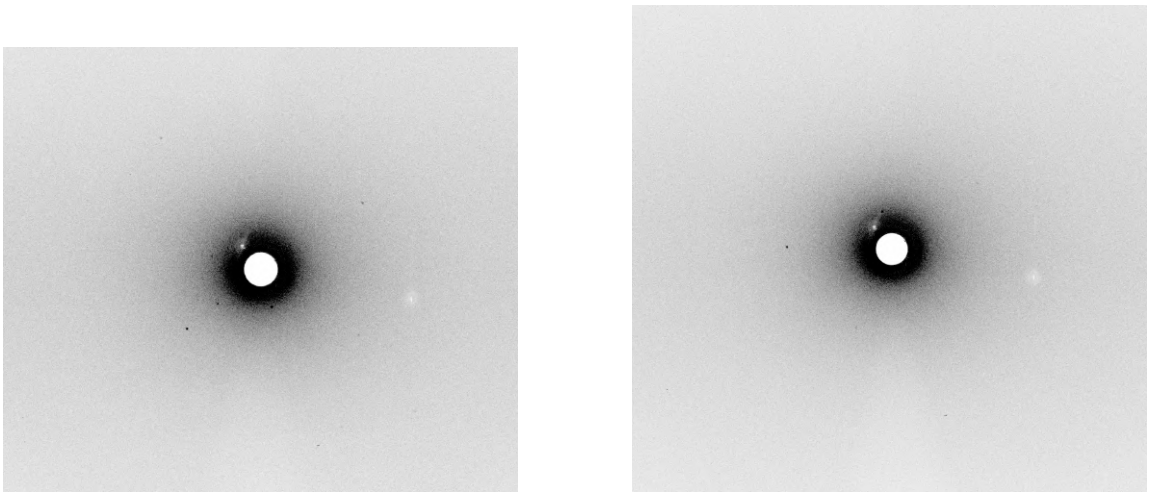

Figure 16: Neutron diffraction obtained from the crystals of perdeuterated E573Q mutant of ECAO.

Table 2: Crystallographic refinement statistics for H/D exchanged E573Q mutant of ECAO
The numbers in parenthesis represent the high resolution shell

	ECAO (E573Q mutant)
Wavelength (Å)	0.985
Resolution range (Å)	83.56 - 1.62 (1.75 - 1.62)
Space group	P212121
Unit cell	79.85 134.9 167.1 90 90 90
Total reflections	123616 (247030)
Unique reflections	235742 (36078)
Multiplicity	5.0 (5.0)
Completeness (%)	98.9 (97.0)
Mean I/sigma(I)	21.5 (5.5)
Wilson B-factor (Å^2)	20.21
R-merge	0.062 (0.)
R-meas	0.055 (0.212)
R-pim	0.025 (0.874)
CC1/2	0.999 (0.947)
CC*	1 (0.99)
Reflections used in refinement	238023 (35768)
Reflections used for R-free	12194 (1750)
R-work	0.2003 (0.3205)
R-free	0.2306 (0.3720)
CC(work)	0.998 (0.811)
CC(free)	0.955 (0.799)
Number of non-hydrogen atoms	13076
macromolecules	11376
ligands	4
solvent	1696
Protein residues	1440
RMS(bonds) (Å)	0.012
RMS(angles) (°)	1.56
Ramachandran favored (%)	96.04
Ramachandran allowed (%)	5.54
Ramachandran outliers (%)	0.42
Rotamer outliers (%)	4.45
Clashscore	3.5
Average B-factor (Å^2)	23.6
macromolecules (Å^2)	22.4
ligands (Å^2)	24.7
solvent (Å^2)	36.7

D.5 Molecular dynamics simulations: ECAO

Radius of gyration (R_g) was plotted for all three systems (wild-type, E573Q and E573Q/I342F mutants) of ECAO in order to monitor any global changes in the dynamics with simulation time.

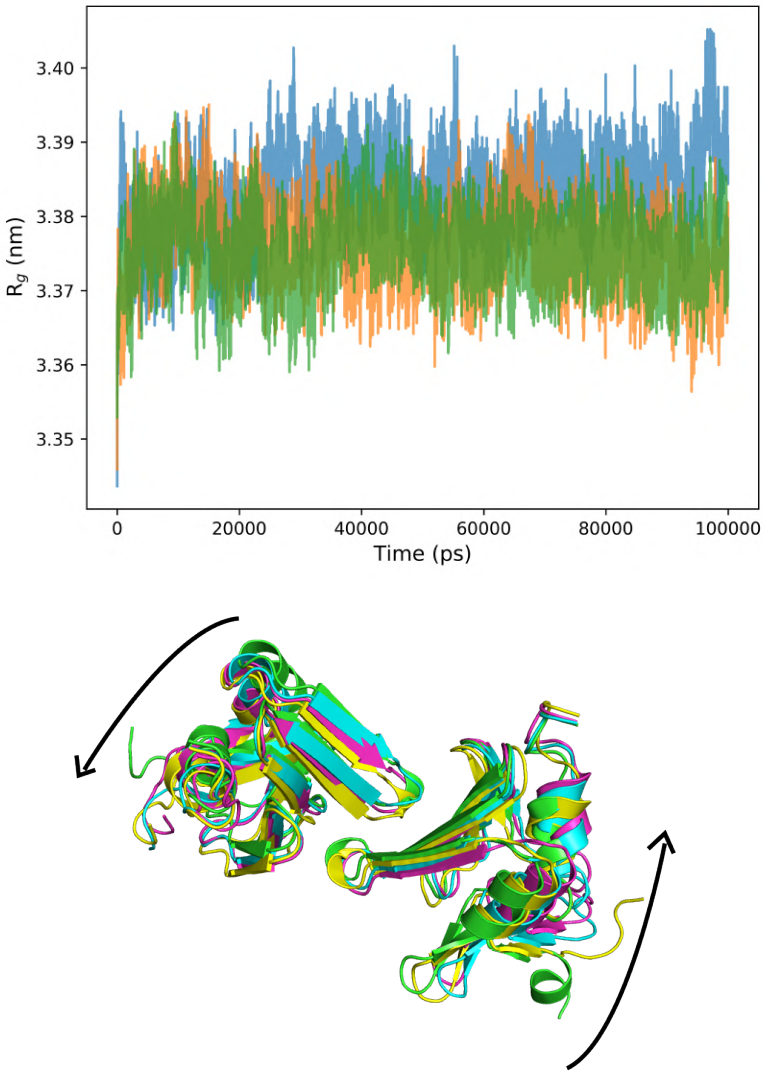

Figure 17: Radius of gyration (R_g) versus time (ps) for wild-type (blue), E573Q (orange) and E573Q/I342F (green) mutants of ECAO (top) and the stalk region of ECAO indicating the twisting motion.

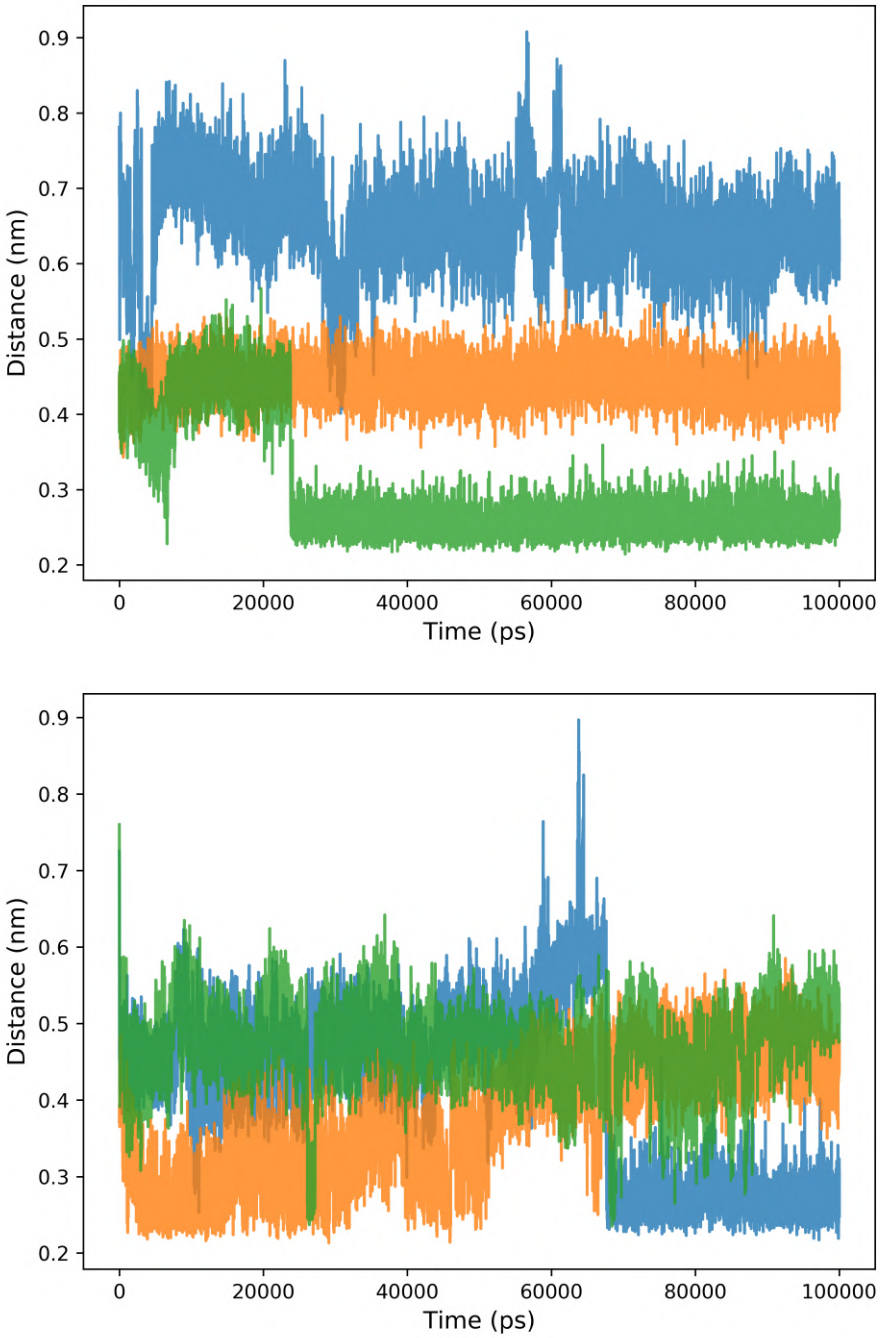

Figure 18: Distance between O4 of TPQ/TYR-466 and Cu(II) of wild-type (blue), E573Q (orange) and E573Q/I342F (green) mutants of ECAO for subunit A (top) and subunit B (bottom).

E Neutron beamtimes

Table 3: A list of all the neutron scattering experiments carried out during the thesis. The DOI (Digital Object Identifier) can be used to access the raw , the reduced data and the logfiles.

Experiment	Sample	Nature of the sample	Concentration (mg/mL) (solution)/ amount (mg) and level of hydration (%)	Wavelength (Å)	Experiment number/DOI	Instrument
QENS experiments on ADC-APO and ADC-LIG, neutron diffraction experiment on perdeuterated wild-type and E573Q mutant of ECAO	ADC, ECAO	Solution for IN5 and IN16B, Crystal in case of LADI-III	50 (ADC-APO and ADC-LIG)	5 (IN5), 6.27 (IN16B), 2.9 to 3.9 (LADI-III)	10.5291/ILL-DATA.8-05-428	IN5 (TOF) IN16B (NBS) LADI-III (test day)
QENS experiments on ADC-APO, ADC-LIG, wild-type and E573Q/I342F mutant of ECAO	ADC, ECAO	Solution	134 (ADC-APO and ADC-LIG), 100 (ECAO)	5 (IN5), 6.27 (IN16B)	10.5291/ILL-DATA.8-05-431	IN5 (TOF) IN16B (NBS)
Vibrational spectroscopy on ADC-APO and ADC-LIG	ADC	Powder	290 (38.09 % hydrated) (ADC-APO), 371 (37.70 % hydrated) (ADC-LIG)	5 (IN5), final energy of 4.5 meV (IN1)	10.5291/ILL-DATA.8-05-434	IN1 (LAGRANGE) IN5 (TOF)
Vibrational spectroscopy on ADC-APO and ADC-LIG	ADC	Powder	290 (38.09 % hydrated) (ADC-APO), 371 (37.70 % hydrated) (ADC-LIG)	2.5, 5, 10 (IN5), final energy of 4.5 meV (IN1)	10.5291/ILL-DATA.8-05-447	IN1 (LAGRANGE) IN5 (TOF)
Vibrational spectroscopy on ADC-APO and ADC-LIG	ADC	Powder	131.1 (38.1 % hydrated) (ADC-APO), 156.1 (38.1 % hydrated) (ADC-LIG)	6.27	10.5291/ILL-DATA.8-05-459	IN16B (EFWS)

F X-ray beamtimes

Table 4: A list of all the X-ray diffraction experiments carried out during the thesis.

Experiment	**Sample**	**Wavelength (Å)**	**Experiment number/DOI**	**Beamline**
Single crystal X-ray diffraction experiments on H/D exchanged crystals of wild-type and E573Q/I342F double mutant of ECAO	ECAO (H/D exchanged)	0.82	BAG proposal	ID23.1 (ESRF)
Single crystal X-ray diffraction experiments on perdeuterated wild-type ECAO	ECAO (perdeuterated)	0.97	MX24717	I03 (Diamond)

G List of hazardous chemicals used

Reagent	GHS hazard	Hazard statement	Precautionary statement
Acetic acid, glacial	GHS02 GHS05	H226, H314	P280, P305+351+3 38, P310
Acrylamide 30%	GHS06 GHS08	H301, H312, H316, H317, H319, H332, H340, H350, H361f, H372	P201, P280, P301+310, P305+351+3 38, P308+313
Ampicillin	GHS08	H334, H317	P280, P261, P302+P352, P342+P311
Ammonium chloride	GHS07, GHS09	H302, H319	P270, P305+351+338
Ammonium persulphate	GHS03 GHS07 GHS08	H272, H302, H315, H317, H319, H334; H335	P280, P305+351+3 38, P302+352, P304+341, P342+311
Ammonium sulphate	GHS07, GHS09	H315, H319, H335	P261, P264, P270, P271, P273, P280, P301+P312, P302+352, P304+340, P305+351+338, P312, P321, P330, P332+313, P337+313, P362, P391, P403+233, P405, P501
Boric acid	GHS08	H360FD	P202, P281, P308+313, P405, P501
Calcium chloride	GHS07	H319	P395+351+338
Cobalt chloride hexahydrate	GHS07, GHS08, GHS09	H302, H317, H334, H341, H350, H360, H410	P201, P202, P261, P264, P270, P272, P273, P280, P284, P311, P303+361+353, P330, P333+313, P342+311, P391P405, P501
Copper chloride dihydrate	GHS05, GHS07	H290, H302, H314	P260, P280, P303+361+353, P305+351+338, P290, P501
Copper sulphate pentahydrate	GHS05, GHS07, GHS09	H301, H315, H319, H400, H410	P264, P273, P280
Di-ammonium hydrogen citrate	GHS07	H319	P395+351+338
Dithiothreitol	GHS07	H302, H315, H319, H335	P302+352, P305+351+33 8
Ethylenediaminetetraacetic acid	GHS07	H319	P305+351+3 38
Ethanol	GHS02	H225	P210

Ferric chloride	GHS05, GHS07	H290, H302, H314, H318	P234, P260, P264, P270, P273, P280, P301+312, P301+330+331, P303+361+353, P363, P304+340, P310, P3221, P305+351+338, P390, P405, P406, P501
Hydrochloric acid >25%	GHS05, GHS07	H314, H335	P261, P280, P310, P305+351+3 38
Imidazole	GHS05, GHS06, GHS08	H301; H314; H361	P260, P281, P303+P361+P353, P301+P330+P331, P305+P351+P338, P308+P313
Isopropanol	GHS05, GHS06, GHS08	H225, H319, H336	P260, P281, P303+P361+P353, P301+P330+P331, P305+P351+P338, P308+P313
Kanamycin sulphate	GHS08	H360	P201, P202, P280, P308+313
Manganese sulphate tetrahydrate	GHS08, GHS09	H302, H411	P260, P273, P314, P391, P501
Monopotassium phosphate	GHS07	H315, H319	P264, P280, P305+351+338, P321, P332+313, P337+313
Nickel sulphate hexahydrate	GHS07, GHS08, GHS09	H302, H332, H315, H317, H334, H341, H350i, H360D, H372, H410	P260, P273, P280, P302+352, P304+340, P342+311
Phosphorous pentoxide	GHS05, GHS07	H314, H330	P260, P264, P271, P280, P284, P301 + P330 + P331, P303 + P361 + P353, P304 + P340, P305 + P351 + P338, P310, P320, P363, P403 + P233, P405, P501
Sodium selenite	GHS06, GHS09	H300, H330, H315, H317, H319, H411	P260, P280, P301+330+331+310, P304+340+310, P305+351+338
TEMED	GHS02 GHS05 GHS07	H225, H302, H314	P260, P264, P271, P280, P284, P301+330+331, P303+361+353, P304+340 P305+351+338, P310, P320, P363, P403+233, P405, P501
Tris	GHS07	H315, H319, H335	P261, P305+351+338
Zinc sulphate heptahydrate	GHS05, GHS07, GHS09	H302, H318, H410	P273, P280, P305+351+338, P313

G.1 GHS, risk symbols and information about hazards

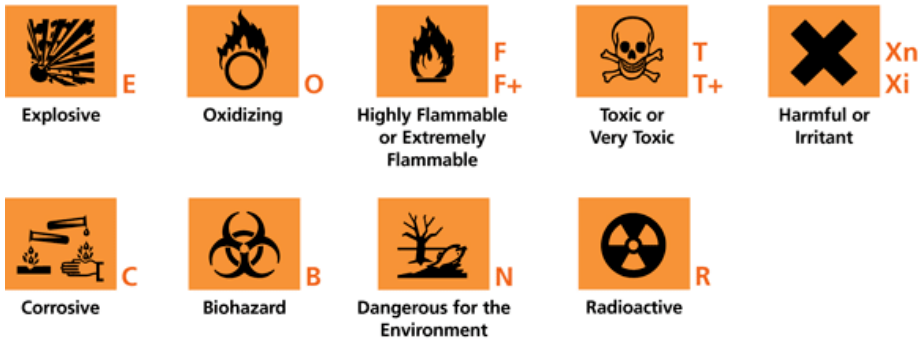

Figure 19: Hazard symbols for the respective labels according to http://www.sigmaaldrich.com.

Figure 20: GHS pictograms according to http://www.sigmaaldrich.com.