Ultrafast Electrons and X-rays as Probe of Biomolecular Dynamics

by

Ganesh Subramanian

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved October 2016 by the
Graduate Supervisory Committee:

John Spence, Chair
Peter Rez
Terry Alford
Uwe Weierstall
Richard Kirian

ARIZONA STATE UNIVERSITY

December 2016

ABSTRACT

The structure-function relation in Biology suggests that every biological molecule has evolved its structure to carry out a specific function. However, for many of these processes (such as those with catalytic activity) the structure of the biomolecule changes during the course of a reaction. Understanding the structure-function relation thus becomes a question of understanding biomolecular dynamics that span a variety of timescales (from electronic rearrangements in the femtoseconds to side-chain alteration in the microseconds and more). This dissertation deals with the study of biomolecular dynamics in the ultrafast timescales (fs-ns) using electron and X-ray probes in both time and frequency domains.

It starts with establishing the limitations of traditional electron diffraction coupled with molecular replacement to study biomolecular structure and proceeds to suggest a pulsed electron source Hollow-Cone Transmission Electron Microscope as an alternative scheme to pursue ultrafast biomolecular imaging. In frequency domain, the use of Electron Energy Loss Spectroscopy as a tool to access ultrafast nuclear dynamics in the steady state, is detailed with the new monochromated NiON UltraSTEM and examples demonstrating this instrument's capability are provided.

Ultrafast X-ray spectroscopy as a tool to elucidate biomolecular dynamics is presented in studying X-ray as a probe, with the study of the photolysis of Methylcobalamin using time-resolved laser pump – X-ray probe absorption spectroscopy. The analysis in comparison to prior literature as well as DFT based XAS simulations offer good agreement and understanding to the steady state spectra but are so far inadequate in explaining the time-resolved data. However, the trends in the absorption simulations for the transient intermediates show a strong anisotropic dependence on the axial ligation, which would define the direction for future studies on this material to achieve a solution.

DEDICATION

This dissertation is a dedication to God, Guru, Parents, extended family, teachers and every single well wisher that has sacrificed for, inspired and blessed me.

Endaro Mahanubhavulu Andiriki Vandanamu …

ACKNOWLEDGMENTS

To my advisor Prof. Spence, Dr. Weierstall, Dr. Kirian, Dr. Zatsepin, Dr. Shah, the STC award and Deanna Clark - you ensured my journey of its achievement.

To Shibom, Chufeng and Yun – you ensured my journey was never lonely.

To Prof. Fromme (ASU), Prof. Schmidt (UWM), Prof. Chapman (DESY), Prof. Cherezov (USC), Dr. Bostedt (APS), Dr. Aoki, Dr. Kodis, David Lowry, Karl Weiss and their corresponding groups – you ensured a happy extended family of collaboration.

To Dr. Mandar Deshmukh - you ensured I embarked on the journey.

To Sampath Sir - you ensured I dreamt of a journey.

To Summer Day, DAV, BITS-Pilani and TIFR – you ensured I was prepared for the journey.

TABLE OF CONTENTS

Page

Page

Page

Page

Page

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

Figure Page

Figure Page

Figure Page

Figure Page

Figure Page

Figure Page

Figure Page

Figure Page

Figure Page

INTRODUCTION

One of the most primal questions that has yet to be solved is 'What is life?' Every society has had its own metaphysical hypothesis to elucidate this. A branch of Indian philosophy, centuries before Christ, proposed that, "From the ether came air, fire from air; water from fire; earth from water, shrubs from earth, lower forms of life from shrubs (larger plants and lower animals) and finally man from other forms of life." (Note: this is translated and not exact). We are familiar with such a progression of logic as 'methodologically reductionist'. Each of the five natural elements are placed in the order of increasing number of sensory perceptions – Ether can be heard (ear); air can be heard and felt (ear and skin); fire can be heard, felt and seen (ear, skin and eyes); water can be heard, felt, seen and tasted (ear, skin, eyes and tongue); and earth can be heard, felt, seen, tasted and smelt (ear, skin, eyes, tongue and nose) before progressing to plants, animals (herbivores and then carnivores) and eventually man. Interestingly, this approach has a very physical, albeit primitive, sense of reality.

Millennia later, as sensory perception extended leaps and bounds with technological advancements, a series of experiments on understanding fermentation, starting with Louis Pasteur and followed by Marcelin Berthelot and Buchner brothers, revealed that (a) microorganisms were involved in the chemical process of converting sugar to alcohol (b) these chemical processes could be performed by entities (proteins) extracted from and outside of the microorganism itself. [1] This birthed the currently accepted notion of 'life' as – "a consequence of numerous chemical process made possible thanks to proteins". [2] The functions of proteins were identified to range from catalysis to transportation to gene expression. But, how could proteins achieve such a variety in functional capacity?

The answer is in the structural complexity. An amino acid sequence (primary structure) could be bound by hydrogen bonds to induce helices/pleats (secondary structure), which eventually could be held together into a tight three-dimensional (3D) geometry (tertiary structure) by a variety of non-covalent interactions. Occasionally there are also cases of additional multi-polypeptide conjunctions (dimers, trimers etc.) to give one multi-subunit protein. To give a

perspective of the complexity – there are 20 naturally occurring amino acids. If there are N positions in a linear amino acid sequence, the number of different ways to arrange them is N^{20} (for a really conservative 10 position sequence, the total number of possibilities of producing an amino acid chain is a hundred billion times billion! notwithstanding the secondary and tertiary structural possibilities for each such case). This immense diversity in protein structure meant an immense possibility in at least a few of them being functionally relevant. Scientists like Emil Fischer, [3] Linus Pauling [4] and Christian Anfinsen [5] went a step ahead to propose and confirm that 'the specific 3D structure of a protein directly determines its function'. The manifestation of the importance of this structure-function relation is the distinct research field of 'Structural Biology'.

Naturally then, the scientific efforts subsequently were aimed at understanding the structure of these proteins and how they correlated with their respective functions. The aim being three-fold: (i) to understand human and other life forms better, (ii) to artificially mimic protein functions (iii) to cure diseases. As much as the structural information is relevant, many functions of the proteins include a reversible or permanent structural reconfiguration. This meant that it was not only important to obtain 'static' protein structures, to completely understand the working of the protein it was imperative to watch the structure evolve, in time, through its duty cycle. Such dynamic structure characterization of the proteins is the elusive golden-egg that the current generation of scientific research is aiming to grasp and in the following sub-section, the major structural characterization techniques used to study proteins will be briefly introduced highlighting the current state of art and the advantages and limitations vis-à-vis each of the techniques.

1 Methods of Protein Structure Determination:

The major protein structure determination techniques are almost all based on the idea of an incident radiation source interacting with the protein sample. This interaction can be elastic (no energy lost/gained) or inelastic (energy loss/gain present). Accordingly the techniques of structure determination can be categorized as scattering/diffraction and spectroscopic. Amongst the various types of such structure determination present, 4 major types are highlighted below.

1.1 X-ray Diffraction

The concept of diffraction is well represented by Bragg's law, which relates the inter-planar spacing (d) to the wavelength of the incident radiation (λ) as given below:

$$2dSin\theta = n\lambda$$

Diffraction of this form is obtained only when the sample is crystalline (possessing long range order). For a 3D object in real space, this corresponds to a Fourier related reciprocal space which is also 3D. A true infinite 3D crystal would diffract intensities that are infinitesimal points. But a finite crystal implies a finite spot volume. X-rays, whose typical probing energies for these studies are 5 – 12 keV (corresponding to about 1 Å in wavelength), would slice through a few such spots in the reciprocal space (curvature of the Ewald's sphere determines the crossing through the 3D reciprocal volume). In an ideal experiment, this same measurement would be repeated for different orientations of the protein crystal with the X-ray beam, until a sufficient number of the reciprocal diffraction spots have been sampled. Then by using methods such as Molecular Replacement, the information from reciprocal space is converted to real space charge densities and the structure of the protein thus deciphered. The typical working steps in an X-ray diffraction (XRD) measurement would be (a) isolation and purification of the protein, (b) crystallization of the protein, (c) Irradiation and data acquisition and (d) diffraction data analysis and reconstruction of real space densities. (Refer [6] for a detailed account of this technique).

In the more recent past, ultra-short (few 10s of femtoseconds), ultra-intense (10^{12} X-ray photons/pulse) and highly coherent X-ray lasers have been utilized to obtain structures of proteins. [7] The burst of research in using this X-ray source is thanks to the diffract-before-destroy methodology of obtaining data before the X-rays damage the sample. With the previous (3rd generation) X-ray sources, radiation damage [8,9] has been a major issue of concern. Serial crystallography at the X-ray laser facilities help not just avoid this problem but also collect data at orders of magnitude higher brightness thereby ensuring potentially faster data collection. Latest results [10,11] demonstrate the ability of the method to solve structures to very high resolutions in space and time.

1.2 Electron Crystallography (Diffraction and Imaging):

Electron crystallography is very similar to X-ray crystallography in the basic idea of diffraction of a probe wave by the crystal lattice planes. There a are a few major differences in the case of an electron probe (as compared to an X-ray probe): (a) a probe electron can diffract off an atomic electron as well as its nucleus, while X-rays diffract purely off the atomic electrons (b) electron beams have the ability to be focused and hence real-space images can be formed (not possible to the same extent with X-rays), (c) the very high energy of the incident beam (100s of keV as compared to 10 keV for X-rays) implies a much flatter Ewald's sphere in the (diffraction) reciprocal space and hence, much larger sampling of complete intensities and (d) since the scattering cross-sections of electrons are a few order of magnitude higher than that of X-rays, correspondingly the kinematic limit of crystal thickness is substantially smaller and therefore this technique is generally most successful with 2D crystals of proteins and more recently with single molecules. [12]

The biggest problem so far with electron crystallography is radiation damage. [13,14] Preliminary solutions to this problem included the use of negative staining [15] (heavy metal immersion). A more contemporary solution is to flash freeze the protein crystals in liquid N_2 or liquid He (Cryo-electron microscopy). This technique uses multiple crystals of the same protein to be imaged and averages to provide a better overall signal-to-noise (SNR) under the low (lower than damage threshold) electron dose. With the advent of direct electron detectors [16,17] and complementary analysis programs, the ability to detect signal with much higher sensitivity and speed helps obtain much better resolution providing a stiff competition to traditional X-ray diffraction.

1.3 Nuclear Magnetic Resonance (NMR)

NMR is based on the response of atoms to an external magnetic field. Those nuclei that have a non-zero spin moment can be aligned parallel/anti-parallel to the external field. The difference in energy of the two possible states typically falls in the frequency of radio waves (10s to 100s of MHz). This can be altered by interaction of one atomic nuclear spin moment with other

nearby electron/nuclear moments thereby causing chemical shifts. [18] Typically since the sensitivity of the difference in high-energy to low-energy spins is small, a couple of options of either (a) increasing the magnetic field or (b) increasing concentration are used to maximize the signal. Generally since a protein has thousands of atoms, there are thousands of resonance peaks that need to be separated to characterize the system. By repetitive measurements, the time difference between each measurement is controlled in such a manner that the correlation of adjacent nuclei presents itself in the spectrum (2D-NMR). [19] This helps provide good structure determination. For large proteins, 3D/4D NMR can be used as an option. Another option widely used is radioisotope labeling using ^{13}C, ^{15}N or ^{2}H. [20]

A substantial difference with the previously mentioned techniques is the non-necessity of crystals. Especially in studying dynamics of the protein in solution, NMR is perfectly suited to mimic/provide the native environment (conventionally performed in the liquid phase as a solution). Typical steps in the measurement of NMR structure of proteins is as follows: (i) preparing the protein solution, (ii) NMR measurement, (iii) signal assessment, (iv) structure determination using either geometry-based or energy-based calculations. A major drawback however in the customized nature of performing these experiments is assigning spectral signatures, which takes a lot more time than other comparable techniques. (refer [21] for details on this technique).

1.4 X-ray Spectroscopy

Complementary to the diffraction methods, the spectroscopic measurements are primarily based on energy lost by the X-rays on interaction with the sample. With electronic excitations, the energy lost by the probe is used by the core-electron to fill any of the unoccupied empty states, each with a probability (upholding the selection rules), as reflected in the X-ray absorption spectra (XAS). The hole (vacancy) left in the core-shell is unstable being close to the nucleus and hence is re-filled by an electron from a relatively higher energy shell (farther away from the nucleus than the existing hole). This gives rise to the X-ray emission spectra (XES), again with probabilities upholding the coupling and selection rules. These X(A/E)S data give insight about the electronic structure of the inspected system. In the specific case of the XAS data, for kinetic energies 20 eV

or more beyond the absorption edge, the ejected photoelectron is completely unbound. This then scatters (singly and multiply) with the neighboring atoms causing resonance at the X-ray energies (wavelengths) that are inversely related to the total scattering path. This represents itself as oscillations in the higher energy region of the XAS spectra. With good models, this can be simulated and with goodness of fit measures, the experimental data can then be correlated to simulations and hence bond-distances and coordinations identified. In this way, the XAS data does also reveal structural information, however due to extreme multiplicity and low signal, is generally localized around the excited atom. For this specific reason, these measurements do not require crystals. However, within the first (few) coordination shells that can be identified, the resolution is on the scale of a tenth of Angstrom or better. [22] In the case of diffraction, this would require an extraordinary perfect crystal. (refer to [23–25] for XAS and [26,27] for XES in detail).

There are many other techniques (http://www.esciencecentral.org/ebooks/advances-in-protein-chemistry/APC_recent_methodologies.php) used to study the dynamics and structure of proteins that include UV-Vis Absorption, Electron Paramagnetic Resonance (EPR), Fourier Transform Infra-Red Spectroscopy (FT-IR), Circular Dichroism (CD), Atomic Force Microscopy (AFM), Dual Polarization Interferometry (DPI), Raman Spectroscopy, Hydrogen Exchange Mass Spectroscopy, Fluorescence Correlation Spectroscopy (FCS), Fluorescence Recovery After Photo-bleaching (FRAP). The above-mentioned techniques are normally coupled with physical models (molecular simulations) and data analysis to give explanations of the characteristic biophysical dynamics of proteins. These measurement schemes (NMR included) are not directly relevant to this dissertation, and hence will not be dealt with in any further detail.

2 Aim and Scope of This Dissertation

As mentioned earlier, one of the ultimate frontiers of current scientific research is to decrypt the structure (dynamics) of biomolecules and relate them to their function. The two factors that determine the fate of such a study is (i) the ability to avoid radiation damage (ii) the ability to probe dynamic changes. This dissertation will address this issue via two parts.

The first part of the dissertation is dedicated to electron as the probe. The first chapter of this part will start with understanding the ability and limitations of electron crystallography in delivering structures of proteins, more specifically in handling dynamic changes. It will then proceed to the second chapter where alternative approaches to tackle the threshold factors (mentioned above) will be suggested. Measurements to study dynamics can be performed in two modes – (a) time domain and (b) frequency domain. The second chapter thus is dedicated to proposing alternatives to support dynamical studies respectively in both these modes.

The dissertation will then proceed to X-ray as the probe in part two. Here, the final chapter of the dissertation will demonstrate an example of using spectroscopy in studying dynamical electronic and structural changes in the metallocofactor Methylcobalamin (MeCbl). Experimental studies on the ultrafast X-ray probes will be supported by measurements using optical photons and simulations using Gaussian (DFT) and FEFF (XAS). As a combination, these experimental and theoretical approaches will together highlight new information underlying the photolysis of MeCbl.

Finally, the dissertation will conclude with the overall conclusions summarizing the results from each chapter and an overview of perspectives for future research.

3 References:

1. Nurse SP. The great ideas of biology. 2003;3(6).

2. Kessel A, Ben-Tal N. *Introduction to Proteins: Structure, Function, and Motion*.; 2012.

3. Fischer E. Einfluss der Configuration auf die Wirkung der Enzyme. *Ber Dtsch Chem Ges*. 1894;27:2985-2993.

4. Mirsky AE, Paulin L. On the structure of native, denatured, and coagulated proteins. In: *Proceedings of the National Academy of Sciences of the United States of America*. Vol 22. ; 1936:439-447.

5. Anfinsen CB. Principles that Govern the Folding of Protein Chains. *Science (80-)*. 1973;181(4096):223-230.

6. Drenth J, Mesters J. *Principles of Protein X-Ray Crystallography: Third Edition*.; 2007.

7. Spence JCH, Weierstall U, Chapman HN. X-ray lasers for structural and dynamic biology. *Rep Prog Phys*. 2012;75(10):102601.

8. Teng TY, Moffat K. Primary radiation damage of protein crystals by an intense synchrotron X-ray beam. *J Synchrotron Radiat*. 2000;7(Pt 5):313-317.
9. Weik M, Ravelli RBG, Kryger G, et al. Specific chemical and structural damage to proteins produced by synchrotron radiation. *Proc Natl Acad Sci*. 2000;97(2):623-628.

10. Tenboer J, Basu S, Zatsepin N, et al. Time-resolved serial crystallography captures high-resolution intermediates of photoactive yellow protein. *Science*. 2014;346(6214):1242-1246.

11. Zhang H, Unal H, Gati C, et al. Structure of the angiotensin receptor revealed by serial femtosecond crystallography. *Cell*. 2015;161(4):833-844.

12. Li X, Mooney P, Zheng S, et al. Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. *Nat Methods*. 2013;10(6):584-590.

13. Henderson R. The potential and limitations of neutrons, electrons and X-rays for atomic resolution microscopy of unstained biological molecules. *Q Rev Biophys*. 1995;28:171-193.

14. Glaeser RM, Taylor KA. Radiation damage relative to transmission electron microscopy of biological specimens at low temperature: a review. *J Microsc*. 1978;112(1):127-138.

15. Brenner S, Horne RW. A negative staining method for high resolution electron microscopy of viruses. *Biochim Biophys Acta*. 1959;34:103-110.

16. Kühlbrandt W. Cryo-EM enters a new era. *Elife*. 2014;3:e03678.

17. Callaway E. The revolution will not be crystallized: a new method sweeps through structural biology. *Nat News*. 2015;525(7568):172.

18. Cavalli A, Salvatella X, Dobson CM, Vendruscolo M. Protein structure determination from NMR chemical shifts. *Proc Natl Acad Sci U S A*. 2007;104:9615-9620.

19. Bax A. Two-dimensional NMR and protein structure. *AnnRevBiochem*. 1989;58:223-256.

20. Tugarinov V, Kanelis V, Kay LE. Isotope labeling strategies for the study of high-molecular-weight proteins by solution NMR spectroscopy. *Nat Protoc*. 2006;1(2):749-754.

21. Downing AK. *Protein NMR Techniques*. Vol 278.; 2004.

22. Corker J, Lefebvre F, Evans J, et al. Catalytic Cleavage of the C-H and C-C Bonds of Alkanes by Surface Organometallic Chemistry: An EXAFS and IR Characterization of a Zr-H Catalyst. *Science (80-)*. 1996;271(5251):966-969.

23. Penner-Hahn JE. X-ray Absorption Spectroscopy. *Compr Coord Chem II*. 2003:159-186.

24. Penner-Hahn JE. X-ray absorption spectroscopy in coordination chemistry. *Coord Chem Rev*. 1999;190-192:1101-1123.

25. Debeer S. X-ray absorption spectroscopy. *Methods Mol Biol (Clifton, NJ)*. 2011;766:165-176.

26. Bergmann U, Glatzel P. X-ray emission spectroscopy. *Photosynth Res*. 2009;102(2-3):255-266.

27. P. G, A. J, Glatzel P, Juhin A. X-ray Absorption and Emission Spectroscopy. *Local Struct Charact*. 2014:89-171.

PART-1: ELECTRON AS A PROBE

MULTIPLE SCATTERING and MOLECULAR REPLACEMENT in TED

1 Introduction

1.1 Transmission Electron Diffraction for Biomolecular Structures

Electron crystallography as a tool to study the structure of proteins, viruses and other similar macromolecular complexes is well established for quite some time now. In fact, common to both X-ray and electron crystallography is the ability to utilize data from three dimensions to interpret the structures. (Although electron crystallography extends itself to include not just electron diffraction but also electron microscope images, for the reminder of the chapter the term electron crystallography will be used synonymously with transmission electron diffraction (as with X-rays). Hence, the computational sophistication that supports X-ray crystallography definitely lends itself into the electron-based techniques. Significant developments in the microscope have led to obtaining structures using electron crystallography to be on par with X-ray crystallography or multidimensional-NMR in resolution, thus enabling this technique as an alternative and unique option to study biomolecular structures.

The uniqueness of this field of study arises from two specific physical properties of electrons – (a) 5-6 times higher elastic scattering power as compared to X-rays or neutrons. (b) The ability to focus an electron beam using electromagnetic optics to form a probe as small as an Angstrom. Combining this with the ability to image using electrons, is an advantage that has no parallels with any other technique. (However, this chapter will not touch upon electron microscopy imaging and will limit itself to focus only on electron diffraction).

While electron and X-ray crystallography share a lot of similarities, one major difference is that while X-rays are scattered by atomic electrons, an electron beam can be scattered by atomic electrons and the nucleus. This is also why, although the mathematical theory for scattering of electrons and X-rays off a crystal is identical in nature, the physical object (scatterer) for an X-ray beam is the periodic arrangement of atomic electrons and hence the diffraction-reconstructured object constitute the electron charge densities while for e-beams, since the

physical object (scatterer) is dominated by the nuclear charge, the diffraction-reconstructed object constitutes coulomb potentials.

Although there are many types of biomolecules that could potentially be studied using electrons, a strong working point for electron diffraction is the ability to form thin two-dimensional crystals (some that occur naturally, while others that need to synthesized). The pioneering first example of a successful electron crystallography application is on bacteriorhodopsin, [1] a naturally occurring 2D crystal. Since then, there have been multiple such attempts to solve structures of 2D crystals of biomolecules that serve as fine examples such as the light harvesting complex (LHC II) [2] from green plants, ATPase of Calcium [3,4] and Sodium, [4] PhoE porin, [5] water-channel proteins isolated from red blood cell membranes, [6] tubulins [7] and many others.

Though this chapter is not intended to be a thorough discussion on the methodology of electron diffraction, a brief overview of the steps involved in this technique is presented in the next sub-section.

2 Electron Crystallography (TED) Procedures:

The Electron Crystallography protein process can be categorized into five sequential steps (Below is mentioned is a generic set of steps followed for proteins. This can be extended to any biomolecule of interest with the suitable changes.) –

2.1 Protein Expression and Purification:

Arguably the toughest nut to crack in the determination of the structure, the single-minded aim of this step is to obtain as much quantity of purified protein as possible. This starts with expressing the protein.

Protein expression is a process that involves '*a combination of an expression vector, its cloned DNA, and the host for the vector that provide a context to allow foreign gene function in a host cell, that is, produce proteins at a high level*' (http://www.biology-online.org/dictionary/Expression_system). Commonly used hosts are bacteria (such as E. Coli),

yeast or even eukaryotic cell lines. Commonly used DNA sources and delivery mechanisms are viruses, plasmids or bacteriophage. The type of expression system is eventually dependent on the gene involved and nature of experiment (which determines the bare minimum quantity of purified protein that could be worked with).

Protein purification is a combination of procedures meant to isolate a specific protein from a mixture (cells or even organisms) that are present during expression. The purification process separates the protein and non-protein parts of the mixture. Sometimes, the presence of multiple proteins makes separation of one protein from all others, typically, the most laborious aspect of protein purification. A typical procedure of purification would involve (i) extracting the protein by disrupting the contents of the host-cell. This can then be followed by (ii) precipitating the protein with Ammonium sulphate. Sometimes, if the precipitation process is not as unique, detergents are used to dissolve the cell membranes to keep alive the (membrane) proteins. In such cases, the solution is then (iii) centrifuged to separate the protein from the rest of the solution by aggregating the proteins together (particles of similar mass are aggregated at the same depth in the vile). As a penultimate step, a variety of chromatography techniques are available to separate the proteins (based on their properties) such as through (iv) size-exclusion, ion-exchange, hydrophobic interaction, metal-binding affinity amongst others. Finally, the aim is to have (v) a concentrated protein solution, which can be done by lyophilization (freeze drying) or ultrafiltration.

2.2 2D Protein Crystal Formation:

Once the protein is purified and concentrated, the next step is to crystallize. Care must be taken to identify and use the right type of detergents (as in the case of membrane proteins). Typically 2D proteins can take stronger detergents to purify the solution since the 2D crystals form much fast than 3D crystals (as required for X-ray crystallography) and hence the scope for damage to the crystal formation is reduced. Crystallization is achieved then by substituting the detergent by lipids forcing reconstitution of the protein environment. The detergent is then removed either by dialysis or chelation or even as simple as lowering the detergent concentration to lower than the critical micelle concentration (CMC) by dilution. It is generally noted that

crossing the CMC slow and steady helps the formation of good crystals since extending this cross-over time allows the protein sufficient time to diffuse through the lipid bilayer and make contacts with neighboring proteins. In the case of dialysis a similar effect can be achieved by lowering the temperature of the dialysis operation.

Some factors that are possibly critical for the formation of high quality 2D crystals are – (i) nature/type of lipid used for reconstitution, (ii) lipid-to-protein ratio, (iii) composition and pH of dialysis buffer, (iv) detergent concentration pre-crystallization, (v) temperature cycles during reconstitution. Such 2D crystals could come of different shapes and facets including vesicular, tubular or sheet-like (as mentioned before).

2.3 TEM Specimen preparation:

There are quite a few techniques used to study 2D biological samples and these have evolved in time.

2.3.1 Negative Stain

The technique of negative staining involves embedding the sample specimen within a thin matrix of salt (such as uranyl acetate) and neutralized with ammonium molybdate. The goal here is to replace the bulk water that the specimen it normally embedded in by a solid matrix of amorphous salt. Although not very effective to obtain high-resolution data, this technique can be used to characterize the morphology (EM imaging) and long-range crystal quality. It can provide crystallographic structures at few 10s of angstrom resolution, which can be used to estimate the surface envelope, molecular shape and even domain organization/sub-unit packing arrangement of large macromolecules. To a considerable extent, this technique can reduce flattening and/or collapse of the biological macrostructure, which happens when water is removed. This technique is prone to radiation damage but the post-damage structure still retains some record of the larger 3D structure of the biomolecule since the stain itself is typically unaffected by the e-beam radiation.

2.3.2 Sugar Embedding

This technique is well known to preserve the high-resolution 'hydrated' structure of crystalline biological specimens. It involves a 2% glucose solution that when embedded can preserve the well-ordered structure of specimen in vacuum, at room temperature. As with negative staining, this technique also prevents the flattening and collapse of the biological macrostructure. Apart from glucose, (neutralized) tannic acid and trehalose also serve the same purpose. A major drawback of this technique however is that the addition of these low-Z compounds vitiates the effect of radiation damage on these crystals since both the crystal and embed are equally susceptible to e-beam induced radiation damage.

2.3.3 Vitrification

A commonly used and preferred alternative to stained samples is embedding the specimen in vitreous ice. This way (as with sugar embedding), the native hydrated structure before freezing is intact. To ensure there is no sublimation, this sample specimen is mounted on a cold stage in the TEM thereby introducing the user to the realm of Cryo-EM, or in this specific case Cryo-ED. To ensure a good quality of crystal, it is important to ensure a few things - the sample is well hydrated during vitrification and that the mechanical stress imbued due to volume change, ice formation or changing thermal expansion coefficients is kept to a minimum. The key to vitrification of thin aqueous films is rapid plunging into liquid ethane at a temperature slightly above its melting point and well below its boiling point. A film of liquid ethane that is generally present once the specimen it taken out of the cryogen generally evaporates leaving very little (if any) cryogen remaining on the sample. Also, liquid ethane does not boil on contact with the sample and hence avoids formation of any gaseous layer. These two advantages make liquid ethane the single preferred candidate to carry out this procedure.

2.4 Data Collection

Electron Diffraction data gives the amplitude information of the scattered e-beam as opposed to an imaging technique, which would offer both the amplitude and the phase of the scattered e-beam. However, the amplitudes acquired through diffraction tend to be relatively

more accurate. In addition, there is an added advantage of operating in the diffraction mode, which is that the diffracted intensities are sample translation invariant.

The data are typically collected on a CCD or on an image plate which is then converted to digital data. The experiment proceeds with mounting the specimen crystal on the sample holder and then orienting the sample holder such that the diffraction arising from the specimen is from a high symmetry. This is taken as the reference orientation. Extreme care is taken in calculating the sample damage electron dose (which is the threshold number of electrons per unit area per second, impinging on the sample that degrades the Bragg diffraction intensity at the resolution of interest, say by 10%) and ensuring that the e-beam exposure is kept within this limit throughout the entire experiment. Once this is noted, the sample now is tilted in calibrated amounts using the goniometer sample stage holder to obtain diffracted intensities on the detector that spans as much of the (reciprocal) Fourier space as possible. While working on the same sample specimen, it is highly likely that the accumulated electron dose might at some instant be sufficient to permanently damage the sample. If and when this is noted, the sample specimen is replaced with a new one with the calibration procedure performed again. The farthest spots from the center on the detector can give a qualitative indication of the 'resolution that the crystal diffracts to'. This entire process is repeated until sufficient numbers of diffraction frames are collected.

2.5 Data Processing and Analysis:

The handling of data is extensive and can be categorized into three parts for convenience. The obtained data need to be (i) indexed, scaled and merged before (ii) obtaining the phase information. Once the phases are obtained, various appropriate procedures are used to (iii) refine the phases and build the new model that represents the sample.

2.6 Indexing, Scaling and Merging

Data processing begins with *indexing* the Bragg diffracted intensities. The first step while performing indexing is to identify the real-space unit cell parameters for the crystal specimen. This also includes identifying the space-group and symmetry types associated with the space-

group. This information is then used to label diffracted (Bragg) spots in each frame by a miller index (reciprocal space notation for diffracting planes of the real-space crystal). Certain unique features can help ease this fingerprinting procedure. For example, reflection symmetries cannot be observed in chiral molecules; thus, only 65 space groups of 230 possible are allowed for protein molecules which are almost always chiral. Since this process requires shuffling between real and reciprocal space coordinates, such a transformation is aided by Fourier transforms (more specifically discrete/fast Fourier transforms).

A full data set may consist of hundreds/thousands of separate images taken at different orientations of the crystal. Hence, after assigning the miller indices for each diffracted spot in each frame, the next step is to merge the intensities of the same spots occurring in various frames and scale them respectively. This is the process of identifying the same diffracted Bragg spots (same miller index) that could appear in multiple frames and add them together. In the process of adding them together, care needs to be taken to ensure that the intensities are scaled since each frame could have had a different exposure time (and hence different electron dose) amongst other variations. This process of index-merge-scale is iterated until the unit cell parameters converge to something meaningful. In these cases, any prior information about the biological system ab initio or through other parallel experiments can offer a significant help. This is also where estimates can be obtained of how many frames of diffraction are required to obtain a complete dataset.

2.6.1 Initial Phasing

In an electron diffraction experiment, the only data collected is the diffracted intensity. This has a representation of the scattered wave amplitude but not the phase. To be able to reconstruct the real space picture, the complex structure factor (which includes the structure factor amplitude and phase) is required. And such phase information is typically obtained through a variety of techniques, a couple of the most common of which are mentioned here – simultaneous imaging and diffraction (or) molecular replacement.

Simultaneous imaging and diffraction is unique to an electron microscope such that it utilizes the property of electrons to be focused, to record not just the diffraction (which offers better estimates of structure factor amplitudes) but also record images in real space, which contain the amplitudes and phases. The images with an e-beam can be assumed as a 2D projection of incident wave, which is scattered and propagated through successive layers of the sample. The image recorded on the detector is the exit wave emanating from the final layer of the sample, modulated by the nuclear potentials for one last time before propagating to reach the detector. By correlating the 'tilt-series' of diffraction frames with their corresponding images, the initial phases for each such frame is identified which is then used to further refine the structure.

Molecular Replacement (MR) is an alternative technique used to obtain phase information without doing experimental phasing. This method is based on prior knowledge from existing similar structures with 20% or greater homology in amino acid sequence. Nevertheless, it is one of the most frequently used methods to determine the phase when some prior structural information is known (such as homologous models).

The principle of the MR is simple. If there is a model known, a priori, whose structure is very similar to the protein of interest, then the 'known' model is used as a primary iteration to obtain a 'final' model that fits the experimentally obtained structure factors (diffracted intensities) of the crystal. In the case of proteins, one can then use translational and rotational motions and modify the phase information to determine the structure factors (which should eventually match). Rebuilding the model for the structure of interest follows this. The similarity between the 'known / search' model and the unknown model can be defined as the similarity in protein (i.e., amino acids) sequences. Common protocols consider a minimum cut-off similarity identity ≥30%, or require belonging to the same class or family of proteins by having some conserved regions in the sequence or similarity in protein folding.

2.6.2 Structure Refinement:

Having obtained initial phases, an initial model can be built. This model is then used to refine the phases, leading to an improved model, and so on. With a starting model of some

atomic positions and Debye-Waller (B) factors (to account for thermal motion of atoms), this model is then refined to fit the observed experimental diffracted intensities. Once this is done, a new set of phases is obtained. A new model is then fit to the new electron density map and a further round of refinement is carried out. This continues until the correlation between the diffraction data and the model is maximized. The agreement between the two is measured using a statistical parameter called the Rfactor. The Rfactor is a statistical parameter widely used in crystallography that measures how close the experimental |F| values are in relation to the crystallographic model. It is expressed as –

$$Rfactor = \frac{\sum \left| \left(|F_{obs}| - |F_{calc}| \right) \right|}{\sum |F_{obs}|} \qquad \text{(a)}$$

where $|F_{obs}|$ and $|F_{calc}|$ are the is the experimental and model structure factor amplitudes respectively. During such refinement procedures, a portion (~10%) of the dataset is consciously not included. The Rfactor measured over the dataset used in refinement is termed as Rwork, while that determined over the data originally left out of refinement is termed Rfree. Both *R* factors depend on the resolution of the data. As a rule of thumb, R_{free} should be approximately the resolution in angstroms divided by 10; thus, a data-set with 2 Å resolution should yield a final R_{free} ~ 0.2.

It may not be possible to observe every atom of the crystallized molecule – it must be remembered that the resulting electron density is an average of all the molecules within the crystal. In some cases, there is too much residual disorder in those atoms, and the resulting charge density for atoms existing in many conformations is smeared to such an extent that it is no longer detectable in the density map. Weakly scattering atoms (Hydrogen) are routinely invisible.

3 Features of Refinement in Electron Crystallography

The previous sub-sections provide a brief overview of the entire process involved in obtaining a bio-molecular structure using the electron diffraction technique. In this sub-section, a

quick recognition of certain common (with Xrays) and different but unique features of refinement in electron diffraction is highlighted below.

3.1 Refinement Improves Values of Parameters Defining the Structural Model

At a resolution of 2.5 Å, the refinement can accommodate an increased number of parameters that are required to model localized water molecules, salt ions and other non-protein features. By including these non-protein features, which are indeed present in the specimen, the refined atomic-resolution model produces improvements in both the Rfactor and the R_{free}. If it is established that the R_{free} parameter is reduced in conjunction with the assignment of new values, it is a good signature of the agreement of model with experimental data.

3.2 Over-Interpreting the Available Data

With the addition of a high number of parameters, there is always the possibility for improvement of Rfactor purely because of the presence of excess parameters. In such cases, R_{free} can be used to monitor actual betterment of the model as opposed to over-interpretation. In the case of the latter although the Rfactor would continue to decrease, Rfree would still be high or increase indicating a false alarm.

3.3 Larger Rfactors for Electrons than X-rays in Crystallography

It is very commonly noted that the X-ray refinement parameter Rfactor is about 0.2 – 0.25. R_{free} is about this value or slightly higher. When the Rfactors are at about 0.3, the ability to ascertain details such as contacts between amino acids or the number and location of water molecules. However, with electron diffraction the Rfactors are normally seen to be larger (0.3 – 0.35) and generally accepted. A possible reason for this discrepancy could be the significantly higher error contributions to the diffraction measurements (such as specimen bending or radiation damage).

3.4 Chemical Bonding Effects:

In X-ray diffraction, it is normal to approximate scattering from each atom in a molecule as that from isolated neutral and free atoms. In reality, chemical bonding ensures that this is not the case and the atoms are far from neutral. (Oxygen atoms would be negative and Carbon atoms would be positive). For example, the development of partial charges on the atoms in the polypeptide backbone is quite pronounced, which lends itself to the high dipole moment of the peptide bond. Such chemical bonding effects affect electron scattering much more significantly and they are typically not modeled for adequately (which is thought as being one of the contributors to the increased Rfactors).

As mentioned earlier, since the electron scattering is strongly influenced by the Coulomb potential of the atoms, the stronger sensitivity of the Coulomb potential feels the altered electron charge distribution due to bonding effects and hence is represented much more in electron diffraction than in X-ray diffraction. This is noticed to be especially true in the lower resolution data sets as seen in the fig. 7 in the work by Chang et.al. [8]

4 Salient Features of Electron Diffraction

In the previous section, a broad and brief outline to the scope of electron diffraction for biological molecules was provided. Moving ahead, this section will provide some detail to the theory of two fundamental concepts that serve as a premise to this technique.

4.1 Kinematic Approximation

4.1.1 What is the Kinematic Approximation and Why is it Important?

Elastic scattering of an e-beam with atoms in a molecule can be split into two regimes – (i) Single (Kinematic) scattering and (ii) multiple (Dynamic) scattering. In the kinematic scattering limit, the relationship between the scattered wave ψ and the Coulomb potential ρ is –

$$\psi_k(S) = F\left\{\widehat{\rho}(R)\right\} \tag{1}$$

If the incident e-beam is assumed as a plane wave (original solution to the Schrodinger equation), then the scattered wave is represented as a small perturbation (introduced by the interaction with the scatterer) that is added to this plane wave. Such a perturbed wave function can be represented using the Born series. The first order term in this series is simply the Fourier Transform of the scattering potential (evaluated on the Ewald sphere). Thus the first Born approximation is identical to the kinematic scattering approximation.

$$|\psi\rangle = |\phi\rangle + \{G_0(E)V\}|\phi\rangle + \{G_0(E)V\}^2|\phi\rangle + \ldots \quad (2)$$

Here, *V* is the Fourier transformed interaction potential and *G*, the Green's operator. Clearly, in retaining the just the first order term, it is clear that the Fourier relation can be established exactly as in eqn. (1). However, in including higher order terms, the ability to Fourier correlate the scattering potential directly to the scattered wave function is lost, just as in kinematic scattering. This then leads into the realm of dynamical scattering where higher order Born series can be retained as is necessary.

However, amplitudes are never measured in most measurements. In specific relation to electron diffraction, the measured quantity is the diffracted intensity which is just the square of the scattered wave amplitude and is given as –

$$I = |\psi_k(S)|^2 = |F\{\hat{\rho}(R)\}|^2 = |V(k)|^2 \quad (3)$$

This is a very important relation since this tells us that when the diffracted intensities are measured, there is a direct relation to the structure factor (scattering potential) amplitudes by using the mathematical operation of the square root. What is also important to note is that, when the first Born approximation is no longer sufficient and the higher order terms need to be included, this reverse reduction from diffracted intensity to the structure factor amplitude is not possible and it becomes more complex. Diffraction experiments are known to provide just the

Coulomb potential (structure factor) amplitudes and not the phase. What is evident here is that, even this is only applicable under certain conditions.

Hence in the success of the electron (and for that matter even X-ray) diffraction technique exists a fundamental assumption that "*the samples have to be within the kinematic scattering regime"*.

What does this statement mean? In the analogy developed, the e-beam scattering can be expressed through a Born series, which is an expansion of an external perturbation. For an atom of a particular element interacting with an e-beam, this interaction is fixed in nature. However, if there are multiple successive such interactions, the overall perturbation exerted on the e-beam is strong and cannot just be approximated to the first order in expansion. By extrapolation, if an e-beam traveling along *z* and interacts with a sample specimen which is strictly 2-dimensional along the *xy* plane, then the probability for multiple successive scattering is directly correlated with increasing the number of sample layers along *z* (the primary propagation direction of the e-beam). Incidentally, it turns out that this is indeed the case and is represented by an equation as mentioned below -

$$I_g = \left(\sigma V_g t\right)^2 \left[\frac{Sin\left(\pi s_g t\right)^2}{\left(\pi s_g t\right)^2}\right] \qquad (4)$$

This equation tells that the Bragg intensity recorded at the g^{th} Bragg diffraction location is not just proportional to the square of the Fourier transformed Coulomb potential, but also proportional to the square of the thickness of the sample. Here s_g, the excitation error, measures deviation from the Bragg condition and $\sigma = \pi/(\lambda V_o) = 2\pi m_e \lambda/h$ is an interaction constant, with V_o the microscope accelerating voltage and λ, the electron wavelength. This applies to the case of small beam divergence angle. In this single-scattering approximation, the intensity at the Bragg condition ($s_g = 0$) varies as the square of sample thickness, and so can be compared with the actual thickness variation predicted by accurate multiple scattering calculations, in order to determine the thickness range of this approximation for various resolutions and beam energies.

For weaker beams, where the Bragg condition is not exactly satisfied, the equation also predicts a thickness-squared dependence. A degree of convergence in the illumination will span a range of excitation errors s_g.

4.1.2 Are There Any Tests for the Breakdown of the Kinematic Limit?

4.1.2.1 Multiple Factors Govern the Importance of Kinematic Limit.

For the kinematic theory to be valid, it is generally assumed that the number of electrons scattered once must be very small (twice, thrice and further should almost be negligible). For example, a hypothetical sample of carbon, nitrogen and oxygen with a mass density of 1.5 g/cc would have a mean free path of about 190 nm. If the intention is to limit the e-beam specimen interaction to not more than 10%, this already limits the sample thickness (at 100 keV) to be less than 20 nm.

Chemical composition is also an important factor. Scattering increases with atomic number and hence the kinematic limit, intuitively can be guessed to be larger for low Z elements as compared to high Z. The scattering of the historically famous example of UF_6 at 40 keV wherein multiple elastic scattering [9] was found to have a major effect is a direct consequence of the high Z of Uranium.

Increasing the incident e-beam energy reduces the scattering cross-section. This implies the probability of scattering is reduced with increase in e-beam energy. In such a case again, intuition would suggest that the kinematic limit would be higher at higher e-beam energies.

Density of arrangement of atoms within a unit cell can also strongly affect the strength and amount of scattering. Dynamical effects become much more significant when individual atoms in the specimen are precisely aligned in the direction of the electron beam (symmetry directions of a crystal). As in this calculation [10] it can be seen that in graphite (made only of carbon), dynamical effects become significant at 5nm!

4.1.2.2 Breakdown of Kinematic Scattering – A Review

The importance of multiple scattering in cryo-EM imaging was first investigated by Grinton and Cowley (1971) [7] for single-particles, following the first tomographic TEM imaging experiments of DeRosier and Klug [11] and others. At 100 keV, Grinton et.al. found false detail in high-resolution TEM images (not diffraction patterns) at 6 Å resolution and 30nm thickness, and severe distortions due to multiple scattering at 50nm thickness of protein. At 3 Å resolution they expected severe distortion for 100 Å thickness. For diffraction from 2D crystals, Glaeser and Downing [12] found that thicknesses up to 100 Å might be used for electron incident beam energies above 100 keV.

An important effect of multiple scattering is to preserve sample symmetry in the diffraction pattern (violating Friedel symmetry). So, a simple test for the presence of multiple scattering is failure of Friedel symmetry. Moreover, all protein crystals are non-centrosymmetric due to the presence of helical structures. For 4nm-thick bacteriorhodopsin at 16 Å resolution and 100 keV, Glaeser and Downing found an Rfactor of 11% between Friedel pairs, increasing to 40% for a 200 Å thick multi-layer. They recommend averaging the pairs to reduce perturbations, and support a thickness limit between 100 and 200 Å of protein at energies above 100 keV for high resolution. A detailed study of the validity domain of the single-scattering (kinematic) approximation for light-element inorganic crystals can be found in Anstis et.al. [13] In experiments however, while Friedel symmetry is well preserved for 2D crystals (because of their relative insensitivity to orientation), it is a less useful test for 3D crystals because of their extreme sensitivity to diffraction conditions and orientation (unless convergent-beam methods can be used). The "kinematic convergent-beam method" has been used to solve an oxide nanocrystal, using the absence of intensity variation within the convergent-beam (microdiffraction) disks as a test for the absence of multiple scattering. [14]

The presence of significant intensity in reflections, which should be absent due to the presence of screw and glide symmetry elements is often used as an indication of multiple scattering. In fact, for an unbent thick perfect crystal, (conditions rarely met in protein crystallography unless a very small beam diameter is used) these reflections remain absent for all

thicknesses and beam energies as a result of an elegant cancellation theorem. [15] Taken together, under the normal conditions of strong multiple scattering, great thickness sensitivity, and specimen imperfections, it is almost impossible to record the same set of Bragg beam intensity ratios on two separate occasions using an electron microscope, unless the sample consists of a monolayer, and a very small beam is used. These TED intensities (unlike XRD intensities) are, therefore, not normally sufficiently reproducible for structure analysis. (Those collected from a particular region of known thickness, however, may be analyzed quantitatively in great detail using the convergent beam method. [16]

At the same time, the sub-nanometer beam diameter of the electron microscope allows 'perfect crystal' data collection from regions smaller than one "mosaic block" for inorganic crystals. Striking images of these mosaic regions (perhaps 50 nm in width) in protein have been directly imaged at high resolution, for the first time recently, by Nederlof et al. [17] (The thickness of these "blocks" is not clear, they may be regions of locally bent crystal, as discussed below). It is noted therefore that by using a sufficiently small beam diameter, many of the effects of defects and disorder could be eliminated as the beam becomes smaller than one "mosaic block" [17] or bent region. However this requires instrumental modifications to allow the beam intensity to be reduced in addition to its size (so that the total number of electrons in the beam is not held constant during focusing), while radiation damage will also limit the size of the smallest beam which can be used (see [18] for estimates of the smallest crystal which can give useful electron diffraction data when subject to damage limits). The simulations which follow do not allow for mosaicity effects in electron diffraction and so refer to the case where a small microdiffraction beam is used (See [19] for a review of TED and an extended discussion of the effect of dynamical perturbations on attempts to solve organic structures by electron diffraction, and the history of this subject).

4.1.3 Molecular Replacement

4.1.3.1 What is the Phase problem?

Elastic scattering under kinematic limit can be expressed through the following equation –

$$\psi_k(S) = F\left\{\hat{\rho}(R)\right\} \tag{5}$$

where Ψ_k, the scattered wave is proportional to the Fourier transform of $\rho(R)$, the shielded Coulomb potential.

Assuming two scatterers, one (scatterer A_1) at at a vector distance R_1 from the origin and the second (scatterer A_2) at a vector distance R_2 from the origin. Assuming an incident plane wave, the scatterers will act as a source of a new spherical wave, the amplitudes of which are modulated by the strength of the scatterers (scattering potential). When the scattered waves travel a substantial distance l_1 and l_2 such that l_1, l_2 >> (R_2 - R_1), if the incident wave is along the k direction and detection along k_0, the path difference between the two scattered waves is –

$$2\pi\frac{(l_1 - l_2)}{\lambda} = (k - k_0)\cdot(R_1 - R_2) \tag{6}$$

Here, the RHS is termed as the phase difference between the two different scattered waves in reciprocal space. At distances much larger than the distance between the scatterers, the scattered wave can be assumed to have a plane-wave front. Using this, the scattered wave function for two scattering points can be written as the superposition of the two separate scattered waves as –

$$\psi_k = \left\{\hat{\rho}(R_1)e^{i(k-k_0)\cdot R_1}\right\} + \left\{\hat{\rho}(R_2)e^{i(k-k_0)\cdot R_2}\right\} \tag{7}$$

Extending this to a continuous distribution of n-scatterers and substituting $k - k_0 = 2\pi S$, this can be re-written as –

$$\psi_k = \sum_{j=0}^{N} \hat{\rho}\left(R_j\right) e^{i\left(k-k_0\right)\cdot R_j} \quad \text{i.e.}$$

$$\psi_k\left(S\right) = \int \hat{\rho}\left(R\right) e^{i\left(2\pi S\right)\cdot R} dR \quad (8)$$

This is a general equation that displays the far-field scattered wave function as a Fourier transform of the scattering potentials (as in eqn. 1). For a detector in the far field, the detected quantity is the scattered intensity, which is the square of the wave function in eqns. (5) or (8). This is the same as eqn. (2) –

$$I = \left|\psi_k\left(S\right)\right|^2 = \left|F\left\{\hat{\rho}\left(R\right)\right\}\right|^2 = \left|V(k)\right|^2$$

As is clear from eqn. (2), the diffracted intensity can be related to the structure factor amplitude (scattering potential) as long as the scattering is within the kinematic limit. However, by squaring the complex structure factor, the phase information of the scatterers is lost – this is commonly referred to as the phase problem in crystallography.

4.1.3.2 Why is this Significant?

To understand the importance of the phase problem, eqn. 6, which gives the relation between the path lengths of the two scatterers, needs to be revisited. In this relation, the LHS gives the path length and the RHS was identified as the phase difference. This implies that the phase difference directly represents the relative (not absolute) positions of scatterers. In protein crystallography, it is widely accepted that about 2/3 of the information in a charge density map comes from structure factor phases and 1/3 from the amplitudes. [20,21]

In using electron diffraction to study protein structures, the information obtained from the experiments (the diffracted intensity corresponding to the structure factor amplitude) accounts only to about 1/3rd of the total information required to reconstruct the real-space structure. More importantly, the rest 2/3rd of the information (the phases) are obtained from models constructed

from prior knowledge about same or similar systems using a technique called Molecular Replacement.

This is significant considering the major contribution to a new structure solution is obtained from a guess based on a model. However, with sufficient amount of checks and safeguards, this has now developed into a very versatile tool to support crystallography based structure identification (refer to [22–27] for a detailed description, review and the success stories related to this technique). Over 60% of the biological structures solved using X-ray crystallography attribute their success to MR [22].

Molecular replacement is basically a six-dimensional problem: the coordinates of the search model transforms into the coordinates of the structure to be solved. This search model is subject to a rotation (by three angles) and a translation (by the three vectors). The more similar, both in primary and tertiary structure, the search model is to the target, the more likely it is that a solution to an MR problem can be found. It has generally been accepted that a 1.5 Å r.m.s deviation between the search model and the target is the lowest limit at which a related structure can be used as a search model. However, the r.m.s deviation indicates the similarity between the model and the target only post MR. The only metric to check similarity before any processing is the sequence identity. One of the generally accepted rules of thumb is that the search model is good if there is at least a 30% match (if not higher) of the model to the target. However, since there may be relative domain movements or loop variations, the aforementioned statement is to be taken as a general guideline and not a strict metric. In fact, molecules with a much lower identity can be good search models in the cases when three-dimensional similarity is retained.

The success of this technique, thus, depends primarily on two factors: the fraction of the asymmetric unit for which there is a suitable model(s) and the deviation (after optimal superposition) between the model and target structures. The deviation (r.m.s) generally increases with decreasing sequence identity, so good models generally have high sequence identity with the target structure.

5 Aim of This Chapter

In pursuit of solving for the structure (static/dynamic) of biomolecules, transmission electron diffraction (TED) is demonstrated as a very capable technique. The aim of this chapter is to identify the limitations in using this technique, by studying the two essential concepts that make electron diffraction – kinematic scattering limit and molecular replacement.

In specific, this is done by addressing two questions –

- How thick can a biological sample be before the kinematic scattering approximation breaks down?
- To what extent is the Molecular Replacement method tolerant to errors in measured diffraction intensity?

6 Results and Discussion

6.1 Kinematic Limit

This section highlights the simulations performed to identify (i) the kinematic (single-scattering) thickness limit in proteins (within which a direct interpretation of the diffracted intensities is possible). All our studies were done with the protein (hen-egg white) lysozyme in the $P4_32_12$ space group and unit cell parameters a = b = 79.1 Å and c = 37.8 Å. The structure factors were obtained from the relativistic Hartree-Fock calculations of Doyle and Turner for neutral atoms, [28] and so take no account of bonding effects. These have been refined in detail against experimental electron diffraction intensities using inorganic crystals [29] where bonding effects are very small, but may be larger for the light C, O, N atoms of a protein, or graphene. [30] Hydrogen atoms were not included. Since multiply scattered TED intensities are known to agree with high accuracy with experimental intensities in convergent beam electron diffraction (CBED) (see [31,32] for reviews), the ZMult (multislice [33] program in [29]) program was used to simulate diffraction intensities as a function of thickness in lysozyme. The analyses assume that Bragg beam intensities are proportional to the X-ray structure factor amplitudes |F| (adjusted by a scaling factor).

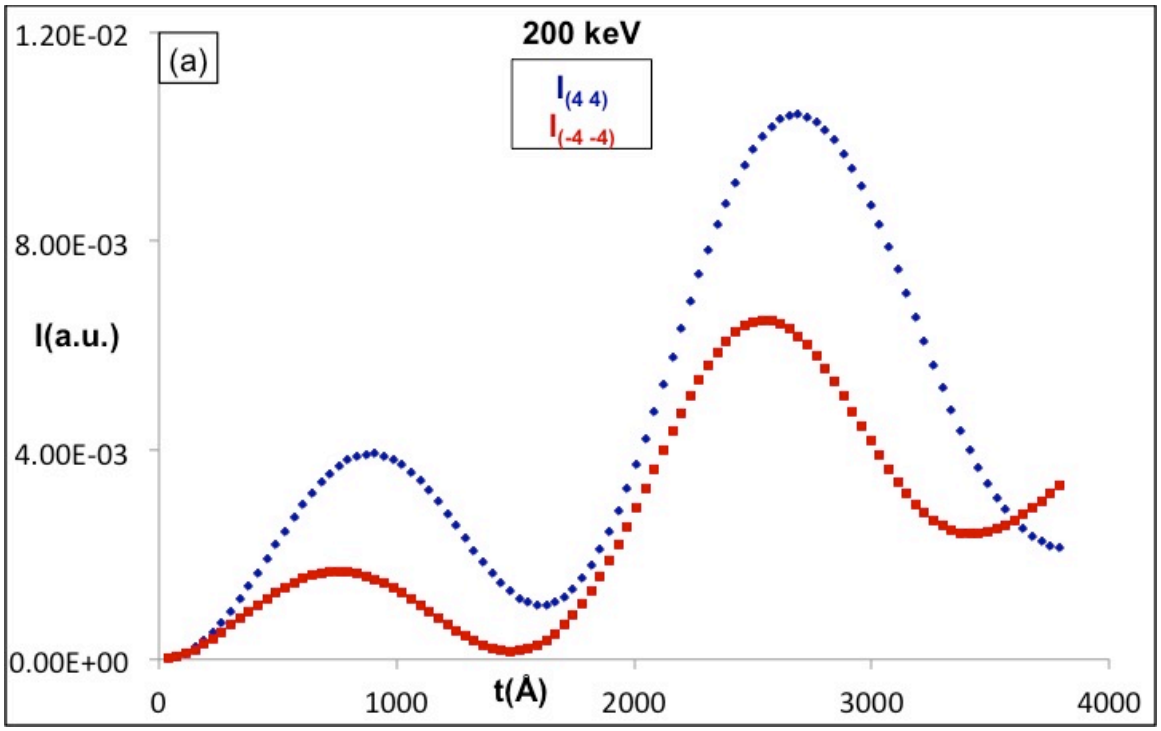


Fig (1.1 a): Diffracted beam intensity vs thickness for lysozyme crystals. At 200 keV diffracted beam (4 4) and its Friedel pair (-4 -4)

First, the variation of diffracted beam intensity with thickness for the lysozyme crystals was studied, which was subsequently used to identify the kinematic limit for the different diffracted beams. Check [19] for similar curves for paraffin layers. The simulations are performed for two conditions – (a) low resolution: (4 4) at about 13.98 Å and (b) high resolution: (23 13) at about 3 Å. The miller indices are provided in two dimensions because the incident electron beam is considered as incident along the z direction, orthogonal to the (001) plane of the lysozyme tetragonal crystal. Figs. 1.1 (a and b) show the characteristic 'pendellosung' oscillations in the intensities of low (4 4) and high (23 13) resolution diffracted beams (approx. 13 Å and 3 Å respectively) and their Friedel pairs at 200 keV. The oscillations of the diffracted beam intensities allow for an oscillation period or an "extinction distance" to be defined. However only in a two-beam theory (or kinematic theory away from the Bragg condition) are the oscillations periodic.

Equation (4) shows that a different extinction distance can be associated with every excitation error or point within the diffraction disk. (A similar result holds in two-beam theory).

In this way, both the convergent-beam and precession methods, [34] by including a large range of orientations around the Bragg condition, average over many of these thickness periods and so reduce the rapidity of the thickness (Pendellosung) oscillations. Figs. 1.1 (c and d) represent the same oscillation trends for the two diffracted beams, but at 400 and 1000 keV respectively.

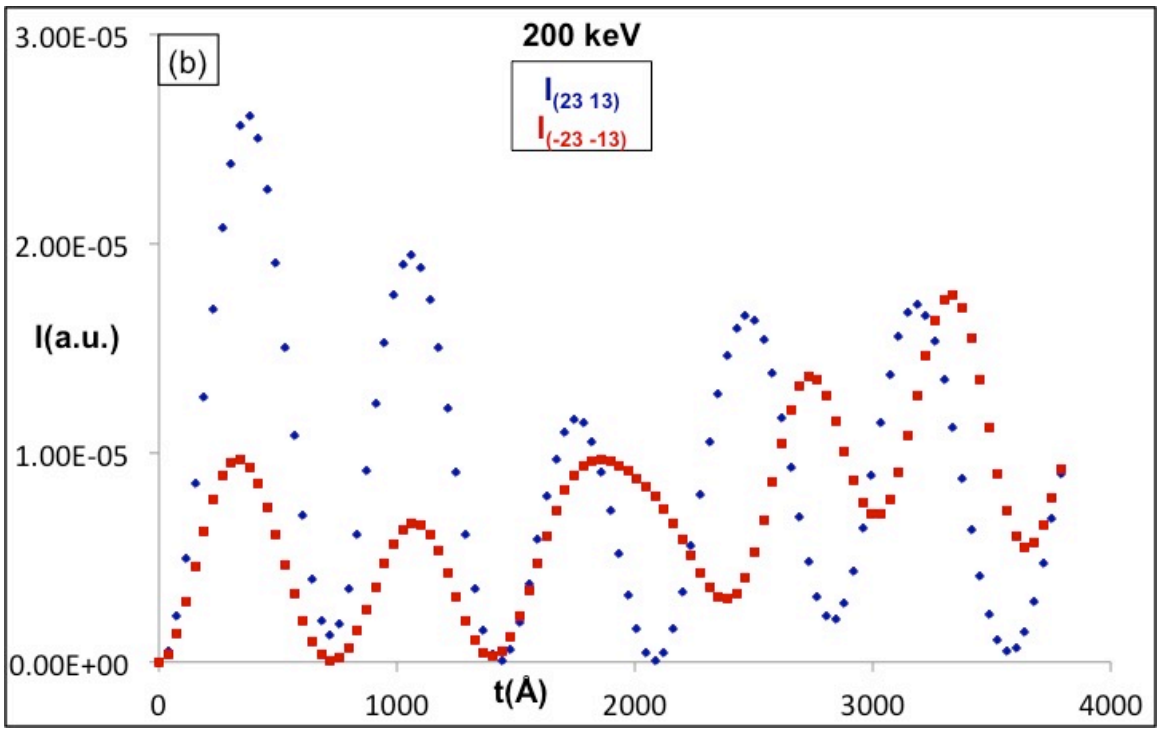


Fig (1.1 b): Diffracted beam intensity vs thickness for lysozyme crystals. At 200 keV diffracted beam (23 13) and its Friedel pair (-23 -13)

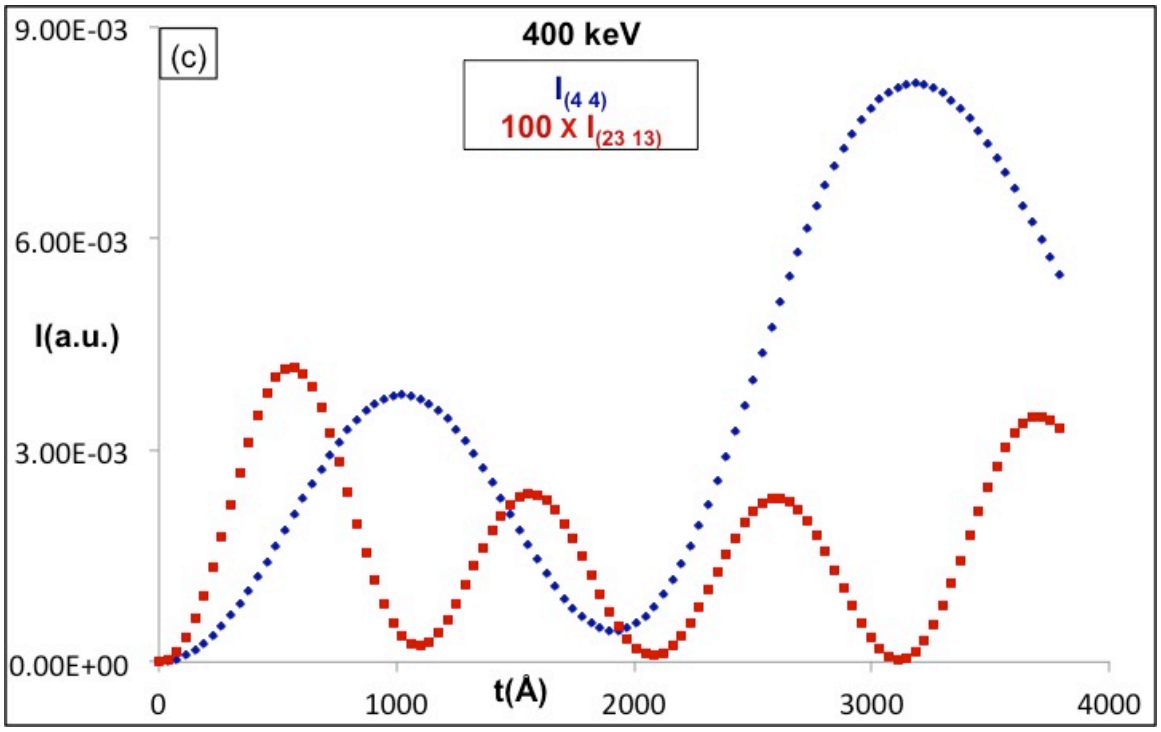


Fig (1.1 c): Diffracted beam intensity vs thickness for lysozyme crystals. At 400 keV, diffracted beams (4 4) and (23 13).

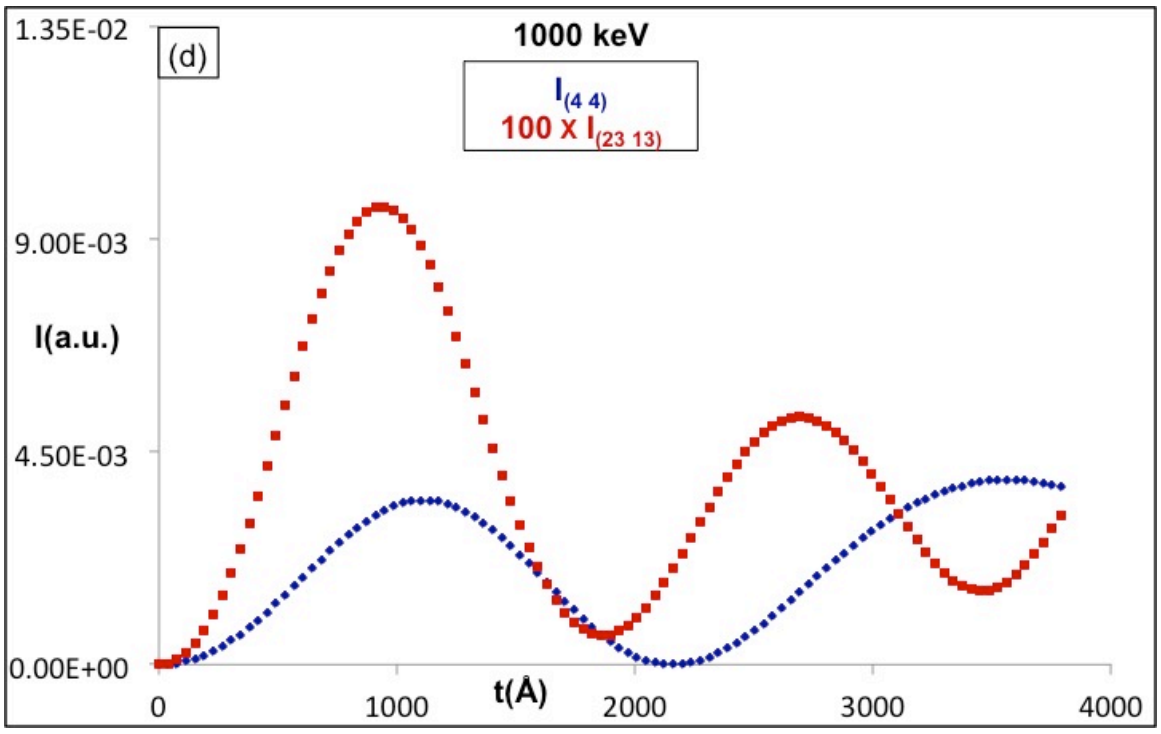


Fig (1.1 d): Diffracted beam intensity vs thickness for lysozyme crystals. At 1000 keV, diffracted beams (4 4) and (23 13).

The increase in pendellosung oscillation periods with increase in energy can also be noticed. This can be linked to the fact that with increasing energies, the extinction distance increases as well. Two-beam calculations for extinction distances of a relativistic electron can be shown to be directly proportional to electron velocity (square root of the electron beam incident energy) as in eqns. 9.22 and 9.25 of Hirsch, Peter Bernhard, et al. *1965, 549 P. Butterworth Inc., 7300 Pearl Street, Washington, D. C. 20014* (1966)). For n-beam models, this is expected to get more complicated, maintaining however a similar trend.

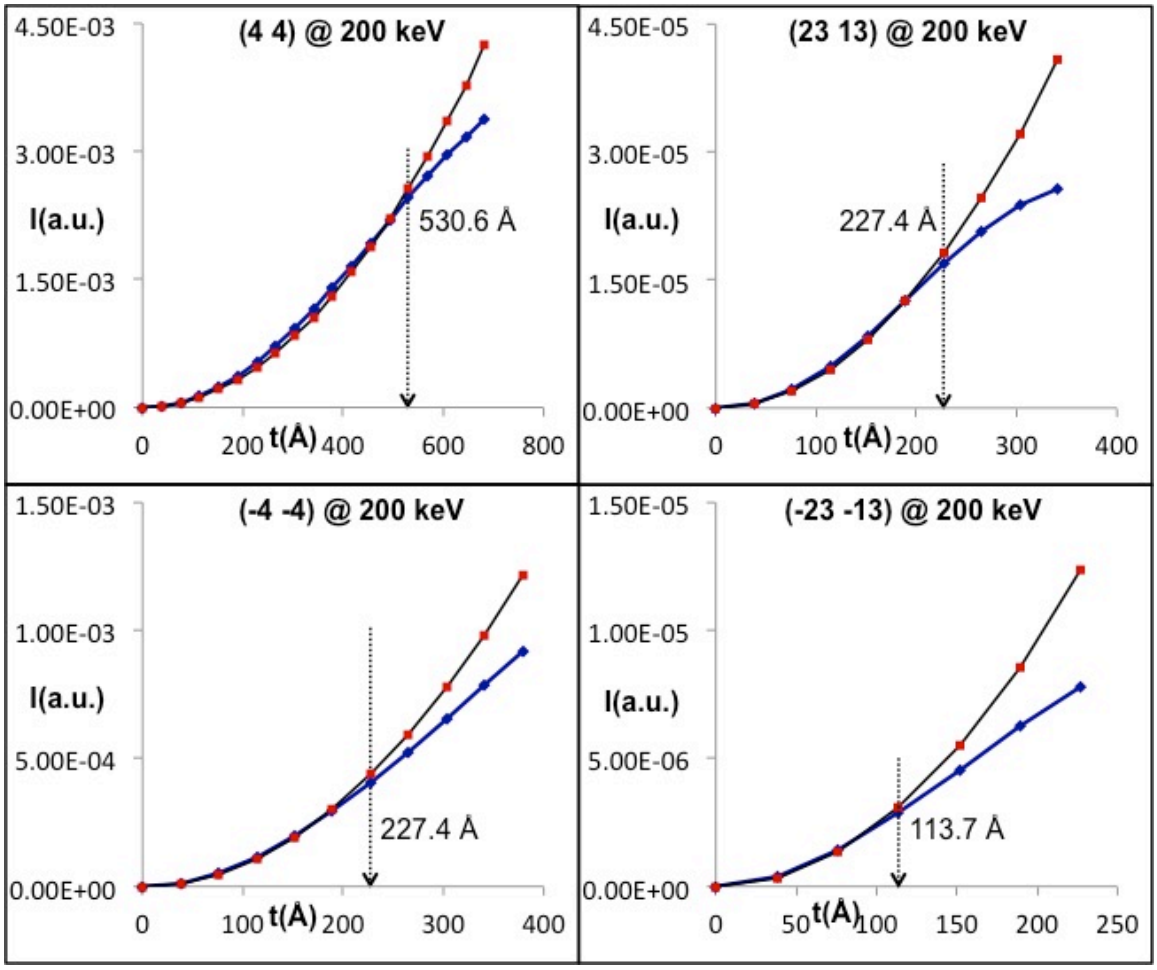


Fig (1.2-a): Quadratic fits (black) to the simulated diffracted intensities (blue) as a function of thickness at 200 keV.

To estimate the kinematic limit, a quadratic curve of the type $y = at^2$ is fit to each of the above curves. As in eqn. (4), since the diffracted intensity is proportional to the square of the sample thickness under the kinematic limit, any deviation from the quadratic behavior is taken as an indication of breaking away from the single scattering regime (as is consistent with the explanation of using the higher order terms in the Born series expansion). The simulated diffracted intensities and corresponding quadratic fits are presented in figs. 1.2 (a and b). The black curve (the quadratic fit) in figs. 1.2 (a and b) proposes diffracted intensities assuming single scattering by the specimen. Clearly, the deviation of the simulated diffracted intensities from the quadratic nature is visible. The early departure from the kinematic behavior of the higher order diffracted spots (23 13) as compared to that of the lower order Bragg spots (4 4) is also visible.

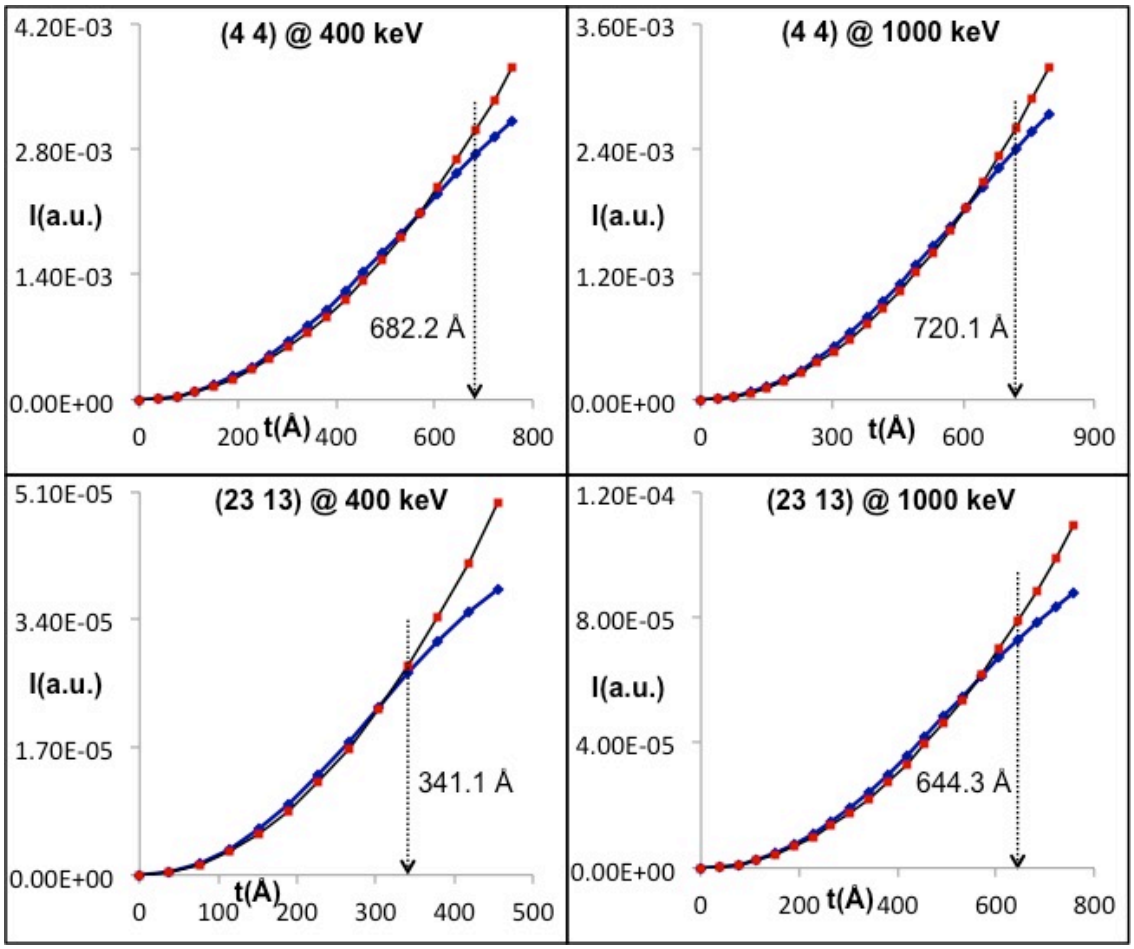


Fig (1.2-b): Quadratic fits (black) to the simulated diffracted intensities (blue) as a function of thickness at 400 keV and 1000 keV.

Allowing 10% error tolerance in the intensities, the kinematic (single-scattering) limit were calculated from the fit curves in figs. 1.2 (a and b). These results along with the fit equations and the corresponding parameters are tabulated in Table (1). At 3 Å resolution (23 13), the kinematic limit was observed to be 227.4 Å at 200 keV. This kinematic limit increases [31] with incident beam energy, thus indicating a possible use of higher incident beam energies to support the study of thicker protein samples that still diffract within the kinematic regime. Although on the one hand this warrants the ability to use thicker crystals of proteins, on the other hand the signal obtained is weaker because of weaker interaction and hence this compromise has to be adjusted for. Moreover, high beam energy increases the probability of knock-on radiation damage (especially for light elements). The reduction in ionization damage, but rapid increase in such "knock-on" ballistic damage, with increasing beam energy has been thoroughly studied and reported in the literature for beam energies up to 3 MeV (see [31,35] for reviews). Hence, using very high electron incident energies as an option for electron diffraction of biomolecules has to be carefully and more rigorously evaluated.

INCIDENT ENERGY	DIFFRACTED SPOT	Coeff. 'a'	Fit Eqn. ($y = at^2$)	KINEMATIC LIMIT (Å)
200 keV	(4 4)	9.11×10^{-9}	$y = (9.11 \times 10^{-9})\ t^2$	569
	(-4 -4)	8.44×10^{-9}	$y = (8.44 \times 10^{-9})\ t^2$	227
	(23 13)	3.52×10^{-10}	$y = (3.52 \times 10^{-10})\ t^2$	227
	(-23 -13)	2.39×10^{-10}	$y = (2.39 \times 10^{-10})\ t^2$	114
400 keV	(4 4)	6.47×10^{-9}	$y = (6.47 \times 10^{-9})\ t^2$	682
	(23 13)	2.39×10^{-10}	$y = (2.39 \times 10^{-9})\ t^2$	341
1 MeV	(4 4)	5.02×10^{-9}	$y = (5.02 \times 10^{-9})\ t^2$	720
	(23 13)	1.91×10^{-10}	$y = (1.91 \times 10^{-10})\ t^2$	644

Table 1.1: Variation of kinematic thickness in lysozyme with incident beam energy

6.2 Molecular Replacement

In protein crystallography (both X-rays and electrons), as mentioned earlier, it is widely accepted that about 2/3 of the information in a charge density map comes from structure factor phases (commonly obtained using the molecular replacement method, using a similar protein as model) and 1/3 from the amplitudes. [20,21] To evaluate this, the analysis started with data from X-ray diffraction first and introduced random errors to the x-ray intensities but maintained accurate phases. The idea was to generate charge density maps with the erroneous intensities and use these density maps to identify the percentage error beyond which, real-space charge density starts to be substantially different from ideal. To generate the electron density maps, the PHENIX program suite is used. This suite allows generating regular electron density maps, omit maps, and difference maps and many others.

X-ray diffraction intensities and phases for hen-egg white lysozyme were obtained from 4ET8.pdb. [36] Keeping the phases unchanged, random errors were introduced to the intensities. At this stage, an omit map was generated. An omit map is type of electron density map that is generated by omitting a certain number of protein residues from the refinement process and calculating/recalculating the phases and structure factor amplitudes of the model. Once the refinement is completed, these residues are added back and plotted. In this case, residues 50-60 fare omitted from the refinement process. Fig. 1.3 shows the omit maps where the charge densities (wireframes) were laid over the ideal structure (sticks) from the model. For a perfect fit these should agree. The red wire-mesh corresponds to ideal (zero-error) omit- map. Thus figs. (1.3a), (1.3b), (1.3c) and (1.3d) correspond to charge density maps from 20% 50%, 70% and 100% error in X-ray intensities (10%, 22%, 34% and 41% errors in |F| respectively) overlaid upon the density map with zero-error.

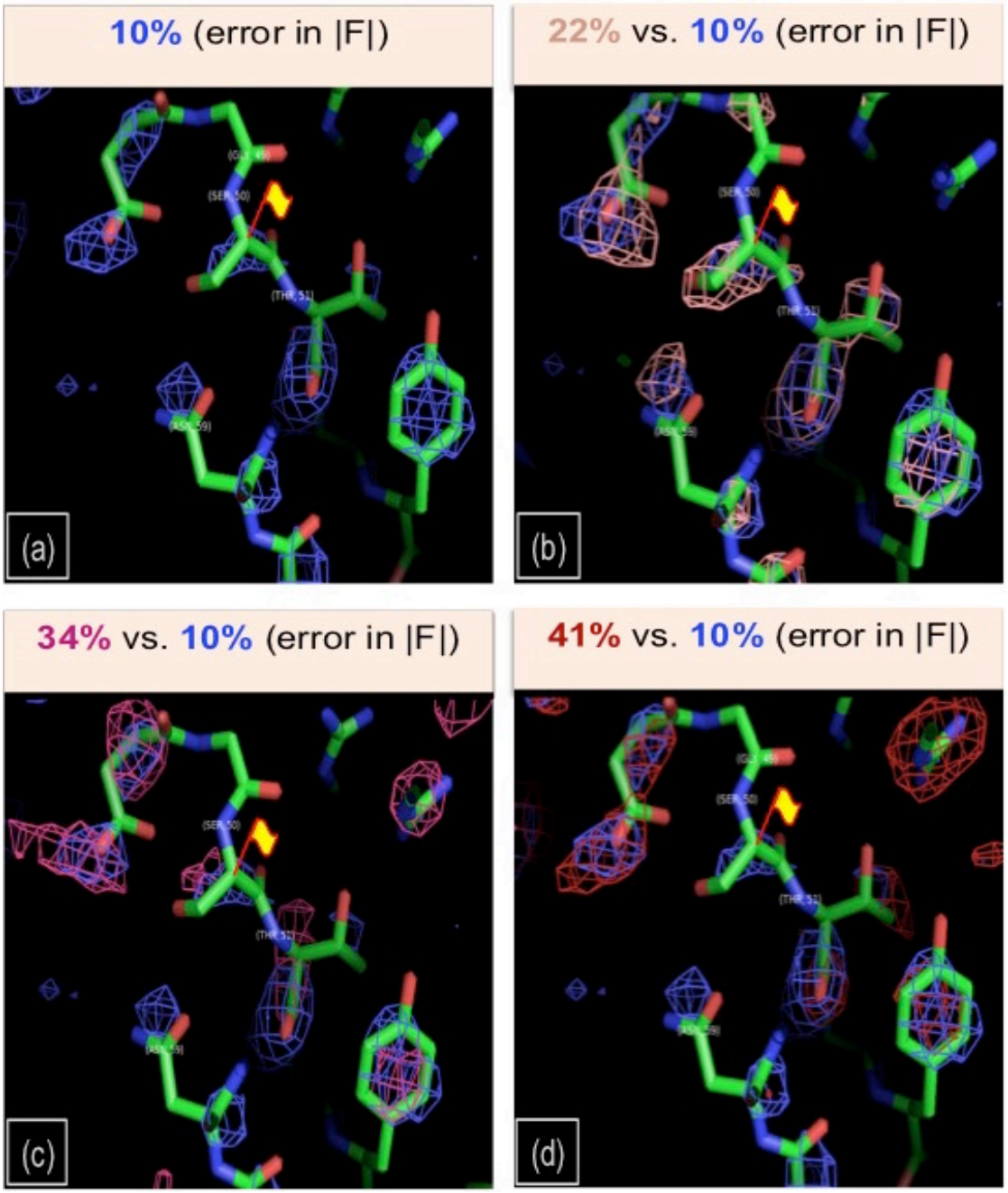


Fig: 1.3: Omit Map for Lysozyme (residues 50-60 omitted from refinement) at *1.5*σ. Each sub-plot contains charge density maps with increasing random errors introduced in the experimental diffracted intensities. The sticks in green are some of the omitted residues. Red wire mesh corresponds to 'zero' error omit map. ***(a)*** 20% error (pink wire-mesh) overlapped with zero error. ***(b)*** 50% error (sky-blue wire-mesh) overlapped with zero error. ***(c)*** 70% error (marine-blue wire-mesh) overlapped with zero error. ***(d)*** 100% error (white wire-mesh) overlapped with zero error.

With increasing error in the intensities, it can be noticed that the charge density (wire-mesh) does not match with the atomic positions. Also, the noise in the data increases with error. For this specific case, monitoring the regions around the residues 50, 51 and 59 (labeled in figs 1.3 a-d), it is noted that at 34% error in |F|, the electron density substantially fails to match the model. In fact, at 41% error (fig. 1.3 d), electron density is almost completely absent around the labeled residues. The threshold error in |F| in this case can, thus, be taken at 34%. However, caution needs to be exercised in extending an interpretation of this kind. In this calculation, the random errors are added to the diffracted intensities (reciprocal space), while the density maps are plotted in real physical space.

Since the real and reciprocal space are related by the Fourier transform, the errors in reciprocal space are distributed at every single sampling point of the real space through the Fourier coefficients. This means that it is not straightforward to establish a correlation between the addition of error in reciprocal space and the manifestation of the same in specific real space coordinates. Nevertheless, for any protein, there are always specific regions that are characteristic of their structural/functional behavior (like active sites of catalysts) and the accuracy/inaccuracy in characterizing these structures can be used as benchmarks to categorize the threshold extent of error. This can, for example, be understood using the following case where the threshold error beyond which, similar (but not identical) amino acid sequences from hen-egg white (HEWL) and turkey-egg white (TEWL) lysozymes could not be differentiated, was identified. It is known that, HEWL possesses His15 (i.e., Histidine residue at position 15 in the sequence) amino acid while the TEWL possesses Leu15 (i.e., Leucine residue at position 15 in the sequence), everything else remaining the same. In such a case, for a model based on turkey-lysozyme and experiments based on hen-lysozyme, one would expect the overlap of the model (sticks) with the experimental charge density (wire-mesh) to not be perfect at that specific amino-acid position falling at residue 15. In Figs. (1.4 and 1.5), an $F_{obs.}$–$F_{calc.}$ difference map is plotted for this exact case. Fig. 1.4 shows the difference density map with the stick models representing both the hen (golden sticks) and turkey lysozymes (silver sticks) while Fig. 1.5 shows the difference density map with the stick model based only on turkey-lysozyme (silver sticks).

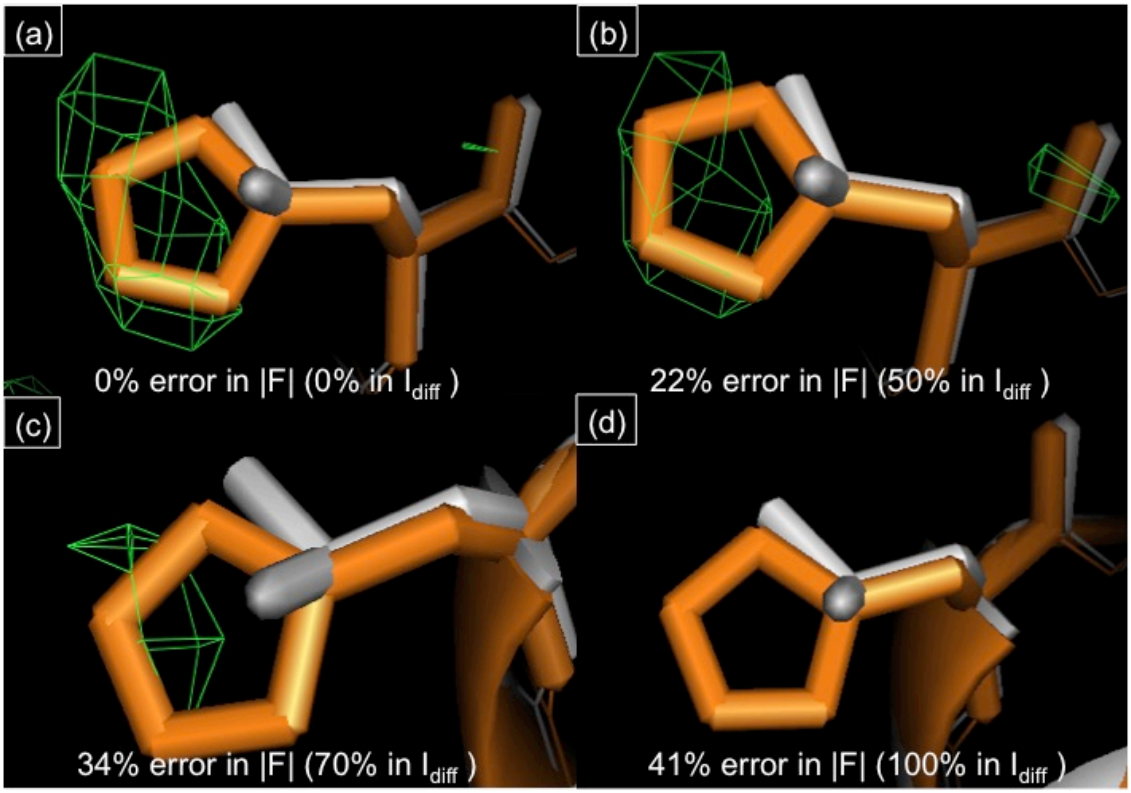


Fig 1.4: F_o-F_c difference maps for TEWL (silver stick - model) and HEWL (golden stick – model) with experimental data obtained from HEWL presented at at 3.0σ. Density outside the model (wire frame) indicates distinguishability, absent with 41% error in |F|.

A difference map is any charge density map that is calculated from the difference between two sets of structure factor amplitudes. As with the previous figure, random errors to the diffracted intensities to varying degrees are added here as well. Fig. 1.4 demonstrates clearly that at residue 15, the density (in green mesh), which is generated from experimental diffraction intensities and MR model based phase, overlaps with the golden sticks (model) corresponding to the hen-lysozyme. However, this is evident until the error in structure factor amplitude is greater than 30% (greater than 75% error in diffracted intensities) and only when the error in the structure factor amplitude is greater than 40% (greater than 100% error in diffracted intensities), does the program lose scope of identifying the correct model for the sample. Fig. 1.5 demonstrates this exact same effect using just the turkey-lysozyme as the model and hen-lysozyme diffraction data.

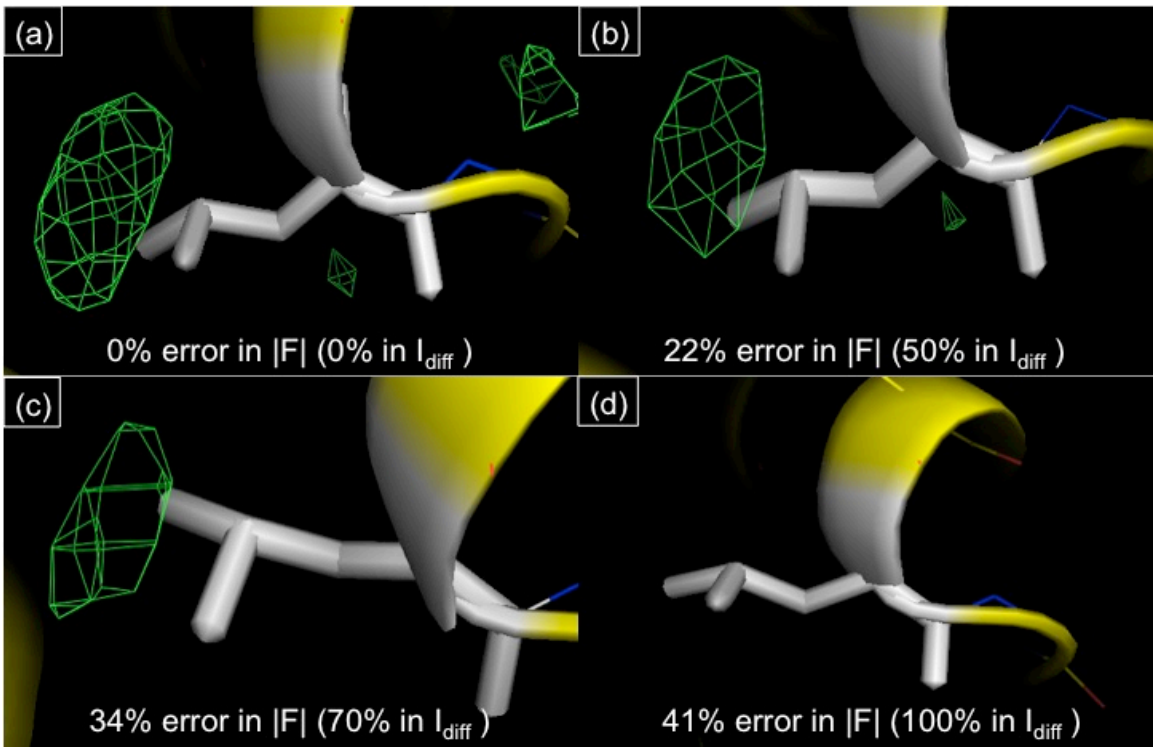


Fig 1.5: F_o-F_c difference maps for TEWL (silver stick - model) and experimental data obtained from HEWL presented at at 3.0σ. Density outside the model (wire frame) indicates distinguishability, absent with 41% error in |F| same as in fig. 1.4.

In the above exercise, the diffracted intensities obtained from 4ET8.pdb file were modified by adding random errors. This was followed by analysis of the extent of error tolerance by generating charge density maps.

A complementary scheme of analysis can be performed by studying the variation of the statistical parameters used to characterize the veracity of such molecular replacement based model refinement. Presented in table 1.2 is a summary of the refinement parameters as a function of the random error introduced to the diffracted intensities to HEWL.

In the above table three parameters namely Rwork, Rfree and CC are presented as a function of the error introduced to the diffracted intensities obtained from 4ET8.pdb. The Rfactor is a statistical parameter widely used in crystallography that measures how close the experimental |F| values are in relation to the crystallographic model. It is expressed as in equation (a) where $|F_{obs}|$ and $|F_{calc}|$ are the experimental and model structure factor amplitudes

respectively. In protein crystallography, a practically accepted limit for Rfactor is 0.3. Wilson finds an extreme upper limit in the R-value of 0.59, for non-centrosymmeteric crystals such as proteins, beyond which the model can have the same symmetry but a completely random atomic arrangement. [37]

Error in intensity (|F|)	Rwork		Δ Rwork (%)	Rfree		Δ Rfree (%)	CC (model vs map)
	Initial	Final		Initial	Final		
0 (0)	0.1835	0.1727	- 5.8	0.1813	0.1894	+ 4.5	0.9340
20 (10)	0.1892	0.1762	- 6.8	0.1953	0.1986	+ 1.7	0.9280
35 (16)	0.2044	0.1931	- 5.5	0.2093	0.2168	+ 3.5	0.9200
50 (22)	0.2226	0.2130	- 4.3	0.2399	0.2510	+ 4.6	0.9010
60 (27)	0.2481	0.2281	- 8.0	0.2412	0.2517	+ 4.4	0.8820
70 (30)	0.2540	0.2296	- 9.6	0.2590	0.2645	+ 2.1	0.8890
80 (34)	0.2864	0.2733	- 4.6	0.3183	0.3224	+ 1.3	0.8380
100 (41)	0.3430	0.3331	- 2.9	0.4368	0.4533	+ 3.8	0.7580

Table 1.2: Refinement parameters as a function of error introduced to HEWL diffracted intensities.

From Table 1.2 it can be observed that both the Rfactor parameters behave similarly to the extent that the initial starting values increase with increasing error. However, post molecular replacement based refinement there is a completely opposite behavior between the two parameters. While the Rwork parameter consistently decreases post refinement suggesting a likely improvement of the model for each case, the Rfree parameters consistently increase post refinement. The increase in the Rfree parameter is an indication that the model is not defining the

system better. So the apparent reduction in the Rwork parameter is possibly due to over-fitting raising a false alarm of progressive refinement. The cross correlation (CC) values comparing the model and density maps generated also indicate a pretty high match although it steadily decreases with increase in random error.

This table suggests a few additional things - (i) Rwork has to be taken in combination with Rfree to better legitimize predictions of models that fit the experimental data. (ii) In spite of substantial error in the diffracted intensities, the refinement still churns out decent Rfactors. The numbers are beyond threshold (0.3) acceptance values only at a 100% error to diffracted intensities! This is a direct evidence of the extent to which the molecular replacement procedure depends on the phases and why error in intensities have to be stupendously high to cause (iii) Correlation parameters are also susceptible to be biased towards the phases much stronger than the diffracted intensities as seen in the table. The decrease in correlation is gradual and again only significant at very large errors in diffracted intensities.

The entire analysis so far presented to study the tolerance to errors in diffracted intensities in molecular replacement has been done with data collected by using X–rays. However, everything else remaining the same, the observations and conclusions can be extrapolated to diffraction data collected using an e-beam as well. Thus a major inference from the above analysis is that in using electron diffraction and molecular replacement to identify biological structures, the technique is substantially less sensitive to the errors in actual experimental measurement (diffracted intensities) if the initial model used is accurate in phases. This can be re-written to state that this technique is so heavily dependent on the accuracy of phases from the model for its success that, in solving ab inito structures of hitherto unsolved proteins the experimental data need not have much significance.

In the previous analysis, it was noted that as long as the protein crystal thickness is within the kinematic (single-scattering) limit, diffracted intensities can be detected as the square of the structure factor amplitudes. When the thickness of sample increases, the e-beam experiences dynamical scattering and the diffracted intensity vs thickness relation deviates strongly from the quadratic behavior. This process can be seen as equivalent to artificially introducing errors in the

ideal single-scattered diffracted intensities. Thus, the Rfactor can be calculated as a function of the crystal thickness to establish a maximum crystal thickness for lysozyme under the kinematic limit, by introducing random errors to the simulated electron diffraction intensities (which are considered analogous to the error associated with deviation from the quadratic behavior in thick dynamical samples).

For this, multiple diffracted beams (90 in total) of hen-egg white lysozyme at 200 keV were used. The structure factor amplitudes corresponding to 37.9 Å (1 unit cell thickness) were considered as the $|F_{obs}|$, with respect to which, the structure factor amplitudes of higher thicknesses ($|F_{calc}|$) were evaluated. The results are plotted in fig. 1.6 as a plot of the Rfactor as a function of the sample thickness.

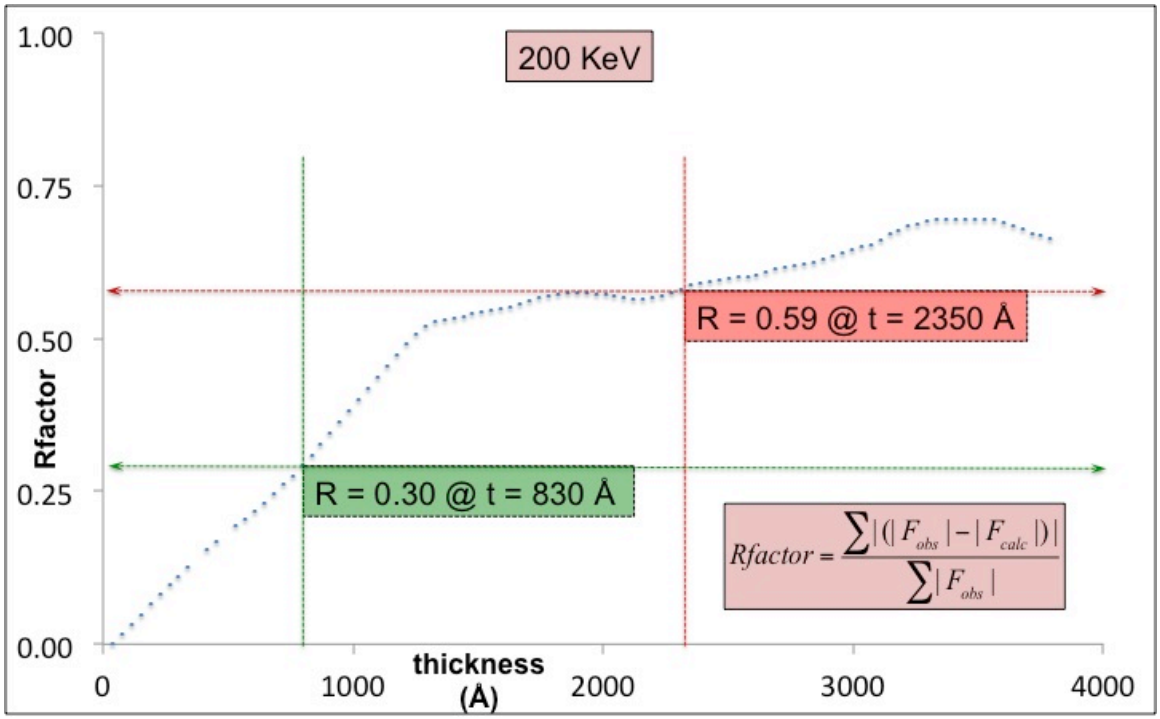


Fig. 1.6: Rfactor vs thickness for lysozyme @ 200keV.

From Fig. 1.6, the maximum thickness of lysozyme crystal that can be studied for R ≤ 0.3 was 830 Å. For thicknesses greater than 2300 Å, R ≥ 0.59. Therefore, beyond 2300 Å thickness, the |F| values obtained from diffracted intensities completely fail to specify the atomic

arrangement of the lysozyme crystal. This calculation was done with (10 10 0) as the highest order diffracted beam, which corresponds to roughly 6 Å resolution. A more rigorous study would include hundreds of diffracted beams, up to the resolution of interest. Such a study could decrease the limit of acceptable thickness within the single-scattering approximation, below that presented here.

Several tests for the absence of multiple-scattering perturbations exist, most of which require knowledge of the sample structure. However, the obsevation of uniform intensity in convergent-beam disks [14] does not, nor does the observation of Friedel symmetry (which is destroyed by mulitple scattering) in samples known to be acentric, such as proteins. Also (as mentioned earlier), for samples whose space-group is known, reflections that are absent as a result of screw or glide planes may occur through multiple scattering. In fact, the Gjonnes-Moodie cancellation theorem [15] shows that these reflections remain absent for all thickness and beam energies, as long as they are oriented along certain specific beam directions, so that a test for multiple scattering based on observation of absent reflections fails, for perfect crystals. For proteins however, it is often a useful test, since the samples are usually bent, and hence a range of orientations is illuminated that is greater than the angular window over which cancellation of multiple-scattering paths occurs. Lastly, experience in other fields [34,38,39] shows that dynamical interactions are reduced in orientations away from prominent zone axes, since, in the 2-beam dynamical theory, the periodicity of beams is inversely proportional to the structure-factor magnitudes, and so increases for weak reflections at the Bragg condition. Although zone axes orientations excite the maximum number of strongly interacting reflections simultaneously, they are needed to synthesize the major high-symmetry projections of the density map. However, by contrast, as shown by equation (1), the short thickness periodicity of weak reflections, which are well away from the Bragg condition, is reduced to $2/s_g$, and so should be avoided.

7 Conclusions

The importance of multiple scattering as a limiting factor in cryo-EM structure analysis of protein microcrystals has been studied. (i) From the multislice simulations for perfectly ordered hen-egg lysozyme, it is observed that (a) the TED single-scattering limit at 3 Å resolution for lysozyme at 200 keV is about 230 Å at a zone-axis orientation for 10% error in intensities. (b) At higher incident beam energy, the kinematic limit increases (about 650 Å at 1 MeV for 3 Å resolution) but at the risk of a high probability of knock-on radiation damage. (ii) By introducing random errors to |F| for lysozyme (while maintaining the correct phases) it is found that (a) at around 35% error in |F|, the charge densities obtained from the erroneous |F| start to substantially deviate from the ideal model, at the region of interest. (b) the ability to distinguish hen and turkey lysozyme is lost at about 40% error in |F|. The high tolerance to error in diffracted intensities (40% error in |F| corresponds to almost 100% error in diffracted intensity) is a consequence of the inherently greater phase dependence of charge density maps. Although this seems to suggest a possible leverage in using TED with such high tolerance to error in intensities, it emphasizes the necessity of accurate phase information (for which one is dependent on prior [36] experiments). (iii) R-factor analysis over a subset of simulated diffracted beams show that the maximum crystal thickness, for perfectly ordered protein crystals, that can be used to maintain a reasonable value ($R < 0.3$) is about 1000 Å.

These results are, however, limited by the fact that (a) although all proteins have similar atomic density and composition, the thickness limit of the TED method, (in this case given for lysozyme) will vary depending on the particular crystal structure. Nanocrystals with a large unit cell produce weaker diffracted beams for the same crystal volume, and a denser reciprocal lattice. A larger tolerable thickness can therefore be expected for larger unit cells, as is confirmed by detailed simulations for inorganic oxides. [31] Also one can refer to [40] for scattering (of x-rays) by nanocrystals that are a few unit cells thick in all three dimensions. (b) this paper takes no account of mosaicity or crystal bending. It applies strictly to microdiffraction patterns obtained from very small regions of perfect crystal, and so subject to the limitations of radiation damage. Theories of mosacity in electron diffraction have been published (see [19] for a review) in which coherent

multiple scattering within mosaic blocks are coupled to incoherent mulitple scattering of intensities between blocks. However experimental evidence [41] suggests that continuous bending is more likely for TEM samples than the tilted block structure assumed in the much thicker X-ray samples, combined with the "secondary scattering" of intensities between layers proposed by Cowley (see [41] and references therein). This can have beneficial effects for structure analysis by TED, since the thickness oscillation periods shown in Fig. 1.1 are proportional to the inverse of the deviation s_g from the Bragg condition (equation 1). As in the precession electron diffraction method, [34] by spanning a range of orientations, each with a definite value of s_g, the effect is to smooth out the thickness oscillations. In a recent paper [42] it is claimed that useful structure analysis can be achieved by TED using protein crystals in the 160 - 400nm range (R ~ 25% at 0.25 nm resolution). Some of this increase beyond the estimate of 83 nm may be due to this bending effect. (c) This work has not considered single-particle imaging. Recent results suggest that useful images at 3.6 Angstrom resolution can be obtained from a virus as thick as 100nm. [43] In the phase-object approximation, single scattering requires a phase-shift of less than ninety degrees as the beam traverses the sample. For a crystal, because molecules stack up in columns in the beam direction, this phase shift (relative to solvent columns) accumulates more rapidly than for single-particle imaging, where the molecules are not ordered. So the tolerable thickness to avoid multiple scattering could be greater for single-particle imaging than for Bragg diffraction.

The challenge of growing protein nanocrystals whose size falls within the domain of validity of the single-scattering approximation has been taken up by several groups. In particular the SONICC method [44] provides a crucial characterisation tool that can identify protein crystals as small as one tenth of a micron, and has been used to enable the growth of protein nanocrystals for both membrane and soluble proteins through microfluidics. [45] The success and versatility of protein nanocrystal growth and characterization techniques, thus, directly relates to the possibility of using TED to solve protein structures.

8 References

1. Henderson, R. *et al.* Model for the structure of bacteriorhodopsin based on high-resolution electron cryo-microscopy. *J. Mol. Biol.* **213,** 899–929 (1990).

2. Kühlbrandt, W., Wang, D. N. & Fujiyoshi, Y. Atomic model of plant light-harvesting complex by electron crystallography. *Nature* **367,** 614–621 (1994).

3. Shi, D., Hsiung, H. H., Pace, R. C. & Stokes, D. L. Preparation and analysis of large, flat crystals of Ca(2+)-ATPase for electron crystallography. *Biophys. J.* **68,** 1152–1162 (1995).

4. Taylor, K. A. & Varga, S. Similarity of Three-dimensional Microcrystals of from Pig Kidney and from Skeletal Muscle Sarcoplasmic Reticulum *. *J. Biol. Chem* **269,** 10107–10111 (1994).
5. Jap, B. K. High-resolution electron diffraction of reconstituted PhoE porin. *J. Mol. Biol.* **199,** 229–231 (1988).

6. Fujiyoshi, Y. *et al.* Structure and function of water channels. *Current Opinion in Structural Biology* **12,** 509–515 (2002).

7. Löwe, J., Li, H., Downing, K. . & Nogales, E. Refined structure of αβ-tubulin at 3.5 Å resolution. *J. Mol. Biol.* **313,** 1045–1057 (2001).

8. Chang, S., Head-Gordon, T., Glaeser, R. M. & Downing, K. H. Chemical bonding effects in the determination of protein structures by electron crystallography. *Acta Crystallogr A* **55 (Pt 2 ,** 305–313 (1999).

9. Glauber, R. & Schomaker, V. The Theory of Electron Diffraction. *Phys. Rev.* **89,** 667–671 (1953).

10. Son, J., Millward, G. R. & Thomas, J. M. The Role of Multiple Scattering in the Study of Lattice Images of Graphitie CarbOns. *Acta Crystallogr. Sect. A Cryst. Physics, Diffraction, Theor. Gen. Crystallogr.* **32,** (1976).

11. De Rosier, D. J. & Klug, a. Reconstruction of Three Dimensional Structures from Electron Micrographs. *Nature* **217,** 130–134 (1968).

12. Glaeser, R. M. & Downing, K. H. High-resolution electron crystallography of protein molecules. *Ultramicroscopy* **52,** 478–486 (1993).

13. Anstis, G. R., Lynch, D. F., Moodie, a. F. & O'Keefe, M. a. n -Beam lattice images. III. Upper limits of ionicity in W 4 Nb 26 O 77. *Acta Crystallogr. Sect. A* **29,** 138–147 (1973).

14. McKeown, J. T. & Spence, J. C. H. The kinematic convergent-beam electron diffraction method for nanocrystal structure determination. *J. Appl. Phys.* **106,** 074309 (2009).

15. Gjønnes, J. & Moodie, a. F. Extinction conditions in the dynamic theory of electron diffraction. *Acta Crystallogr.* **19,** 65–67 (1965).

16. Huang, W. J., Zuo, J. M., Jiang, B., Kwon, K. W. & Shim, M. Sub-ångström-resolution diffractive imaging of single nanocrystals. *Nat. Phys.* **5,** 129–133 (2009).

17. Nederlof, I., Li, Y. W., Van Heel, M. & Abrahams, J. P. Imaging protein three-dimensional nanocrystals with cryo-EM. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **69,** 852–859 (2013).

18. Glaeser, R. M. Review: electron crystallography: present excitement, a nod to the past, anticipating the future. *J. Struct. Biol.* **128,** 3–14 (1999).

19. Dorset, D. L. *Structural electron crystallography*. (Springer Science & Business Media, 2013).

20. Read, R. J. [7] Model phases: Probabilities and bias. *Methods Enzymol.* **227,** 110 (1997).

21. Lattman, E. & DeRosier, D. Why phase errors affect the electron function more than amplitude errors. *Acta Crystallogr. A.* **64,** 341–4 (2008).

22. Scapin, G. Molecular replacement then and now. in *Acta Crystallographica Section D: Biological Crystallography* **69,** 2266–2275 (2013).

23. Rossmann, M. G. The molecular replacement method. *Acta Crystallogr. Sect. A* **46,** 73–82 (1990).
24. Brünger, A. T. Extension of molecular replacement: a new search strategy based on Patterson correlation refinement. *Acta Crystallogr. Sect. A* **46,** 46–57 (1990).

25. Rossmann, M. G. Molecular replacement - Historical background. *Acta Crystallogr. - Sect. D Biol. Crystallogr.* **57,** 1360–1366 (2001).

26. Abergel, C. Molecular replacement: Tricks and treats. in *Acta Crystallographica Section D: Biological Crystallography* **69,** 2167–2173 (2013).

27. Evans, P. & McCoy, A. An introduction to molecular replacement. *Acta Crystallogr. D. Biol. Crystallogr.* **64,** 1–10 (2008).

28. Doyle, P. A. & Turner, P. S. Relativistic Hartree-Fock X-ray and electron scattering factors. *Acta Cryst (1968). A24, 390-397 [doi10.1107/S0567739468000756]* **3,** 1–8 (1968).

29. Zuo, J. M. & Spence, J. C. H. *Electron microdiffraction*. (Springer Science & Business Media, 2013).

30. Meyer, J. C. *et al.* Experimental analysis of charge redistribution due to chemical bonding by high-resolution transmission electron microscopy. *Nat. Mater.* **10,** 209–215 (2011).

31. Spence, J. C. H. *High-resolution electron microscopy*. (OUP Oxford, 2013).

32. Zuo, J. M. Measurements of electron densities in solids: a real-space view of electronic structure and bonding in inorganic crystals. *Reports Prog. Phys.* **67,** 2053–2103 (2004).

33. Cowley, J. M. *Diffraction Physics*. (Elsevier, 1995).

34. Vincent, R. & Midgley, P. A. Double conical beam-rocking system for measurement of integrated electron diffraction intensities. *Ultramicroscopy* **53,** 271–282 (1994).

35. Reimer, L. & Kohl, H. *Transmission Electron Microscopy*. *Book* (2008). doi:10.1007/978-0-387-34758-5

36. Boutet, S. *et al.* High-resolution protein structure determination by serial femtosecond crystallography. *Science (80-.).* **337,** 362–364 (2012).

37. Wilson, A. J. C. Largest likely values for the reliability index. *Acta Crystallogr.* **3,** 397–398 (1950).

38. Marks, L. D. & Plass, R. Atomic structure of Si (111)-(5x 2)-Au from high resolution electron microscopy and heavy-atom holography. *Phys. Rev. Lett.* **75,** 2172 (1995).

39. Own, C. S.-Y. System Design and Verification of the Precession Electron Diffraction Technique. (Northwestern University, 2005).

40. Kirian, R. A. *et al.* Femtosecond protein nanocrystallography—data analysis methods. *Opt. Express* **18,** 5713 (2010).

41. Dorset, D. L. Electron diffraction intensities from bent molecular organic crystals. *Acta Crystallogr. Sect. A Cryst. Physics, Diffraction, Theor. Gen. Crystallogr.* **36,** 592 – 600 (1980).

42. Nannenga, B. L., Shi, D., Leslie, A. G. W. & Gonen, T. High-resolution structure determination by continuous-rotation data collection in MicroED. *Nat. Methods* **11,** 927–30 (2014).

43. Liu, H. *et al.* Atomic structure of human adenovirus by cryo-EM reveals interactions among protein networks. *Science* **329,** 1038–43 (2010).

44. Kissick, D., Wanapun, D. & Simpson, G. Second-order nonlinear optical imaging of chiral crystals. *Annu. Rev. Anal. …* **4,** 419–437 (2011).

45. Abdallah, B. G., Kupitz, C., Fromme, P. & Ros, A. Crystallization of the large membrane protein complex photosystem i in a microfluidic channel. *ACS Nano* **7,** 10534–10543 (2013).

OUTRUNNING RADIATION DAMAGE and HOLLOW-CONE TEM

1 Introduction

Using hard X-ray pulses of a few tens of femtoseconds duration, it has recently been shown that it is possible to outrun radiation damage, that is, to collect a high-resolution diffraction pattern from elastic scattering, before the incident pulse (and the detection process) terminates, and prior to the vaporization of the sample due to the subsequent photoelectron cascade. [1] For a review, and applications to protein crystallography of this "diffract-and-destroy" method, see. [2] In recent work, [3] the structure of the important angiotensin GPCR protein (controlling blood pressure) and its drug ligand has been solved by this method to a resolution of 2.9 Å. This ability to record snapshot diffraction at room temperature without radiation damage at near-atomic resolution from protein nanocrystals has recently also produced a "molecular movie" of the photo cycle, using pump-probe methods, from Photoactive Yellow Protein, at 1.6 Å resolution. [4] These studies, using nanocrystals, take advantage of the coherent amplification of Bragg scattering, which, if the angular range of the Bragg condition spans many detector pixels increases scattered intensity by a factor of N^2 over that of one molecule, for N molecules in the crystal (e.g. by 1E6 for a nanocrystal just 10 molecules on a side). For the more difficult problem of single-molecule imaging (e.g. one virus particle per shot), which avoids the need for crystallization, the best resolution obtained to date in a single diffraction pattern (a 2D projection, not a 3D reconstruction) is about 10 nm. [5]

The question naturally arises as to whether this capability might also be possible using electron diffraction or imaging. A detailed comparison of X-ray and electron imaging techniques for biology has been given by Henderson (1994)**.** [6] Electron sources, specimen interactions, and electron optics provide several advantages (see [7] for a review) - the field-emission source to provide a fully spatially coherent beam of very high brightness, the availability of aberration-corrected lenses capable of atomic resolution, and above all, a scattering cross section at least 1000 times

greater than that for X-rays. It is noted, however, that a typical field-emission source produces only about 50 electrons per picosecond, insufficient to form an image in a single shot. Recent dramatic advances in fast electron sources have made it possible to generate over a million electrons in a high energy electron pulse of about 100 fs duration. Operating at MeV energies, one has the further advantage of minimizing multiple elastic scattering artifacts for sample thicknesses of 100-1000 nm typical of a whole cell. Finally, a new generation of direct-injection electron area detectors has recently revolutionized cryo-electron microscopy, with single-electron detection and can further reduce the demands on the sources for ultrafast electron imaging.

In this chapter, the number of electrons scattered by a large bioparticle illuminated by a femtosecond electron pulse is estimated first, to determine whether or not electrons can outrun damage. Then, a new imaging mode of much higher efficiency is proposed for time-resolved electron microscope full-field imaging, rather than diffraction.

2 Can Electron Beams Outrun Damage for Single-Particle Imaging?

In a previous study [8] of electron scattering with a continuous beam in the STEM (Scanning Transmission Electron Microscopy) mode from individual bioparticles (not crystals, and not cryo-em imaging), the number of electrons scattered by the beam, focused from a field-emission tip to 0.5 nm diameter onto protein in a thin ice film, was calculated for the case where the dose to the sample fell below the damage threshold for protein (about 5 electrons per square Angstrom). The results were compared with (linear) Cryo-EM phase-contrast imaging, where real-space images are combined from many identical particles, and the phase problem avoided. For STEM, using a 100 kV beam, it was found that, if the sample were thin enough (less than a few hundred angstroms thick) to avoid multiple scattering perturbations, then less than one electron was scattered when using this beam diameter, which defines the image resolution. The electron cross section falls off with increasing beam energy by a factor S^2, as discussed below, while the onset of multiple scattering perturbations for electron scattering in protein is delayed to greater thickness with increasing beam energy (see [9] for details).

Here, the case where a 3 MeV electron beam from a photocathode is considered such that it can be focused down to 500 micron diameter beam at the sample with pulse duration of 100 fs. The effect of pulse duration on the development of radiation damage for hard X-rays has been extensively studied, since it determines the X-FEL resolution in reconstructing the images. [10] For the nanocrystal studies referenced above, a resolution of less than 2 Angstroms has been obtained using 50 fs pulses. The detailed mechanism for damage will differ between electrons and X-rays, since X-rays are not scattered by the nucleus, unlike electrons, which scatter from both atomic electrons and the slower-moving nucleus. Therefore a 100 fs electron-beam pulse is considered, which provides a conservative estimate of the pulse duration required to outrun damage using electron beams, and for which experimental data exists. For such a system, [11] experimental measurements give a current density 'j' of 5.3E8 e^-/cm^2 at 3MeV for a 100 fs pulse into a 500 micron diameter 'd' focused beam, with beam divergence 5E-5 rad, much smaller than the Bragg angle for silicon and providing adequate spatial coherence. The corresponding number of electrons per pulse is -

$$M = j \cdot (\frac{\pi d^2}{4}) = 1.04E6 \text{ electrons per pulse at 3MeV}$$

It is initially assumed that the number of electrons per pulse is only weakly dependent on the transverse beam diameter and mostly affects the electron bunch length, as discussed in [12] due to longitudinal space charge effects. In fact this beam is always larger than the single particle. The particle assumed for the calculations here will be similar to a large icosohedral virus, whose molecular weight might be 300 MDa, and diameter 2r = 90 nm, thin enough in an ice film to avoid multiple scattering perturbation when using a 3 MeV beam.

Such an idealized spherical sample of 90nm diameter is chosen, assumed to consist of uniformly distributed protein with density 1.3 gm/cm^3, or r_o = (1/15) effective carbon atoms/ $Å^3$. [13] (Any increased scattering from DNA packed in the genome of a similar virus is assumed to be compensated by structural inhomogeneity). For protein of average composition, the average electron cross-section per effective carbon atom for elastic scattering is s = 0.003 A^2 [8] for 100 kV

electrons integrated over an angular range extending to 100 mrad. The total number of elastically scattered electrons per shot from one such large macromolecule (beam larger than molecule) is -

$$N_S = M\left(\frac{\pi r^2}{\pi R^2}\right)\left(\frac{\sigma n}{\pi r^2}\right)S_R^2$$

where n = 4/3 p 450^3 / 15 = 25,446,900 is the number of effective carbon atoms in the macromolecule, R is the radius of the beam, and r = 45nm the radius of the macromolecule. The beam-energy scaling factor S_R, set to unity at 100 kV, is obtained from

$$S = \frac{\lambda}{\sqrt{1 - v^2/c^2}}$$

with S_R^2 = $S_{(3\ MeV)}/S_{(100kV)}$ = 0.4. Two cases are studied, with beam diameters D = 2R = 500 microns (as for the measurements in, [11] and R = r, for the idealized case where the beam could be focused down to the size of a virus (the largest viruses have diameter about 0.45 microns).

For R = 500 microns

$$N_S = 1.04E6\left(\frac{0.003\times 25E6}{\pi\times(250E4)^2}\right)0.4$$

i.e. $$N_S = 0.00159$$

Thus when focusing a 500 micron diameter beam on an approximately 0.1 micron "virus" less than one electron is noticed to have scattered per 100 fs shot. In the idealized case where space-charge would not limit focusing of this beam, i.e. when R = r can be assumed,

$$N_S = 1.04E6\left(\frac{0.003\times 25E6}{\pi\times(450)^2}\right)0.4$$

i.e. $$N_S = 49043$$

which is sufficient for reconstruction of a single 2D projection (in the absence of appreciable background), and for the orientation determination needed to allow merging of multiple shots into a 3D image. For comparison, [14] it is noted that the maximum total number of 1.8 kV X-rays elastically scattered from a single giant Mimi Virus (diameter 0.45 microns, 1.2E6 base pairs in the genome) using the LCLS XFEL (presumably a direct hit) was about 1.7E6 (resolution 32 nm, 70fs pulses, 1E12 photons incident per pulse).

These calculations assume that the pulse duration needed to out-run radiation damage for electron beams is about the same as that for X-ray beams. In fact, at MeV energies, the dominant damage process for electron beams is not ionization, but the much slower ballistic "knock-on" process. The rate of atomic displacements due to this process is

$$R = \frac{N\nu\sigma E_a \varphi}{4E_d}$$

where N is the atomic density, v the reduced mass factor (0.3), σ the atomic displacement cross section (approximately 1 barn), φ the integrated flux, E_a the incident beam energy and E_d the displacement energy, taken to be 280 eV. The number D of irreversible atomic displacements per atom is R/N. For the conditions used above, D ~ 0.1 % diplacements per atom is obtained at 100 fs. Allowing 1% displacements per atom would then increase the number of elastically scattered electrons N_s by an order of magnitude in both the above estimates.

From these calculations, several conclusions may be drawn –

i) As has seen from the above discussion of the STEM case where the incident beam is continuous in time (CW) and focussed down to sub-nanometer dimensions, single-particle imaging at high resolution of single biological particles by coherent diffractive imaging is not possible using CW illumination (even at a "dose" of about 5 electrons per square Angstrom) below the damage threshold. [8]

ii) If, by contrast, radiation damage is avoided by using femtosecond pulses in the diffract-and-destroy mode to out-run the damage, the important challenge (limited by coulomb interactions in the beam) is the formation of a sufficiently small beam, about three orders of magnitude smaller than has currently been demonstrated.

The success of cryo-EM imaging in electron microscopy for imaging proteins relies on several factors, including the linearity of phase contrast (proportional to the scattering factor rather than its square, as for these cross sections), the merging and sorting of many images of similar particles (each unrecognizably noisy), each recorded below the damage threshold, and the absence of the phase problem when dealing with real-space images formed by a magnetic lens.

iii) The niche which coherent MeV-energy single-particle diffraction, if successful, might provide in biology would have several aspects.

Firstly, the ability to reconstruct images of large bioparticles, such as whole cells, for which multiple scattering precludes cryo-em imaging at kilovolt energies. Since all cells are different, merging such data for 3D reconstruction in the diffract-and-destroy mode does not seem possible, unless local regions of identical structure are merged first, then assembled into a three-dimensional typical cell structure.

Heterogeneity in virus structure is a subject of active research. Heterogeneity increases with size in biology, and this limitation on the size and thickness of samples affects X-ray and cryo-electron microscopy in the same way by complicating 3D merging of data. Recently, 2D XFEL snapshots of whole-cells have been reconstructed in hydrated form so that they are alive at the time the femtosecond snapshot is taken. [5] Advanced algorithms have been developed which are capable of sorting 2D diffraction or real-space imaging data by both orientation and a limited number of distinct conformations (see, for example, [15] and references therein).

Other aspects of MeV fast diffractive imaging which distinguish it from cryo-electron microscopy are the rapid diffusion of chemical species into a nanoparticle, which would facilitate the imaging of chemical reactions using a stream of very small particles, the absence of radiation damage, the important ability to study irreversible processes in the diffract-and-destroy mode (rather than being limited to repeating processes which can be studied by stroboscopic methods, where space-charge effects may be avoided by using one electron per pulse [16]), the fact that the particle size is comparable with the absorption length of an optical pump laser in pump-probe studies, and the much greater amount of data obtainable. Systematic calculations, however, need to be performed to identify ideal electron incident energy in the MeV range required to establish

practical feasibility of diffractive imaging using electron pulses (The 3 MeV energy used here was only to demonstrate calculations of detection feasibility under current state of the art).

3 Reciprocity Theorem and Hollow-Cone TEM

In the previous section, the possibility of using electrons and outrunning radiation damage using current electron diffraction instruments has been discussed. Fast electron imaging (as opposed to diffraction) using a pulsed electron beam generated at a photocathode in a transmission electron microscope (TEM) under "full-field" imaging conditions also have been developed by several groups (see [17] for a review). Although the formation of a multi-pixel image within a picosecond time interval using a scanning transmission electron microscope (STEM) is not possible due to the serial readout mode of the instrument, the formation of sub-angstrom resolution STEM images under continuous illumination (CW) [7,18] does establish that high resolution is possible using *incoherent detection* over a large (annular) detector area. A fully coherent field-emission point source is used and focused to form the scanned probe in this transmission geometry, shown in figure 1(a). Here the reciprocity principle is applied to show that a large *incoherent annular source***,** such as a photocathode, can be used in the HCTEM mode (reciprocal to STEM) to produce a fast high resolution TEM image, in which many pixels are detected simultaneously, as shown in figure 1(b). The annular hollow cone illumination mode has successfully been demonstrated [19] in many publications for near-atomic resolution (see [7] for a review). The large incoherent source that can be used for such HCTEM then will produce far more electrons per TEM snapshot image pixel than for current fast TEM designs. This is due to the fact that current designs rely on coherent illumination, where the required collimation of the source greatly reduces image intensity.

3.1 Inferences from Reciprocity Theorem :

Let there be two systems: A and B. Let us define, electric and magnetic fields and currents associated with each of the systems: E_a, B_a, J_a and E_b, B_b, J_b respectively as in Fig. 2(a) below:

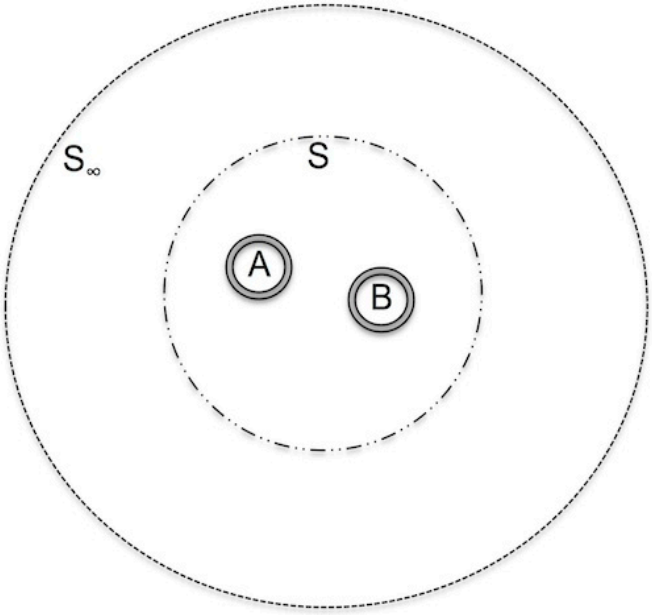


Fig: 2(a): Schematic of the two sources A and B and the surface enclosing them at infinite distance S_∞ and closer S_1.

The reciprocity theorem can be stated as

$$\int_{V_B}\left[\overrightarrow{E_a}\cdot\overrightarrow{J_b}\right]dV=\int_{V_A}\left[\overrightarrow{E_b}\cdot\overrightarrow{J_a}\right]dV \qquad \ldots \qquad (I)$$

Note: Equation (I) is obtained for the case of the enclosed surface being at infinity (S_∞) for the LHS. But reducing the envelope surface from infinity to any finite size (the barest minimum being that which just contains the two systems) will still hold this solution (I) true.

There are certain inferences that can be identified from the above equation:

1. Equation (I) implies that the power experienced by a current J_b at B due to an electric field (E_a) with its source at A is exactly equal to that experienced by a current J_a at A, due to an electric field (E_b) with source at B. In a sense, this is an alternative way of expressing the conservation of energy.
    a. For the case of an electron microscope this is equivalent of saying that the number of electrons reaching the detector (B) in a unit time originating from the source (A) is the exact same as the number of electrons reaching reciprocal detector (A) originating from source (B), everything else in the system remaining the same.
2. This power (at A or B) is the integrated Intensity over the volume of the respective systems, each independent of the other. So this theorem inherently holds good for point as well as finite-sized systems A and B.
    a. For the case of an electron microscope, this is equivalent of saying that the size of the source and detectors need not be points. They can well be finite sized and the reciprocity would hold good as in Fig (6).
    b. Larger the STEM-detector area, lower the reciprocal (hollow cone) TEM-source brightness required to obtain the same amount of signal at the TEM-detector.
    c. The only physical variable that obeys the reciprocity is Power (for same total exposure times - Energy). For the case of a point source/detector, the energy entering/exiting the source/detector is the same as its intensity. Therefore the intensities of the point source-detector system A-B are reciprocally related. However, for a finite source detector, this is not true.

3.2 Schematic Representation of Reciprocity Theorem for HC-TEM:

Below is presented a systematic extension of the reciprocity theorem from a single point source and point detector to the geometry of interest in this research. The extension of argument (looking through the STEM mode, TEM is equivalent) is as follows –

- 1a: Single Point Source, Single Point Detector, On Axis, No Optics
- 1b: Single Point Source, Single Point Detector, On Axis, With Optics
- 2a: Single Point Source, Single Point Detector, Off Axis – position ‘1’, With Optics
- 2b: Single Point Source, Single Point Detector, Off Axis – position ‘2’, With Optics
- 3a: Single Point Source, Dual Point Detectors, Off Axis – position ‘1’ and ‘2’, With Optics
- 3a: Single Point Source, Multiple Point Detector, Off Axis - on an Annulus, With Optics

This argument is shown with incident and detected photon counts per single point source/detector.

- 4a: Single Point Source, Single Point Detector, 1-electron in, 1-electron out
- 4b: Single Point Source, Off Axis Single Point Detector, 1-electron in, 1-electron out
- 5: Single Point Source, Off Axis Dual Point Detectors, 2-electrons in, 2-electron out
- 6: Extended Source, Extended Detector, With Optics

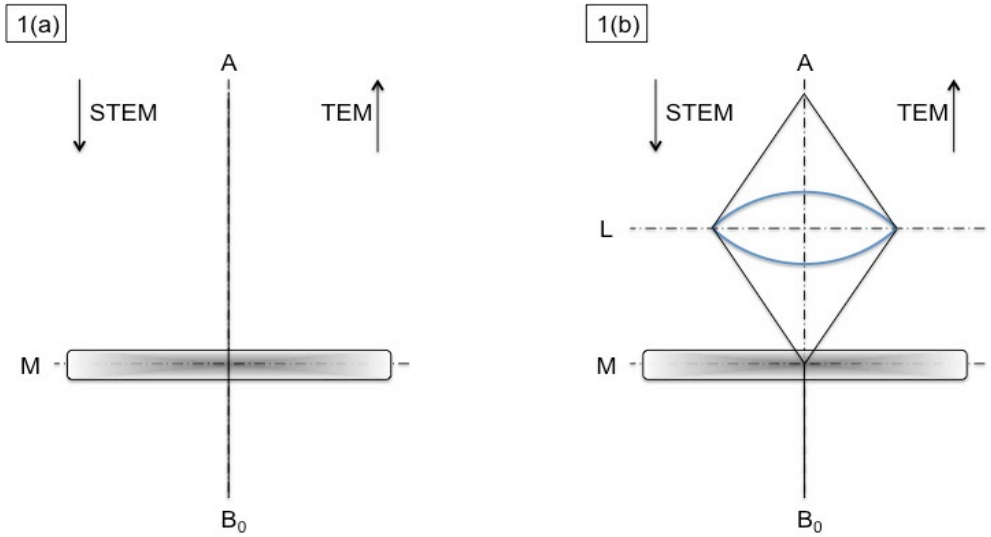


Fig. 2.1: Reciprocal equivalence for STEM and TEM shown (a) without and (b) with lenses.

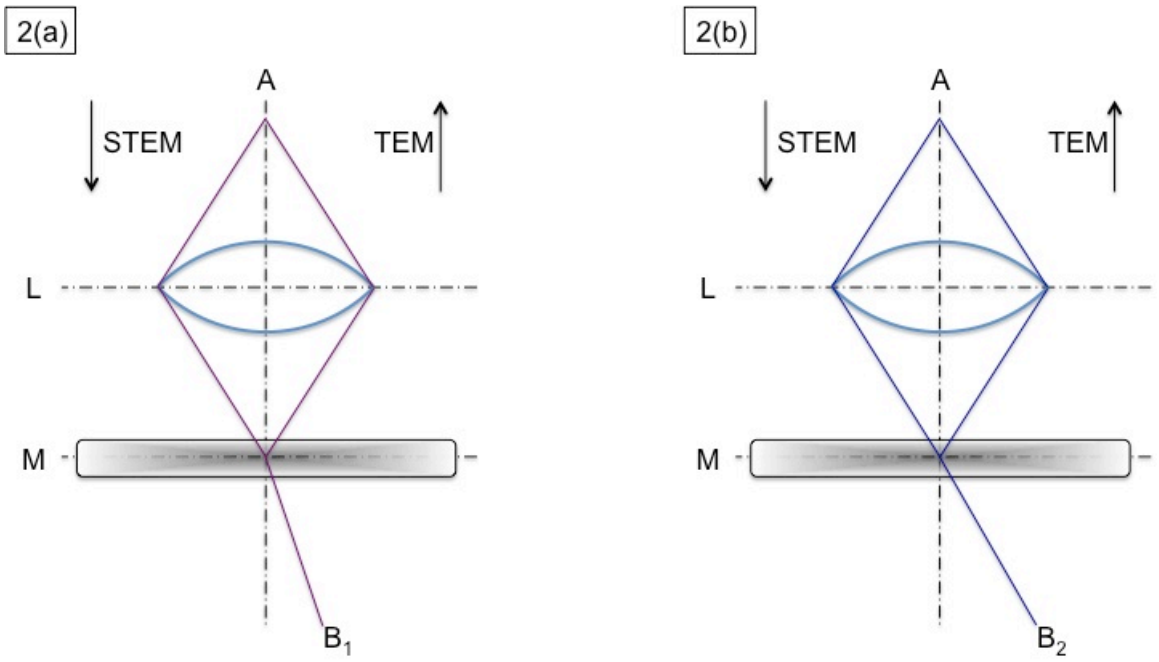


Fig. 2.2: Reciprocal equivalence for STEM and TEM shown (a) at one detector point off-axis B_1 (b) at a second off-axis point B_2.

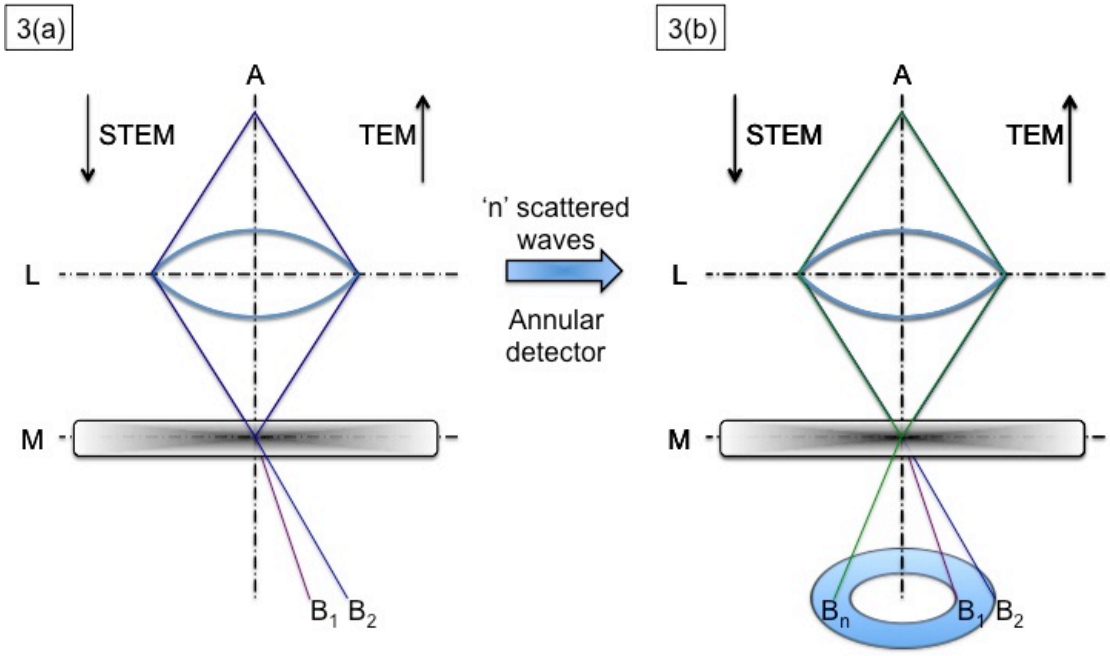


Fig. 2.3: Reciprocal equivalence for STEM and TEM shown (a) for two different detection points B_1 and B_2 (b) for n different detection points B_1, B_2 … B_n into an annular region.

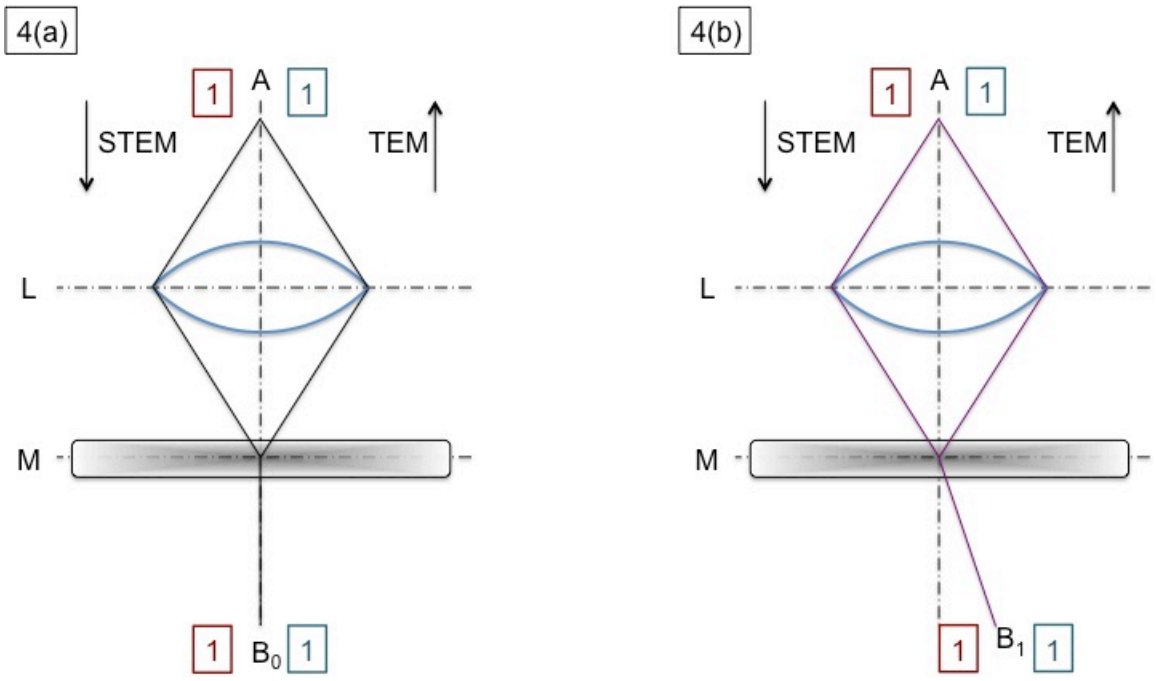


Fig. 2.4: Reciprocal equivalence for STEM and TEM with input and output number of electrons shown (a) for on-axis detection point B_o (b) for off-axis detection points B_1.

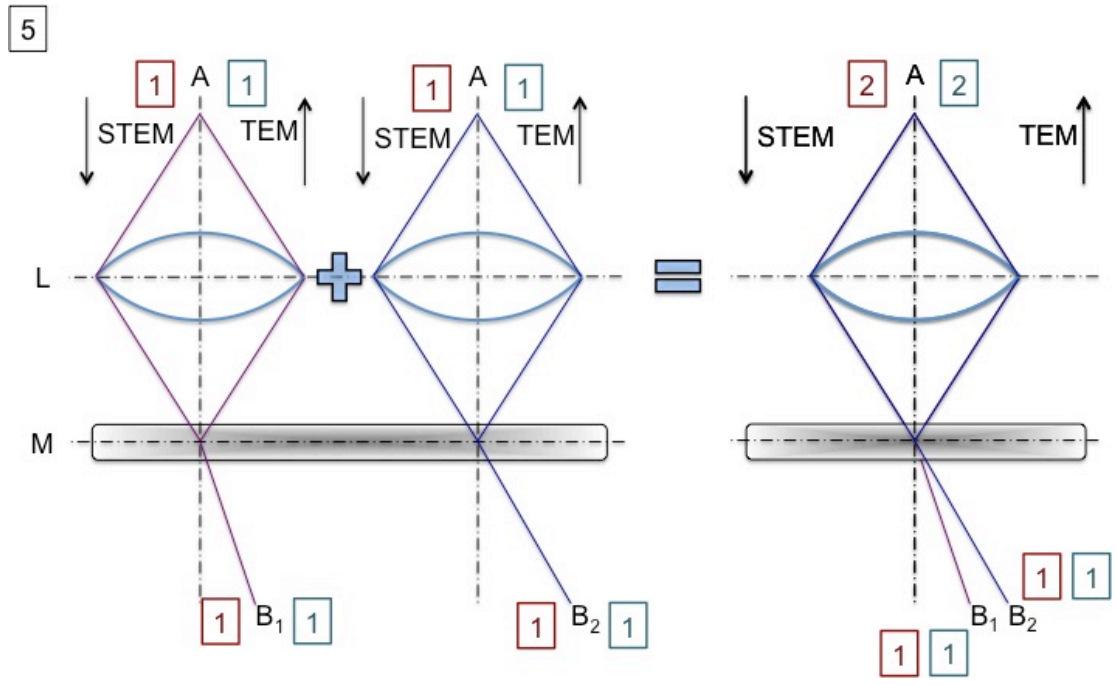


Fig. 2.5: Reciprocal equivalence for STEM and TEM with input and output number of electrons shown for two different detection points in a serial (LEFT) and parallel (RIGHT) acquisition.

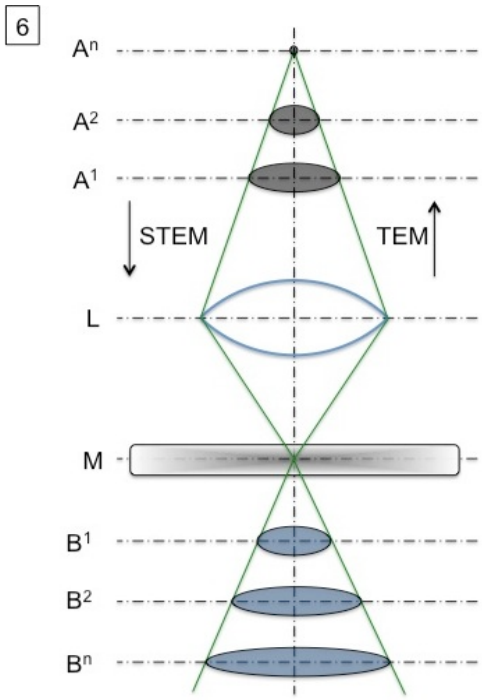


Fig. 2.6: Reciprocal equivalence for STEM and TEM showing the various different positions of detectors and sources that are possible while still maintaining the reciprocity.

3.3 Hollow Cone-TEM for Full Field Femtosecond Imaging

Figure 2.7 illustrates the reciprocity principle, first described by Helmholtz [20] and applied to electron microscope imaging in the STEM and TEM modes by Cowley [21] and later Zeitler and Thomson, [22] and to elastic scattering (for complex amplitudes) and inelastic scattering (for intensities) in the presence of multiple scattering, by Pogany and Turner. [23]

In these papers it is shown that, under reciprocal aperture conditions, the dark-field image produced in the STEM arrangement at left (dark-field STEM, or ADF STEM, not high-angle Annular Dark Field STEM, which uses quasi-elastic scattering) is identical to that produced in the HCTEM arrangement at right (also a dark-field image), apart from scaling factors. The reciprocity theorem, based on time-reversal in the wave equation, states that, for elastic scattering in any optical system, the complex amplitude at a point detector D due to a point source S is equal to the amplitude which would be detected at S if the same source were moved to D.

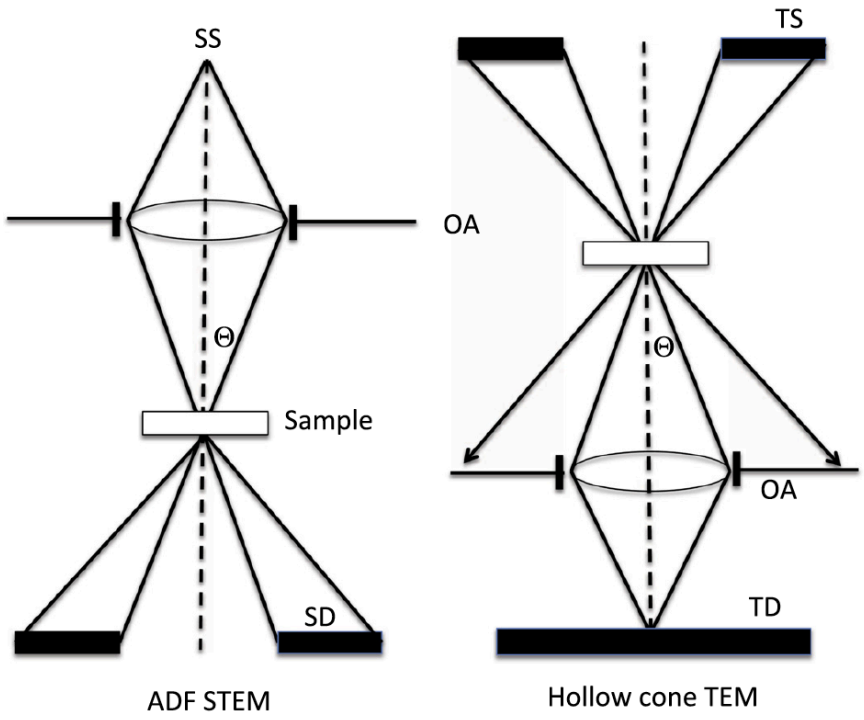


Fig. 2.7: Reciprocal ray diagrams for STEM (left, 7(a)) and TEM (right, 7(b)). Under reciprocal aperturing conditions (identical objective lenses, source brightness etc.) the image produced by scanning the beam in STEM is identical to that produced simultaneously at every pixel in the TEM arrangement. Electron flow is down the page in both cases. STEM source SS, TEM annular source TS, detectors SD and TD. All sources and detectors are considered ideally incoherent. Objective aperture OA

This interchange of point source and point detector may be combined with the linearity property of the wave equation to apply the theorem to the case of extended incoherent sources and detectors, as shown in figure 2.7. In particular, the theorem can be applied to relate the image pixels successively read out in STEM to the many image pixels recorded simultaneously in full-field TEM. This can best be understood for the STEM and TEM case by noting that the STEM image could be obtained by scanning the sample (rather than the beam). Each of these sample positions describes exactly the reciprocal imaging conditions occurring simultaneously at each TEM full-field pixel. The images in figure 2.7 are therefore only identical in form if the pixel area A_{TEM} in the HCTEM image is equal (after allowance for magnification factors) to the source size A_{STEM} in the STEM case. Similarly, the annular STEM detector must subtend the same angles as the annular TEM source in HCTEM (reciprocal aperturing conditions).

In addition to relating the form of the STEM and TEM images, the theorem also gives a useful relationship between the intensity in these two images, for sources of a given brightness (defined as electron beam flux per unit solid angle per unit area of source size). A comparison of recording times for images of a given number of pixels clearly depends on the scan speed of the STEM. Consider first the idealized case where the brightness of the STEM field-emitter was equal to that of the annular emitter in HCTEM, which subtends the same angles as the STEM annular detector. Then, *under reciprocal aperturing conditions, the number of electrons arriving per unit time at each pixel of the TEM image would be the same as that arriving over the whole annular STEM detector, for one STEM probe position* (and hence contributing to one dark-field STEM image pixel).

While the annular hollow cone mode has successfully been demonstrated in many publications showing near-atomic resolution, the formation of an annular photocathode of high brightness creates many difficulties. Shaping the photocathode drive-laser generating the electrons at the cathode is one possibility to obtain the hollow cone source. Alternatively several groups have experimented with arrays of photofield emitters, each of which has a brightness similar to a field-emission tip used in STEM, but whose density per unit area is limited. [24,25] However, when comparing the STEM counts at any one scan position with any HCTEM pixel, the HCTEM count rate will be reduced by the ratio of photocathode to field emitter brightness.

Therefore it is important to evaluate the brightness ratio between field emitter and photoemitter sources to understand the limitation of this technique. The field-emission sources used in STEM produce a few hundred electrons per picosecond from a source whose diameter is typically 1.5 nm, while typical modern fast electron photoguns can generate $>10^6$ electrons per 100 fs, from an extended photocathode of tens of micron dimensions. [11,26] The transverse brightness ratio is therefore $> 10^6/10^2 \times (15\text{ nm} / 15\text{ mm})^2 \sim 10^{-4}$. The *maximum* brightness of a flat photocathode is achieved when one uses the maximum intensity of the laser allowed by space charge effects. At this limit, brightness depends on the applied electric field E_0 (which is why RF photoinjectors with fields as high as 100 MV/m have an advantage) which determines the source surface charge density (which must be kept below e_0E_0). Photoemission brightness is also

inversely proportional to the excess kinetic energy of the electrons (the difference between laser photon energy and cathode work-function). [27] By carefully tuning the laser frequency and the (flat) photocathode material properties it is possible to obtain excess kinetic energies lower than 50 meV, which would improve the above ratio by an order of magnitude or so.

The formation of a hollow cone illumination beam (and the sharpness of the image of this annular ring in the back-focal plane of the lens OA in figure 1(b)) will thus be affected by space-charge effects (which decline at higher energy as electric and magnetic field effects tend to cancel) and by electron-optical design considerations. In particular, electrons extracted from an annular photoemitter tend to "fill-in" to the central hole, due to space charge effects as they propagate from the metal surface to the sample. This effect can be minimized through the use of suitable condenser optics, and a suitable central blocking aperture. Preliminary simulations with such a system, using an RF photoinjector outputting 5 MeV electrons, [28] show the possibility of preserving up to 90 % of the initial brightness in the final illuminating cone (10-60 mrad angular aperture) converging on the sample.

4 Conclusions

The possibility of diffraction by electrons before destruction to obtain snapshot diffraction without radiation damage from single biological particles (such as a viruses) has been discussed. The study shows that -

(i) A stationary cold field-emission STEM probe scatters less than one electron from a bioparticle prior to either the damage-dose threshold or the 100 fs pulse duration needed to outrun damage.

(ii) If a partially coherent high-energy beam from a larger photocathode can be focussed to the size of a large virus (about 0.1 micron), a sufficient number of electrons are elastically scattered within 100 fs to allow orientation determination of the bioparticle and image reconstruction from many shots, without damage. This beam size is about three orders of magnitude smaller than obtained in current practice, which is limited by space-charge effects even at beam energies of several MeV.

(iv) The use of a hollow-cone mode (HCTEM) for fast TEM full-field imaging is proposed, since it allows use of a large incoherent emitter (such as a photocathode), yet can in principle provide atomic resolution imaging. The reciprocity theorem indicates that, under reciprocal aperturing conditions, the electron count at every HCTEM image pixel, recorded simultaneously, is equal to that recorded at one equivalent STEM pixel (for one position of the scanned beam), but reduced by the ratio of the brightness of the HCTEM photocathode to that of the STEM "point" source. However space-charge effects, which cause the annular HCTEM source to fill-in the central hole as it propagates, must be controlled, perhaps through the use of higher accelerating voltage and optimized electron-optical design. [29] The HCTEM mode thus appears ideal for time-resolved imaging using electrons.

5 References

1. Chapman, H. N. *et al.* Femtosecond diffractive imaging with a soft-X-ray free-electron laser. *Nat. Phys.* **2,** 839–843 (2006).

2. Spence, J. C. H., Weierstall, U. & Chapman, H. N. X-ray lasers for structural and dynamic biology. *Rep. Prog. Phys.* **75,** 102601 (2012).

3. Zhang, J. *et al.* Agonist-bound structure of the human P2Y12 receptor. *Nature* **509,** 119–22 (2014).

4. Tenboer, J. *et al.* Time-resolved serial crystallography captures high-resolution intermediates of photoactive yellow protein. *Science* **346,** 1242–6 (2014).

5. van der Schot, G. *et al.* Imaging single cells in a beam of live cyanobacteria with an X-ray laser. *Nat Commun* **6,** (2015).

6. Henderson, R. The potential and limitations of neutrons, electrons and X-rays for atomic resolution microscopy of unstained biological molecules. *Q. Rev. Biophys.* **28,** 171–193 (1995).

7. Spence, J. C. H. *High-resolution electron microscopy*. (OUP Oxford, 2013).

8. Rez, P. Comparison of phase contrast transmission electron microscopy with optimized scanning transmission annular dark field imaging for protein imaging. *Ultramicroscopy* **96,** 117–24 (2003).

9. Subramanian, G., Basu, S., Liu, H., Zuo, J. M. & Spence, J. C. H. Solving protein nanocrystals by cryo-EM diffraction: Multiple scattering artifacts. *Ultramicroscopy* **148,** 87–93 (2015).

10. Barty, A. *et al.* Self-terminating diffraction gates femtosecond X-ray nanocrystallography measurements. *Nat. Photonics* **6,** 35–40 (2012).

11. Muro'oka, Y. *et al.* Transmission-electron diffraction by MeV electron pulses. *Appl. Phys. Lett.* **98,** 251903 (2011).

12. Musumeci, P., Moody, J. T. & Scoby, C. M. Relativistic electron diffraction at the UCLA Pegasus photoinjector laboratory. *Ultramicroscopy* **108,** 1450–1453 (2008).

13. Huldt, G., Szoke, A. & Hajdu, J. Diffraction imaging of single particles and biomolecules. *J. Struct. Biol.* **144,** 219–227 (2003).

14. Seibert, M. M. *et al.* Single mimivirus particles intercepted and imaged with an X-ray laser. *Nature* **470,** 78–81 (2011).

15. Dashti, A. *et al.* Trajectories of the ribosome as a Brownian nanomachine. *Proc. Natl. Acad. Sci. U. S. A.* **111,** 17492–7 (2014).

16. Lahme, S., Kealhofer, C., Krausz, F. & Baum, P. Femtosecond single-electron diffraction. *Struct. Dyn.* **1,** (2014).

17. LaGrange, T. *et al.* Nanosecond time-resolved investigations using the in situ of dynamic transmission electron microscope (DTEM). *Ultramicroscopy* **108,** 1441–1449 (2008).

18. Pennycook, S. J. & Nellist, P. D. *Scanning Transmission Electron Microscopy: Imaging and Analysis*. (Springer Science & Business Media, 2011).

19. Saxton, W. O. TEM observations using bright field hollow cone illumination. *Optik (Stuttg).* **49,** 505–510 (1978).

20. Helmholtz, H. *Handbuch der physiologischen*. *Optik* **9,** (1867).

21. Cowley, J. M. Image Contrast in a Transmission Scanning Electron Microscope. *Appl. Phys. Lett.* **15,** 58 (1969).

22. Zeitler, E. & Thomson, M. G. R. Scanning transmission electron microscopy. *Optik (Stuttg).* **31,** 258 (1970).

23. Pogany, a. P. & Turner, P. S. Reciprocity in electron diffraction and microscopy. *Acta Crystallogr. Sect. A Cryst. Physics, Diffraction, Theor. Gen. Crystallogr.* **24,** 103–109 (1968).

24. Spence, J. C. H., Vecchione, T. & Weierstall, U. A coherent photofield electron source for fast diffractive and point-projection imaging. *Philos. Mag.* **90,** 4691–4702 (2010).

25. Hernandez Garcia, C. & Brau, C. A. Pulsed photoelectric field emission from needle cathodes. in *Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **483,** 273–276 (2002).

26. Musumeci, P. *et al.* Capturing ultrafast structural evolutions with a single pulse of MeV electrons: Radio frequency streak camera based electron diffraction. *J. Appl. Phys.* **108,** 1–6 (2010).

27. Bazarov, I. V., Dunham, B. M. & Sinclair, C. K. Maximum achievable beam brightness from photoinjectors. *Phys. Rev. Lett.* **102,** (2009).

28. Li, R. K. & Musumeci, P. Single-Shot MeV Transmission Electron Microscopy with Picosecond Temporal Resolution. *Phys. Rev. Appl.* **2,** 024003 (2014).

29. Chatelain, R. P. *et al.* Space-charge effects in ultrafast electron diffraction patterns from single crystals. *Ultramicroscopy* **116,** 86–94 (2012).

1 Introduction

1.1 Spectroscopy – Inelastic Scattering

When electrons (of an electron beam) interact with the atoms in a sample specimen, they interact through the Coulombic forces. The nature of interaction can be of two types. The first kind is when some of the electrons are scattered such that the kinetic energy and momentum are conserved and the only change is the direction of the momentum of these incident electrons. This is elastic scattering (Fig. 3.1 a) and which accounts to the diffraction process (which has been described in the previous chapter.

The other kind of interaction is when a part of the incident electron energy is transferred to the atoms/atomic electrons of the specimen (Fig. 3.1 b and c). In such cases, the total energy is conserved along with the momentum.

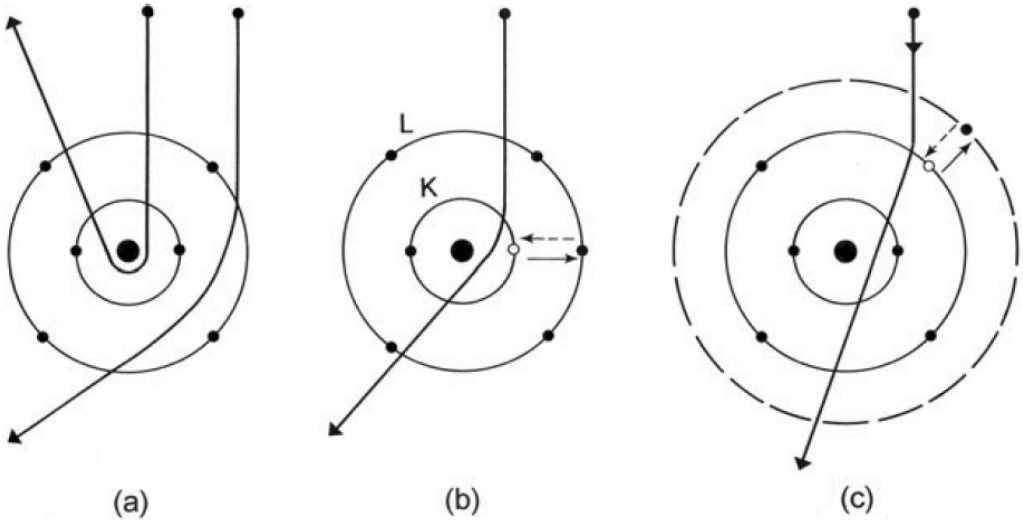


Fig. 3.1: Schematic of electron scattering off an atom. (a) is the representative diagram for elastic scattering off a nucleus. (b) is a diagram showing core (inner-shell) atomic-electron excitation and de-excitation (depicted using broken arrows). (c) is a similar representation for the interaction with outer-shell electrons.

Some of the processes are understood by single electron excitations while some others are through multi-electron excitations. Such interaction requires the presence of electron states that are unoccupied, which are typically above the Fermi energy.

The ground state energy of inner-shell electrons is typically few hundreds to few thousands of eV apart from Fermi Energy (band gap) of a material. Since the unoccupied states lie only above the Fermi energy, an inner shell electron can make a transition to the higher unoccupied state only by absorbing energy from the incident electrons, which is at least equal, or beyond the difference of the occupied ground initial electron state and the unoccupied final state. In an electron microscope, such inner-shell inelastic scattering angle is typically < 1 mrad. At the end of such an event, the atom is left ionized (excited) state with a vacant state in the inner shell termed as a 'core hole'. Since this is not the neutral lowest energy state of the atom, it will relax through a de-excitation process with an outer/inner shell electron undergoing a downward transition to fill the vacant core hole and thereby releasing the excess energy as an X-ray (fluorescence) or another atomic electron (Auger emission).

The outer-shell inelastic scattering however, typically involves more than one electron (from many atoms) in the sample (especially with solid state samples). Collective excitations of the type known as plasma resonance (valence electrons) travel in the form of a longitudinal wave. The essential requirement for such excitations is that the participating electrons interact with each other and share their energy.

1.2 The Energy-Loss Spectrum

When fast moving electrons interact with a material, the Coulombic (electrostatic) forces due to the atoms present in the material scatter the incident electrons. Such scattering causes the incident electrons to change the direction of their momenta and, many a time, also lose energy. By attaching a spectrometer to the side facing the transmitted electron beam (Fig. 3.2), that can record and resolve the energy of the electron beam impinging on it, there is the potential of studying various different natures of interactions of the probe (electron) with electrostatic forces within the material (atom and atomic-electrons). This process is described as Electron Energy

Loss Spectroscopy (EELS). It is so named because it studies the energy lost by the incident electrons in interacting with matter as it comes out as the transmitted beam.

Incident electrons as they travel through a sample can go through various experiences. A majority of the electrons do no interact with any of the atoms and thus pass through unscattered. This is termed as the 'Zero-loss' regime. Of those electrons that interact with the material, the two major contributors to the scattering of the incident electrons are (a) the nucleus and (b) atomic-electrons. The nucleus is a dense bundle of charge and hence the deflection off a nucleus is generally very large. In other words, these are typically responsible for very high scattering angles of the incident beam. This is christened as 'Elastic Scattering'. This kind of scattering leads to the formation of a diffraction pattern (periodic arrangement of the nuclei is reflected by the interference of the electron waves that are scattered off the nuclei). In addition to this, the lattice vibrations (collective mechanical motion of the atom centers - called as phonons) also scatter the incident electrons. But such interactions result in very small energy losses (one-hundredth of an eV or lesser). These kind of electron-phonon interactions have been difficult to study using EELS for many reasons. Recent developments in instrumentation, though, may have broken through this barrier. This will be discussed in a forthcoming section.

Albeit being a dense bundle of charge, the nucleus is small in comparison to the vast regions of space present around it. (Typical nuclear radius is 10^{-5} A° while typical atomic radius is 1 A° and interatomic distances in a crystal are generally about 2-3A°). Surrounding the nucleus are the atomic-electrons and the incident electrons can interact with these electrons too. These sorts of interactions typically involve a considerable energy exchange and hence are termed 'Inelastic Scattering'. As mentioned in the previous section, all atomic-electrons do not behave similar while interacting with the incident beam. In a typical electron-electron interaction, the atomic-electron absorbs energy from the incident beam to excite and make a transition to the unoccupied electronic states above the Fermi level. To compensate for the 'hole' now created in such an excitation, a de-excitation process soon follows in which an electron from a higher electronic state than that of the initial state of the excited electron, drops down.

Excess energy (if any) will thus be liberated as x-rays or the kinetic energy of another atomic electron (Auger emission or ionized-electron that backscatters from the surrounding atoms). Inner-shell electrons have ground-state energies that are a few hundred to thousands of eV lower than the Fermi level. The energy loss recorded on the spectrometer is few 100s of eVs and this regime is termed as the 'Core-Loss' regime.

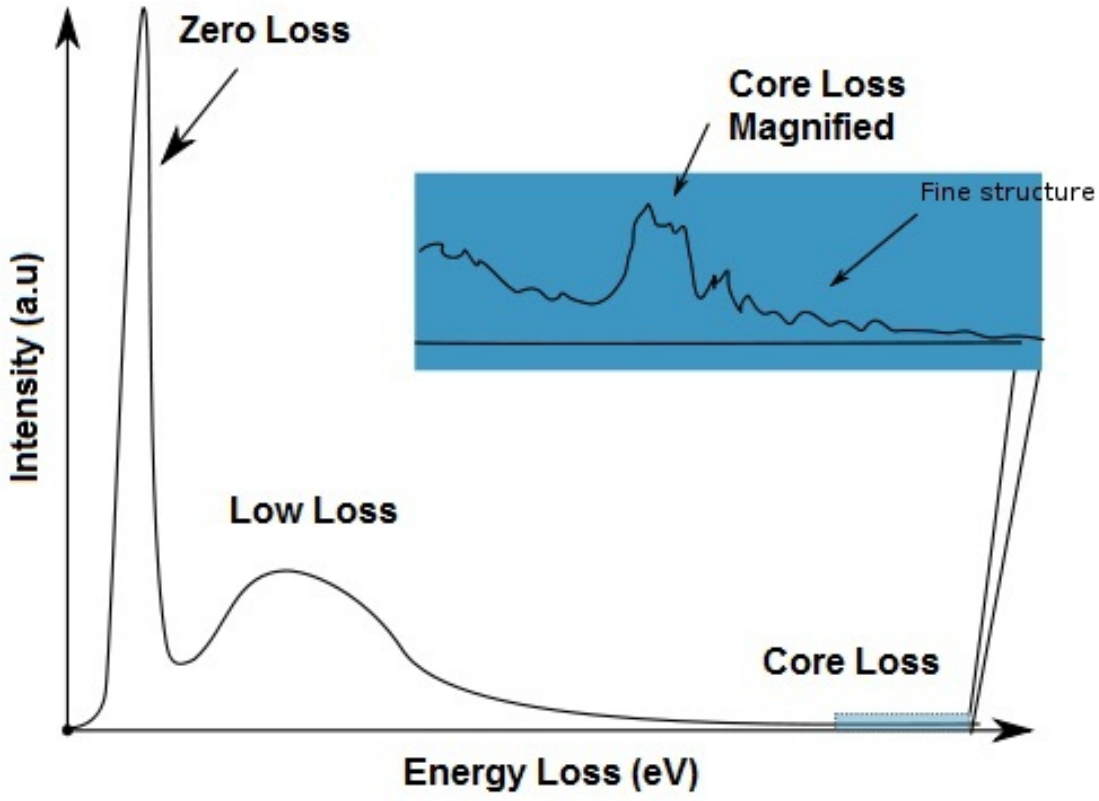


Fig. 3.2: Representative image of an EEL spectrum. The different regions are marked appropriately and the core-loss region magnified in the inset. Since the EEL spectrum is cumulative the background of the spectrum increases with energy loss.

Inner-shell excitation gives rise to relatively low scattered intensity (due to the low cross section) and therefore has a mean free path that is long compared to the specimen thickness. The probability that a fast electron produces more than one inner-shell excitation is therefore negligible.

Outer-shell electrons also undergo similar single-electron transitions as long as they have enough energy (obtained from the incident beam) to excite across the energy gap (for semiconductors and insulators) and other higher electronic states (for metals). This also includes sharing excess energies amongst collectively excited electrons as mentioned before. Both these processes are found to occur at energy losses at around 5-30 eV. This is termed as the 'Low Loss' regime.

1.3 Bethe Theory: (this section is inspired from Egerton, R. F. *Plenum Press, New York* 229, no. 241 (1986): 357-361).

To explain the details of inelastic scattering of incident electrons by an atom, let us assume an initial state wave function Ψ_0 and a final state wave function Ψ_f. The differential cross-section for such a transition can be written as –

$$\frac{d\sigma_n}{d\Omega} = \left(\frac{m_0}{2\pi\hbar^2}\right)^2 \left(\frac{k_1}{k_2}\right) \left| \int V(r)\psi_0\psi_n^* \exp(iq\cdot r)\, d\tau \right|^2 \qquad (1)$$

where $q = k_0 - k_1$ is the momentum transferred to the atom, r the coordinate of the incident electron, V(r) the interaction potential and $d\tau$ the volume element within the atom. The interaction potential is a combination of the coulomb attraction by the nucleus (first term) and the incident electron – atomic electron repulsion (second term) and can be expressed as –

$$V(r) = \left[\frac{Ze^2}{4\pi\varepsilon_0 r}\right] - \left[\left(\frac{1}{4\pi\varepsilon_0}\right)\sum_{j=1}^{Z}\frac{e^2}{|r-r_j|}\right] \qquad (2)$$

Selection rules for transitions dictate that the initial state and final state wave functions are orthogonal. This implies that the nuclear contribution to the integral is non-existent. Thus,

inelastic transitions are characterized by strict electronic (atomic-electrons) contributions to the interaction potential, while both atomic electrons and the nucleus characterize elastic scattering. A point worth noting here is that, since in electron – electron interactions, mass of the incident and atomic electrons are very similar, there is a considerable scope for energy transfer.

Combining the above two equations, the differential cross-section can then be written as

$$\frac{d\sigma_n}{d\Omega} = \left(\frac{4\gamma^2}{a_0^2 q^4}\right)\left(\frac{k_1}{k_0}\right)\left[\left\langle \psi_n \left| \sum_j e^{(iq\cdot r_j)} \right| \psi_0 \right\rangle\right]^2 \qquad (3)$$

where the last term is the 'inelastic form factor' or the 'dynamical structure factor' and is defined as the square of the transition matrix element defined within the bra-ket notion. Here, k_1/k_0 can be assumed as unity for all cases when the transition energy is much smaller than the incident energy (typically true for incident energies greater than 100 keV).

Expanding the last term (dynamical structure factor) gives the dipole approximation –

$$\left\langle \psi_n \left| \sum_j e^{(iq\cdot r_j)} \right| \psi_0 \right\rangle$$

$$= \left\langle \psi_n \left| \sum_j \left\{1 - 2\pi i q\cdot r_j + \ldots\right\} \right| \psi_0 \right\rangle$$

$$= \left\langle \psi_n \left| \sum_j \{1\} \right| \psi_0 \right\rangle - \left\langle \psi_n \left| \sum_j \{2\pi i q\cdot r_j\} \right| \psi_0 \right\rangle$$

$$= \left\langle \psi_n \left| \sum_j 2\pi i \left(q\cdot r_j\right) \right| \psi_0 \right\rangle$$

Thus equation (3) can be reduced to equation (4) to give the differential cross-section for inelastic electron scattering under the dipole approximation as –

$$\frac{d\sigma_n}{d\Omega} = \left(\frac{4\gamma^2}{a_0^2 q^4}\right)\left(\frac{k_1}{k_0}\right)\left[\left\langle \psi_n \left| \sum_j 2\pi i\left(q \cdot r_j\right) \right| \psi_0 \right\rangle\right]^2 \qquad (4)$$

However, equation (4) is valid only under certain conditions, which can be explained using fig. 3.3 as shown below.

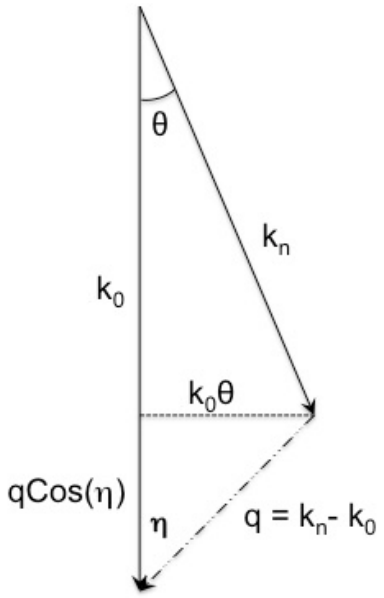


Fig. 3.3: Schematic of the conservation of momentum for inelastic scattering of electrons by atomic electrons.

Using the above geometric sketch, it is possible to solve for the energy and momentum conservation as shown below:

- Conservation of Momentum –

$$k_n^2 = k_0^2 + q^2 - 2k_0 qCos\left(\eta\right) \qquad (i)$$

- Conservation of Energy –

$$\Delta E = \frac{h^2}{2m}\left(k_0^2 - k_n^2\right) \qquad (ii)$$

Combining equations (i) and (ii) –

$$\Delta E = \left(\frac{h^2 k_0 q}{m}\right) Cos(\eta) \qquad \text{(iii)}$$

From the inelastic scattering geometry the following relation can be presented –

$$q^2 = \{k_0\theta\}^2 + \{qCos(\eta)\}^2 = k_0^2\{\theta^2 + \theta_E^2\} \qquad \text{(iv)}$$

where the characteristic inelastic scattering angle is given as -

$$\theta_E = \left(\frac{\Delta E}{\gamma m v^2}\right) \qquad \text{(v)}$$

Substituting eqn. (iv) in eqn. (4) for 'q' the momentum transfer, the differential cross-section under the dipole approximation is modified as –

$$\frac{d\sigma_{0,n}}{d\Omega} = \frac{\lambda^2}{\pi^2 a^2}\left[\left\langle \psi_n \left| 2\pi i (q \cdot r) \right| \psi_0 \right\rangle\right]^2 \left(\frac{1}{\theta^2 + \theta_E^2}\right) \qquad \text{(5)}$$

The characteristic angle for inelastic angle θ_E is typically 0.1 mrad for ΔE = 20 eV at an incident energy of E_0 = 100 keV for inner shell excitations and about 0.1 rad for outer-shell (valence) excitations. At such low scattering angles, the main angular dependence in eqn. (5) comes from the 'Lorentzian' type last term $(\theta^2 + \theta_E^2)^{-1}$. This Lorentzian term gives the shape to the energy loss spectrum. However, this Lorentzian function tends to have long tails and hence can contribute to the background in an electron diffraction pattern. Additionally, the above equation when expressed as a double-differential inelastic scattering cross section, adding a differential with energy, is equivalent to the absorption spectra obtained from light/X-ray excitations.

1.4 Dielectric Formulation

The equations derived in the previous section are readily applicable to single atoms, gases and even inner shell excitations of compounds and solids. However, in dealing with outer-shell (valence) electrons, their wave functions are more complex because they involve bonding effects. Moreover, as discussed earlier, this is also the regime where collective excitations involving many atoms become significant. In such case, an alternative formalism is usually utilized that describes the interaction of incident electrons with a sample in terms of the dielectric response function. This is especially significant because the same response function describes the interaction of photons with a solid, and hence this formalism permits the direct comparison of EELS data with optical (UV-Vis) absorption data.

Assume the incident electron post interaction with sample is scattered and this inelastically scattered electron moves with a velocity 'v' and coordinate 'r' along the 'z' direction. Assuming this is a point charge that generates a spatial and time-dependent electrostatic potential that can be described as –

$$\rho(r,t) = -e\delta(r - \upsilon t) \tag{vi}$$

Fourier transforming this relation in space and time -

$$\rho(k,w) = -e\delta(k\upsilon - \nu) \tag{vii}$$

Calculating the electric field acting on this point charge –

$$\varepsilon\{\Delta\phi(r,t)\} = -\{\rho(r,t)\} \tag{viii}$$

Fourier transforming this again –

$$4\pi^2 k^2 \{\varepsilon(k,\omega)\}\{\Phi(k,w)\} = \{\rho(k,\omega)\} \tag{ix}$$

The electric field corresponding to this field –

$$E(k,\omega) = -2\pi ik\Phi(k,\omega) \tag{x}$$

The stopping power (dE/dz) is the backward force on the inelastically scattered electron in its direction of motion, which is also energy loss per unit path in the medium and is given as –

$$-\left(\frac{dW}{dz}\right) = -\left(\frac{eE\upsilon}{\upsilon_z}\right) = e^2\left(\frac{\upsilon}{\upsilon_z}\right)\int\left\{\frac{ik\delta(k\upsilon-\nu)}{2\pi k^2\varepsilon(k,\omega)}\right\}e^{-2\pi i(k\upsilon-\nu)t}d^3kd\nu \tag{xi}$$

Since the electron moves in the z-direction, integrating over k_z of the previous equation –

$$-\left(\frac{dW}{dz}\right) = \left(\frac{+ie^2}{2\pi\upsilon_z^2}\right)\int\left\{\frac{\nu(d\nu)(d^2k_\perp)}{k^2\varepsilon(k_\perp,\omega)}\right\} \tag{xii}$$

Since energy loss is a real quantity, the real part of the above equation (xii) is retrieved as –

$$-\left(\frac{dW}{dz}\right) = \left(\frac{e^2}{2\pi\upsilon_z^2}\right)\int\left\{\mathrm{Im}\left(\frac{-1}{\varepsilon(k,\omega)}\right)\right\}\left\{\frac{\nu(d\nu)(d^2k_\perp)}{k^2}\right\} \tag{xiii}$$

The imaginary part of complex dielectric function is known as the energy-loss function and this provides complete description of the response of the medium through which the electron travels. The energy loss per unit path can also be related to the double differential cross-section (per atom) for inelastic scattering as –

$$\left(\frac{dE}{dz}\right) = \iint(n_aE)\left(\frac{d^2\sigma}{d\Omega dE}\right)d\Omega\,dE \tag{xiv}$$

For small scattering angles, equating (xiii) and (xiv), this eventually reduces to –

$$\left(\frac{d^2\sigma}{d\Omega dE}\right) \approx \left(\frac{1}{\pi^2 a_0 m_0 v^2 n_a}\right)\left\{\mathrm{Im}\left(\frac{-1}{\varepsilon(k,\omega)}\right)\right\}\left\{\frac{1}{\theta^2+\theta_E^2}\right\} \qquad (6)$$

where θ_E is the characteristic angle as explained before. Equation (6) contains the same Lorentzian angular dependence and the v^{-2} factor as the Bethe equation in equation (5).

In the small-angle dipole region, $\varepsilon(k,\omega)$ varies very little with 'k' and hence can be replaced by the optical dielectric function (relative permittivity of the specimen at an angular frequency 'ω') $\varepsilon(0,\omega)$. An energy-loss spectrum that has been recorded using a small collection angle can be compared directly to optical data.

The underlying idea of the dielectric theory can be understood in the following manner. The optical properties of a material can be described either by a complex refractive index ($n + i\kappa$) where 'κ' is the absorption coefficient, or by a complex permittivity ($\varepsilon = \varepsilon_1 + i\varepsilon_2 = (n + i\kappa)^2$). At small energy losses, *Im(-1/ε) is proportional to both ε_1 and ε_2*, while at large energy loss, ΔE, ε_2 is small and $\varepsilon_1 \approx 1$ so that "*Im(-1/ε) becomes proportional to ε_2*".

1.5 Extraction of the Energy Loss Function

The method of logarithmic deconvolution described in the work by Johnson & Spence (1974) for removal of multiple energy-loss processes also takes full account of removal of the elastic peak background from EELS spectra. This is important because of the broad base of the 'Zero Loss Peak' which is the strongest peak containing the non-scattered majority of the incident probe. To do so it is necessary to record a separate elastic peak with the sample removed from the beam. In this way, artifacts in the beam energy distribution due to aberration correctors can be removed by deconvolution. According to the dielectric formulation of the energy-loss problem

(Daniels et al., 1970; Egerton, 2011), the EELS intensity is related to the energy-loss function by the following expression, for small scattering angles (very similar to eqn. 6):

$$S(E) = \left(\frac{2I_0 t}{\pi a_0 m_0 v^2} \right) \left\{ \mathrm{Im}\left(\frac{-1}{\varepsilon(E)} \right) \right\} \ln\left\{ 1 + \left(\frac{\beta}{\theta_E^2} \right)^2 \right\} \qquad (7)$$

where I_0 is the zero-loss intensity, a_0 the Bohr radius, m_0 the rest mass of electron, and v the speed of the electron. An angular integration over a small collection semi-angle β is assumed, which has the value of 20 mrad, θ_E is the characteristic scattering angle.

Eqn. (1) immediately gives the energy-loss function from EELS data, however, the prefactor is undetermined if I_0 is unknown.

1.5.1 Kramers Kronig Analysis:

Application of the causality condition to this system leads to the Kramers–Kronig relationship between the real and imaginary parts of linear response functions:

$$\mathrm{Re}\left[\frac{1}{\varepsilon(E)} \right] - 1 = -\frac{2}{\pi} P \int_0^\infty \mathrm{Im}\left[\frac{-1}{\varepsilon(E)} \right] \frac{E'}{E'^2 - E^2} dE' \qquad (8)$$

The integrand has a divergence around E' = E that makes use of the Cauchy principal value necessary. There are several ways to avoid this divergence numerically:

Approach 1: Cut out the singularity. Sample the integrand with a sufficiently small interval to form a discrete substitute and perform the integration omitting the singularity . This results in significant error since the integrand increases rapidly when approaching the singularity.

Approach 2: Add a zero-value part to the integrand to avoid the divergence around the singularity. This gives -

$$\mathrm{Re}\left[\frac{1}{\varepsilon(E)}\right]-1=-\frac{2}{\pi}P\int_0^{\infty}\frac{\mathrm{Im}\left[\frac{-1}{\varepsilon(E')}\right]E'-\mathrm{Im}\left[\frac{-1}{\varepsilon(E)}\right]E}{E'^2-E^2}dE' \qquad (9)$$

Equation (9) is obtained by adding the term $-\left(\frac{1}{\pi}\right)P\int_{-\infty}^{+\infty}\frac{\mathrm{Im}\left[-1/\varepsilon(E)\right]}{E'-E}=0$ to equation (8) and assuming that $\mathrm{Re}\left[\frac{1}{\varepsilon(E)}\right]-1$ is even and $\mathrm{Im}\left[\frac{-1}{\varepsilon(E)}\right]$ is odd. Then equation (9) has a removable isolated singularity, since $\lim_{E'-E}\left[\frac{\mathrm{Im}\left[\frac{-1}{\varepsilon(E')}\right]E'-\mathrm{Im}\left[\frac{-1}{\varepsilon(E)}\right]E}{E'^2-E^2}\right]$ = finite. This makes the Cauchy principal part unnecessary, and the result is much smoother without a divergence at the singularity.

Approach 3: Fourier series method for KKA (Johnson, 1975). Because of causality, the odd and even parts of time-dependent response are related as follows:

$$p(t)=\mathrm{sgn}(t)q(t)$$

$$p(t)=FT^{-1}\left\{\mathrm{Re}\left[\frac{1}{\varepsilon(E)}\right]-1\right\}$$

$$p(t)=FT^{-1}\left\{i\times\mathrm{Im}\left[\frac{1}{\varepsilon(E)}\right]\right\}$$

where $p(t)$ and $q(t)$ are, respectively, the even and odd parts of the time-dependent response function, which is the FT of $\left[\frac{1}{\varepsilon(E)}\right]-1$. Here sgn($t$) is a sign function that takes the value of 1 for $t>0$ and −1 for $t<0$. By using equation (4) and FT, $\mathrm{Re}\left[\frac{1}{\varepsilon(E)}\right]$ can be calculated without evaluating the Cauchy principal value of an improper integral.

The complex dielectric function is then obtained using:

$$\varepsilon(E)=\varepsilon_1+i\varepsilon_2=\frac{\mathrm{Re}\left[\frac{1}{\varepsilon(E)}\right]+i\,\mathrm{Im}\left[\frac{-1}{\varepsilon(E)}\right]}{\left\{\mathrm{Re}\left[\frac{1}{\varepsilon(E)}\right]\right\}^2+\left\{\mathrm{Im}\left[\frac{-1}{\varepsilon(E)}\right]\right\}^2} \tag{10}$$

1.5.2 Normalization

Before computing the complex dielectric function using equation (10), normalization is needed because it may not be possible to determine the zero-loss intensity I_0 and hence the prefactor α in equation (7). To implement normalization, equation (9) is used at $\omega=0$ and a general expression for the refractive index of materials (Tan et al., 2005). For most insulators, such as those studied here, the imaginary part of the dielectric constant ε_2 is negligible, therefore

$$\mathrm{Re}\left[\frac{1}{\varepsilon(0)}\right]=\frac{1}{n^2}=1-\frac{2}{\pi}P\int_0^\infty\left(\frac{\mathrm{Im}\left[-1/\varepsilon(E')\right]E'}{E'^2}\right)dE' \tag{11}$$

where n is the refractive index of the material. Together with equation (7), equation (11) gives the prefactor α, so that the normalized complex dielectric function can be obtained from equation (10).

1.6 Common Techniques Used to Obtain Dielectric Function

A variety of techniques offer the scope to study the dielectric function. Typically these are sorted by the regimes of energy that they probe. Fig. 3.4 represents the energy-scale and corresponding types of radiation.

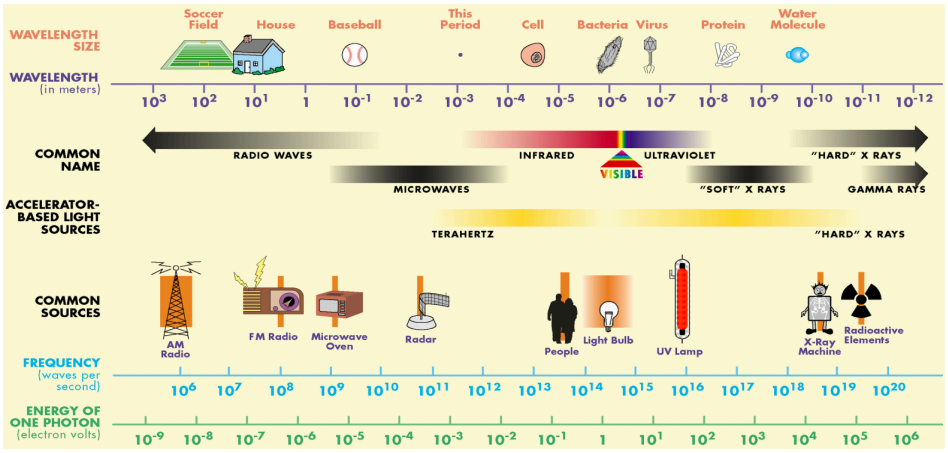


Fig. 3.4 Radiation regime with Energy (frequency) and Wavelength (Source courtesy: http://chemwiki.ucdavis.edu/Textbook_Maps/General_Chemistry_Textbook_Maps/Map%3A_Chem1_(Lower)/05._Atoms_and_the_Periodic_Table/5.2%3A_Light,_Particles,_and_Waves)

1.6.1 Terahertz Regime:

Terahertz (THz) is the regime of electromagnetic radiation that spans the 10s to 100s of micrometer wavelengths. THz spectroscopy started as efforts to generate and detect ultra-short electrical transients propagating through a transmission line. However, it was realized that these transmission lines were also radiating short bursts of electromagnetic radiation. A radical change occurred in 1988/1989 when these radiated pulses were propagated through free space from a generator to a detector. This is when, researchers started to think of utilizing this far-IR light source and detector pair by placing samples in between them. While there exist methods for determining the index of refraction when using conventional far-IR sources and detectors (refer [1] for a direct comparison of the two techniques), a great majority of the far-IR results present the frequency-dependent absorption coefficient and not the refractive index. In comparison, THz spectroscopy provides a convenient method for determining the complex permittivity, even for steady state studies.

A variety of materials have been studied using this technique. Semiconductor carrier dynamics as a function of lattice variations have been extensively studied [2–4] using this technique. THz spectroscopy is also very popularly used to study large structural correlations and changes in biological systems. From works ranging on the smallest of components as the DNA bases [5] to complete proteins [6,7]. A comprehensive review of the technique and its capabilities are presented in [8] with the aim of studying dynamics.

1.6.2 Infra-Red Regime:

Infrared regime, or more specifically the near infrared regime broadly spans the 0.8-micrometer to 5-micrometer regime. Optical and acoustic phonons and vibrational modes of organic bonds are typically studied in this regime. There are multiple ways of having such studies, some of which are – (Fourier transform) Infra red spectroscopy, Raman Spectroscopy and Ellipsometry. The difference between an IR and Raman measurement is for a transition to be Raman active there must be a change in the polarizability of the molecule during the vibration. This means that the electron cloud of the molecule must undergo positional change. On the other hand, for an IR detectable transition, the molecule must undergo dipole moment change during vibration. So, when a molecule is symmetrical, e.g. O_2, IR absorption lines cannot be observed, since the molecule cannot change its dipole moment. It has been observed that molecules with a strong dipole moment are typically hard to polarize. On the other hand, spectroscopic ellipsometry is known as an excellent technique for determination of the complex dielectric functions of materials, and avoids inaccuracies due to extrapolation into experimentally inaccessible spectral regions as necessary for Kramers-Kronig analysis of reflectivity data. The measured signal is the change in polarization as the incident radiation (in a known state) interacts with the material structure of interest (reflected, absorbed, scattered, or transmitted) and depends on the thickness as well as the complex refractive index of the material. The polarization change is quantified by the amplitude ratio, and the phase difference.

Here again, examples of materials studied range from solid state semiconductors and insulators, [9,10] to semi-metals, [11] and from metamaterials [12] to biologically relevant systems [13–15].

1.6.3 Optical and Ultraviolet Regime:

The most extensively studied electromagnetic radiation regime is the optical regime, which corresponds to the 400 – 700 nm wavelength scales. Absorption spectroscopy refers to spectroscopic techniques measure the absorption of radiation, as a function of frequency or wavelength, due to its interaction with a sample. The sample absorbs energy, i.e., photons, from the incident beam. The intensity of the absorption varies as a function of frequency, and this variation is the absorption spectrum. The frequencies where absorption lines occur, as well as their relative intensities, primarily depend on the electronic and molecular structure of the sample. Electronic transitions [16,17] correspond to a change in the electronic state of an atom or molecule and are typically found in the visible region and occur about the highest occupied molecular orbitals (HOMO) and the lowest unoccupied molecular orbitals (LUMO) and in the vicinity. When working in the UV regime, this is extended from 400 nm – 100 nm. In such cases, moderately lower lying electrons are excited to more unoccupied orbitals higher than the LUMO levels. UV regime becomes specifically relevant with wide band gap semiconductors [13,18] and insulators to probe around their band gaps and for regular semiconductors and metals to probe their higher unoccupied and relatively deeper core levels. However, the one clear distinction between the THZ and infrared on the one hand and Optical and UV on the other hand is that, while the former probes nuclear and vibrational motion, the latter probes electronic excitations. Hence, together, the former tend to give the phonon band structure and the latter give the electronic band structure for solids. This is why the dielectric response tends to be such a powerful tool since it provides the details of the band structure that forms the basis of all physical properties of the material (Note: For inorganic/organic molecules, liquids and gases, the band structure can be substituted instead by the molecular orbital picture).

1.7 EELS as a Tool for Dielectric Response

The low-loss spectra of EELS probe exactly the same region as those mentioned in the previous section. The primary difference between all the aforementioned techniques and EELS is the difference in the probes, photon vs. electron. A major advantage in working with

the EELS technique is the possibility to work on extremely focused probes (nm to few Å) that can excite locally on the sample as opposed to the photon probes which typically have much broader spot sizes (100s of μm, recently efforts are on to focus them to about few μm or even smaller). However, the energy resolution that is offered by an electron probe is typically one to two orders of magnitude larger than that of the IR or UV-Vis probes. Nevertheless, the EELS low-loss spectrum is very commonly used to study the dielectric response. For example, it has been used to study 2D semiconductors including graphene [19] where they researchers find that the commonly assigned π and π* transitions do not reflect plasmon behavior and dichalcogenides [20] that have electronic conduction properties ranging from metallic to semiconducting to insulating that unanimously all these materials see a red-shift of the plasmon peaks as the thickness (number of 2D layers) in the sample reduces, from a bulk sample to a strictly 2D single layered sample. The technique has also been used widely to characterize organic molecules [21,22] and to characterize the electronic properties of organic thin films and semiconductors [23,24]. Special care is taken to use 30 keV or lesser incident beam energy and/or maintain a substantially lowered exposure time to ensure that these light element (H, C, N, O) compounds do not suffer radiation damage and thereby display altered physical properties. EELS low-loss has also been used, sparingly however, to study the dielectric response of biological molecules such as nucleic acid bases [25–27]. Since spectroscopy does not require a long range order, there have also been examples of studies done on non-crystalline samples [28,29] such as glass or carbon.

Although EELS in the low-loss regime has been used extensively, as mentioned above, the limiting factor is to find the combination of a high energy and space resolution. With the ability to focus photons not being as effective so far, the question then arises as to whether electron microscopes can be the elusive alternative.

1.8 NiON UltraSTEM™100 MC at ASU

The latest addition to the list of TEMs at ASU is the Nion UltraSTEM™100MC. [30] The Nion is a stand-alone STEM instrument and the latest versions of their designs have C3 and C5 aberration correctors. In this version (the latest of them all) however, the addition to the

instrument is a mono-chromator (MC). The schematic of the instrument is presented in fig. 3.5. The presence of the MC reduces the FWHM of the ZLP by more than a factor of 10 and as a routine now, the instrument offers 10-20 meV energy resolution.

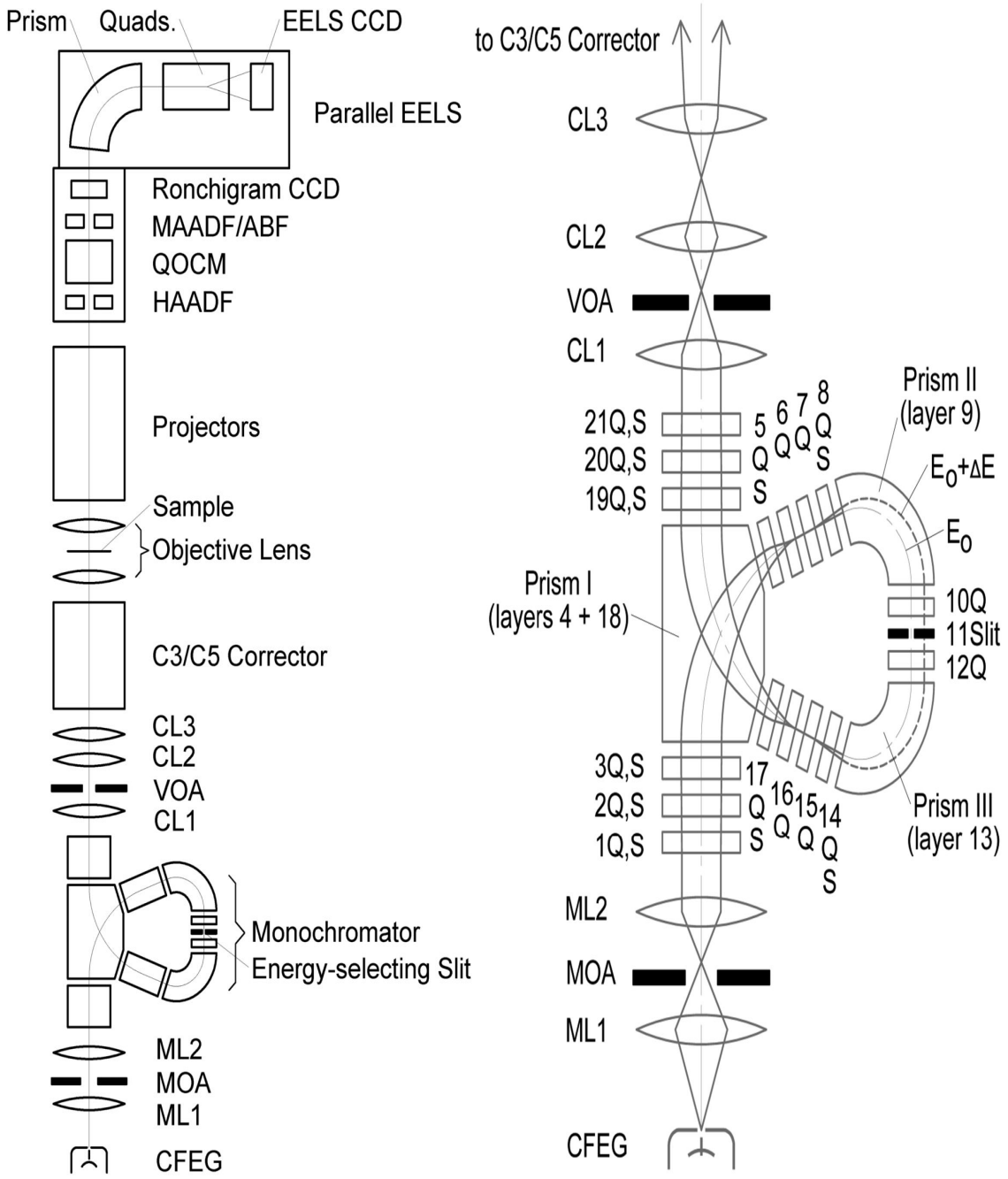


Fig. 3.5: LEFT: Schematic of the NiON UltraSTEM^{TC}100MC installed at ASU. RIGHT: Schematic of the cross-section of the MC showing the lenses and apertures. (Figure courtesy [30])

This is a big step ahead because the instrument is unique in being able to offer 20 meV energy resolution coupled with sub-nm spatial resolution. In a sense, this clubs together the positives of IR/Raman spectroscopy (ability to probe phonon modes) with STEM (ability to offer <3A° spatial resolution). Fig. 3.5 (LEFT) gives the overall schematic of the instrument and Fig. 3.5 (RIGHT) details the arrangement of the monochromator.

Briefly, the gun (CFEG) is located at the bottom. The electron beam passes the MC aperture into the slit where the energy selection is made. After narrowing down the beam, it passes through a set of condenser lenses and then the aberration correctors. At this point the beam interacts with the sample and passing through the projector lenses it is received by various detectors. As a typical STEM instrument, it supports ABF, MAADF and HAADF. There is a CCD placed to view the ronchigram and past all that is the ELS detector.

The instrument was first installed at ASU in 2013 and was state-of-the-art in the simultaneous spatial and energy resolution that it provided. One of the first demonstrations of the ability to probe phonon modes at high spatial resolution was with $SiSiO_2$ [31] where the use of aloof-EELS which is a systematic study on moving the electron beam farther away from the sample region into vacuum was also demonstrated. The idea here is to utilize the sufficiently excite vibrational states in matter farther away from the region of interest thanks to the delocalized (10s – 100s of nm) nature of vibrational excitations. This way, the background signals are reduced significantly and most importantly radiation damage limited dose and exposures are increased. Using this improvement, researchers tracked the presence of water (hydration) on/within catalysts which can routinely be done using IR spectroscopy and other techniques, but limited only to large samples. Here the researchers [32] were able to detect the OH bond signatures (approx. 450 meV) on/within (60nm MgO) nanoparticles.

With one step leading to another, very recently in fact, has been the first successful demonstration of extending vibration spectroscopy to biological materials within high spatial and spectral resolution. In the work, [33] the scientists have investigated on the biogenic guanine crystals and successfully compare (with IR absorption) and present the C-H, N-H and C=O vibrational signatures with the aloof-EELS technique using the NiON UltraSTEM MC.

1.9 Frequency Domain Measurement of Protein Dynamics

Dynamics can be measured both in the time and the frequency domain. The example of time domain measurements with electron probe was discussed in the previous chapter. A frequency domain measurement measures the response of a system to an external probe in time. Where a time-domain measurement would require multiple measurements to map the response of the material to the external probe, a frequency domain measurement would capture this very essence in one measurement. This subsection will provide a brief explanation of some basic concepts that deal with the correlation between time and frequency domain measurements.

1.9.1 Relation Between Frequency and Time Domains

Frequency and time are conjugate variables that are inversely related of each other. This is exactly similar to the relation between real-space coordinate (position) and reciprocal space coordinate (wave vector (or) momentum) that was introduced in the first chapter. Hence to move from the time domain to the frequency domain or vice versa, the variables undergo the Fourier transform. This is very commonly used in dealing with signals of electromagnetic radiation and some of the most common relations of the transforms are presented in table 3.1.

Some very important and commonly used transforms are shaded blue. The ‘delta’ function as an example would mean an instantaneous pulse in the time domain. This in the frequency domain translates to something that is constant. The ‘hat’ (top hat) function is the most basic approximation of the nature of a pulsed electromagnetic radiation. It represents a radiation pulse, which is switched ON for a specific period of time at a single value and then switched OFF. It is also a good second step as an extension from the delta function. The frequency domain transformed representation in this case is referred to as the ‘sinc’ function, which is a sinusoidal function damped by an inverse (x^{-1}) function.

Another very important time domain signature is the exponential decay. This is very commonly used to represent electromagnetic fields in the near-field optics. Typically in far-field optics, the wave nature of radiation can be approximated by the ray behavior. The ‘hat’ is a perfect example of such an approximation. However, in near-field optics since the dimensions (in

length and/or time) are comparable to that of a period of the wave, the rise and/or fall of a signal cannot be approximated as a 'step' or 'hat' anymore. Traditionally, one of the most common types of decay is the 'exponential decay' (Arrhenius type). When transformed, the exponential decay becomes a 'Lorentzian' function. The Lorentzian function is surprisingly similar to equations describing a simple harmonic oscillator (which will be revisited in next section).

	Time Domain $x(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} X(\Omega)e^{j\Omega t}d\Omega$	Frequency Domain $X(\Omega) = \int_{-\infty}^{\infty} x(t)e^{-j\Omega t}dt$	
HAT	$x(t) = \begin{cases} 1, & \lvert t\rvert < T \\ 0, & \text{otherwise} \end{cases}$	$X(\Omega) = \frac{2\sin(\Omega T)}{\Omega}$	SINC
	$x(t) = \frac{1}{\pi t}\sin(Wt)$	$X(\Omega) = \begin{cases} 1, & \lvert\Omega\rvert < W \\ 0, & \text{otherwise} \end{cases}$	
DELTA	$x(t) = \delta(t)$	$X(\Omega) = 1$	CONSTANT
	$x(t) = 1$	$X(\Omega) = 2\pi\delta(\Omega)$	
	$x(t) = u(t)$	$X(\Omega) = \frac{1}{j\Omega} + \pi\delta(\Omega)$	
	$x(t) = e^{-at}u(t), \quad a > 0$	$X(\Omega) = \frac{1}{a+j\Omega}$	
	$x(t) = te^{-at}u(t), \quad a > 0$	$X(\Omega) = \frac{1}{(a+j\Omega)^2}$	
EXPONENTIAL DECAY	$x(t) = e^{-a\lvert t\rvert}, \quad a > 0$	$X(\Omega) = \frac{2a}{a^2+\Omega^2}$	LORENTZIAN
	$x(t) = \frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}$	$X(\Omega) = e^{-\frac{\Omega^2}{2}}$	

Table 3.1: Fourier transform relations (courtesy: http://allsignalprocessing.com/the-fourier-transform/)

1.9.2 Excitation Dynamics

Spectroscopy is the study of excitation of a system of particles (electrons, atoms etc) from state A to state B. This means that an excitation is accompanied by a variation in the population of the involved states. The study of change in population as a function of time is done using 'rate equations' and such a study is defined as 'dynamics'. Although such population

(depopulation) schemes can get very complicated, the most basic and one of the most widely seen types of rate equations is christened '1st order'. This can be derived as follows:

Consider that a system has its entire population in state A (initial state) and that post interaction with an external stimulus it undergoes a transition to state B (excited/final state). To describe the nature of transition (excitation), the most basic type would be that which is only a function of the population of the system and is independent of the nature of the probe. Zeroth order dynamics corresponds to the case when the dynamics is independent of the population in either states (A or B).

First order dynamics, corresponds to the case when the excitation is dependent only on the population of the initial state. Since the excitation involves removing particles from state A and depositing into state B, this can be written mathematically as –

$$-\left(\frac{d[A]}{dt}\right) = k[A] \qquad \text{(a)}$$

where [A] is the number/concentration of particles in state A. The negative sign implies population removal and not formation. This can be rearranged and solved to obtain an equation as shown below –

$$[A]\Big|_{t=0}^{t=t} = e^{-kt}\Big|_{t=0}^{t=t} \qquad \text{(b)}$$

As mentioned earlier, since the initial state is occupied completely and slowly has a population erosion, if the initial population is assumed as A_0 then –

$$[A]_t = [A]_0 e^{-kt} \qquad \text{(c)}$$

This equation is an exponentially decaying function (where k is termed as the time constant) exactly the same as in table 3.1. (Table 3.1 had a function that was exponentially

decaying for both positive and negative values of the argument. Here, since negative time does not have a physical meaning, the equation represents just one half of the region. However, the functions are otherwise identical).

If one were to be interested in the dynamics corresponding to the formation of population in state B, it would be

$$-\left(\frac{d[A]}{dt}\right)=\left(\frac{d[B]}{dt}\right)=k[B] \qquad \text{(d)}$$

meaning that the rate of depopulation of state A is the same as the rate of populating state B.

Going back, table. 1 suggests that the Fourier related variable would to equation (c) would be a Lorentzian function.

1.9.3 Drude-Lorentz Oscillator Model for Excitations:

In thinking of electronic transitions, it is easier to take a reference of the classical harmonic oscillator model (assuming 1-dimensional and along 'x') whose wave equation is described as –

$$m\frac{d^2x}{dt^2}=-m\omega_0^2x+qE_x(t)+m\gamma\frac{dx}{dt} \qquad \text{(e)}$$

where the first term in the RHS is the restoring force, the second term is the driving force (written as a sinusoidal electrical force field) and the third term is the damping force. If a plane wave solution to the wave function x(t) is assumed, then it would be of the form –

$$x(t)=ASin(\omega t)+BCos(\omega t) \qquad \text{(f)}$$

and in which case solving differential equation for x(t) would give –

$$\omega^2 x(t) + i\omega\gamma x(t) + \omega_0^2 x(t) = \frac{q}{m} E_x \qquad \text{(g)}$$

This can be rearranged as –

$$x = \frac{qE_x}{m}\left[\frac{1}{(\omega^2 - \omega_0^2) + i\omega\gamma}\right] \qquad \text{(h)}$$

Now, for an electric field induced dipole, the polarization can be defined as –

$$N\left\{q(\delta r_x)\right\} = Nq(x) = P_x = \left\{\varepsilon_0 \chi\right\} E_x \qquad \text{(i)}$$

Since the polarization 'P' is directly proportional to dipole distance 'x',

$$\frac{\left\{\varepsilon_0 \chi\right\} E_x}{Nq} = \frac{qE_x}{m}\left[\frac{1}{(\omega^2 - \omega_0^2) + i\omega\gamma}\right] \qquad \text{(j)}$$

Substituting relative permittivity for susceptibility, and defining $w_p^2 = Nq^2/m\varepsilon_0$ as the plasmon frequency, the equation reduces to –

$$\varepsilon_r = 1 + \frac{\omega_p^2}{\left(\omega^2 - \omega_p^2\right) + i\omega\gamma} \qquad \text{(k)}$$

Finally, assuming a complex permittivity $\varepsilon_r = \varepsilon_{Re} - i\varepsilon_{Im}$, the above equation can be segregated into real and imaginary permittivities as –

$$\varepsilon_{\mathrm{Re}} = 1 + \frac{\omega_p^2\left(\omega^2 - \omega_p^2\right)}{\left(\omega^2 - \omega_p^2\right)^2 + \omega^2\gamma^2} \quad \text{and} \quad \varepsilon_{\mathrm{Im}} = \frac{\omega_p^2\left(\omega\gamma\right)}{\left(\omega^2 - \omega_p^2\right)^2 + \omega^2\gamma^2} \qquad (l)$$

LHS of equation (l) is the complex permittivity split into its real and imaginary components, same as obtained in equation (10) post the KKA. Also, the RHS of equation (l) is of exactly the same form as that in table. 1 termed the 'Lorentzian function'. These matches have two very important significances –

(1) LHS similarity (eqns. l and 10) tells us that the KKA analysis can be compared with the equations obtained here and one could build a model-based approach to identify the various parameters w_0 and γ. In such a model, 'w_0' would correspond to the central frequency (energy) of the transition (the energy difference between the initial (ground) and final (excited) states which is also the same as the energy lost by the incident electron beam in interacting with the sample to inelastically scatter and create the excitations). On the other hand, γ is the damping factor

(2) RHS similarity (eqns. l and table. 3.1 'Lorentzian function) tells us the when Fourier transformed, eqn. l would have the same form as an exponential decay, which is the same as the first order dynamics as in eqn. (c). Equation l a few additional terms suggesting that there is more than just an exponential decay. This is true as equation (e) shows that there is a 'restoring force' term and a 'driving force' term apart from the 'damping force'. Matching coefficients, it is straightforward to figure out that w_0 comes from the restoring force and w from the driving force. Therefore, it is evident that eqn. l has a sinusoidal oscillatory portion that is ideally infinitely persistent, although in reality is damped thanks to the γ term. Now linking this back to eqn. (d), 'γ' is equivalent of 'k' the time constant and so if Fourier transformed and plotted in time, γ would be represented as the exponential decay envelope over the sinusoidal oscillation.

What is the purpose of establishing all these relations? Getting back to the start of this section, the idea behind establishing these relations was to relate information obtained in the time domain and the frequency domains. Now that the tools are developed the advantage of the frequency domain measurement can be established through a simple example – same as the one used in section 1.9.b.

Let us assume there is a system with all its particles in state A. Let us also assume that on interacting with an external probe, the system has some of its particles excited to state B. If it is assumed that the reaction is 1^{st} order, then to know everything about this process, there is just one relation that needs to be established – the variation of initial state population as a function of time. This can be done in two ways –

(1) Time domain measurement: In this type of measurement, an initial external stimulus would initiate the excitation and subsequently a second external stimulus would probe the population. This is the classic pump-probe methodology. To obtain the dynamics using this experimental design, this experiment would have to be repeated for various different time-delays between the initial external stimulus and the second external stimulus. Each such measurement would give the population at the initial state at different points in time since initiating the excitation and hence would map out the mathematical nature of the excitation.

(2) Frequency domain measurement: In this type of measurement, an external stimulus will scan the frequency (energy) to identify a 'peak' at the appropriate frequency at which the excitation occurs. However, since this is not a permanent or infinitely long lasting excitation, the shape of the 'peak' would be modified. Just by identifying the position of this peak (w_0) and the shape of the peak causing it to die and not rise again (γ), the same information obtained above is also obtained here. There on, by performing an Inverse Fourier transform, the time domain information can be obtained rather easily.

Of course, the schemes can get really complex with multiple excitations overlapping on each other and the excitations not necessarily following a 1^{st} order dynamics. With modifications to the equations presenter earlier, in theory, all such cases can be attended to.

2 Aim of This Chapter

In the pursuit of solving for the structure (static/dynamic) of biomolecules, frequency domain spectroscopy using electron probe has been reviewed and established as a successful technique. In this chapter, the aim is to identify the possibilities that are thrown open thanks to the new monochromated NiON UltraSTEM^{TM} MC100 installed at ASU.

In specific, this is done by addressing the following question –

- Is it feasible to actually use the NiON UltraSTEM MC to obtained hitherto impossible spatial and spectral resolutions and use this instrument to study the structure/dynamics of biological molecules?

3 Observation

3.1 Time-Energy Uncertainty and the NiON UltraSTEM

The aim here is to first understand the atomic and electronic response of proteins to fast electrons in the time domain by FT of the energy-dependent dielectric function. Since the energy resolution of the new monochromated EELS systems is about equal to kT at room temperature, the driving energy for biological processes, this may provide similar information on molecular interactions to that given by Raman and infrared absorption spectroscopy, but with the advantages of higher spatial resolution, and the possibility of obtaining angle resolved spectra. To obtain an estimate of the time resolution and time range, the Heisenberg uncertainty principle is used:

$$\Delta t \approx \frac{h}{\delta E} = \frac{4.14 eV.fs}{\delta E} \qquad \text{(I)}$$

and

$$\delta t \approx \frac{h}{\Delta E} = \frac{4.14 eV.fs}{\Delta E} \qquad \text{(II)}$$

where δ and Δ stand for resolution and range, respectively. Equation (I and II) give the energy– time domains conversion table (Table 3.2).

Table 1. Resolution and Range Conversion Between Energy and Time Domains.

Energy Resolution	Time Range	Energy Range	Time Resolution
5 eV	0.828 fs	30 eV	0.138 fs
1 eV	4.14 fs	100 eV	0.0414 fs
0.3 eV	13.8 fs	300 eV	0.0138 fs
0.1 eV	41.4 fs	1,000 eV	0.00414 fs
50 meV	82.8 fs	—	—
10 meV	414 fs	—	—
4.14 meV	1 ps	—	—
4.14 μeV	1 ns	—	—
4.14 neV	1 μs	—	—

Table 3.2: Resolution-Range in the Energy-Time Fourier relations

From the table, it can be seen that the time resolution from an EELS spectral range of 1 keV is 0.004 fs (1 fs = 10^{-15} s). This would allow ultra-fast dynamic processes to be identified if a sufficient number of spectrum pixels were available to collect such a spectrum at millivolt energy resolution. The time range that was achieved with the Nion UltraSTEM at 10 meV energy resolution is then about 400 fs. Note that the Debye period for atomic vibration is estimated to be around 137 fs, from the Debye temperature of 350 K for hexagonal ice at room temperature, and the Bohr model gives an estimate of 0.15 fs as the revolution period of an electron in a hydrogen atom. This suggests that it might be possible to identify several cycles of atomic vibration and even to resolve electron motion in atoms in the time domain.

3.2 Comparison of Different KKA Approaches

KKA was applied to the input data energy loss function using three different approaches as mentioned in the previous section. The energy loss function Re ε(E) calculated using different approaches from the EELS of hexagonal ice is shown in Figure 3.6. In spite of minor difference in magnitude, the shapes of the three curves in Figure 3.6 are almost the same. This difference can be attributed to the errors of the three approaches. Errors in Approach 1 and 2 mainly result from the divergence at the singularity, as well as the finite integration range. The error in Approach 3 is

caused by the limited bandwidth and the causality constraint, which may be stronger than in traditional KKA. Nevertheless, the three different approaches do not seem to cause large differences.

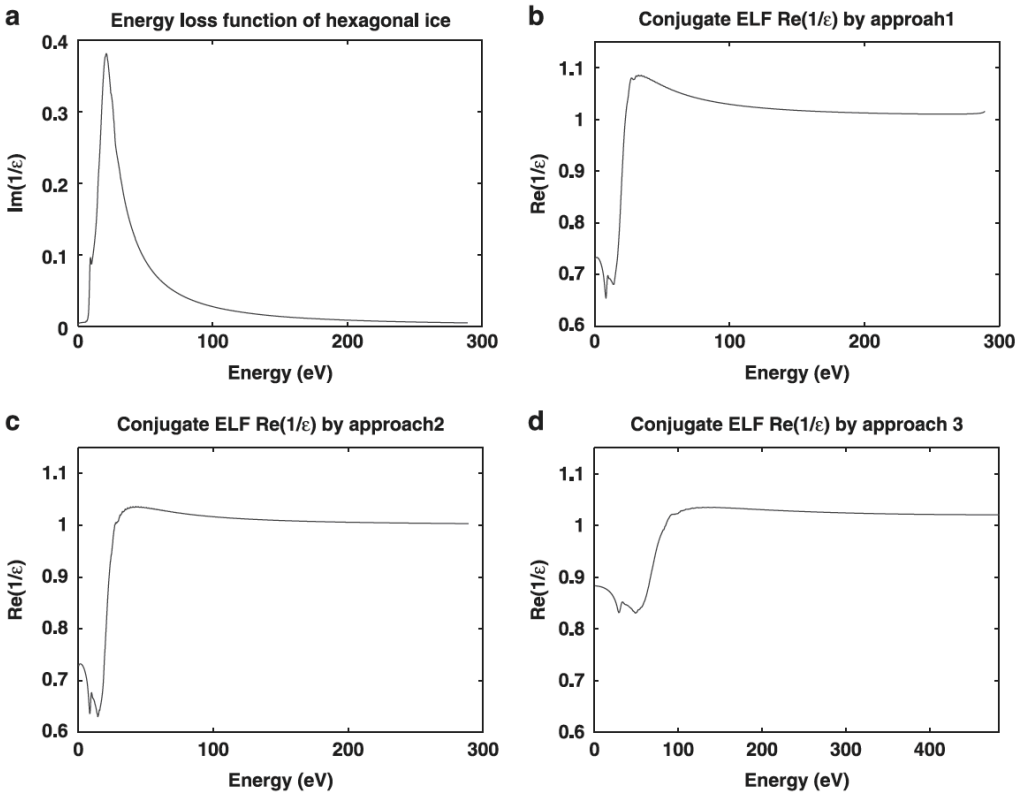


Fig. 3.6: Comparison of the energy loss function from different KK analysis approaches (a) ELF of hexagonal ice. Conjugate ELF from (b) approach 1; (c) approach 2 and (d) approach 3

3.3 Hexagonal Ice and Protein

Using this technique, the time dependent dielectric function for hexagonal ice and protein were extracted from the experimental data reported by Sun et al. (1993), as shown in figures 3.7, 3.8, and 3.9.

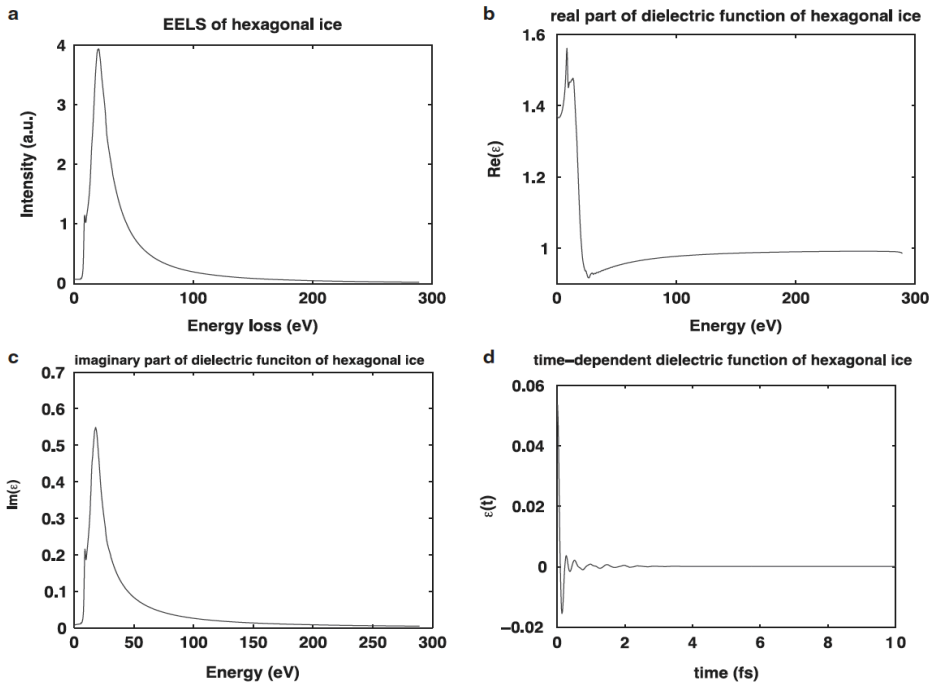


Fig. 3.7: Hexagonal Ice: (a) EELS spectra [from Sun et al. (1993)]. Dielectric function of Hexagonal Ice (b) real part (c) imaginary part (d) time-dependent.

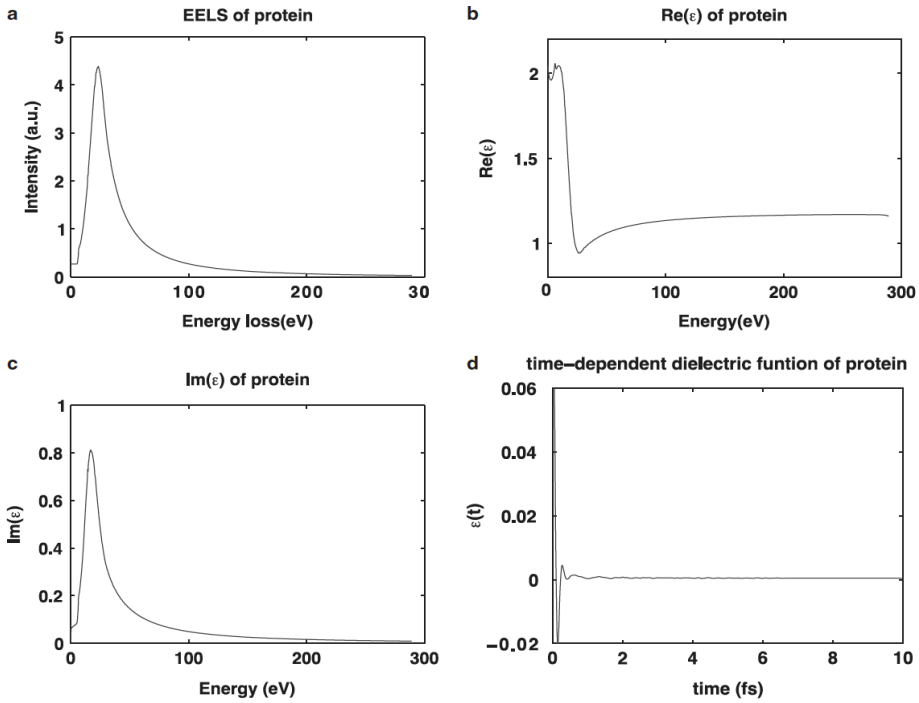


Fig. 3.8: Protein: (a) EELS spectra [from Sun et al. (1993)]. Dielectric function of Protein (b) real part (c) imaginary part (d) time-dependent.

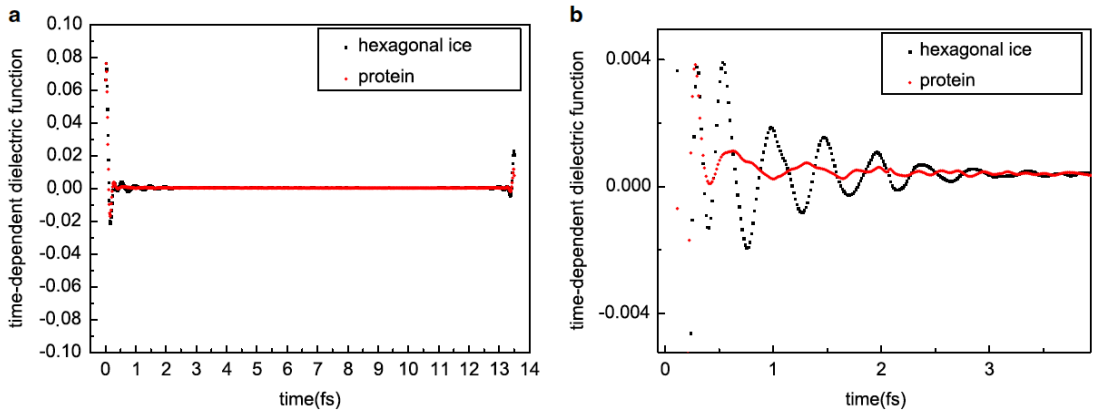


Fig. 3.9: Comparison of time-dependent dielectric functions for Hexagonal Ice (black) and Protein (red): (a) Full range in time (b) first few fs.

Steep edges in the $\varepsilon_1(E)$ curve, and peaks in $\varepsilon_2(E)$ positioned at 6.7 eV (protein), 9.1 eV (ice), 20.4 eV (ice), and 23.4 eV (protein) are found. The 20.4 and 23.4 eV peaks correspond to plasmon absorption in ice and protein, respectively. The 9.1 eV peak is believed to result from transitions to the unoccupied states above the band gap of ice. The 6.7 eV peak can be attributed to the excitation of π electrons in the protein. All these peaks at different energies contribute to the damped oscillation of the time dependent dielectric function for different frequency components.

The time dependent dielectric function of hexagonal ice is damped to negligibly small values within 4 fs while that of protein is even faster, about 3 fs (fig. 3.9). The dominant oscillation in the first few femtoseconds is the plasmon behavior, since it has the largest oscillator strength and the frequency is also consistent with the plasmon energy. Protein obviously has a stronger damping effect than hexagonal ice. The tails appear due to the error resulting from the limited bandwidth of the EELS data, when applying the KKA. Phonons are not detected with this energy resolution of 0.3 eV, as its typical oscillation period is 137 fs for ice.

Thus, given the dielectric function from EELS, it is possible to calculate the electron response in matter for different forms of external field source functions D. When extended to the thermal energy range using millivolt EELS, this may be helpful in understanding the way in which

biomolecules are able to take up energy through their various excitations. The results can also provide us with an estimate of the lifetime of beam-induced excitations in biological specimens. Excitation by a pulse briefer than this lifetime is not expected to show effects of radiation damage in the resulting elastic scattering, however, irreversible damage may occur at longer times.

3.4 BaF_2 EELS

The Nion UltraSTEM™ 100MC installed at Arizona State University (ASU) has the unique capability of obtaining 40 meV energy resolution currently at subnanometer spatial resolution. This resolution will soon be further improved by installation of new power supplies (a resolution of 12 meV is demonstrated by Krivanek et al., 2013). The design is thus aimed at enhancing the study of the low-loss region. The low-loss EELS has been obtained from the alkali halide BaF_2 using this new Nion instrument at ASU. For this specific experiment, the full width at half maximum of the zero loss peak (ZLP) was at 100 meV so as to obtain the necessary range on the abscissa. The beam diameter was ~0.15nm (point scan mode). As is observed from fig. 3.10, the low loss region of BaF_2 is rich with a variety of peaks. These peaks are identified by comparing fig. 3.10 with fig. 7 and table 1 of the work by Frandon et al. (physica status solidi (b) 53.2 (1972): 565-575). With a band gap of around 10.7 eV, peaks immediately following the band gap are excitations (d and e) corresponding to the fluorine 2p states lying at the top of the valence band. Following that are the relatively low lying Ba 5p core state inter-band transitions (f-h) with the bulk plasmon excitations (i) at around 26 eV. More interestingly, the pre band gap exciton peak (b) at 7.8 eV and a color center (a) at around 3.2 eV can be observed. These color centers are formed by anion vacancies created by the displacement of F atoms from lattice positions by the electron beam, leaving behind electrons in their place. Thus, extracting the time dependent dielectric function permit probing electronic excitations from localized regions in space on a femtosecond time scale.

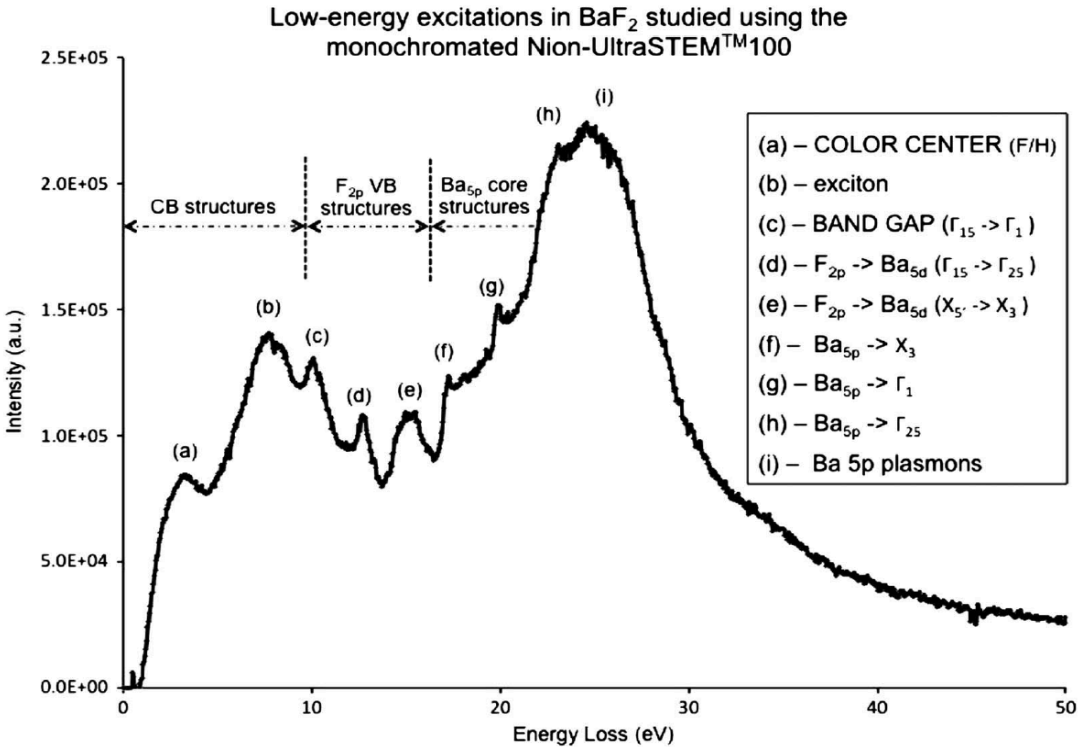


Fig. 3.10: Low-loss electron energy-loss spectra (EELS) data from BaF2 acquired using the new monochromated Nion UltraSTEM 100 at Arizona State University. The convergence semi-angle for the experiment was 30 mrad and the EELS collection angle 15 mrad. Probe size was 0.15 nm and full-width at half-maximum of the zero-loss peak was 100 meV. The pre band gap, Fluorine 2p and Ba 5p electronic transition features are highlighted as peaks (a–i), notably (a) the color center peak and (b) the band-gap exciton peak.

Figure 3.11 (a) shows the inelastic component of the EELS after removal of the ZLP using the "reflected tail" model. Note that there is a steep and sudden rise from zero intensity, which is reflected as a sharp spike in the e1 function in fig. 3.11 (b). This is purely an artifact of the sudden increment from zero in the EELS data. Figure 3.11 (c) shows the imaginary part of the dielectric function in which the color center and the band gap excitons are identified (amongst the various other valence and core level transitions). A comparison of the ε_1 and ε_2 obtained here with that obtained by Frandon et al. (1972, figure 6) using EELS at 20 kV in a dedicated magnetic spectrometer can be made to observe that the peak positions here are in agreement with theirs.

But since the data here were acquired with a probe of near atomic dimensions, the color centers (a) and the band gap excitons (b) can also be observed in addition other features. The extracted time dependent spectrum for BaF_2 is shown in figure 3.11 (d). The time evolution of ε(t) here contains contribution from all nine peaks identified initially from the EELS.

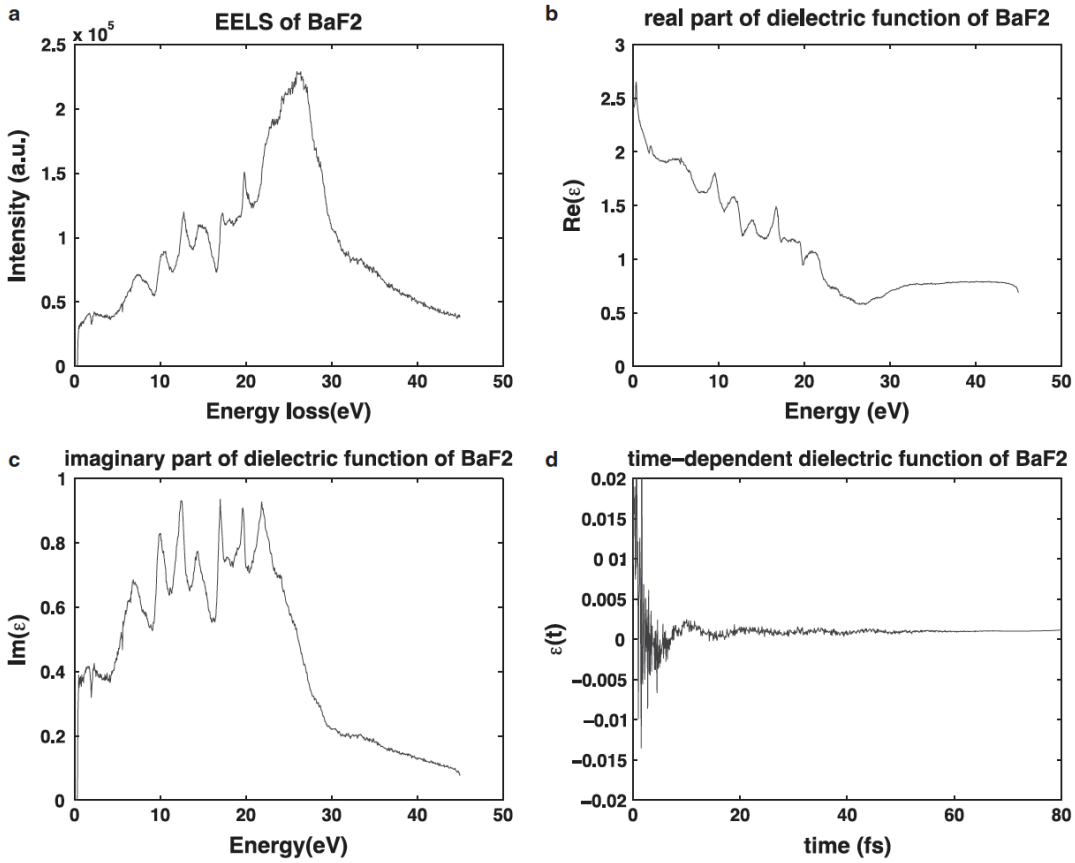


Fig. 3.11: BaF_2 (a) EELS after deconvolution. BaF_2 dielectric function (b) real part; (c) imaginary part; and (d) time-dependent.

For clarity, fig. 3.12 gives the time evolution of the dielectric function for a more limited energy range of 9 eV. For this, the spectrum has been linearly extrapolated from 9 eV to intersect the abscissa at 27 eV so as to meet the requirements for the KKA. In this case, only the first two excitations [color center (a) and bandgap excitons (b)] are included. It can then be clearly seen, comparing figs. 3.11 (d) and 3.12, that the low frequency components (with longer wavelength) in fig. 3.11 (d) arise as a contribution from excitations (a) and (b). Note also that when comparing

the two figures at t = 10 fs, the addition of the other excitations causes the overall amplitudes to be damped.

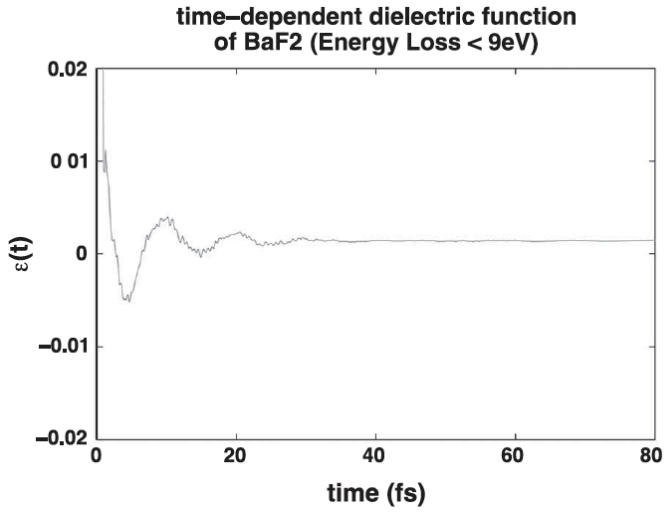


Fig. 3.12: Time-dependent dielectric function of BaF2 obtained over an energy-loss range of 9 eV. This encompasses only the peaks (a) and (b) from Figure 8. To avoid an abrupt termination, the electron energy-loss spectra data has been extrapolated linearly from 9 eV, intersecting the abscissa at 27 eV.

Coupled oscillators behave different from the EELS spectra that demonstrate multiple oscillators that are activated simultaneously, each unrelated to the other (no phase interaction). Hence deconvoluting the above ELS spectrum of multiple excitations using Lorentz-oscillator based models is not directly possible and can only be performed one at a time..For excited states with different momentum transfers, angle-resolved EELS can be used to significantly increase the strength of the relevant peak with respect to the others. Peak extraction can also be acheived using modeling approaches. Here, the nature of the peak is identified using its decay time. By tilting the beam toward or away from Bragg condition, signals corresponding to certain excited states maybe enhanced (channeling effect), which might be applied to separate localized states such as color centers.

4 Discussion

Time resolved spectra from EELS data has been extracted using logarithmic deconvolution followed by KKA and Fourier analysis. Based on a Drude-Lorentz model and Maxwell's constitutive relations, the time dependent dielectric function was interpreted as the impulse response of the electrons, ions, or phonons to a fast passing electron beam. A Drude-Lorentz model with two oscillators was tested with varying relative strengths and damping coefficients using this technique.

The extracted time dependent dielectric function showed typical damped oscillator behavior, and the effect of the relative strengths, as well as that of the damping coefficients was consistent with the definitions of the Drude Lorentz model [equation (8)]. Because of the superior time resolution (0.138 fs for EELS with a 30 eV energy range), this time resolved technique provides opportunities for the study of ultrafast dynamics in materials, and has higher time resolution than that obtainable by current X-ray free electron lasers (Spence et al., 2012). The analysis is, however, based on a one-dimensional configuration coordinate (even when many effective oscillators contribute), unless angle and energy resolved data can be used.

EELS data from hexagonal ice and protein (Sun et al., 1993) were used to extract the electronic response to 100 keV fast electrons. It was found that the electronic oscillators are weaker and damped more rapidly in protein (3 fs) than in hexagonal ice (4 fs), which probably indicates that protein is more easily damaged within a few femtoseconds by fast electrons than hexagonal ice. This provides one estimate of the X-ray pulse duration that can be used to outrun radiation damage inside biological samples.

A similar analysis has been applied to color centers (induced by beam damage) and excitons in BaF_2. To extend this technique into the infrared region corresponding to phonons or molecular vibrational modes, the EELS data from amorphous silicon films published by Schroder et al. (1978) have also been analyzed, with an energy resolution of about 3 meV and beam energy of 30 keV, and from titanium oxide (Anatase; Geiger & Katterwe, 1976) in unpublished work. In the future, it is expected that the high sub-nanometer spatial resolution and further improved energy resolution to perhaps 10 meV for the Nion UltraSTEM (Krivanek et al., 2009),

combined with the timeresolved response technique, will be used to study local electronic or vibrational modes, such as those associated with defects in materials (dislocation cores, point defects, color centers). Despite the delocalization that limits spatial resolution to about 50 nm (Egerton 2013, personal communication), millivolt EELS compares favorably with infrared absorption spectroscopy that typically has millimeter spatial resolution.

5 Conclusions

A proof of principle has been demonstrated that the time evolution of the dielectric response function may be extracted from EELS data with sub-femtosecond time resolution. By comparison with pulsed beam methods, the CW method presented here, provides a much stronger signal, limited only by radiation damage. (The number of electrons emitted by a field emitter is a few hundred per picosecond.) It is observed that electronic oscillations in protein are damped more rapidly than in hexagonal ice. The time evolution of electronic excitations due to color centers in BaF_2 has been extracted from millivolt EELS data.

Much of the biomolecular nuclear modes of vibration are populated in the THz energies. However with high Z elements such as Mn or Fe or Co, which are commonly found as catalytic centers, the vibrational modes of the metal-ligand bond can be in the few 10s of meV as can be seen in the work by Chu et al. (*Biochemistry* 39.47 (2000): 14371-14376). In principle this would be a standout experiment with the capability of the NiON where the energy resolution is available to measure the vibrational mode in the TEM. Additionally, with the near Å angstrom spatial resolution, this study can be done by scanning a single molecule for the strong scatterer (which, in the case of an atom such as Mn would distinguish itself amongst other weakly scattering C, N and O atoms) and then measuring the EELS on a local scale. One could get more ambitious such that if the dose is controlled to below damage thresholds, it would also be possible to perform a site-specific EELS (K-edge electronic transition) on the ligand atoms bound to the metal (a capability that is unparalleled vis-a-vis XAS or IR) to study how the electronic states of the ligand accommodates itself to the reaction about the metal center.

6 References

1. Han, P. Y. *et al.* A direct comparison between terahertz time-domain spectroscopy and far-infrared Fourier transform spectroscopy. *Journal of Applied Physics* **89,** 2357–2359 (2001).

2. Zhang, W., Azad, A. K. & Grischkowsky, D. Terahertz studies of carrier dynamics and dielectric response of n-type, freestanding epitaxial GaN. *Appl. Phys. Lett.* **82,** 2841–2843 (2003).

3. George, P. A. *et al.* Ultrafast optical-pump terahertz-probe spectroscopy of the carrier relaxation and recombination dynamics in epitaxial graphene. *Nano Lett.* **8,** 4248–4251 (2008).
4. Ulbricht, R., Hendry, E., Shan, J., Heinz, T. F. & Bonn, M. Carrier dynamics in semiconductors studied with time-resolved terahertz spectroscopy. *Rev. Mod. Phys.* **83,** 543–586 (2011).

5. Fischer, B. M., Walther, M. & Uhd Jepsen, P. Far-infrared vibrational modes of DNA components studied by terahertz time-domain spectroscopy. *Phys. Med. Biol.* **47,** 3807–3814 (2002).

6. Castro-Camus, E. & Johnston, M. B. Conformational changes of photoactive yellow protein monitored by terahertz spectroscopy. *Chem. Phys. Lett.* **455,** 289–292 (2008).

7. Xu, J., Plaxco, K. W. & Allen, S. J. Collective dynamics of lysozyme in water: Terahertz absorption spectroscopy and comparison with theory. *J. Phys. Chem. B* **110,** 24255–24259 (2006).

8. Schmuttenmaer, C. A. Exploring dynamics in the far-infrared with terahertz spectroscopy. *Chem. Rev.* **104,** 1759–1779 (2004).

9. Ashkenov, N. *et al.* Infrared dielectric functions and phonon modes of high-quality ZnO films. *J. Appl. Phys.* **93,** 126–133 (2003).

10. Schubert, M., Tiwald, T. E. & Herzinger, C. M. Infrared dielectric anisotropy and phonon modes of sapphire. *J. Appl. Phys.* **61,** (2000).

11. Li, Z. Q. *et al.* Dirac charge dynamics in graphene by infrared spectroscopy. *Nat. Phys.* **4,** 6–9 (2008).

12. Zhang, S. *et al.* Experimental demonstration of near-infrared negative-index metamaterials. *Phys. Rev. Lett.* **95,** (2005).

13. Silaghi, S. D. *et al.* Dielectric functions of DNA base films from near-infrared to ultra-violet. in *Physica Status Solidi (B) Basic Research* **242,** 3047–3052 (2005).

14. Fecko, C. J., Eaves, J. D., Loparo, J. J., Tokmakoff, a & Geissler, P. L. Ultrafast hydrogen-bond dynamics in the infrared spectroscopy of water. *Science* **301,** 1698–1702 (2003).

15. Jones, A. R. *et al.* Ultrafast Infrared Spectral Fingerprints of Vitamin B 12 and Related Cobalamins. *J. Phys. Chem. A* **116,** 5586–5594 (2012).

16. Aspnes, D. E. & Studna, A. A. Dielectric functions and optical parameters of Si, Ge, GaP, GaAs, GaSb, InP, InAs, and InSb from 1.5 to 6.0 eV. *Phys. Rev. B* **27,** 985–1009 (1983).

17. Lauret, J.-S. *et al.* Ultrafast carrier dynamics in single-wall carbon nanotubes. *Phys. Rev. Lett.* **90,** 057404 (2003).

18. Lim, S. G. *et al.* Dielectric functions and optical bandgaps of high- K dielectrics for metal-oxide-semiconductor field-effect transistors by far ultraviolet spectroscopic ellipsometry. *J. Appl. Phys.* **91,** 4500–4505 (2002).

19. Nelson, F. J. *et al.* Electronic excitations in graphene in the 1-50 eV range: The π and π + σ peaks are not plasmons. *Nano Lett.* **14,** 3827–3831 (2014).

20. Johari, P. & Shenoy, V. B. Tunable dielectric properties of transition metal dichalcogenides. in *ACS Nano* **5,** 5903–5908 (2011).

21. Roth, F., Bauer, J., Mahns, B., Büchner, B. & Knupfer, M. Electronic structure of undoped and potassium-doped coronene investigated by electron energy-loss spectroscopy. *Phys. Rev. B - Condens. Matter Mater. Phys.* **85,** (2012).

22. Roth, F., Mahns, B., Büchner, B. & Knupfer, M. Dynamic response and electronic structure of potassium doped picene investigated by electron energy-loss spectroscopy. *arXiv* 13 (2011). doi:10.1103/PhysRevB.83.144501

23. Roth, F. & Knupfer, M. Electronic excitation spectrum of doped organic thin films investigated using electron energy-loss spectroscopy. *J. Electron Spectros. Relat. Phenomena* **204,** 23–28 (2015).

24. Mahns, B. *et al.* Electronic properties of spiro compounds for organic electronics. *J. Chem. Phys.* **136,** 124702 (2012).

25. Isaacson, M. Interaction of 25 keV Electrons with the Nucleic Acid Bases, Adenine, Thymine, and Uracil. II. Inner Shell Excitation and Inelastic Scattering Cross Sections. *J. Chem. Phys.* **56,** 1813 (1972).

26. Isaacson, M. Interaction of 25 keV Electrons with the Nucleic Acid Bases, Adenine, Thymine, and Uracil. I. Outer Shell Excitation*. *J. Chem. Phys.* **56,** 1803 (1972).

27. Hainfeld, J. & Isaacson, M. The use of electron energy loss spectroscopy for studying membrane architecture: A preliminary report. *Ultramicroscopy* **3,** 87–95 (1978).

28. Garvie, L. A. J., Rez, P., Alvarez, J. R. & Buseck, P. R. Interband transitions of crystalline and amorphous SiO2: An electron energy-loss spectroscopy (EELS) study of the low-loss region. *Solid State Commun.* **106,** 303–307 (1998).

29. Berger, S. D., McKenzie, D. R. & Martin, P. J. EELS analysis of vacuum arc-deposited diamond-like films. *Philos. Mag. Lett.* **57,** 285–290 (1988).

30. Krivanek, O. L., Lovejoy, T. C., Dellby, N. & Carpenter, R. W. Monochromated STEM with a 30 meV-wide, atom-sized electron probe. *J. Electron Microsc. (Tokyo).* **62,** 3–21 (2013).

31. Krivanek, O. L. *et al.* Vibrational spectroscopy in the electron microscope. *Nature* **514,** 209–12 (2014).

32. Crozier, P., Aoki, T. & Liu, Q. Detection of Water and Its Derivatives on Individual Nanoparticles using Vibrational Electron Energy-Loss Spectroscopy. *Ultramicroscopy* **169,** 30–36 (2016).

33. Rez, P. *et al.* Damage-free vibrational spectroscopy of biological materials in the electron microscope. *Nat. Commun.* **7,** 10945 (2016).

PART – 2: X-RAY AS PROBE

1 Introduction

1.1 B_{12} Biochemistry and Catalysis:

B_{12} coenzymes are medium-sized molecules (about 1300 Daltons) whose reactivity is enhanced and specific to the apoenzymes that they work with. (A coenzyme is a low molecular mass adduct that determines the type of reaction, while an apoenzyme is a high molecular mass protein that determines the substrate specificity and reaction rate. A holoenzyme is the complete and fully functional combination of the two. Refer fig. 5.1 for a cartoon showing the same.)

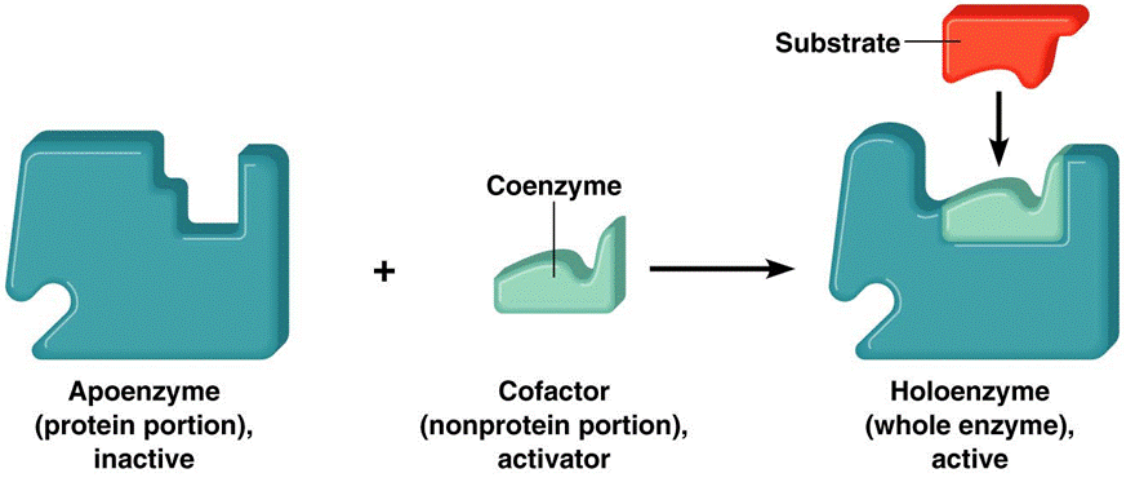


Fig. 5.1: Schematic showing the Apoenzyme, Cofactor and Substrate in an enzymatic process. (Figure courtesy: http://classes.midlandstech.edu/carterp/courses/bio225/chap05/lecture2.htm)

The pathway for absorption, transport and cellular uptake of cobalamins is presented in fig. 5.2. Briefly, the cobalamins are bound by saliva through the food taken in by humans and enter the cell with the help of multiple transport proteins to be finally released as a cob(I)alamin square planar complex.

The metabolic role of vitamin B_{12} (cobalamin, Cbl) is targeted to two pathways – MeCbl (Methylcobalamin), the cofactor of methionine synthase in the cytoplasm and AdoCbl (Adenosylcobalamin), the cofactor of methylmalonyl coenzyme A mutase in mitochondrion.

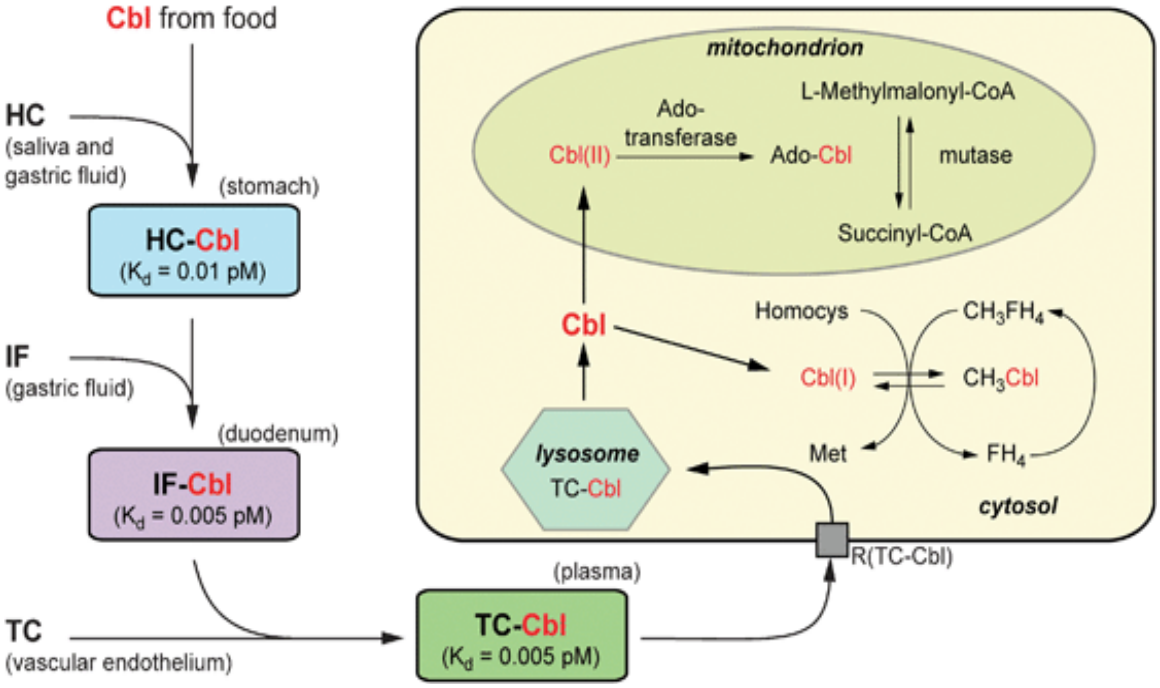


Fig. 5.2: Schematic outline of cobalamin uptake and transport in mammals. (Figure courtesy: Chem. Soc. Rev., 2011, **40**, 4346-4363). Here: R(TC-Cbl) = receptor that mediates cellular uptake of the TC-Cbl complex; Homocys = homocysteine, Met = methionine, FH_4 = tetrahydrofolate, CH_3FH_4 = methyltetrahydrofolate.

The catalytic activity of transferring a methyl group is important in human, animal and bacterial metabolism and is carried out by methyl-corrinoids bound by their corresponding enzymes. B12-dependent, enzyme controlled methyl group transfer reactions are key steps in methylations of homocysteine to methionine. The catalytically active corrinoids in such methyl transferases are protein-bound methyl-Cob(III)alamin with a de-coordinated nucleotide base, and Cob(I)alamin. This formation of methionine by methylation of homocysteine is important in human, animal and bacterial metabolisms. Such a catalysis (represented in fig. 5.3) that aids the transport of the methyl group occurs by a sequential two-step mechanism: (1) abstraction of the cobalt-bound methyl group of MeCbl by homocysteine with the formation of methionine and of an enzyme-bound cob(I)alamin and (2) abstraction of the methyl group from mono/poly glutamate forms of CH_3-THF by the highly nucleophilic cob(I)alamin. In this way, the catalytic cycle is

completed and MeCbl is reconstituted. Synthesis of THF affords the generation of purines and pyrimidines, eventually required for DNA and RNA synthesis.

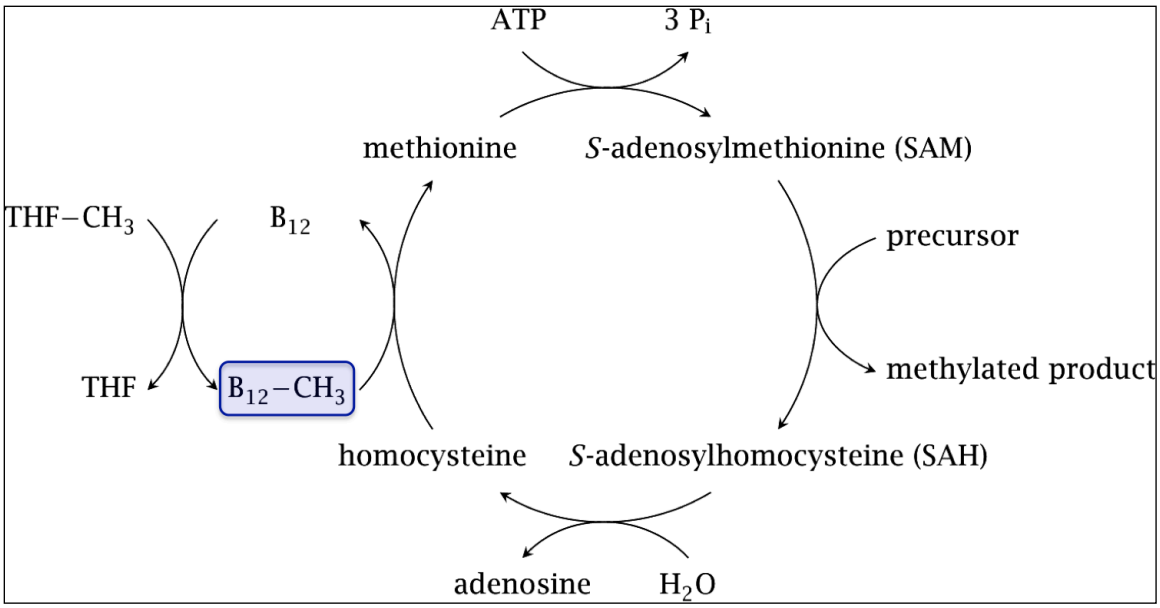


Fig. 5.3: Role of MeCbl in methyl-transfer catalysis. (courtesy: http://watcut.uwaterloo.ca/webnotes/Metabolism/C1CobalaminMethylations.html#).

1.1.1 B_{12} Deficiency and Diseases:

The human liver can store vitamin B_{12} that last a few years. However, there are still B_{12} related diseases that are seen to affect humans. One such case is the lack of intrinsic factor (IF) which results in the malabsorption of the B_{12} vitamin in the stomach. In such a case of malabsorption, the subsequent deficiency of the vitamin results in deficiency in methionine synthase which blocks the folate metabolic pathway and results in the accumulation of the N^5-CH_3-THF derivative. The absence of (non-methylated) THF prevents the formation of the other THF derivatives that are required for the purine and thymidine biosynthesis pathways and thereby impairing the synthesis of DNA, which is identified as pernicious (megaloblastic) anemia.[1,2]

Continuing with the methyl transferase activity, deficiencies in vitamin B_{12} leads to elevated levels of homocysteine. However, homocysteine is known to be highly reactive and thus binds (thio-ligation) to proteins and potentially causing their degradation. Homocysteine also binds to albumin and hemoglobin. When homocysteine binds to lysyl oxidase, an enzyme responsible for proper maturation of the extracellular matrix proteins collagen and elastin, there is an observed detrimental effect. Production of defective collagen and elastin has a negative impact on arteries, bone and skin and the effects on arteries are believed to be the underlying cause for cardiac dysfunction associated with elevated serum homocysteine. [3,4] The increased risk for thrombotic episodes associated with homocysteinemia, such as deep vein thrombosis, is due to homocysteine serving as a contact activation nucleus for activation of the intrinsic coagulation cascade.

Inside the mitochondrion relating to the other catalytic pathway, Methylmalonyl-CoA acts as a competitive inhibitor (or even substitute) to malonyl-CoA in fatty acid biosynthesis. The myelin sheath that is in a continual flux is thus affected by the Methylmalonyl-CoA induced inhibition, resulting in branched-chain fatty acids that can severely modify the architecture of the normal membrane structure of the nerve cells. Deficiency in vitamin B_{12} directly correlates to the increased quantities of Methylmalonyl-CoA which results in the demylenation. Hence B_{12} deficiencies are also associated with neurological complications. [5,6]

Consequently, MeCbl (and cobalamins in general) are rigorously tested as potential therapeutic drugs to ameliorate a variety of diseases and related syndromes: promoting the regeneration of injured nerves, improving nerve conduction in patients of diabetic neuropathy and even treating/slowing cognitive decline arising from Alzheimer's disease. [1,2,7]

1.2 B_{12} Chemistry

(This section is majorly inspired in content from Kaim, Wolfgang, Brigitte Schwederski, and Axel Klein. *Bioinorganic Chemistry --Inorganic Elements in the Chemistry of Life: An Introduction and Guide*. John Wiley & Sons, 2013)

1.2.1 The Corrin Ring

Tetrapyrroles are a class of chemical compounds that contain four pyrrole rings held together by direct covalent bonds or by one-carbon bridges (=(CH)- or $–CH_2$-) units), in either a linear or a cyclic fashion. A pyrrole is a five-atom molecule where four of the ring atoms are carbon and one is nitrogen. Cyclic tetrapyrroles are at least partially unsaturated tetradentate macrocyclic ligands which in their deprotonated forms, can tightly bind metal ions through the lone electron pairs on nitrogen atoms facing the center of the macrocycle ring.

Cyclic tetrapyrroles having three/four one-carbon bridges include (Fig. 5.4):

- Porphyrins, including heme, the core of hemoglobin (Fig. 5.4a).
- Chlorins, including those at the core of chlorophyll (Fig. 5.4b).
- Corrins, including the cores of cobalamins, when complexed with a cobalt ion (Fig. 5.4c).

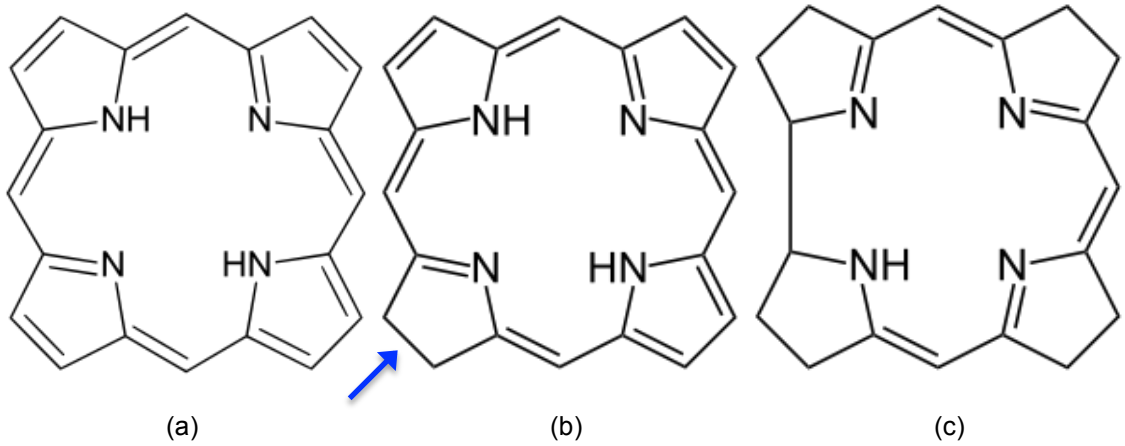


Fig. 5.4: (a) Porphyrin macrocyle (https://en.wikipedia.org/wiki/Porphyrin), (b) Chlorin macrocyle (https://en.wikipedia.org/wiki/Chlorin), (c) Corring macrocyle (https://en.wikipedia.org/wiki/Corrin)

Some of the best-known bio-inorganic compounds contain tetrapyrrole macrocyles –

(i) Chlorophylls - Mg is the central atom and reduced (partially hydrogenated and substituted) porphyrin – chlorin ring, is the cyclic tetrapyrrole.

(ii) Cobalamins – Co is the central atom and a partially conjugated corrin ring (one ring member less than a prophyrin) is the form of the cyclic tetrapyrrole ligand.

(iii) Hemoglobin, Myoglobin, Cytochromes and Peroxidases – Fe is the central atom and a substituted porphyrin is the cyclic tetrapyrrole ligand form.

(iv) Siroheme - Fe is the central atom, an isobacterichlorin ring as the cyclic tetrapyrrole ligand.

There are quite a few characteristics of these tetrapyrrole rings and they are presented below –

(i) The (near) planar structure of the ring is very stable (the bond lengths, angles and torsional angles are in the range generally encountered with sp^2-hybridized carbon and nitrogen centers).

(ii) Tetradentate chelate ligands, after deprotonation carrying single/double negative charge, can bind labile metal ions (A compound in which metal-ligand bonds are easily broken is refered to as "labile").

(iii) These macrocyclic ligands have a cavity that is highly selective about the sizes of the coordinated metal ion. The conjugated double bonds make the macrocyclic ligands very rigid and typically metals with ionic radii in the range of 60 – 70 pm are best suited to fit into the central cavity.

(iv) Cyclic tetrapyrroles have intense absorption bands in the visible electromagnetic radiation regime thanks to the substantially conjugated π system. Also, the uptake (reduction) and release (oxidation) of electrons by the conjugated π system are facilitated by the narrowing/spreading of the π orbital gap thereby stabilizing the resulting anion/cation. These two characteristics render the tetrapyrrole macrocyles as essential in important biological energy transformations such as photosynthesis and respiration.

Tetrapyrroles are tetradentate ligands (single molecular entity that contributes four ligating sites to the same metal ion). For an octahedral arrangement that requires six coordinating ligands, the tetrapyrrole leaves two axial coordination sites available. This makes an octahedral arrangement favourable amongst the macrocycles for catalytic activity by utilizing the two free ligating sites – one for binding the substrate and another for the regulation of the catalytic activity. A few examples are presented that highlight the useful functions of these open axial ligating positions in biologically relevant tetrapyrrole complexes.

(a) In hemoglobin, molecular Oxygen that is reversibly coordinated for transportation to tissues, is the substrate while an amino acid such as histidine coordinated in the opposite axial position is functionally useful.

(b) Chlorophylls utilize multiple coordinative interactions of the magnesium ions to form aggregates in a controlled fashion as seen in antenna pigments or in the photosynthetic reaction centers. While the Mg^{2+} centers act as bifunctional acceptors for electron-pair containing ligands, the carbonyl groups of the chlorophyll molecules function as the electron-pair donors.

(c) Cobalamins contain primary alkyl groups coordinated directly to the Cobalt center. The cleavage of the metal-alkyl bonds to yield radicals is possibly affected by changes in the lower axial coordination.

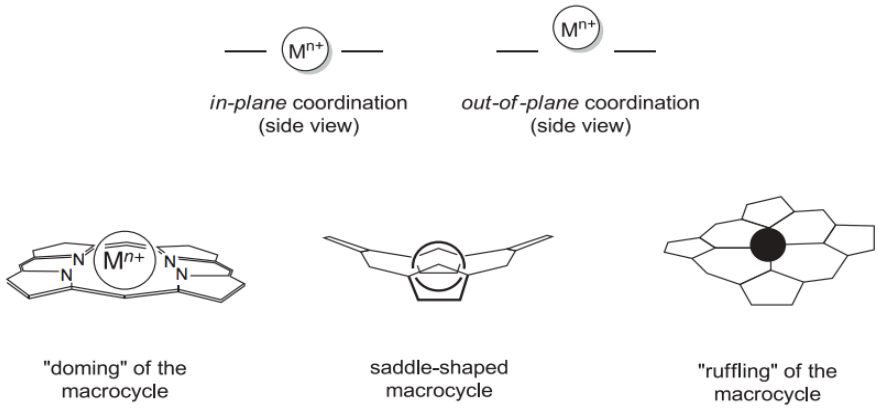


Fig. 5.5: Typical geometrical deviations for complexes of tetrapyrrole macrocycles. (figure courtesy: fig. 2.9 Kaim, Wolfgang, Brigitte Schwederski, and Axel Klein. *Bioinorganic Chemistry - -Inorganic Elements in the Chemistry of Life: An Introduction and Guide*. John Wiley & Sons, 2013).

Octahedral symmetry of the overall complex can be affected by tetragonal distortion because of the strong di-anionic tetrapyrrole ligands that cause a splitting of the d-orbitals of the coordinating transition metal (specifically relevant to the case of the cobalamins). Additionally, the equatorial ligand field strength of planar tetrapyrrole di-anions stabilize the low-spin configurations of the metal centers rather than their high-spin equivalents. However, there are examples in the case of deoxy-hemoglobin and deoxy-myoglobin which feature a high-spin Fe(II) center.

High spin Fe^{2+} is known to have a relatively larger ionic radius, which implies that the metal does not fit into the macrocycle cavity. Instead, there is an out of plane formation of the complex resulting in a reduced ligand-field effect from the porphyrin macrocyle. (Fig. 5.5 shows all the different geometrical variations of tetrapyrrole complexes). A very similar effect is also seen with Ni^{2+} in coenzyme F430. For a detailed study of the various complex formations refer to the study by Munro et.al. [8]

1.2.2 Reduction and Oxidation

Cobalamins in their ground state have a octahedral coordination (fig. 5.6).

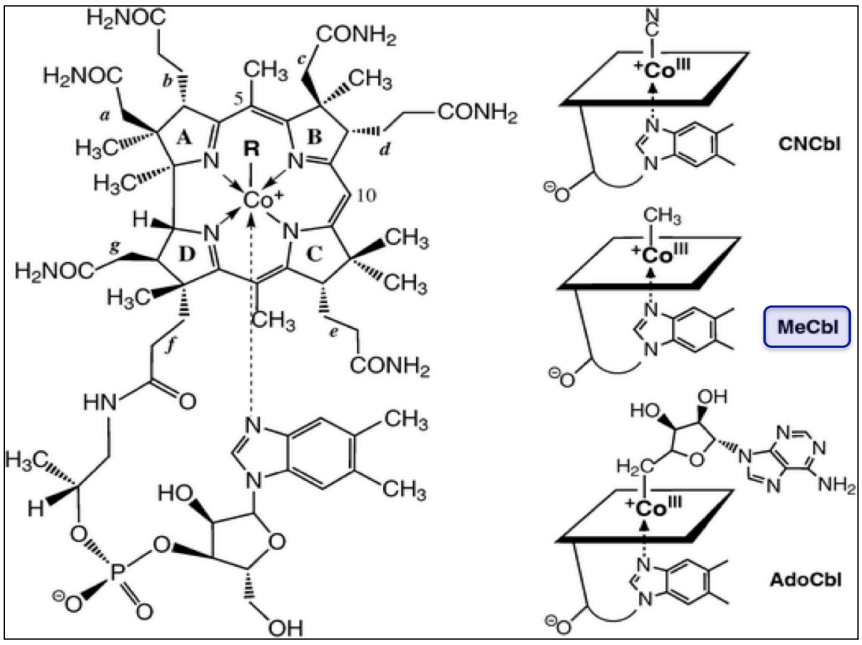


Fig. 5.6: Structure of the cobalamins and the different axial ligands. (figure courtesy: DOI: 10.1039/C1CS15118E)

The equatorial plane consists of the corrin ring and the lower axial spot bound by a dimethylbenzimidazole (DBI) that links back to the ring. The upper axial spot is open to be ligated through multiple options, a few of which are presented in fig. 5.6. When the upper axial ligand is a methyl group, the resulting complex is the Methylcobalamin. In this configuration, the cobalt ion is trivalent (d^6) with the three counter ions offered by the corrin mono-anion, an axial anionic group (in this case methyl) and the negatively charged phosphate that connects the neutral DBI base to the ring.

The kind and number of axial ligands tend to determine the redox potentials corresponding to the one-electron reduction steps that are possible when starting from a 6-coordinate structure. Such a stepwise one-electron reduction of the metal (from a d^6 to a d^8 configuration) is noticed to be accompanied by a tendency towards decreased axial coordination (as in fig. 5.7 below).

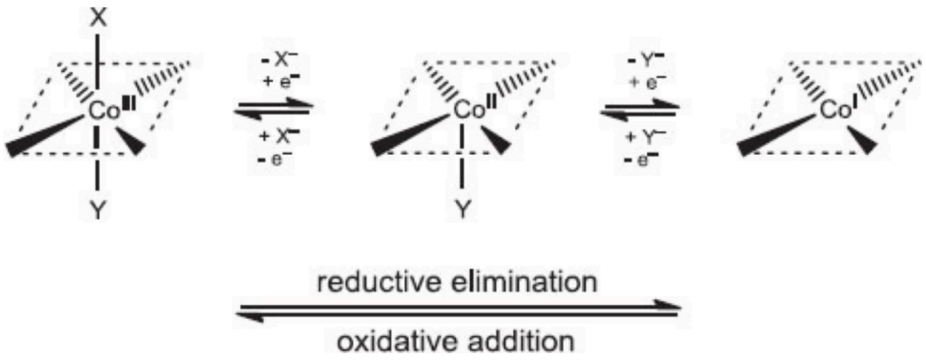


Fig. 5.7: Oxidation and Reduction and the corresponding modifications in coordination of the cobalamin complexes. (figure courtesy: 3.3 Kaim, Wolfgang, Brigitte Schwederski, and Axel Klein. *Bioinorganic Chemistry --Inorganic Elements in the Chemistry of Life: An Introduction and Guide*. John Wiley & Sons, 2013).

One-electron reduction of Co(III)-methylcobalamin results in a decreased Co-C bond strength and corresponding rate enhancement for the bond-cleavage due to a half-filled antibonding $\sigma^*(Co\text{-}CH_3)$ orbital (d_z^2 component). For a d^8 system such as Co(I), the sterically less favorable square-planar configuration is favored due to the non-occupation of the strongly antibonding $d_{x^2-y^2}$ orbital. Incidentally, the stabilization of the Co(I) state is an interesting characteristic of the Cobalt-corrin system.

1.2.3 Co-C Bond Cleavage

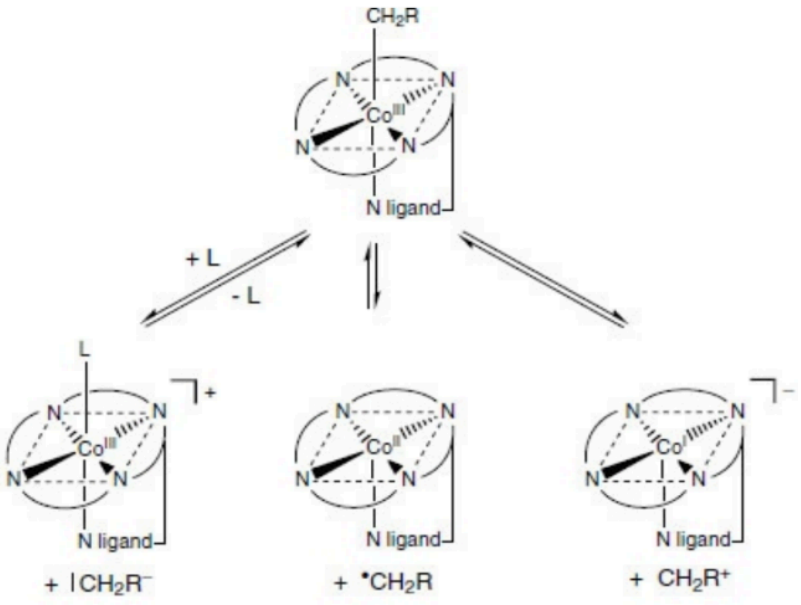
The reactivity of alkyl cobalamins is characterized by the fact that the reactive alkyl groups are made available in a 'controlled' fashion for follow up reactions. There are two formal options for a Co-R bond cleavage (a) heterolytic (which signifies both the electrons forming the bond going one way, to one atom or the other, upon cleavage) and (b) hemolytic (the electrons forming the bond are shared one each to each participating atom upon cleavage, which may be induced by the interaction of the coenzymes with apoprotein and substrate.

With respect to the Cbls, heterolytic bond cleavage can be further divided into two sub-categories. The first is when both the bond electrons upon cleavage are withdrawn by the methyl ligand hence forming a low-spin Co(III) and a carbanion ($^{-}CH_3$) (involving substitution by water). The other can lead to Co(I) and a carbocation ($^{+}CH_3$) in which case both the bond electrons retract to the cobalt atom. The latter is a d^8 configured metal center. This d^8 Co(I) is a σ-electron rich supernucleophile (it has a high affinity towards σ electrophiles because of the filled antibonding d_z^2 orbital) and has a typical behavior of 'oxidative addition'. The carbanion and carbocations do not float as free ions. In the presence of a reaction partner or a polar solvent (water), they are transferred in the transition state of the reaction. The third alternative formation of a paramagnetic Co(II) in a low-spin d^7 configuration and a primary alkyl radical is the homolytic bond cleavage in which the bonded electrons are split one to each fragment.

As in fig. 5.8, all the three alternatives are possible. The nature of the product is decided by the axial coordination or non coordination (base-off) of the heterocyclic ligand, the nature of the substrate and the redox potential of the reaction pathway. If the axial base is not present, the carbanionic mechanism is realized at potentials greater than 0V (vs. the normal hydrogen electrode NHE). For the carbocation Co(I) cleavage to occur, the potentials need to be lower than -0.9V. Homolysis resulting in the single unpaired electron and the methyl radical is a plausible solution in the potential range of 0V to -0.4V, which is also the physiologically relevant regime of electrochemical potentials.

1.3 Why the Curiosity?

The homolytic cleavage of the Co-C bond of AdoCbl is enhanced by roughly 10 orders of magnitude in the presence of the substrate protein. [9] Similarly, the heterolytic Co-C bond cleavage of MeCbl in the presence of methionine synthase is enhanced by 5 orders of magnitude. [10] Naturally, the question is to identify the reason for this substantial increase in reactivity. More specifically put, the first question of interest is – How/Why does the Cobalt-Carbon bond weaken permitting cleavage?



type of reaction:	heterolysis	homolysis	heterolysis
metal configuration in the product	d^6 low-spin, stable, inert	d^7 low-spin, 1 unpaired electron $(d_{z^2})^1$	d^8, "super-nucleophilic" $(d_{z^2})^2$
alkyl ligand, eliminated as:	"carbanion", nucleophilic	primary alkyl radical, very reactive	"carbocation", electrophilic
app. electrochemical potential equivalent[a]:	> 0 V	0 to -0.4 V	< -0.9 V

[a]In biochemistry, all redox potentials are generally referenced to the normal hydrogen electrode (NHE)

Fig. 5.8: Different alternative routes for Co-alkyl bond cleavage and the nature of the reaction, configuration of the product and energetics. (figure courtesy: 3.4 Kaim, Wolfgang, Brigitte Schwederski, and Axel Klein. *Bioinorganic Chemistry --Inorganic Elements in the Chemistry of Life: An Introduction and Guide*. John Wiley & Sons, 2013).

The trans-effect is referred to the weakening of a ligand (bond) in the trans position to another specific ligand. A representative figure (fig. 5.9) is shown below highlighting the difference between the cis and the trans effect.

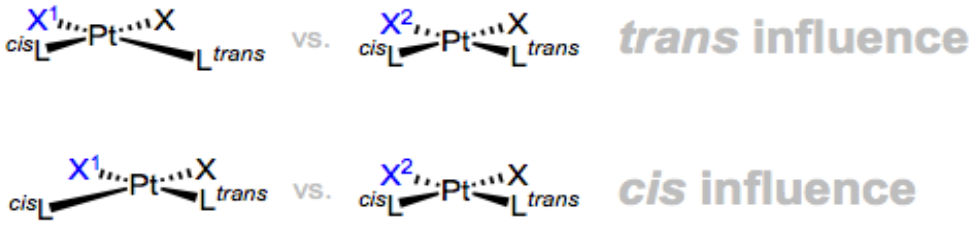


Fig. 5.9: The cis-effect vs the trans-effect in organometallic chemistry. (figure courtesy: https://organometallicchem.wordpress.com/2012/05/01/the-transcis-effects-influences/)

With fig. 5.9 as reference, it can be seen that in the presence of a central metal atom (in this case Platinum, Pt) the modification of the ligand on the top left from X^1 to X^2 creates a response in the coordination complex. In the case of a cis-influence, the ligand adjacent to the X^1(X^2) at the bottom left position is modified in bond-length (and hence strength). The trans-effect refers to the case when the ligand on the other (sometimes opposite) end of the X^1(X^2) ligand, in this case bottom right, is modified in bond-length (strength). For the case of an octahedral coordination compound (such as Vitamin B_{12}), the trans-effect can be visualized as the effect of on axial ligand (say lower axial) on the other (upper axial). There are different types of trans-effects, a couple of commonly noted effects are (a) steric and (b) electronic.

In the case of both MeCbl and AdoCbl, there have been a strong temptation to attribute the exorbitant increase in reactivity of either of these cofactors in the presence of the substrate proteins to the trans-effect (changing the lower axial base from DBI to histidine, as is for MeCbl, would lower the Co-C bond strength thereby making it easier to cleave the bond). [11–13] Several research attempts tried to study this relation as with the study of Krautler, 1987. However, bond stretching studies using IR and Raman [14–16] did not show much validity in such a claim. Additionally, they also show that the catalytic contribution of the presence of a modified lower axial ligand is possibly represented much after the ground state.[14,16]

An additional question that arises is how/why is the reactivity and chemistry behind the Co-C bond in AdoCbl and MeCbl different. (As a reminder, AdoCbl is expected to undergo homolytic bond cleavage and 10 orders of magnitude increase in reactivity in the presence of the protein, while MeCbl is expected to undergo heterolytic bond cleavage with enhancement of up to 4 orders of magnitude in reactivity, in the presence of the substrate protein).

1.4 Prior Experiments

One of the experimental needs in trying to understand chemical reactions is to probe before, after and during the reaction. The before and after portions of the study are relatively straightforward to measure using a variety of techniques termed as a 'steady-state' measurement. Studying systems during a reaction, however, is not as straightforward. In the case of studying the functional behavior of MeCbl, there are two broad approaches that can be pursued – (i) Diffraction: (a) Crystallize the cofactor and the protein substrate and study the reactions through X-ray diffraction (crystallography) by studying the reaction in situ. (b) Flash freeze the cofactor and substrate during the reaction and different stages and study the X-ray structure as a pseudo steady-state measurement. (ii) Optical pump Spectroscopy: Optical-pump at a specific wavelength to excite the cofactor (independently or with the substrate) and then use a probe (say a white optical probe) to study the response to the electronic excitation. The advantage of the former is that the entire system is captured during the reaction, except that it is forced to be a crystalline and often not exactly similar conditions as naturally present in body. Thus the compromise here is to take it out of its natural habitat and occurrence. The latter retains the samples in their natural environment. The caveat however is that the reaction is initiated by an electronic excitation, which is typically further from real (exceptions include photoactive or photosynthetic proteins such as PSII, PYP). In the case of MeCbl, the reaction is driven by a redox potential of the interacting species in the human body and optical excitation is a very crude approximation of the reaction. Nevertheless, this technique is common and popular thanks to high accuracy and ultra short timescales that can be probed. In the forthcoming paragraphs, a brief literature review is presented of certain select types of measurements.

1.4.1 Photolysis and Optical Absorption

Photolysis of Methylcobalamin (and other B_{12} cofactors) have been carried out very early into the 1950s. The preliminary studies focused on identifying the nature of products formed in exposure to sunlight or light of equivalent wavelength ranges. Hogenkamp (1966) performed one such important work where the photodecomposition of MeCbl was studied under aerobic, mildly aerobic and anaerobic conditions. [17] The work highlighted that under aerobic conditions the photolysis yielded aquocobalamin and formaldehyde while under mildly aerobic conditions, the reaction is slowed forming aquocobalamin, formaldehyde and the non oxidized products of ethane and methane. The oxidized products were attributed to have formed through radical coupling, while the non-oxidized products were attributed to hydrogen/methyl abstraction. The experiments were revisited by Schrauzer et al. (1968) who also extended the scope to include other alkylcobalamins and alkylcobaloximes. [18] In their work, they note that the anaerobic photolysis produces the same set of products. However, the non-oxidized products, ethane and methane, were noticed to be formed predominantly by methyl radical dimerization and ligand hydrogen abstraction or methyl radical reduction respectively. They also noticed that higher alkyl-cobalamins yield unsaturated and un-oxidized products, while modifying the axial base to strongly interacting groups such as CN^- completely blocked the production of ethane altogether. In 1973, Taylor et al. came up with the first experiments that measured photolysis of alkylcobalamins as a function of incident wavelength and pH. All prior work had been with white light illumination. [19] They established that in general, most alkyl-Cbls had product quantum yields that were smaller in the base-off form (acidic pH = 1) as opposed to the base-on form (neutral pH = 7) except for MeCbl. MeCbl had quantum yields at acidic pH that were similar to neutral pH below 450 nm wavelength and higher beyond. Also, the photolysis rates for MeCbl were wavelength dependent with faster photolysis at neutral pH when illuminated at 510 nm as opposed to 450 nm when the reverse was observed.

In 1979, Endicott and Netzel made one of the earliest attempts to perform transient absorption to study flash photolysis of MeCbl and AdoCbl. They note bleach at 565 nm with an increased absorbance at 474 nm in the few ps regime, which over a few 100 ps goes through an

increased absorbance at about 565 nm and bleach at 474 nm. Eventually, constant absorbances were observed for 1ns and longer timescales. [20] In the early 90s, the group of Mark Chance worked on the cobalamins in both the optical and the X-ray fields. One of their most important works analyzes the quantum yields of various cobalamins with continuous wave excitation, which shows that MeCbl is less sensitive than AdoCbl to the nature of solvent and hence the solvent caging when comparing experiments performed in water and ethylene glycol. [21] Geminate recombination (recombining of the radical Co(II) and CH_3 pair) is noticed to be at a comparable rate to the cage escape. In a related work, Lott et al. (1995) identified that such geminate recombination in alkylcobalamins is dependent on the geometry of the alkyl radicals. In MeCbl, they see the recombination to be reduced but yet active, and in combination with some magnetic field measurements [22,23] note that the possible reason for the difference in behavior in comparison to AdoCbl is a diffusion based recombination as opposed to inter-system crossing (triplet-singlet) or geometry of the radicals. [24]

In transitioning from steady state to transient optical absorption, a substantial amount of the work in time-resolved pump-probe optical absorption has been done by the group of Prof. Roseanne Sension from University of Michigan. In 1997, they came out with an article that established the character of the metastable intermediate of photolysis of MeCbl as having a Co(III) characteristic. They also established in this article, the division between prompt homolysis and metastable photoproduct and being roughly 1:3. [25]

The transient (time-resolved) absorption is an experimental technique that uses an optical laser to pump excite the sample and set it into the reaction. Subsequently, a broadband (white light) probe is used to obtain the optical response of intermediate species by probing at fixed time-delays post the pump excitation. In a transient difference absorption spectrum, the difference between the excited (intermediate) and the ground state optical absorption is displayed. Prior works on the cobalamins were performed to study the nature of the reaction intermediate in the photolytic Co-C bond cleavage across the range of ultrafast timescales and are presented in figs. 5.10 and 5.11.

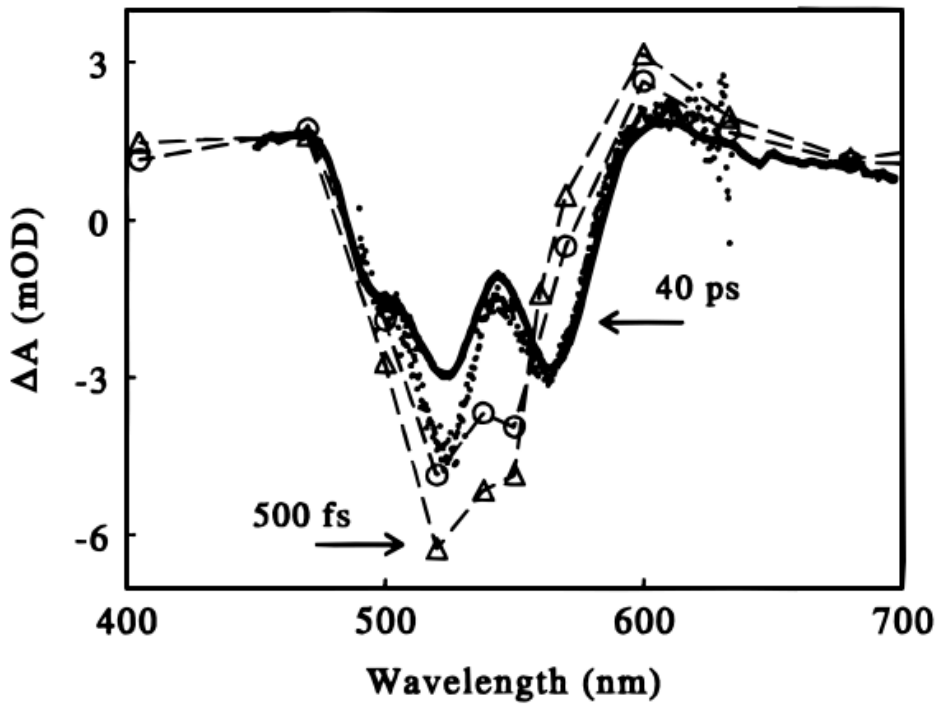


Fig. 5.10: Transient difference absorption spectrum in the photolysis of MeCbl from 500 fs - 40 ps. 500 fs ⇔ Triangles, 1 ps ⇔ Open circles, 5 ps ⇔ dotted line, 40 ps ⇔ solid line (figure courtesy: Walker et al., J. Am. Chem. Soc., Vol. 120, No. 15, 1998, figure 4)

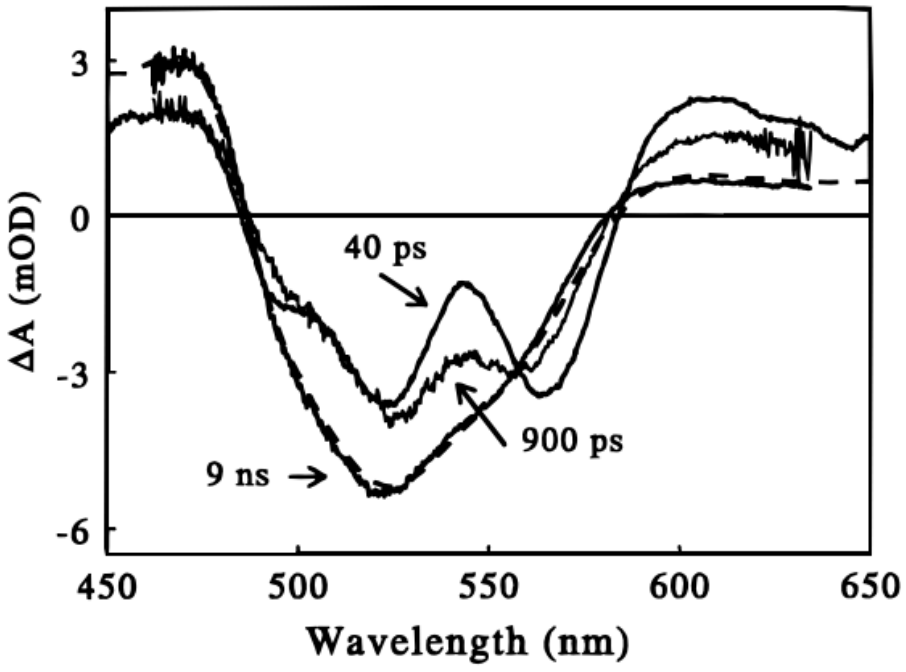


Fig. 5.11: Transient difference absorption spectrum in the photolysis of MeCbl from 40 ps - 9 ns. (figure courtesy: Walker et al., J. Am. Chem. Soc., Vol. 120, No. 15, 1998, figure 5

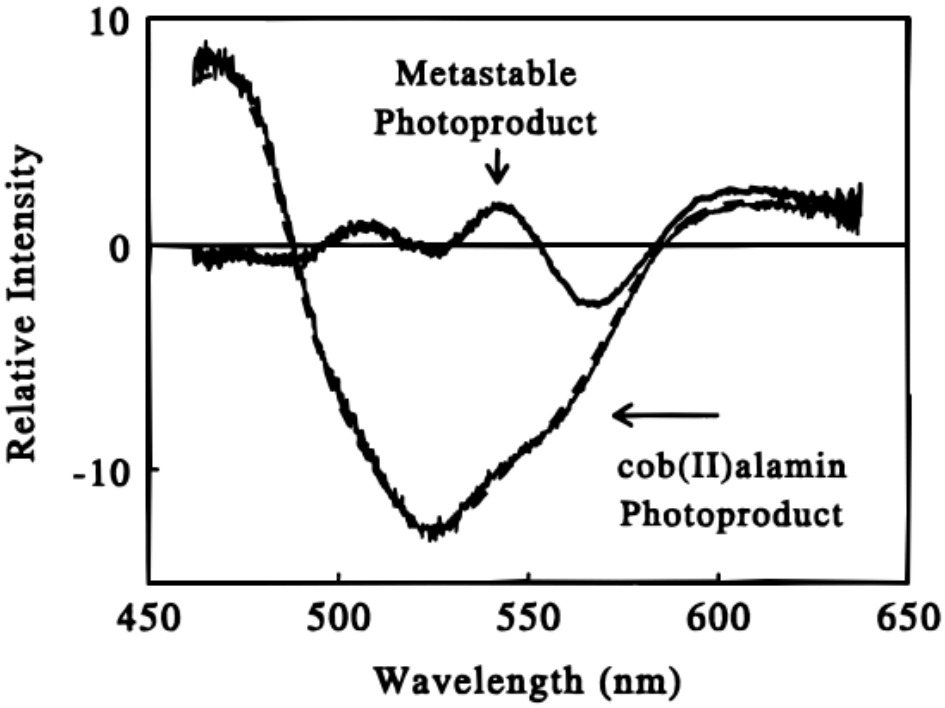


Fig. 5.12: Species associated difference spectra after SVD. Dashed line is the steady state difference spectrum. (figure courtesy: Walker et al., J. Am. Chem. Soc., Vol. 120, No. 15, 1998, figure 7)

It can be observed that at about 9 ns (solid line), the spectrum configures to the bleach consistent with steady-state measurements describing a homolytic bond-cleavage (dashed line) with a Co(II) final product.

Subsequently, the group performed addition experiments of the same type, this time varying the incident wavelength. They identify a crucial difference between AdoCbl and MeCbl which is that the photolysis of MeCbl is wavelength dependent as opposed to AdoCbl which is not. [26] To summarize the results, they observe that laser excitation at 400 nm results in partitioning between prompt hemolysis (25%) and metastable product formation (75%). On the other hand, excitation at 520 nm does not account to prompt homolysis at all and forms 100% metastable product. This metastable product is then seen to revert back majorly to ground state (approx. 85%).

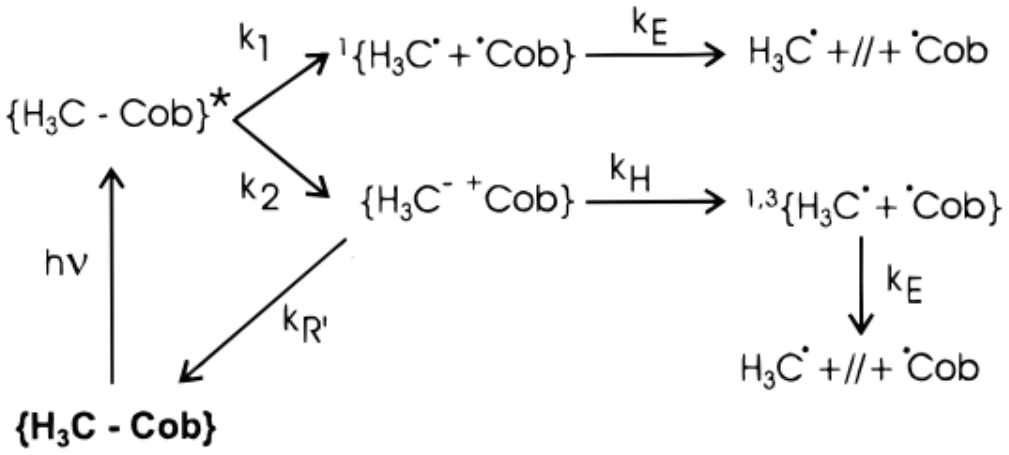


Fig. 5.13: Schematic of the Photolysis of MeCbl with excitation at 400 nm. k_1 corresponds to the prompt homolysis while k_2 corresponds to metastable product formation. (figure courtesy: Shiang et al., J. Phys. Chem. B 1999, 103, 10532-10539, figure 10).

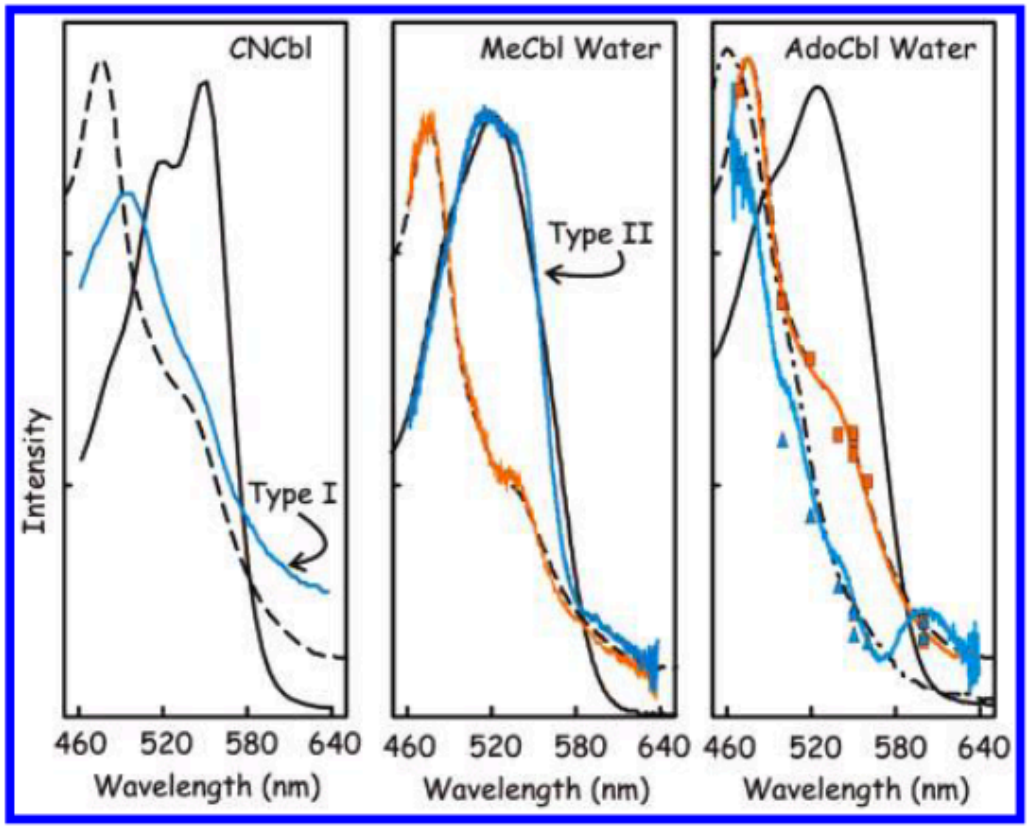


Fig. 5.14: Absorption spectra of intermediate states during the photolysis of CNCbl, MeCbl and AdoCbl. Black solid line is the ground Co(III) state. Blue solid line is the intermediate state. Orange solid line is the product. Black dotted line is the chemically reduced Co(II) state. (figure courtesy: Harris et al., J. Am. Chem. Soc. 2007, 129 , 7578-7585)

The group then progressed to study the effect of solvent on the photolysis mechanism and geminate recombination for multiple alkyl-cobalamins. [27] They note that the lifetime of the excited state is longer by a factor of two than that in water for all the cobalamins. They propose that this might be due to the variation in solvent polarity and hence the solvent friction on the reaction. Regarding recombination within the solvent cage, they observe no significant variation in the geminate recombination rates. Finally they mention that the cage escape in MeCbl being substantially larger than all the other alkyl-cobalamins is probably because of the dissipation of excess energy deposited by the 400 nm incident photon. In a related work, [28] the group extends the understanding of the relation of the solvent and axial ligand type to the electronic structure of the metastable intermediate state. In addition to the increase in lifetime with ethelyne glycol that they noted earlier, they note that the spectral features are independent of the alkyl group.

They characterize the excited-state spectrum (fig. 5.14) as very similar to the ground state Cob(III)alamin spectrum for MeCbl. The only modification is that of a similarity to a carbanion type species of the excited state rather than a covalently bound species. This, in turn, complements the schematic presented in fig. 5.13 where the metastable state is composed of a Cobalt cation and a methyl anion. This state is recognized as a MLCT (metal to ligand charge transfer) state with the charge transfer from the Cobalt to the alkyl ligand ($3d_z^2$ to $2p_z$). (In nonalkyl-cobalamins the reverse, an LMCT (ligand to metal charge transfer) is observed). Fig. 5.14 shows the absorption spectra for the ground state, excited state and the final product state. From the central panel, the intermediate state spectrum clearly overlaps with the ground state spectrum confirming the Co(III) assignment.

More recently in 2010, the group performed the same experiments on base-off (lower axial bond detached) B_{12} cofactors. [29] (Fig. 5.15 shows the ground state pH7 and pH2 spectra for a select few Cbls.) They notice that the absence of the lower axial nitrogen modifies the electron structure thereby opening the reaction to a rapid non-radiative decay. For MeCbl this implies a competition between reforming the ground state and the final Co(II) product which starkly differentiates itself from the competition between cage escape and geminate recombination for the base-on form.

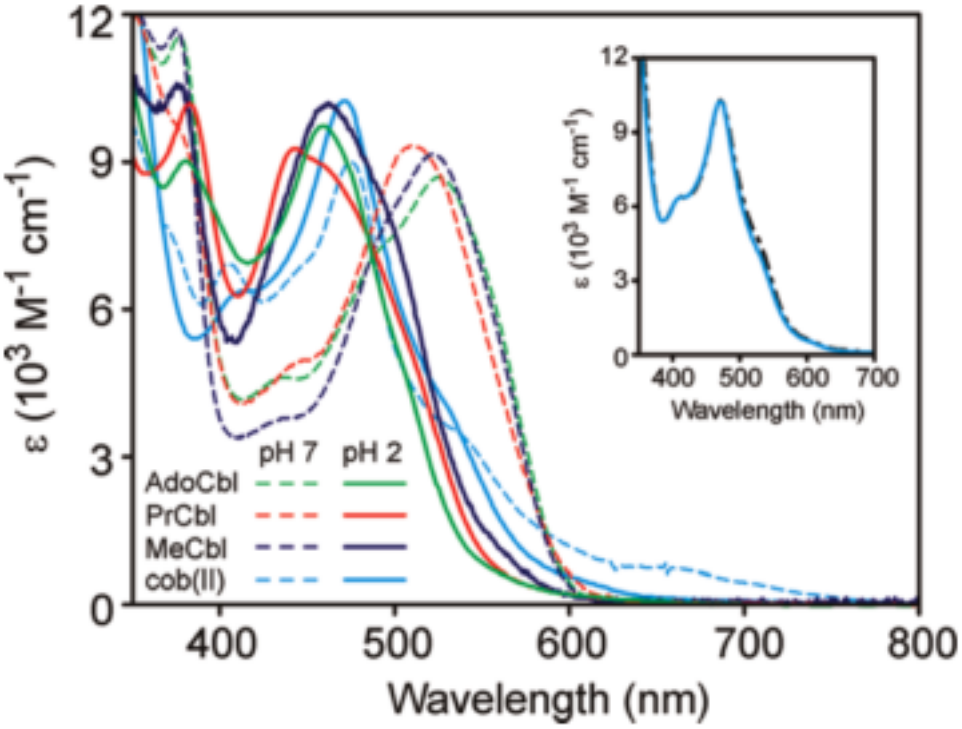


Fig. 5.15: Absorption spectra of AdoCob/PrCob/MeCob/Cob(II)alamins in pH7 (neutral, base-on) and pH2 (acidic, base-off) ground state. (figure courtesy: Peng et al., J. Phys. Chem. B, Vol. 114, No. 38, 2010)

Although not directly related to these experiments, in 2012 there was a study of the vibrational fingerprints of post optical pump excitation of the various Cbls (Me-, Ado-, OH-, CN-). In this work by Jones et al. [30] They notice that although many of the spectral fingerprints are similar between MeCbl and AdoCbl, substantial differences occur in the certain vibrational relaxation modes which shift to substantially higher frequencies (anywhere between 6 – 24 cm^{-1}) in the few 100 ps regime which is attributed to the strong interactions of the adenosyl radical post cleavage with the coring ring. With MeCbl, there are weaker shifts notices (3-5 cm^{-1}) at about 1650 cm^{-1}, which is then attributed to the amide groups in the periphery of the ring. Overall, the kinetics of the relaxation however are noticed to be broadly in the same timescales as with the UV-Vis measurements. [25] This work was done with excitation at 525 nm and hence the entire MeCbl population is considered converting to the metastable intermediate.

1.5 Theoretical Investigations:

There have been quite a few theoretical investigations on the photolysis of MeCbl. The Kozlowski group have been involved in working out the DFT based electronic structures for both alkyl [31–36] and non-alkyl cobalamins [31,35,37–40] since early 2000s. Thomas Brunold's group [41–45] have also contributed in the theoretical investigation of Cobalamins. In this section, however, only the latest work by Lodowski et al. (2014) in [46] will be considered as the most relevant and updated and will be reviewed. Regular convention in DFT calculations is to name the ground state of the molecule as S_0 and the electronic excited state as S_i (I = 1 … n) in the order of increasing energy. Indeed, being electronic excited states, they also contribute to the excitations that form the optical absorption spectrum of the ground state of a molecule.

In the above mentioned work, Lodowski et al. perform TD-DFT to study the S_1 stable intermediate (electronic first excited state) during the electronically (optically) excited photolysis of MeCbl. This was done by generating a potential energy surface (PES) as a function of two reaction coordinates – (i) Co-C bond and (ii) Co-N_{ax} bond distances. These calculations were performed using both DFT and TD-DFT with the BP86 functional and the triple zeta TZVPP potentials for Co, C and N. These were implemented using the Turbomole program. In one of their previous works, they compare multiple types of functionals (hybrid, GGA and range separated) [36] and identify BP86 to be most suitable. The molecule is truncated to reduce the computation time and is approximated by substituting the DBI with an imidazole (Im). Since the phosphate anion is not a part of the molecule, an overall charge of +1 is added to the molecule (instead of maintaining it as neutral). In these simulations, they observe that (i) the prior experimentally proposed MLCT like character of the S_1 state holds good in theoretical studies as well, although it might not be a unique contribution to the excitation. (ii) The possibility of a spin crossover from a singlet S_1 state to a triplet $^3(\sigma_{Co\text{-}C}$ to $\sigma^*_{Co\text{-}C})$ has been ruled out for the base-on MeCbl considering the extremely low probability for such a transition to take place, especially considering it is spin-forbidden.

With excitation at 520 nm they observe that (fig. 16 pathway A) –

(i) S_2 and S_3 states are most likely populated (thanks to the large oscillator strength)

(ii) There is a fast relaxation to S_1 state. $S_1 \Leftrightarrow \{Im\text{-}[Co(III)corrin]\text{-}Me\}^+$

(iii) The energy minimum of the S_1 state tends to have an MLCT type character.

(iv) Radical pair generation proceeds starting from the MLCT state, through a ligand field (LF) state.

(v) The LF state has a weakly bound lower axial ligand. LF $\Leftrightarrow \{Im\ldots[Co(III)corrin]\text{-}Me\}^+$

(vi) It has a doubly occupied d_z^2 orbital.

(vii) Its spectral features resemble the one-electron reduced MeCbl.

(viii) To move from the LF state to the radical pair (products), the system has to cross two barriers

   a. S_1 LF state barrier

   b. Radical escape from the solvent cage

(ix) Elongation of the Co-C takes place before the detachment of the Co-N_{axial}.

With excitation at 400 nm they observe that (fig. 5.16 pathway A) –

(i) S_7, S_8 or S_9 states are most likely populated.

(ii) These states have a LF character.

(iii) Here again, there is a fast relaxation to S_1 state.

(iv) The surplus energy (400nm vs 520 nm excitation) is sufficient to overcome the MLCT-> LF barrier.

(v) The LF state has a weakly bound lower axial ligand. LF $\Leftrightarrow \{Im\ldots[Co(III)corrin]\text{-}Me\}^+$

(vi) The two options of photolysis can be seen as –

   a. Internal conversion to ground state

   b. Dissociation of the Co-C bond.

(vii) Two possible pathways for the dissociation of the Co-C bond

   a. prompt singlet

   b. delayed triplet (via an intersystem crossing).

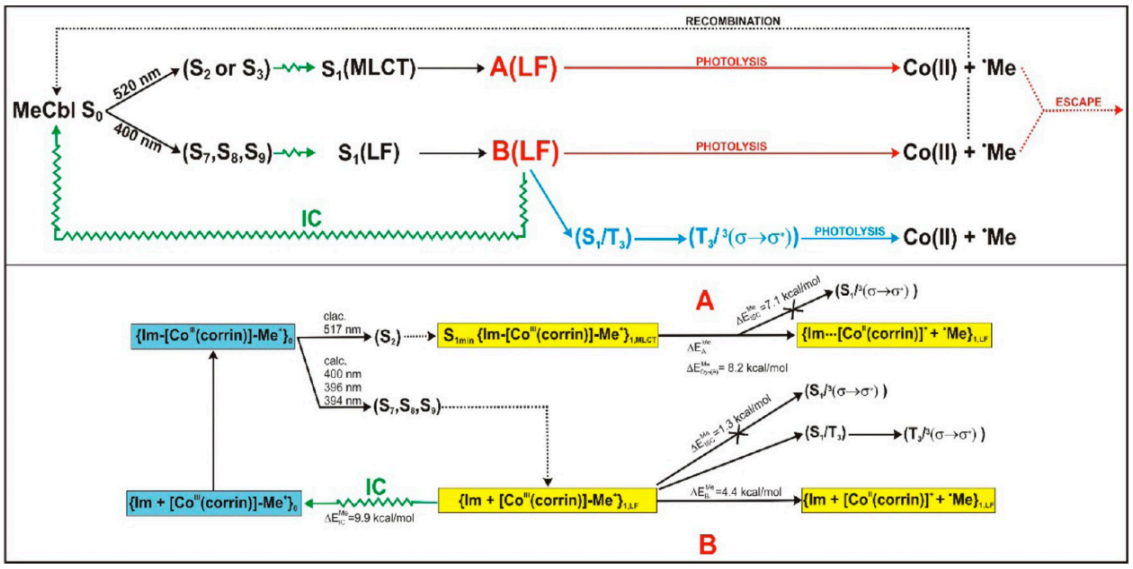


Fig. 5.16: Schematic of the reaction pathways for photolysis of MeCbl from TD-DFT calculations with excitation at 520 nm (Pathway A) and 400 nm (Pathway B). (figure courtesy: Lodowski et al., J. Phys. Chem. A 2014, 118, 11718−11734).

1.6 X-ray Absorption Studies:

In this section, the use of XAS as a tool to characterize the electronic and geometric structure of the Cobalamins is reviewed. Mark Chance et al. did a substantial amount of work starting with a work on CNCbl, AdoCbl and the reduced form of the Cobalamins. From EXAFS studies [47] they infer that in breaking the Co-C bond, the equatorial nitrogens are not affected. However the lower axial nitrogen bond is strengthened. Subsequently, they continued to use EXAFS to study the trans-effect on the Co-C bond scission. [48] In this study they notice that the nonalkyl-cobalamins such as CN-Cbl and OHCbl have a reduced Co-N_{axial} bond distance. However, rather than a trans electronic effect, they suggest this could possibly be due to steric repulsion between the corrin ring and DBI. They propose that in going from a 6-coordinate Co(III) to a 4-coordinate Co(I) square planar species, an intermediate 5-coordinate Co(II) species is formed with the cobalt moving out of the corrin ring towards DBI so as to reduce steric repulsion.

The Chance group then went ahead to study the Cobalt-corrin macrocyles in the presence of proteins. In their work on the Iron-Sulfur protein in Acetyl Coenzyme A, the inactive (not contributing to the catalysis) Co(II) intermediate state is identified to have a square planar

coordination [49] as opposed to the commonly believed square pyramidal (5-coordinate). This turned out to be a direct evidence to favoring of the base-off Co(II) intermediate structure in the catalytic cycle of methyltransfer of MeCbl going from Co(III) to C(I). They followed up this work by studying the final state structure (both EXAFS and pre-edge XANES) and noted that the Co(I) bound to protein final state is also a square planar 4-coordinate species. [50] Although this was expected, this helped establish a systematic argument for the catalytic behavior. The non-invovlement of the DBI ligand in the inactive Co(II) state of the cofactor greatly enhances its reduction to the final state of Co(I), again with no involvement of the DBI ligand. This is crucial since the presence of the DBI would induce stability to the Co(II) and create methyl radicals through homolytic dissociation rather than heterolytic. By ensuring the non-ligation of the DBI, the cofactor maintains a strongly nucleophilic Cobalt center that facilitates re-methylation to get back to the Co(III) state. In their subsequent work, they proceeded to measure time-resolved XAS using a nanosecond laser pump at 532 nm to excite the base-off MeCbl and study the product of photolysis as a function of time ranging from few ns to few ms. [51] The photoproduct in this reaction shows a substantial increase in the intensity of the rising edge while the pre-edge peak intensity is substantially reduced consistent with the formation of a 4-coordinate square planar species. This work confirms the presence of the square planar type product both with and without the enzyme protein in the functioning of the Cobalamins. It also highlights the importance of the enzyme protein in maintaining a base-off configuration for MeCbl so as to ensure a nucleophilic final state that could return back to the ground state Co(III) with the re-addition of methyl ligand.

There is another important work by Champloy et al. (2000) in studying the effect/occurence of X-ray induced radiation damage in Cobalamin cofactors that are free and/or protein (enzyme) bound. [52] They systematically studied the effect of X-rays on the glass phase 100K samples of MeCbl, AdoCbl, CNCbl and AqCbl with and without the proteins. In this study, they noticed that CNCbl and AqCbl undergo radiation-induced reduction at the metal center. The argument is that the reduction potentials to move from Co(III)/Co(II) is -0.04 V for AqCbl, -0.2 V for CNCbl, and less than -1.5 V for alkylCbls. Hence the possibility for reduction is strongest for AqCbls and CNCbls, which is confirmed by the experiments. However, since this experiment was

performed in a glassy state, the ability for the CN^- or the OH^- ions to diffuse away from the cofactors is substantially reduced. A possible scenario then is a pseudo octahedral (6) coordination with a reduced cobalt (II) ion and an elongated Co-N_{axial} bond length. More importantly they noticed that the cofactors of MeCbl and AdoCbl seemed to be stable to within experimental error, both free and protein bound, to X-ray damage.

2 Aim of This Chapter

Most experiments studying the photolysis of MeCbl have been done using IR or UV-Vis Absorption Spectroscopy. While there is definitely a great wealth of material information that has been obtained from these experiments, the conclusive evidence of the entire process, especially the intermediates and transients, is still elusive. Major limitations in comprehending the photolytic process seem to arise from the lack of experiments capable of combining high time resolution and specificity/detail in the electronic structure or crystallographic structure.

Inner shell spectroscopy gives specific information on the nature of an element, oxidation state, local coordination of the central atom and spin state of the system.

This chapter aims to use the experimental technique of time resolved X-ray absorption spectroscopy (TR-XAS) coupled with DFT and FEFF simulations to study the nature and characteristics of the intermediate states that are formed at short timescales (ps-ns) during the photolysis of MeCbl.

3 Materials, Experiment and Analysis Details.

3.1 Materials:

The samples for the experiments (MeCbl powder) were purchased from Sigma Aldrich. They were dissolved in distilled water for regular experiments. For pH-controlled experiments, pH7 was obtained using HEPES buffer and pH2 using conc. HCl. Steady state optical experiments were performed with MeCbl solution in 5 mm (path length) cuvettes. Transient optical absorption measurements were performed using both 5 mm and 2 mm cuvettes. Once

transferred into the cuvettes, the samples were degassed under Argon. Degassing is a process of maintaining higher pressure of inert gas (such as Ar) within a closed sample container to flush out any oxygen. Simply put, it ensures an anaerobic environment. For the transient measurements, additionally a magnetic stir bar was used to mixing the solution to create some turbulence and avoid repeated exposure of the same region under optical excitation. For X-ray measurements, the degassing was performed in larger reservoirs and then transferred into the reservoir at the experimental set-up using an airtight syringe. The X-ray experiments were performed in an air-tight setup under a slightly higher nitrogen pressure to ensure the least exposure to oxygen.

3.2 Steady State UV-Vis Absorption:

The steady state measurements of UV-Vis absorption were performed using the Shimadzu UV-2559 UV-Vis Spectrophotometer.

3.3 Transient Optical Absorption:

The transient absorption measurements were performed using two instruments for two different (partially overlapping) time-ranges.

The femtosecond transient absorption apparatus is a kilohertz (up to sub-hertz frequencies, vide infra) pulsed laser source and a pump-probe optical setup. Laser pulses of 100 fs at 800 nm were generated from an amplified, mode-locked titanium-sapphire laser system (Millennia/Tsunami/Spitfire, Spectra Physics). A portion of the laser pulse was used to pump an optical parametric amplifier (OPA, Spectra Physics) to generate the pump (excitation) pulses. These are then selected using a mechanical chopper. The remaining portion of the laser pulse was sent through an optical delay line and focused on to a 3 mm sapphire plate to generate a white light continuum for the probe beam. This white light continuum was then compressed by prism pairs (CVI) and then passed through the sample. The polarization of the pump beam with respect to the probe beam was set to the magic angle (54.7°). The intensity of the pump beam was adjusted using a neutral density filter. A 300 line grating spectrograph is used to disperse the white probe onto a charge-coupled device (CCD) camera (DU420, Andor Tech.). The final resolution of the spectrum was about 2.3 nm over a dynamic range of 300 nm. The instrument

response function was ca. 150 fs. The time-range that this instrument covered as pump-probe time-delay was 100 fs to 6 ns.

The second instrument was used to cover the time-delay range of 500 picoseconds to 1 millisecond. This data was collected using EOS spectrometer built by Ultrafast Systems (IRF ~ 800 ps.). The excitation pulse was using the set-up of OPA as explained above.

Data analysis for the transient optical absorption was carried out using locally written (ASUFIT) software developed under the MATLAB environment. The transient absorption or fluorescence change curves were fitted over a selected wavelength region simultaneously as described below to obtain decay-associated-spectra (DAS).

$$\Delta A(\lambda, t) = \sum_{i=1}^{n} A_i(\lambda) \mathrm{e}^{\left(\frac{-t}{\tau_i}\right)}$$

where the observed absorption (or fluorescence) change at a given wavelength and time-delay is given as $\Delta A(\lambda, t)$. 'n' is the number of kinetic components used in the fitting. The decay-associated-spectra is given by the plot of $A_i(\lambda)$ (amplitude spectrum of the ith component) as a function of wavelength with a lifetime of τ_i for the specific kinetic component. The analysis procedures have been reviewed elsewhere [53] in literature. Errors associated with the lifetimes obtained from the transient absorption measurements were ≤ 5%.

3.4 Steady State and Transient XAS:

With the help of insertion devices, such as the dual inline undulators, the X-ray probe pulses are extracted from the APS storage ring electron bunches at a flux of about 10^{12} X-ray photons. The X-ray energies can range anywhere between 6 and 30 keV. The X-ray pulse-train has three different options of arriving at the sample. For this experiment, the 'hybrid' mode of setup was used where the X-ray pulses were about 3.2 μs apart. Albeit a majority of the X-ray intensity being lost and better time resolutions being compromised for, the advantage of working in hybrid mode is that the single pulse is 4 times as intense than a single X-ray pulse from the 24-bunch mode (much larger than a single pulse intensity of the 324 bunch mode).

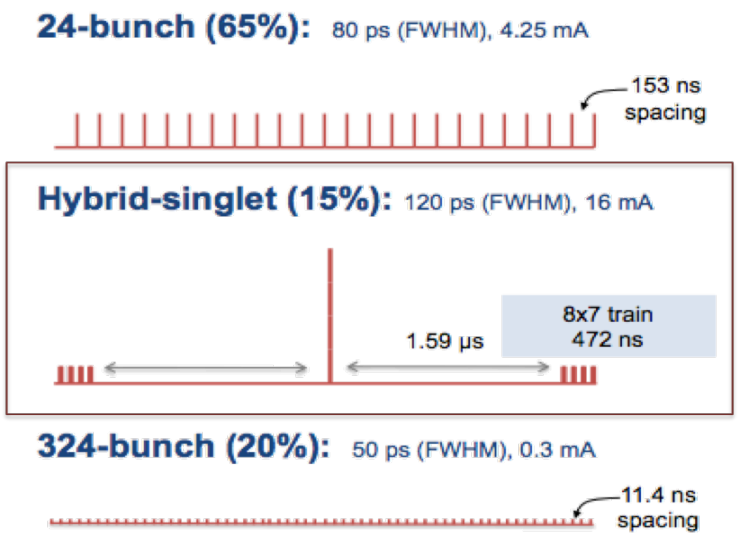


Fig. 5.17: X-ray pulse-trains at APS. (Figure courtesy: http://www.aps.anl.gov/Xray_Science_Division/Conference/052012/posters/Argonne%20National%20Laboratory_Young.pdf)

Refer figs. 5.17 and 5.19 for the different types of X-rays pulse trains and pump-probe timing schematic at beamline 11-ID-D (APS). The typical concentrations of samples studied here (liquid jet) is about 1-5 mMol. The mode of detection is in fluorescence. Since it becomes important to ensure a sufficiently high fraction of the total molecules in the X-ray beam to have been excited (into the excited state) by the pump-laser photons; notwithstanding the constraints coming from solubility of the sample in a solvent and the small amounts of sample available in total; fluorescence yield is the most preferred choice of detection.

The fluorescent X-rays were detected using two avalanche photodiodes (APD), both at 90º to the X-ray sample plane. Soller slits (including a Z-1 filter) were placed, one on each side, between the sample and APD. This (Soller slit) was used to prevent elastic scattering signals from saturating the detector and to prevent fluorescence from other elements (Z-1 filter) being detected. These detected signals are digitized from both the APD detectors by fast analyzer cards (Agilent). The automated data digitization system at 11-ID-D can capture all X-ray pulses after excitation which helps get much better signal-to-noise and resolution of the spectral features.

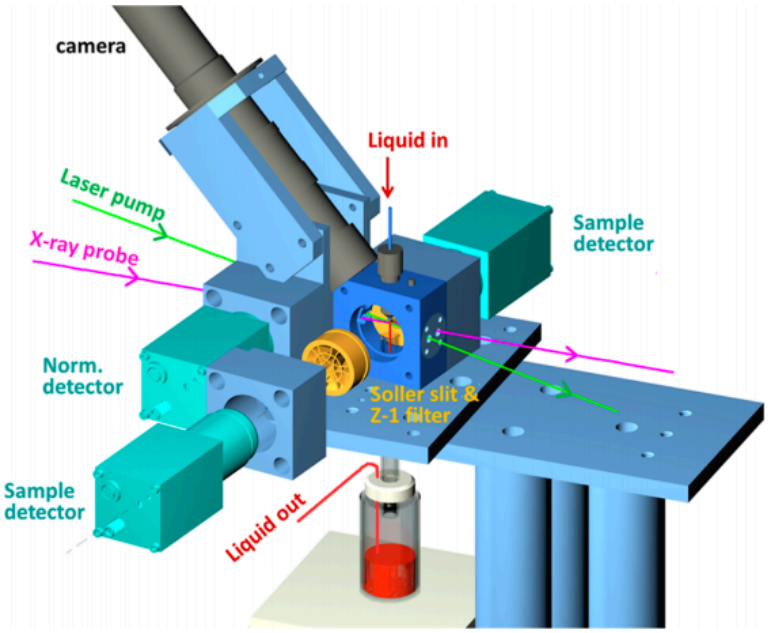


Fig 5.18: Transient X-ray Absorption set-up at APS 11-ID-D. (figure courtesy: Lin Chen and Xioayi Zhang, *J. Phys. Chem. Lett.*, **2013**, *4* (22), pp 4000–4013)

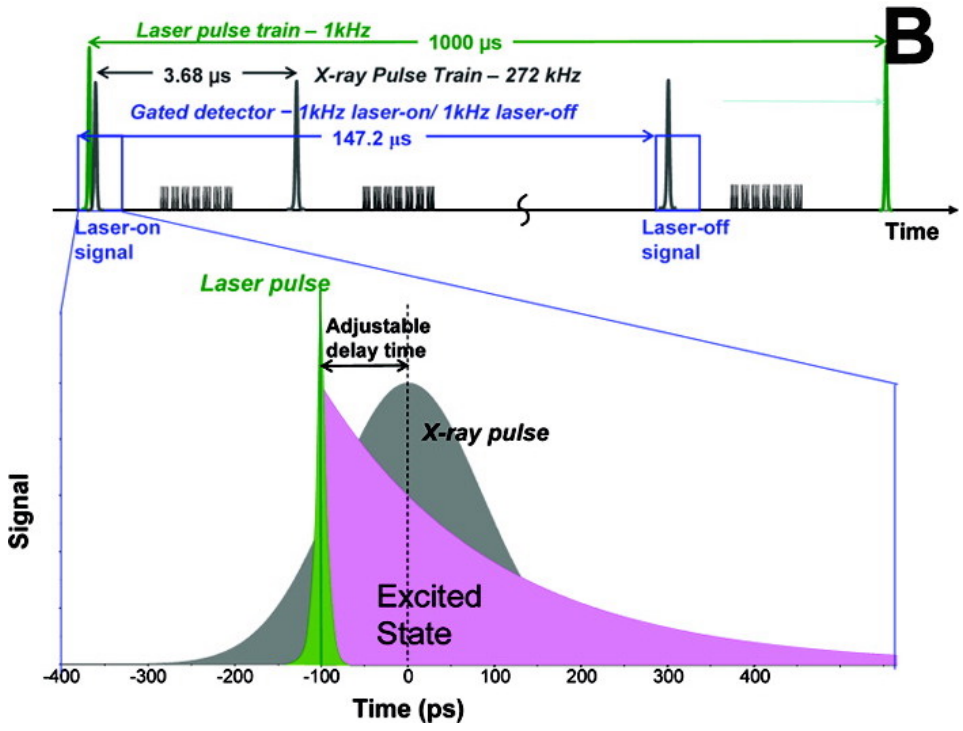


Fig 5.19: Pump-probe timing set-up of the transient X-ray absorption measurements at 11-ID-D. X-ray pulses were extracted under the hybrid-timing mode. Output signals were electronically gated, one coinciding with the laser pulse (laser-ON spectrum) and for every subsequent X-ray 3.6 µs in time interval. The laser-ON pump-probe time delay can be set to any value 100 ps and higher. (Figure courtesy: Chen et al., *J. Am. Chem. Soc.*, **2007**, *129* (31), pp 9616–9618).

The sample is provided to the X-ray beam using stainless steel nozzle with the sample in solution circulated by a pump, to form a cylindrical jet (~550 μm in diameter). The pump laser is placed at an angle of about 20° and overall both the pump and probe perpendicular to the liquid jet containing the sample. A schematic of the set-up is shown in fig. 5.18.

Since the sample was highly sensitive to the presence of Oxygen, it was flown into a free-flowing 500 μm cylindrical jet inside an airtight aluminum chamber and continuously degassed with nitrogen. For the time-resolved measurements, the sample was pumped at 400 nm wavelength using a regenerative amplified laser with 10 kHz repetition rate 5 ps fwhm pulse length and laser power of 300 mW.

The X-ray and laser beam were spatially overlapped with an X-ray spot size of 100 μm (V) × 400 μm (H) and laser spot size of 170 μm (V) × 550 μm. The liquid flow rate was adjusted to ensure that the sample was refreshed for every subsequent pump-optical pulse. The time-delay between the laser and X-ray pulses was varied with a combination of a programmable delay line (PDL-100A-20NS, Colby Instruments) and a digital delay line (v850 4-channel VME digital delay/pulse generator, Highland Technology).

Data analysis for the XAS measurements were performed with the Demeter package [54] – predominantly Athena. The energy scale normalization is performed using the cobalt metal standard for all scans individually. Data in energy space are pre-edge corrected, normalized, deglitched (if necessary), and the background corrected. For the XANES (time resolved experiments) analysis, data from similar time points and sample types were averaged together to provide merged and averaged spectrum with better SNR..

4 Experiments and Observation

4.1 UV-Vis Steady State and Transient Absorption:

Optical absorption experiments were carried out at the ASU Inter-disciplinary Science and Technology Building-5 and Biodesign Building A. The first measurement was to confirm the

nature of the spectrum by comparing it with established spectra from literature. [29,35] Figs. 5.20 and 5.21 show the ground (steady) state spectrum of MeCbl at pH7 and pH2 respectively.

The optical spectrum of the corrin ring with the cobalt can be split into three regions: 500-600 nm (red), 400-500 nm (green) and 300-400 nm (blue) following the style of Firth et al. [57]. The absorption bands in the red, green and blue regions are denoted as α/β, D/E and γ respectively. These conventions were first created for dicyanocobalamin. The most up to date denomination of the transitions for MeCbl are as follows: (a) α and β correspond to to π/d -> π* and d/π -> π* respectively (where π corresponds to the delocalized corrin ring electrons and d correspond to the cobalt (metal) 3d electrons). (b) D and E parts of the band are assigned to d/π -> π*/d and d/π -> d_z^2/ $\sigma^*_{(Im\text{-}Co\text{-}Me)}$ type transitions although the E part of the band is not visible in the absorption spectrum. (c) The γ bands are also a mixture of excitations that involve the corrin ring, the metal d and the axial ligand as d/π -> π* type and d/π -> d_z^2.

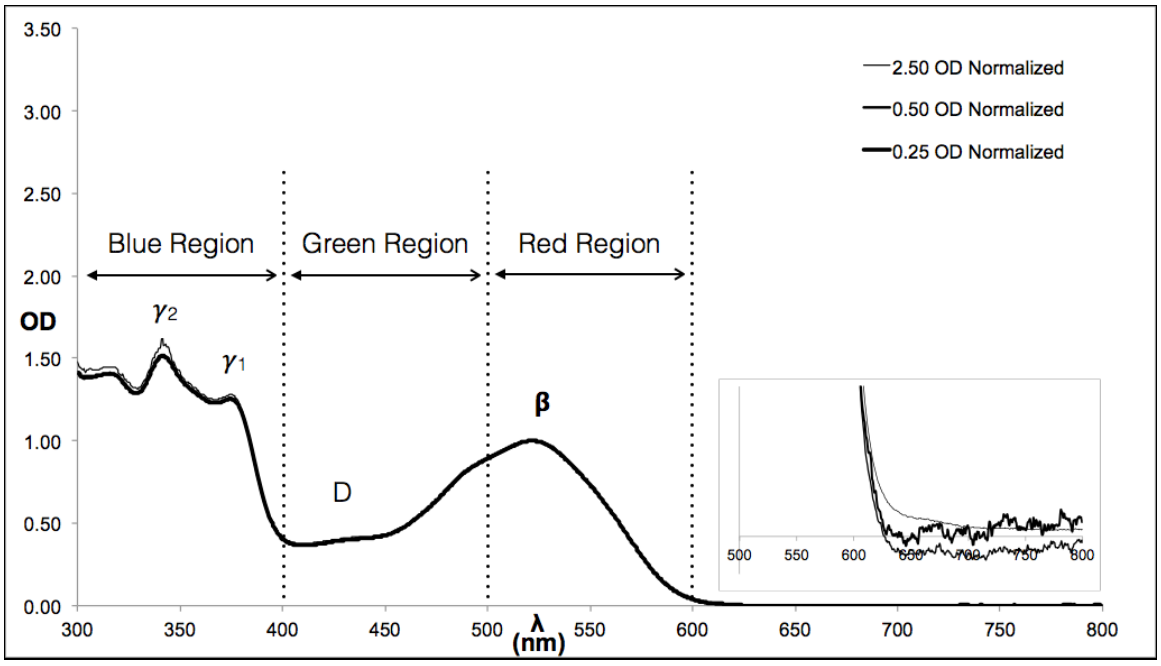


Fig. 5.20: MeCbl in water solution at three different Optical Densities (2.5, 0.5 and 0.25) normalized to OD = 1 at pH7. Inset shows the background with the Signal to Noise with dilution.

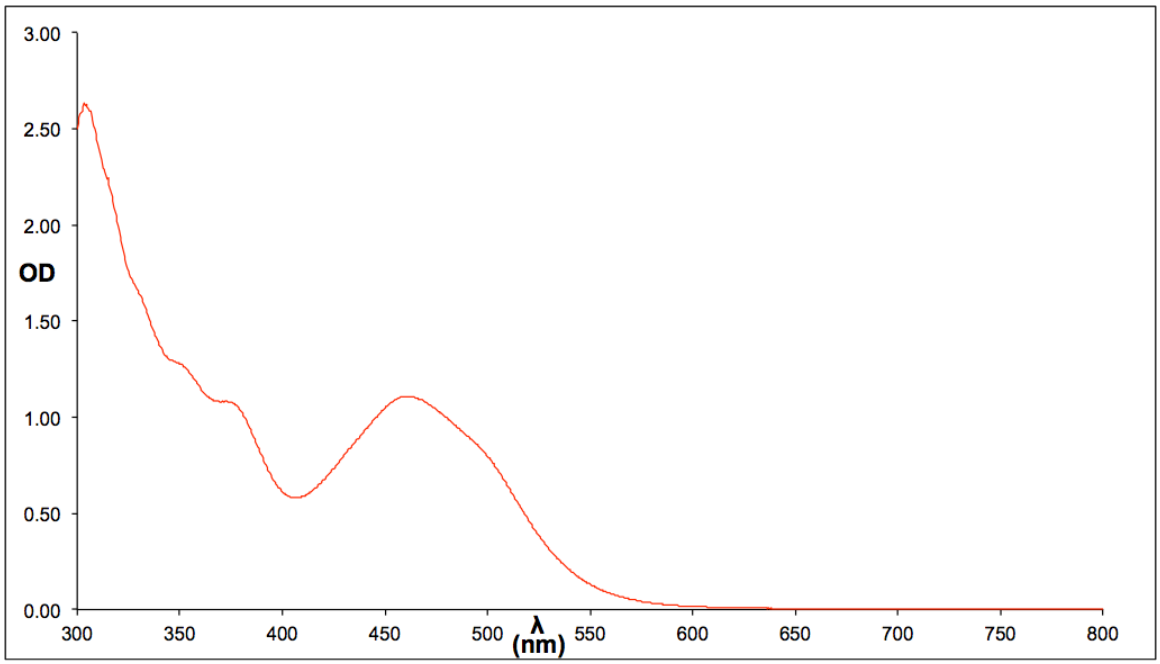


Fig. 5.21: MeCbl in water solution at pH2. Clear visible shift in the α band to shorter wavelengths.

Since this dissertation does not detail the study of optical transitions in MeCbl, deconstructing the nature of excitations of optical absorption in specific details to compare and contrast with prior work is avoided and only the signatures are identified for qualitative comparisons.

For MeCbl at pH7 (fig. 5.20), the α peak (approx. 550 nm) is subdued and lower in intensity as compared to the β peak (approx. 520 nm). The γ band on the other hand consists of 3 peaks (approx. 320, 340 and 375 nm), which are roughly of the same intensities. There is a very weak presence of the D/E band at about 420 nm whose reduced presence is generally attributed to strong bands on either (lower and higher energy) sides of the peak.

The acidic solution of MeCbl in water (pH2) is considered the base-off form where the lower axial DBI is detached from the Cobalt leading to a pseudo 5-coordination. This is expected to cause a reduction in the oxidation state of the cobalt thereby shifting the α and/or β bands to higher energies. Hence the observation of the peak shift to (approx.) 470 nm. Another important change is the sharp rise in one of the peak intensities contributing to the γ band at approx. 300

nm. This has been noted to be anomalous to both dicyanocobalamin (where the strongest γ band peak is close to 400 nm) and MeCbl (where the γ band peaks are roughly of similar intensities). [57] A point to be noted here is that the accuracy of the spectrophotometer to measure absorbances with OD > 1 is not trustworthy to the extent of making quantitative conclusions. This is because of the measurement being in transmission mode, which implies that for an OD > 1 the amount of signal transmitted is very weak. However, depending on the excitation wavelength for the transient optical and X-ray studies, that corresponding peak/band is set to about 1 OD inadvertently shooting the γ band ODs well over 1. Nevertheless, in cases of substantial changes in the trends of peak intensities and energies, qualitative arguments could still be made using the same as data source. This will be discussed in the forthcoming paragraphs on optical studies as and when required.

4.1.1 Radiation Damage Test

This sub-section will be devoted to the measurements done to identify and test for pump-laser damage. The results presented here served as benchmark to identify tolerable pump-lase dose for the X-ray measurements. The damage measurements by were done in two kinds by measuring the steady-state spectra before and after –

Type (a): a pump-probe transient absorption measurement with a 1 kHz 100-femtosecond laser at 520 or 400 nm at an average energy of 100 μJ.

Type (b): a direct exposure to 10 Hz nanosecond 500 nm laser for stipulated time intervals and at an average energy of 0.5 or 1.5 mJ.

The first measurement performed was on pH7 MeCbl solution in water. The steady-state measurements were done before and after a pump-probe measurement (Type a) with irradiation for 30 minutes. As can be seen in Fig. 5.22, the before and after laser exposure spectra were substantially different. The E peak in the green region was seen to be of stronger intensity and energy shifted.

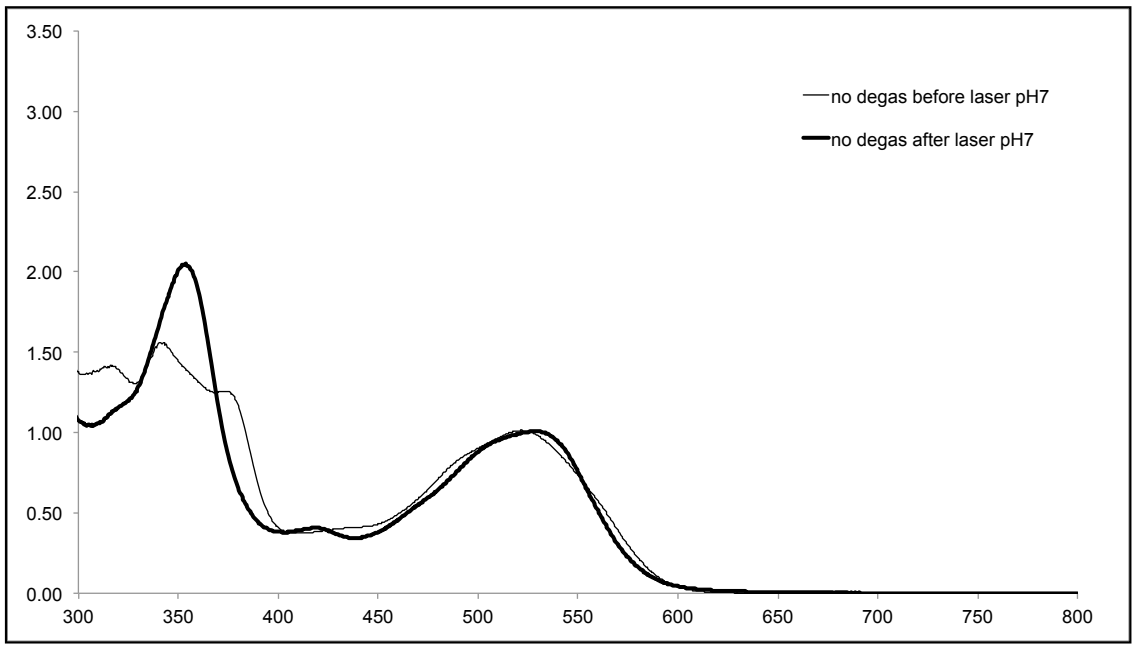


Fig. 5.22: Radiation damage test-1: MeCbl in water solution. Stead-state UV-Vis absorption before and after 520 nm laser exposure of Type (a) for 30 minutes.

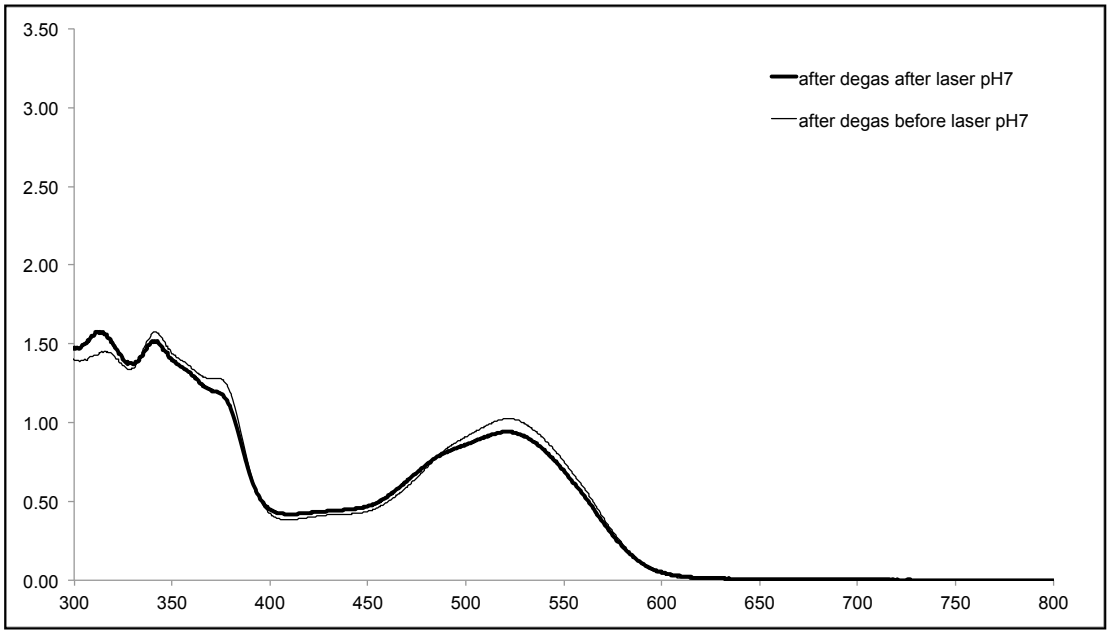


Fig 5.23: Radiation damage test-2: MeCbl in water solution degassed under Ar environment for 2.5 hours. Stead-state UV-Vis absorption before and after 520 nm laser exposure of Type (a) for 45 minutes. No observable change in the qualitative shape of the spectrum or peak energy positions. Modified intensities within 10% of original.

The appearance of an additional shoulder peak at about 470 nm was also noticed. In the blue region, one of the three γ band (blue region) peaks saw a substantial increase in the intensity (compared to the otherwise similarly intense peaks) and the disappearance of another. Finally, there was clear indication of the shift of the β peak to lower energy and near vanishing of the shoulder (550 nm). This indicated the possible formation of a strongly modified alternate species with laser pump energy. In the aim of performing transient X-ray studies this would not be a good inference since the sample damage occurred fast and causing major changes.

As an alternative, the same experiment was repeated with the solution degassed for 2.5 hours with constant Argon bubbling. Fig. 5.23 demonstrates the effect of Ar degassing for a solution that was laser-pump irradiated as in Type (a) and clearly there was a marked difference in the damage effect.

The anomalous peaks were not to be found anymore across the spectrum. There were some noticeable changes as with the reduction in the intensity of the β peak and two of three peaks in the γ band. The third peak in the γ band that showed an increase in intensity was however the peak at 320 nm as opposed to the central peak (at 340 nm) when the solution wasn't degassed with Argon. The presence of Oxygen is known to take the photolysis reaction to products and through reaction mechanisms that are different from anaerobic photolysis. [19,25] Hence degassing becomes a mandatory procedure and the radiation damage studies confirm the importance of degassing the MeCbl solution.

The photolysis of MeCbl even under strict anaerobic conditions forms irreversible products. With time, this product concentration in the cuvette would increase and hence the steady-state absorption spectrum would reflect this change in concentration of the different species present. The net change in this measurement was seen to be less than 10% deviation from original peak intensities with no change in the qualitative spectral features or peak energy positions and hence was taken to be the signature of a well-conditioned experiment.

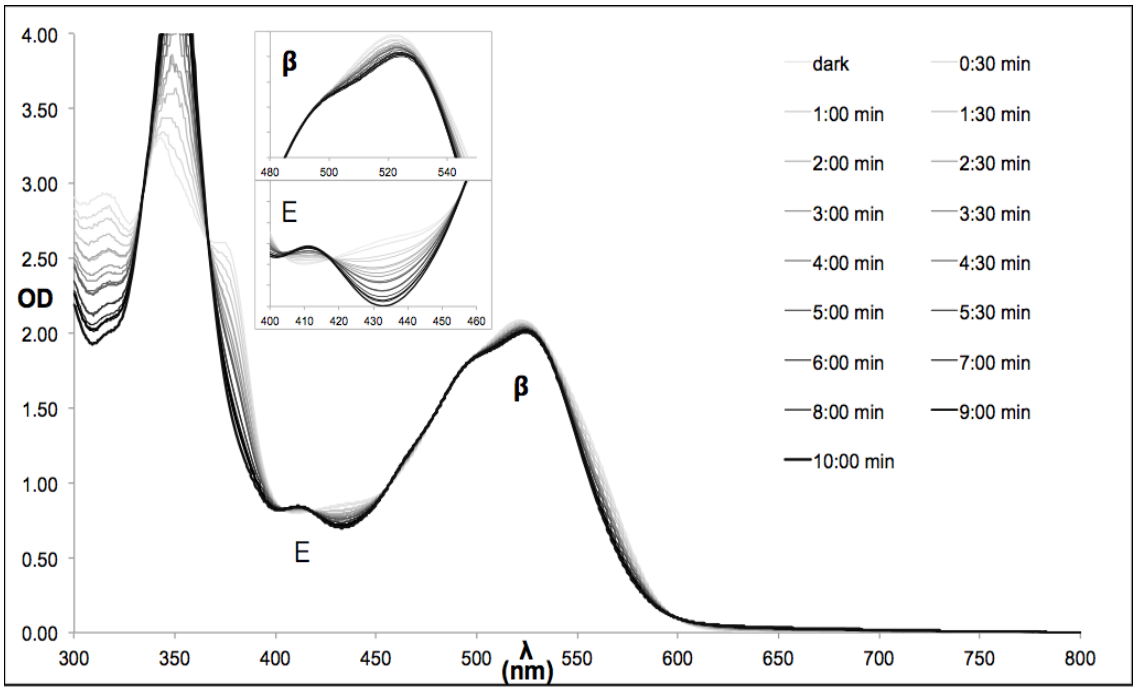


Fig 5.24: Radiation damage test-3: MeCbl in water non-degassed. Stead-state UV-Vis absorption before and after 500 nm laser exposure of Type (b) in time steps as mentioned. Inset shows the changes in peak β and peak E. The time evolution is from grey to black.

The previous experiments (figs. 5.22 and 5.23) were performed to test the pump-laser induced damage to the MeCbl-water solutions were performed using Type (a) pump-laser excitation. However, an average energy of 100 μJ would be at least an order of magnitude lower than the average pump-laser energy at the X-ray measurements. Hence the same experiments were repeated using the Type (b) scheme of irradiation for the non-degassed and degassed MeCbl-water solutions whose UV-Vis absorption spectral evolution in time is presented in figs. 24 and 25 respectively. The pump-laser damage experiments performed in the Type (b) excitation scheme show the evolution of the damage with time. With the sample non-degassed before laser-exposure, it is clear that the same trend as with Type (a) excitation is present. The β peak was seen to shift to higher energy (inset fig. 5.24) and the peak E did show stronger and sharper feature. Most importantly, the central peak at about 340 nm of the γ band substantially increased in intensity with increase in exposure time.

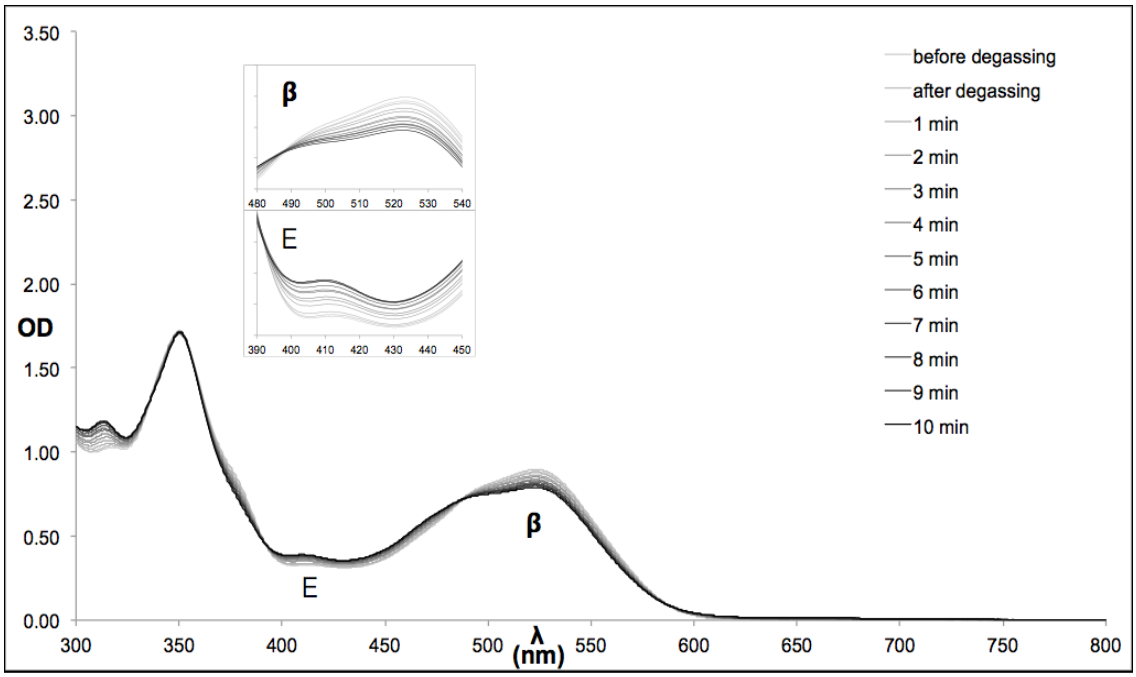


Fig 5.25: Radiation damage test-4: MeCbl in water degassed for 2 hours. Stead-state UV-Vis absorption before and after 500 nm laser exposure of Type (b) in time steps as mentioned. Inset shows the changes in peak β and peak E. The time evolution is from grey to black.

Similarly, the damage studies conducted on the degassed sample (Fig. 5.25) show features similar to that with Type (a) scheme of laser excitation. Differences in the absolute OD values in the two measurements arise from the difference in the starting concentration of MeCbl in water. In theory, 1 minute of exposure under Type (b) amounts roughly to the same average power as with the laser-pump during the X-ray measurements. Also to be noted is the fact that during the X-ray measurements, the sample will be flown as a liquid jet and hence continuously refreshed and recirculated as opposed to the cuvette based measurements with the UV-Vis measurements, which therefore accumulate damage much faster.

As a precaution, experiments were performed with very intense pump-laser (1.5 mJ/pulse) excitation to record and ensure no inexplicable products were formed as a result of non-linear processes or heating. As shown in Fig. 5.26, this hypothesis was confirmed and the spectra showed that the only effect of turning up the laser intensity was faster damage of the same kind as before.

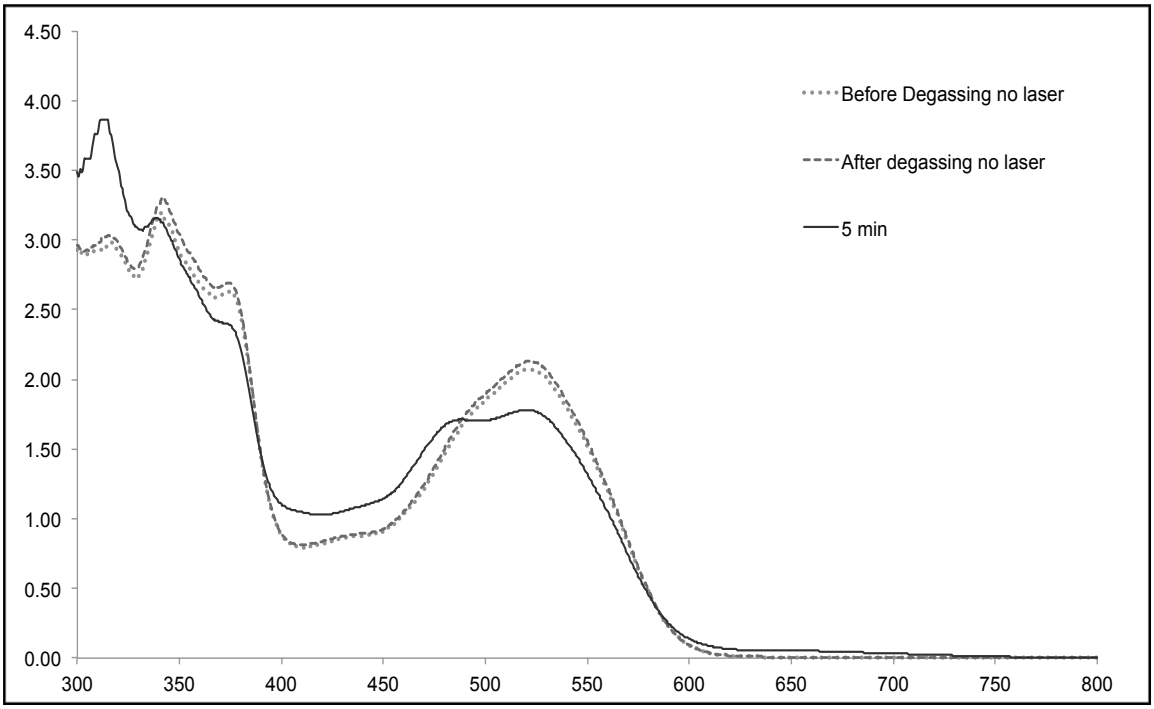


Fig 5.26: Radiation damage test-5: MeCbl in water solution degassed for 2 hours. Stead-state UV-Vis absorption before and after 500 nm laser exposure at (1.5 mJ/ pulse) of Type (a) for 5 minutes. Strong signatures of change with laser exposure but exactly as seen before in fig. 25.

4.1.2 Pump-Probe Transient Optical Absorption:

Transient measurements were performed on the MeCbl in water solution first at 520 nm laser excitation and then subsequently at 400 nm excitation. The aim of these experiments was to hopefully reproduce the spectral features and time constants as with prior work [25,26] and if not to identify the reasons for discrepancies.

The first measurements were at 520 nm excitation and the decay associated spectra (DAS) and the evolution assisted decay spectra (EADS) for these experiments are provided in fig. 5.27 and 5.28 respectively. The data from the femtosecond measurements and slower microsecond measurements were coupled together in the representation. The comparison from prior literature can be done with figs. 5.10 and 5.11 in this chapter.

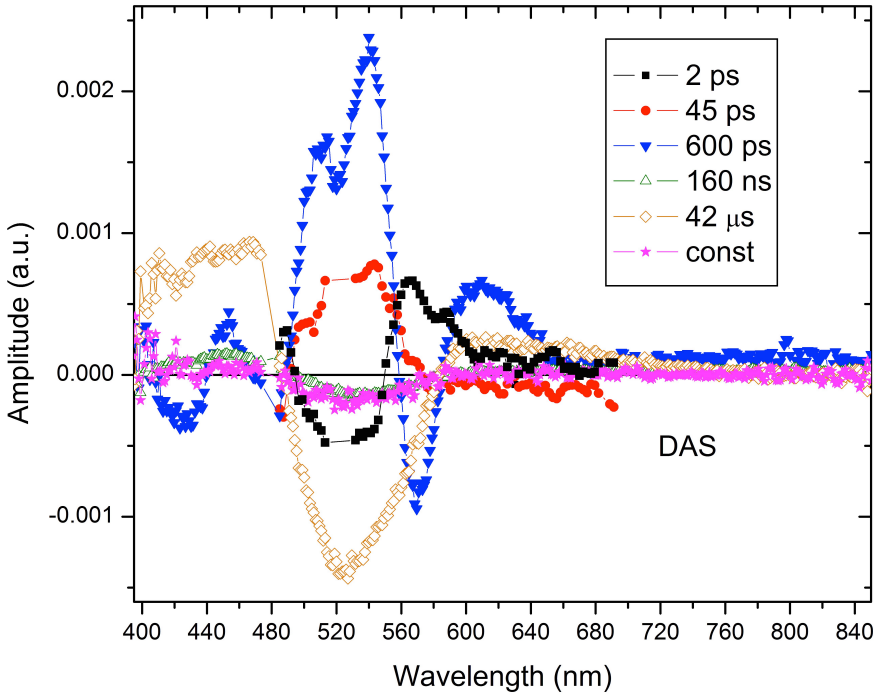


Fig 5.27: MeCbl (pH7) Photolysis @ 520 nm - Decay Associated Spectra: MeCbl in water solution degassed for 2 hours. All the different components and their time constants are listed.

The DAS shows six components through the reaction with the earliest having a 2 ps time constant. The most common way to look at the DAS spectra is to think of them as the absorption spectrum of a particular reaction intermediate. Hence each time constant corresponds to a unique and metastable excited state. In the case of MeCbl, this can be interpreted as follows: Upon excitation at 520 nm (β band), the ground state undergoes bleach and hence the first excited state spectrum displays bleach at the excitation wavelength (2 ps spectrum). However, the increases amplitudes at about 560 – 600 nm indicates strengthening of the α peak. This is followed by a reverse trend leading to formation at the β band and relaxation at the α band with a time constant of 45 ps. This trend is intensified to a substantial formation at the β band and a bleach at the α band and E peak. Once the nanosecond regime is entered, the sample seems to relax with a bleach at the incident wavelength that slowly disappears.

However there is a final product that has an infinite theoretical lifetime, which is indicated in the pink spectra as constant. Comparing with the prior literature information on MeCbl, the above information can be understood as - following an excitation at 520 nm, the sample first

enters into an MLCT excited state in the few 100 fs timescales, which decays into the LF state with a 2 ps half life. This LF state however could have some major structural rearrangements coupled thanks to the bond elongation along the Co-C and/or Co-N_{axial} bond. Such structural rearrangements are followed by electronic rearrangements as well (which could include bond breaking) which occur with the 600 ps half-life. Eventually, a majority of the sample relaxes to the ground state, which happens in the few 100 ns to 10s of μs regime. That which does not relax to the ground state and forms irreversible product forms the final infinitely long component.

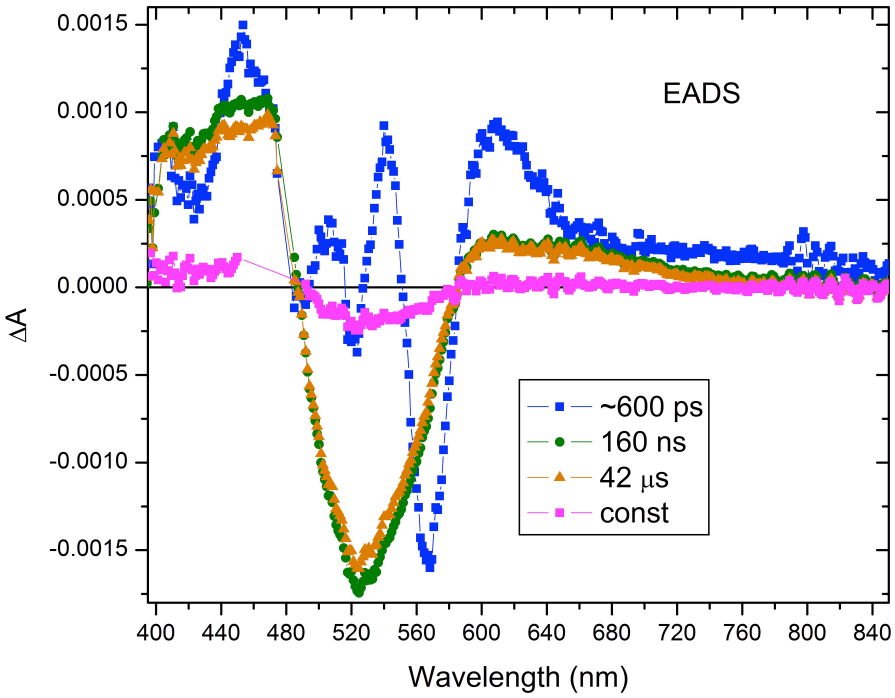


Fig 5.28: MeCbl (pH7) Photolysis @ 520 nm – Evolution Associated Difference Spectra: MeCbl in water solution degassed for 2 hours. Components > 100 ps are listed.

Fig. 5.28 displays the EADS spectra, which can be compared with fig. 5.11 in this chapter. EAD spectra for strictly sequential processes refer to the difference absorption spectra of each of the intermediate state (as opposed to the DAS as in fig. 5.27 which is more reflective. The similarity in the spectral changes (difference absorbance) in both the ps timescales and the ns timescales is convincing. This confirms that the sample and experimental conditions are as similar as possible. The time constants from our experiments are in the same order of magnitude

but not exactly the same as with the prior work. In their work, Shiang et al, [26] report time constants of 1.5 ps, 17 ps and 900 ps whereas we noticed time constants of 2 ps, 45 ps and 600 ps (not accounting for the ns and longer lived intermediates/products).

Since the experiments at the X-ray facility were actually conducted at 400 nm excitation, the above experiments were repeated with 400 nm pump-laser excitation. But since the ultrafast (<100 ps) states cannot be probed at synchrotrons, the transient optical measurements were only performed with the ps-ms instrument. This measurement was performed both for pH7 and for pH2 and the data is presented in Fig. 5.29.

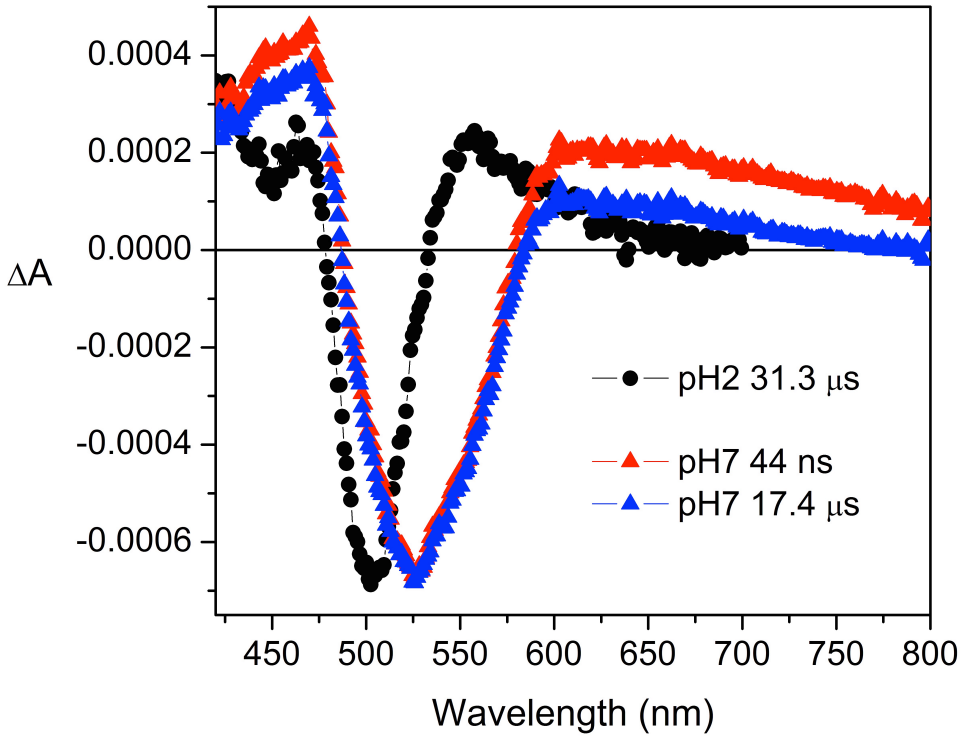


Fig 5.29: MeCbl (pH7 and pH2) Phololysis @ 520 nm – Evolution Associated Difference Spectra: MeCbl in water solution degassed for 2 hours. Only components > 100 ps are listed.

With the pH7 samples excitation at 400 nm show very similar qualitative spectral features with 520 nm excitation wavelength. The only change observed is in the half-lives of the intermediate species. This is expected for two reasons – (i) higher energy of excitation implies more residual energy and hence the modified time to recombine. (ii) In the specific case of MeCbl, it is known to have a competition between geminate recombination within the solvent

cage and solvent cage escape, with the latter being stronger at higher laser energies. [26] In the case of pH2 however, only one long lived component (other than the irreversible product) is observed and this is explained by assumption that in the absence of the lower axial DBI to Cobalt bond, recombination proceeds singly with the relaxation of the Co-C bond (rather than both the Co-C and Co-N_{axial} bonds relaxing in the case of pH7 [58]).

With these measurements, the set of preliminary optical experiments required to test and calibrate the MeCbl samples were completed. The next step was to perform the actual X-ray experiments at beamline 11-ID-D at the APS, which will be described in the following section.

4.2 X-ray Absorption

4.2.1 Important Prior Research Highlights

Of the majority of prior research performed in measuring XANES of the Cobalamins or similar complexes, a couple of results are important and relevant in the context of this chapter. Champloy et al. [52] perform a systematic set of experiments to measure X-ray induced damage in Cobalamin (free factors and protein bound) and identify that only CNCbl and Aquo/OH-Cbl are prone to radiation damage under X-ray illumination.

Figs. 5.30 and 5.31 (reproduced from Champloy et al.) show the XAS of all the free cofactors and the AqCbl cofactor undergo damage related spectral modification to end up similar to the reduced Co(II)B_{12} XANES spectrum respectively. Fig. 5.30 will be taken as a reference in this chapter to compare the experimental spectra obtained here with literature and Fig. 31 serves as a reference to show the nature of reduction that accompanies X-ray induced radiation damage in AqCbl to be very similar to that of chemically reduced Co(II)B_{12}. MeCbl (both free and protein bound), however, was found to be intact and X-ray stable. Another important work on the cobalamin family is also one of the very rare works involving time-resolved XAS measurements. However, this work also provides for some fine calibration standards for experimental XANES spectra of 6-, 5- and 4- coordinate cobalamin (II) complexes, which is reproduced here in fig. 5.32.

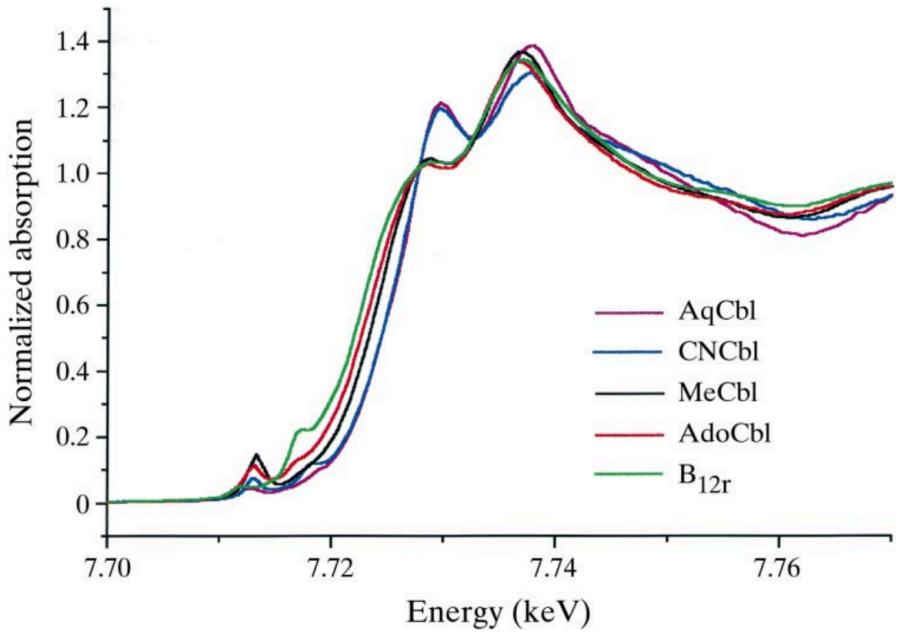


Fig. 5.30: Experimental XANES of Aq/CN/Me/Ado/Co(II)-Cbl solution before extensive X-ray irradiation. (figure courtesy: fig. 2 Champloy et al., *Journal of synchrotron radiation* 7.4 (2000))

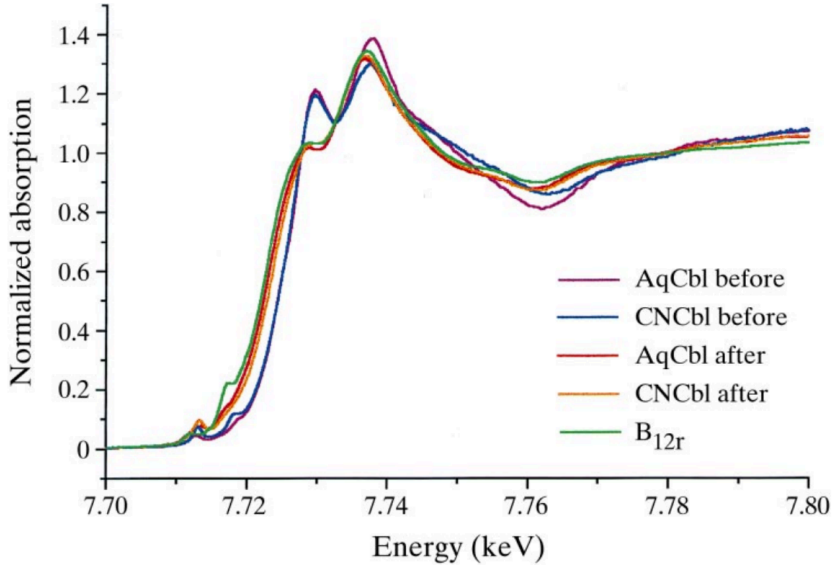


Fig. 5.31: Experimental XANES of Aq/CN-Cbl solution before and after extensive X-ray irradiation. (figure courtesy: fig. 2 Champloy et al., *Journal of synchrotron radiation* 7.4 (2000))

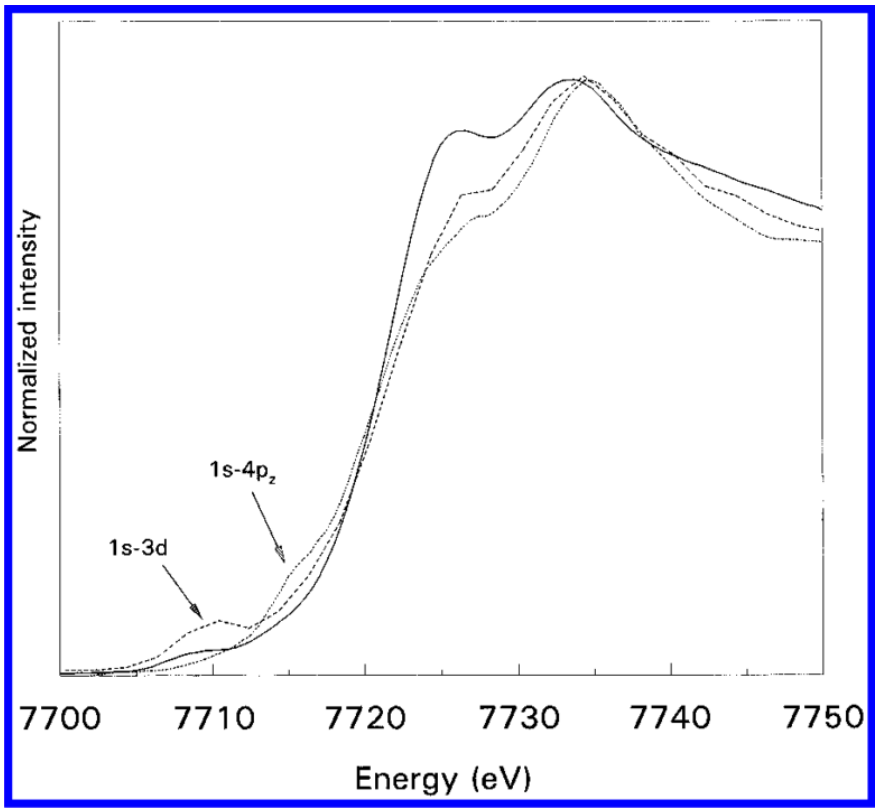


Fig. 5.32: XANES spectra of (a) 4-coordinate Co(II) form of *C. thermoaceticum* – dotted line, (b) 5-coordinate base-off Co(II) MeCbl – dashed line and (c) 6-coordinate chemically reduced base-off Co(II) – solid line. (figure courtesy: fig. 2 Scheuring et al., J. Phys. Chem. 1996, 100, 3344)

4.2.2 Experimental XAS - APS

X-ray absorption measurements were performed at the 11-ID-D beamline at the Advanced Photon Source (APS). As mentioned earlier, the X-rays were generated in the hybrid mode of operation. The energy of the incident X-ray post the insertion device was controlled by a monochromator upstream of the sample. The cobalt K-edge was measured using a reference (foil) to be at 7.7089 keV. The XANES scans were performed with an overall energy range of 100 eV while the XAFS (XANES + EXAFS) scans were measured over an energy range of 600 eV. In both these cases, the entire energy range was split into four regions, each having a distinct sampling size depending on the resolution required. The details are presented in fig. 5.33.

	REGION 1		REGION 2		REGION 3		REGION 4		REGION 5	
	ΔE	δE	ΔE	δE	ΔE	δE	ΔE	δE	ΔE	Δk
XANES	-50 eV -10 eV	5 eV	-10 eV -3 eV	1 eV	-3 eV 17 eV	0.5 eV	17 eV 50 eV	1 eV		
XAFS	-100 eV -10 eV	5 eV	-10 eV -3 eV	1 eV	-3 eV 17 eV	0.5 eV	17 eV 50 eV	1 eV	50 eV 500 eV	0.05 $Å^{-1}$

Fig. 5.33: Table list of XANES and EXAFS sampling range and intervals where ΔE is the range and δE is the step size.

Experimental XANES of the MeCbl solution at both neutral (pH7) and acidic (pH2) conditions is presented in fig. 5.34. Clearly there are a few regions in the spectrum which show qualitative differences in the spectral features. The most prominent in them is the pre-edge peak intensity (7710 eV). The pre-edge peak intensity in a K-absorption spectrum is strongly modulated by the presence/absence of centrosymmetry in the molecule. Energetically, this is where the 3d metal orbitals lie (if seen through the DOS perspective) implying a 1s-3d transition at the pre-edge which is plain dipole forbidden. But 1s-3d is a quadrupole transition that is much weaker in strength than a dipole 1s-4p excitation that forms the absorption edge. In a molecule, however, the orbitals are not atomic. Reflecting this is fact that the 3d metal orbitals are often hybridized with the 2p/3p orbitals of the bonding ligands and with 4p of the metal itself (3d-4p mixing). Therefore, the pre-edge obtains a partial p-character. (Refer to Wong, Joe, et al. *Physical Review B* 30.10 (1984): 5596 for a detailed explanation of the origin of the pre-edge peak.)

In this case, acidic MeCbl is seen to have a substantial pre-edge peak confirming the notion that in the acidic form, the lower-axial Co-N_{axial} bond formed by the DBI is detached. The five-coordinate species then possibly assumes a square pyramidal structure, which is strongly non-centrosymmetric thereby contributing this pre-edge intensity. A point to note is that the Octahedrally coordinated neutral MeCbl also shows a pre-edge intensity and this is generally attributed to the Co-C upper axial bond that breaks the ideally symmetric nitrogen coordination around the cobalt center.

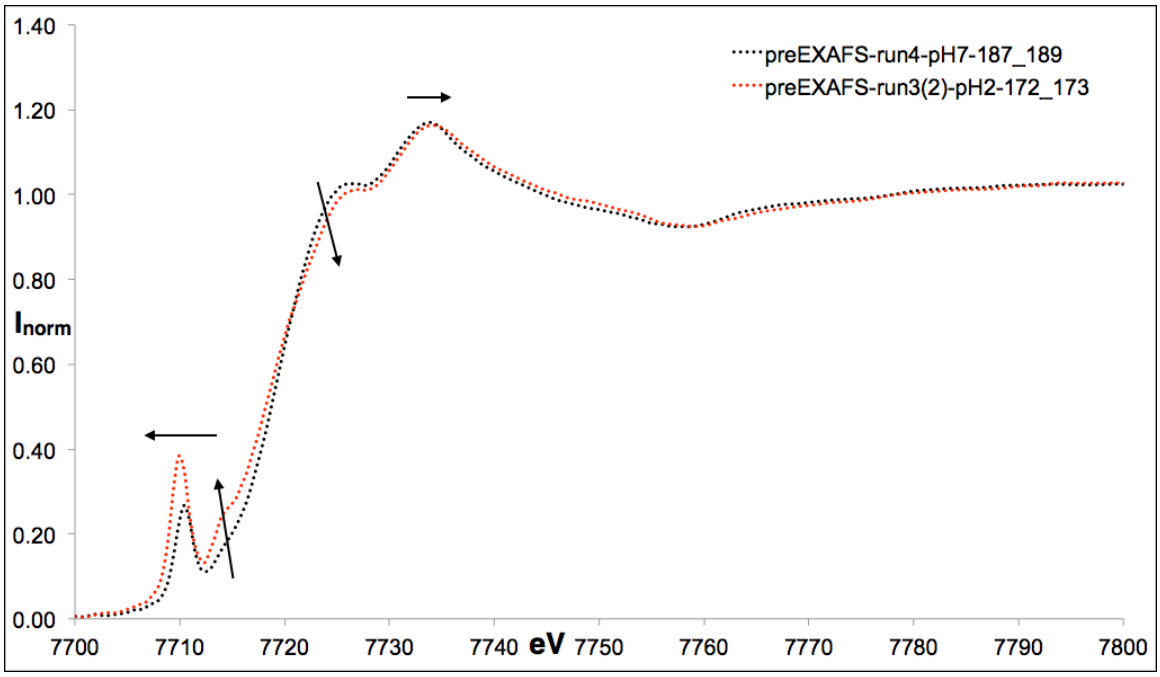


Fig. 5.34: Experimental XANES of MeCbl solution buffered to pH7 (black) and pH2 (red).

The second interesting feature is observed at the rising portion of the absorption edge (approx. 7715 eV). This is identified as the 1s-4p_z+SD (shakedown) transition. [49] This transition is ideally most prominent with square planar geometries where both the axial ligation sites are vacant. The five-coordinates species does not (and is not expected to) show a very strong feature. Yet, the experimental measurement shows that it is significant enough to be distinguished from the octahedral (neutral MeCbl) species. With change in coordination, the bond lengths and strengths possibly get altered reflected by a small energy shift of the peaks: the pre-edge peak shifts red and the strongest absorption feature (7735 eV) shits blue in moving from the neutral MeCbl to acidic MeCbl.

Figs. 5.35 and 5.36 shows the XANES measured from MeCbl before (pre) and after (post) laser illumination. The plotted data have been processed, deglitched and normalized using ATHENA. Interestingly, most of the features that change in the MeCbl XANES spectra post-laser illumination are very similar to both pH7 and pH2. In both cases there is a substantial reduction in the pre-edge peak intensity. In the neutral MeCbl (pH7), there is an additional red-shift of the peak post laser exposure. The second noticeable change is a blue-shift of the entire absorption

edge, again common to both cases. The third change is increase in the intensity of the shoulder peak (7725 eV) on the edge, which is also common to both cases. However, the amount of change seems to be larger in the acidic MeCbl than with neutral. Finally, the strongest absorption feature is noticed to shift blue post laser illumination.

There are a couple of analyses that can support these changes. Starting with the strongest change, the reduction in the pre-edge peak could be attributed to gaining stronger centrosymmetry. For the 5-coordinate acidic MeCbl, this could be a shift back to octahedral coordination. However, the neutral MeCbl also undergoes such a pre-edge intensity drop.

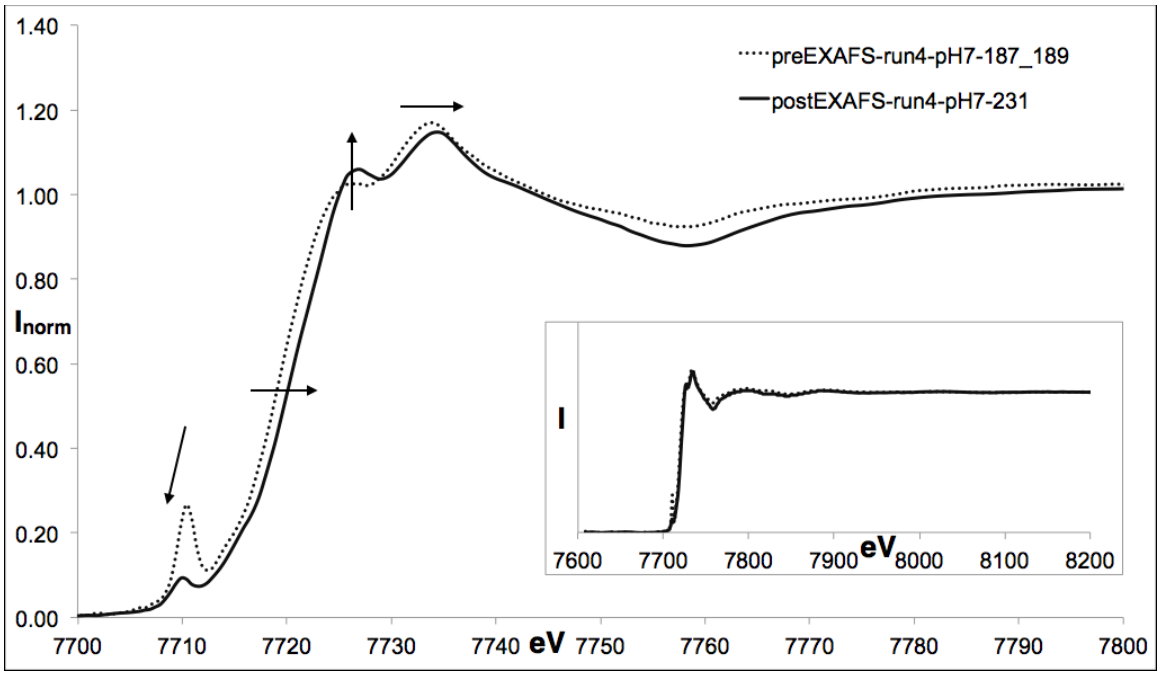


Fig. 5.35: Steady-state experimental averaged XANES (inset: XAFS) of pH7 MeCbl (a) before (pre) – dotted line and (b) after (post) – solid line, pump-laser exposure.

This could mean two options – (a) Formation of a product that is 4-coordinate and square planar with both the axial bonds (the Co-C and the Co-N_{axial} bonds) detached and hence the final solution post laser exposure contains a mixture of the reactant (6/5-coordinate) and the product (4-coordinate) thereby causing a reduction in the pre-edge intensity or (b) Formation of a pseudo 6-coordinate species with Co-C bond cleaved and Co-N_{axial} bond possibly detached as well, however weakly ligated by water molecules from the solvent. In this way there is still an

octahedral coordination, but with the possible Co-O and Co-N_{axial} or double Co-O, the centrosymmetry is reinstated to a large extent.

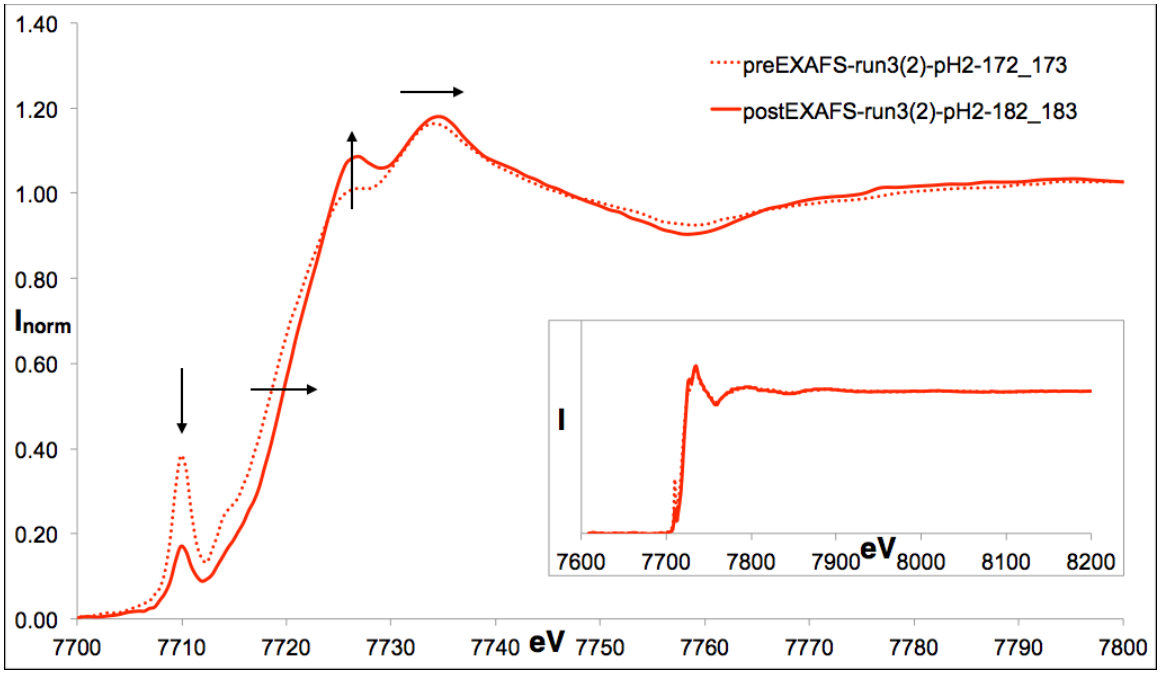


Fig. 5.36: Steady-state experimental averaged XANES (inset: XAFS) of pH2 MeCbl (a) before (pre) – dotted line and (b) after (post) – solid line, pump-laser exposure.

(Note: Co-C is still largely a covalent bond since both the atoms involved in the bond have strong tendencies to be cations whereas Co-O strongly leans towards forming ionic bonds with a clear and strong polarization of the bond electrons thanks to the electronegativity difference being substantially higher). The formation of a square planar product however must hike the rising edge (approx. 7720 ev) portion of the XANES spectrum to correlate with the free axial ligating sites. This is however not seen in the experimental spectra.

A highly relevant literature in this context are the XAS experiments performed on hydroxocobalamin by Giorgetti et al. by electrochemically reducing the compounds to various oxidation state and analyzing their XANES and EXAFS features. [59] In fig. 4 of their manuscript, they notice that the intensity of feature C (shoulder on the edge at 7725 eV in the current work) varies across pH and oxidation state. They notice the feature to be most intense relative to the strongest absorption feature (D) at neutral pH and oxidation state of 3. Referring to this work, the

impression seems to be the formation of a hydro/aquo-Cbl product that 6-coordinated rather than 4-coordinated square planar species.

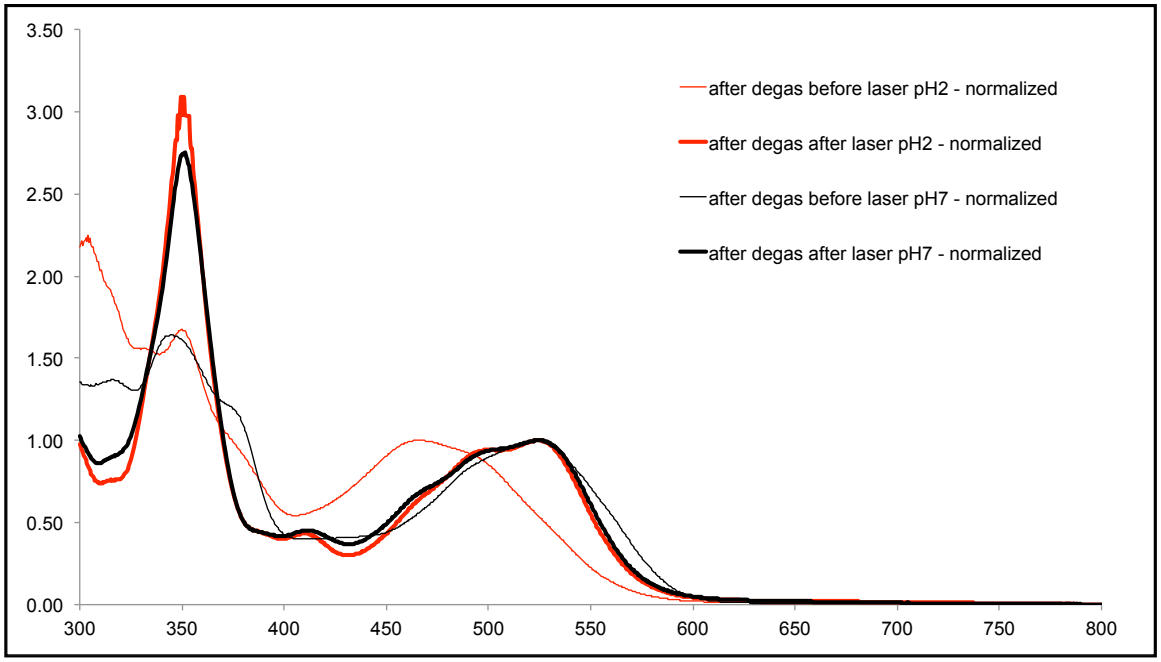


Fig. 5.37: Steady-state UV-Vis absorption spectra MeCbl in water (a) before laser illumination at pH7 – thin black line, (b) after laser illumination at pH7 – thick black line, and (c) before laser illumination at pH2 – thin red line and (d) after laser illumination at pH2 – thick red line.

Fig. 5.37 is the UV-Vis absorption spectrum of the neutral and acidic MeCbl samples before and after laser illumination (measured right after experiments at the X-ray instrument with the help of the Shimadzu spectrophotometer in the Sector 11/12 chemistry lab). Strikingly similar to the X-ray spectra, the final optical spectra of both the neutral and acidic MeCbl solutions seem to be comparable. What is also noticeable is that the final (after laser) spectra do not resemble either of the initial spectra, thereby clearly indicating the formation of some new species. Prior work reported in literature has shown the tendency of air-exposed (aerobic) photolysis of MeCbl to form aquocobalamin as a product. [21] There is a possibility that with time, the solution was exposed and albeit the nitrogen flow, slowly yet steadily, the solution reacted with oxygen to form aquocobalamin. An alternate possibility is the formation of the Aq/OH^- ligation even under anaerobic conditions through diffusive modes with the water present in the solvent.

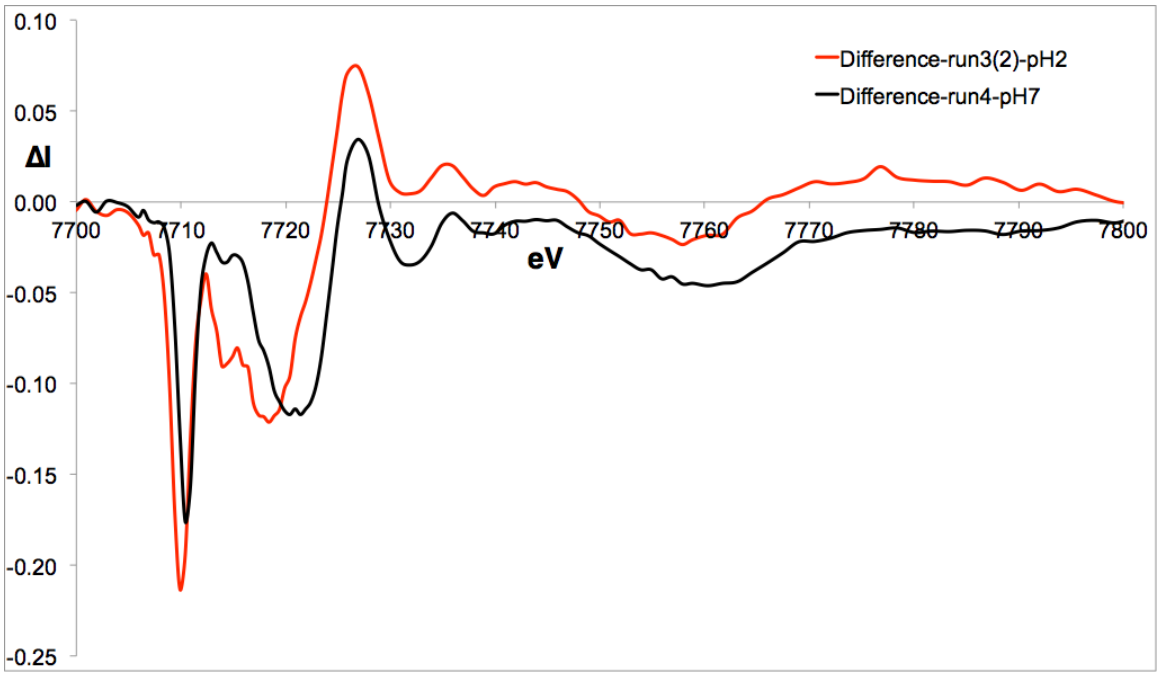
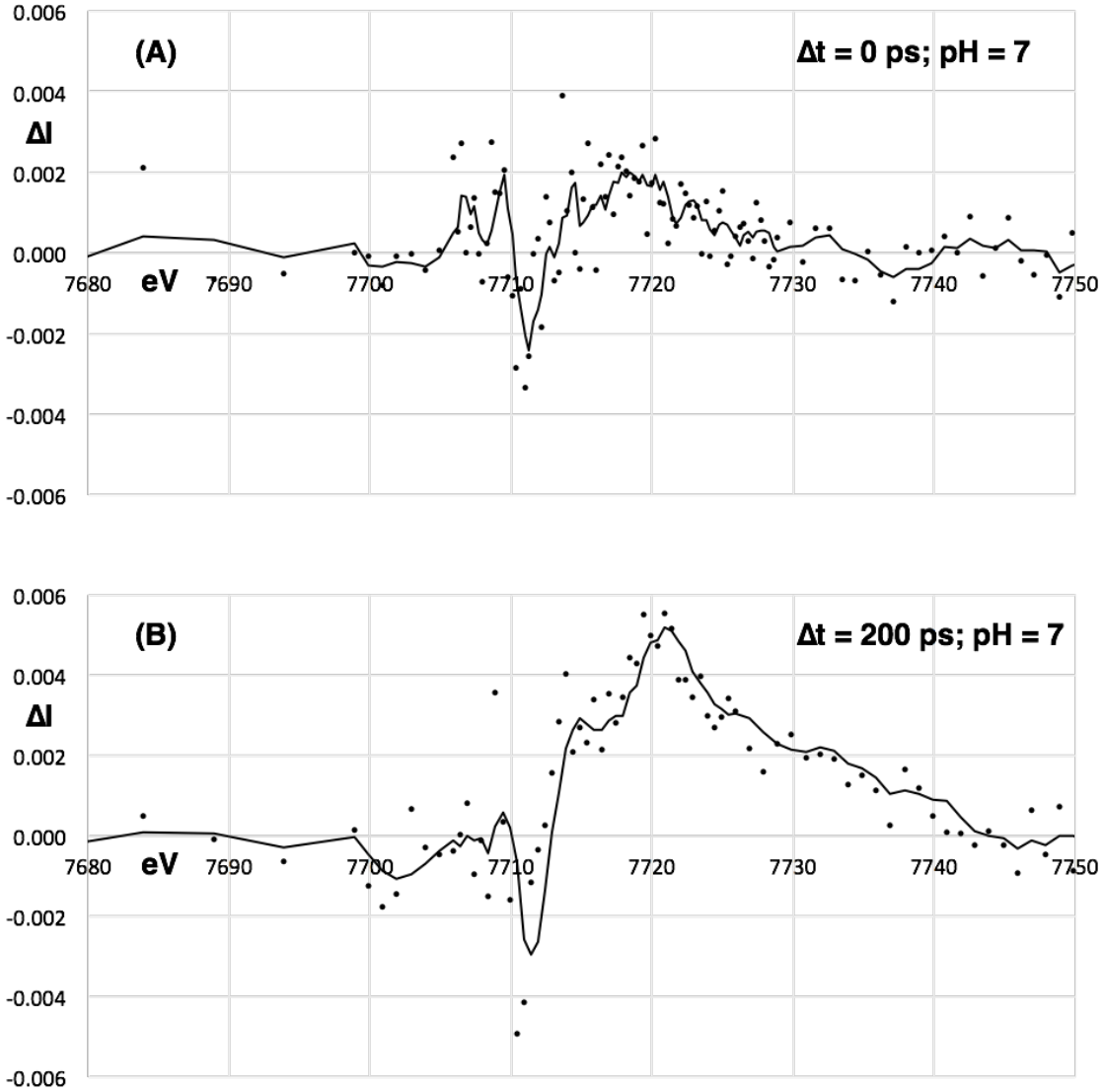


Fig. 5.38: Difference absorption spectra of reactants and products measured at pH7 and pH2 of MeCbl.

Reference spectra of aquocobalamin [60–62] from other literature shows a strong resemblance to the spectra in Fig. 5.37 tilting towards the assumption that a 6-coordinated AqCbl product might be the major component in the postEXAFS spectrum with both pH2 and pH7. With no additional experimental information in hand, the claim of the water ligation will be held only to a qualitative match with other experimental data. This issue will be revisited in the next section on theoretical analysis of the XAS data.

Fig. 5.38 shows the difference absorption of the pH7 and pH2 solutions of MeCbl. Although the absorption spectra have been presented in the earlier figures, the difference spectra are more relevant to be compared with the time-resolved data (which will also be presented as difference absorption spectra) in the forthcoming. The final product in these experiments (both neutral and acidic pH conditions) seems very similar and hence the difference spectra for both the cases are qualitatively similar as well. The most important features are reduced intensity at the pre-edge (7710 eV) and reduced intensity on the rising edge with an inflexion just below the shoulder peak (7725 eV). The shoulder has an increased intensity contrary to the other regions.

The next set of figures represent the experimental time-resolved XAS difference maps for different pump and probe time delays. The time-resolved XAS measurements were performed by modifying the pump laser arrival time with respect to the probe X-ray. The pump laser fluence was measured to be 300 mW with pulse duration of 900 fs (FWHM) arriving at 10 kHz repetition rate. Spot size of this pump-laser beam 500 – 600 μm in diameter incident at a 20° angle with the X-rays. Four time-delays were chosen (fig 38 a-d) – 0 ps, 200 ps, 500 ps and 20 ns.



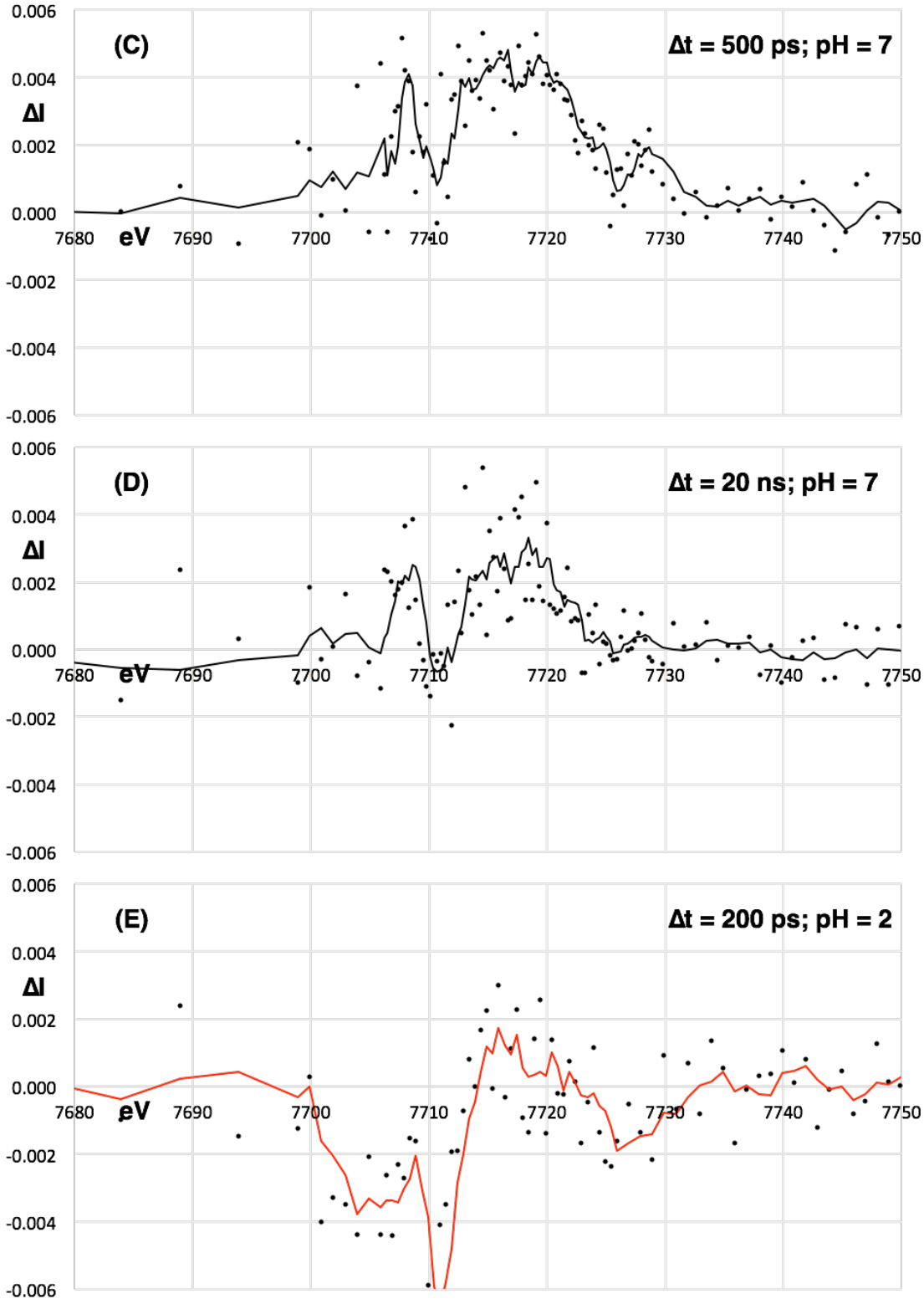


Fig. 5.39: TR-XAS difference spectrum of MeCbl. The dots correspond to experimental data and the solid line corresponds to a moving (4 point) average fit to the data.

The 0 ps (fig. 5.39 a) measurement is under the assignment of the pump having a complete overlap with the probe in time. The 200 ps time delay was also measured with a pH2 MeCbl solution (fig. 5.39 e) to fingerprint the difference between the base-on and base-off photolysis mechanism. The 200 ps pH7 data was averaged over 36 scans, while the rest of the data were obtained by averaging over 10 scans and hence the difference in the signal to noise in the measured data.

The data show a few emerging trends. Figs. 5.39 a and b show a clear reduction in the pre-peak (7710 eV) edge intensity which seems to have recovered moving to longer times as in figs. 39 c and d. Additionally, the entire rising edge, shoulder and strongest absorption peak (from about 7710 eV to 7740 eV) has an increased absorption than the ground state at the 200 ps time delay. This interpretation becomes ambiguous with the other data due to the lack of good signal quality. With the 200 ps pH2 measurement, there is clearly a reduced pre-edge intensity as with the pH7 counterpart. However, spectral data in the region post the pre-edge up to the strongest absorption peak is less certain again due to not sufficient quantity of data although the data trend suggests that there was possibly an increased intensity only around the rising-edge portion of the spectrum. These features might suggest a few things – (i) reduced pre-edge intensity could possibly suggest moving towards stronger centrosymmetry. (ii) increased intensity all through the rising edge and the shoulder leading up to the strongest absorption peak at 200 ps (fig. 5.39 b) could suggest reduction of the metal center.

The analysis of XAS and their relative difference spectra is not possible purely by qualitative fingerprinting based on prior knowledge of certain features from model compounds. Typical model systems are symmetric in their coordination, have ligand field strengths that are generally isotropic and limited in the number of atoms considered constituting the entire complex (amongst other such assumptions). However, in moving to more realistic frameworks, such as MeCbl, the analyses become complicated by the breakdown of these simplistic assumptions (like the appearance of the pre-edge feature albeit having an octahedral coordination). The best way to approach such situations is to avail DFT based modeling. DFT and TD-DFT help reduce the

number of possible solutions by picking out those that are energy conserved and minimized which also match the experimental data (here the absorption spectrum).

4.3 DFT and FEFF Simulations

This section will detail the use of DFT and TD-DFT to obtain energy minimized structures of model ground state and intermediate complexes to mimic the MeCbl photolysis process using simulations. The FEFF code was subsequently used to generate the XAS spectra to compare with and make meaning of the experimental XAS data.

To verify the experimental XAS data obtained from APS, we modeled the sample using DFT to optimize the structure before feeding to the FEFF code for the XAS simulations. Since DFT is a time and computationally expensive process, truncations were applied to the MeCbl molecule to utilize the smallest possible structure that would still reproduce some standard (optical absorption) experimental signatures of the sample faithfully. Prior work [33,36,46] has shown that a truncated structure with (a) the ring cut short to the 15 carbons and terminated with hydrogen atoms as appropriate (b) the lower axial DBI approximated to an imidazole (c) the lower axial imidazole detached from the ring and (d) charge balance compensation by having the molecule under an overall positive charge (absence of the phosphate counter ion that would otherwise have balanced the charge) has been reasonably successful in replacing the entire MeCbl molecular structure for DFT simulations and this work sticks to the same prescription.

For the DFT and TD-DFT calculations, the non-hybrid (GGA) Becke-Perdew (BP86) exchange-correlation functional was used. The 6-31(G)-d level of basis set was utilized for the structure optimization. The Gaussian 09 [63] code was used for the DFT and the TD-DFT calculations.

The FEFF code [55] uses multiple scattering approach in the real space to simulate the X-ray absorption spectra of materials. The code requires an input of the atomic structure (as crystal with unit cell parameters or as a molecule with real space coordinates for each atom). The type of calculations required are mentioned through the use of CARDS that define the parameter space. The calculations were done with full multiple scattering enabled with a effective radius from the

central atom (Cobalt) of 6 Å. The self-consistent fields were calculated for a radius of 5 Å from the central atom (large enough to include up to the second nearest neighbor shell). An imaginary broadening of the spectrum was added as 1 eV across to account for instrumental broadening of the spectral lines. As for the self-energy, the Hedin-Lundqvist model (default with FEFF) was used. The free atom potentials are calculated using a relativistic Dirac-Fock atom code for each atom as if isolated in space. Then, the scattering potentials are calculated by overlapping free atom densities in the muffin tin approximation, and finally including the Hedin-Lundqvist self-energy for excited states. For calculating the density of states, FEFF uses the LDOS card, which can calculate the angular momentum projected density of states for each element whose potential has been declared. For the calculations, the DOS energy broadening was set to 0.1 eV.

Eight model structures (and their variants) were selected as representatives of the ground state and intermediate structures as shown in fig. 5.40 (a,b and c). The models have been created based on the ground state structures and the most probable intermediate/product structures from prior experimental predictions. All models have been truncated to a shorter and smaller version of the complete MeCbl molecule as has been performed in prior literature.[33,46] In such a truncation, the lower axial DBI which connects to the corrin ring is substituted by an imidazole that is disconnected to the ring.

V0 represents the possible structure for ground state pH7. V1 and V6 represent the possible structures for the pH2 MeCbl ground state. In the pH2 (acidic) MeCbl case and for the product structures, there is ambiguity in whether or not there is water ligation at (a) the photolytically cleaved Co-C bond position and (b) at the possible Co-N_{axial} bond cleavage. The latter comes with four independent possibilities – (i) water solvation effects, (ii) explicit water ligation, (iii) water as an oxo bridge to the imidazole nitrogen and (iv) no water ligation. Since the V6 sub-models all converged to the same structures, only V6c was used as a representative case.

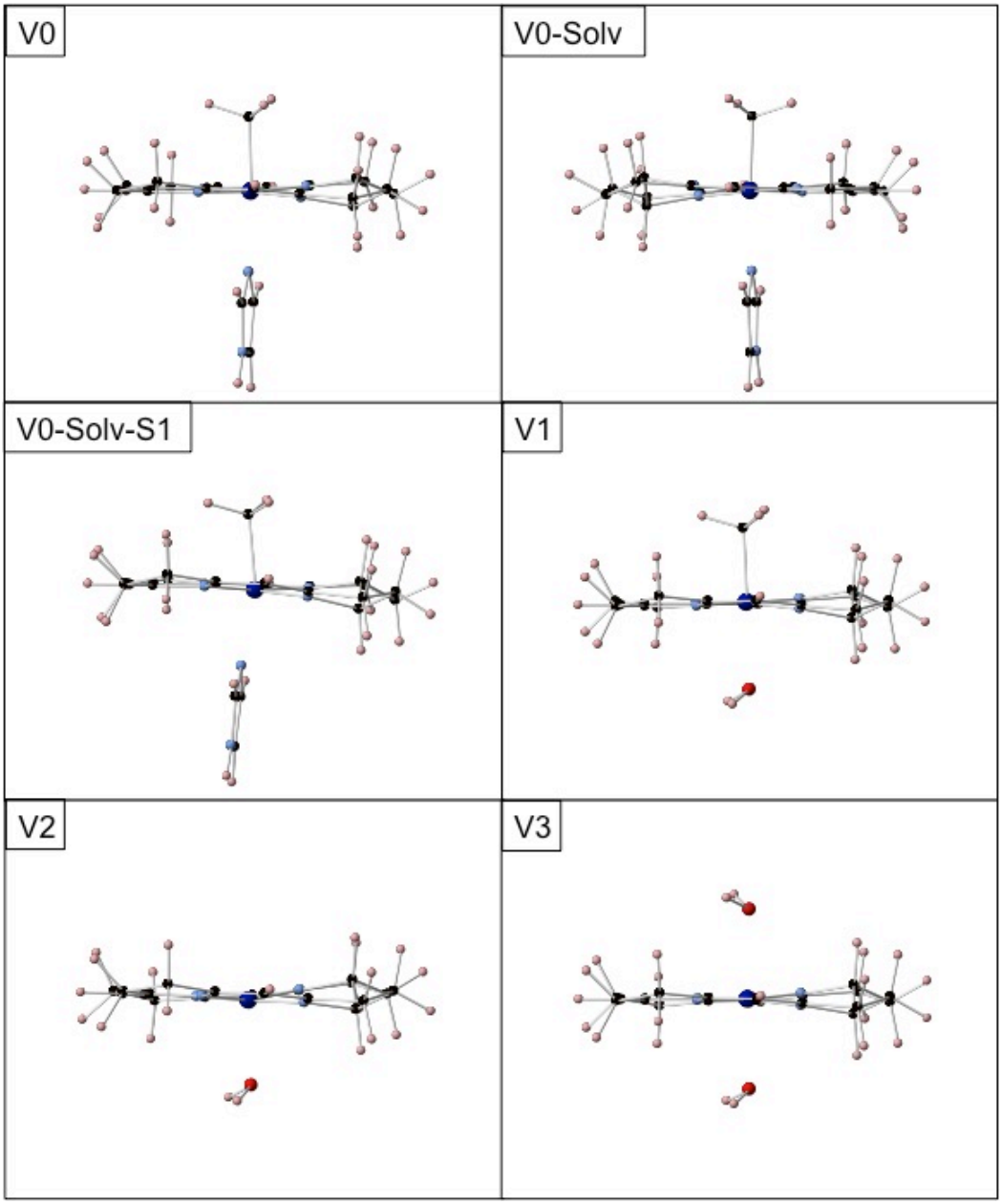


Fig. 5.40 a: DFT optimized MeCbl model structures. V0, V1 and V6 representing ground state pH7 and pH2. V2, V3, V4, V5, V7 and V8 represent the intermediate or final product. Here V0 is the pH7 MeCbl ground state. V0-solv is the ground state structure with a solvation sphere of water around it. V0-Solv-S1 is the TD-DFT optimized structure of the first excited state of V0 with a solvation sphere around it. V1 is the pH2 MeCbl ground state structure. V2 is an intermediate/product state with both axial carbon and nitrogen detached and instead substituted by one water ligand. V3 is an intermediate/product state with both axial carbon and nitrogen detached and instead substituted by two water ligands.

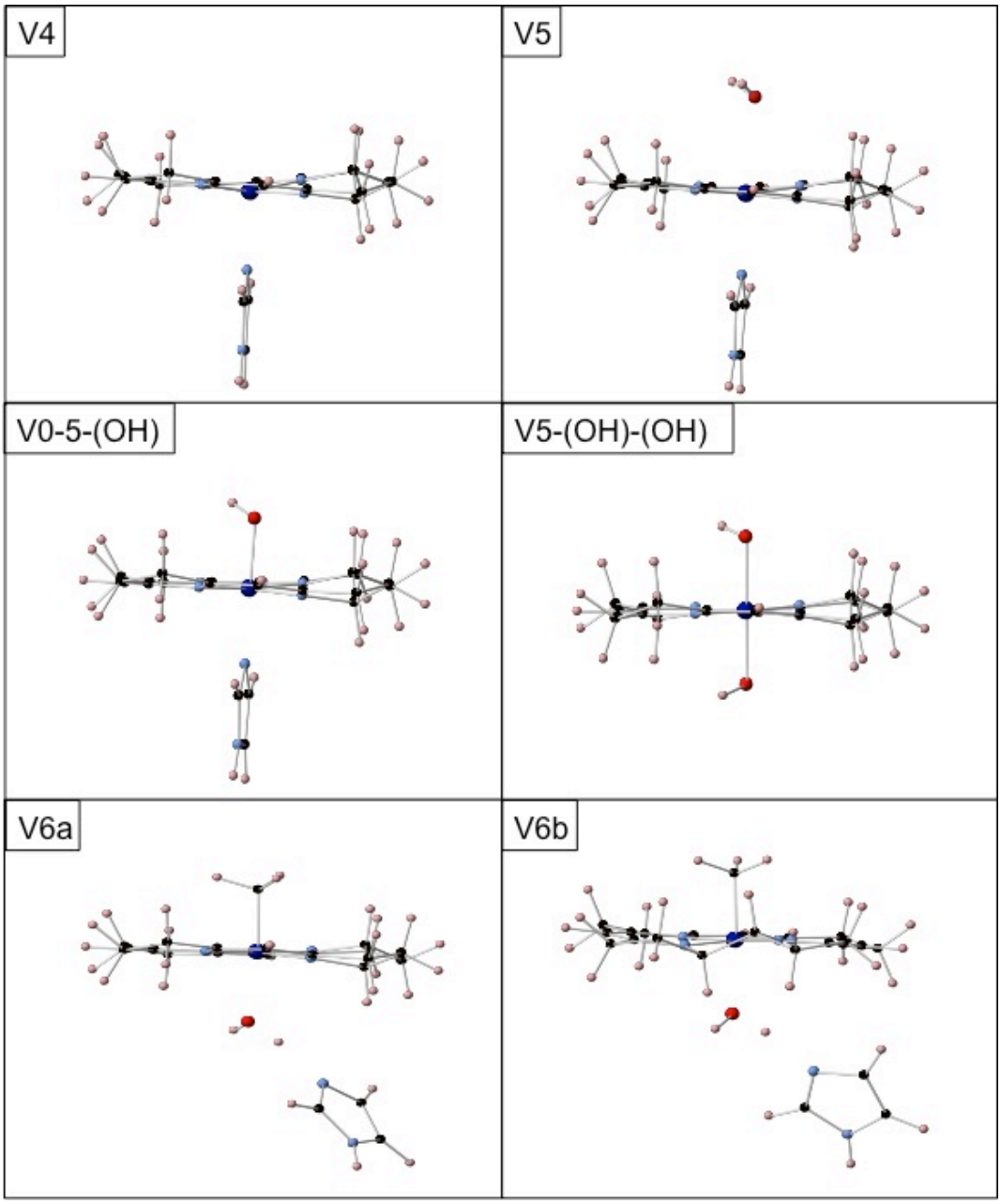


Fig. 5.40 b: DFT optimized MeCbl model structures. V0, V1 and V6 representing ground state pH7 and pH2. V2, V3, V4, V5, V7 and V8 represent the intermediate or final product. V4 is an intermediate /product state with upper-axial carbon detached and the lower-axial nitrogen intact. V5 is an intermediate /product state with upper-axial carbon detached, but substituted by a water ligand and the lower-axial nitrogen intact. V5-OH is an intermediate/product state with upper-axial carbon detached, but substituted by a hydroxyl ligand and the lower-axial nitrogen intact. V5-OH-OH is an intermediate/product state with both the axial carbon and nitrogen detached, but substituted instead by two hydroxyl ligands. V6a is ground-state pH2 MeCbl model with upper-axial carbon intact, but lower axial nitrogen ligated to the cobalt through a water bridge in conformation-1. V6b is same as V6a but with the axial nitrogen in conformation-2.

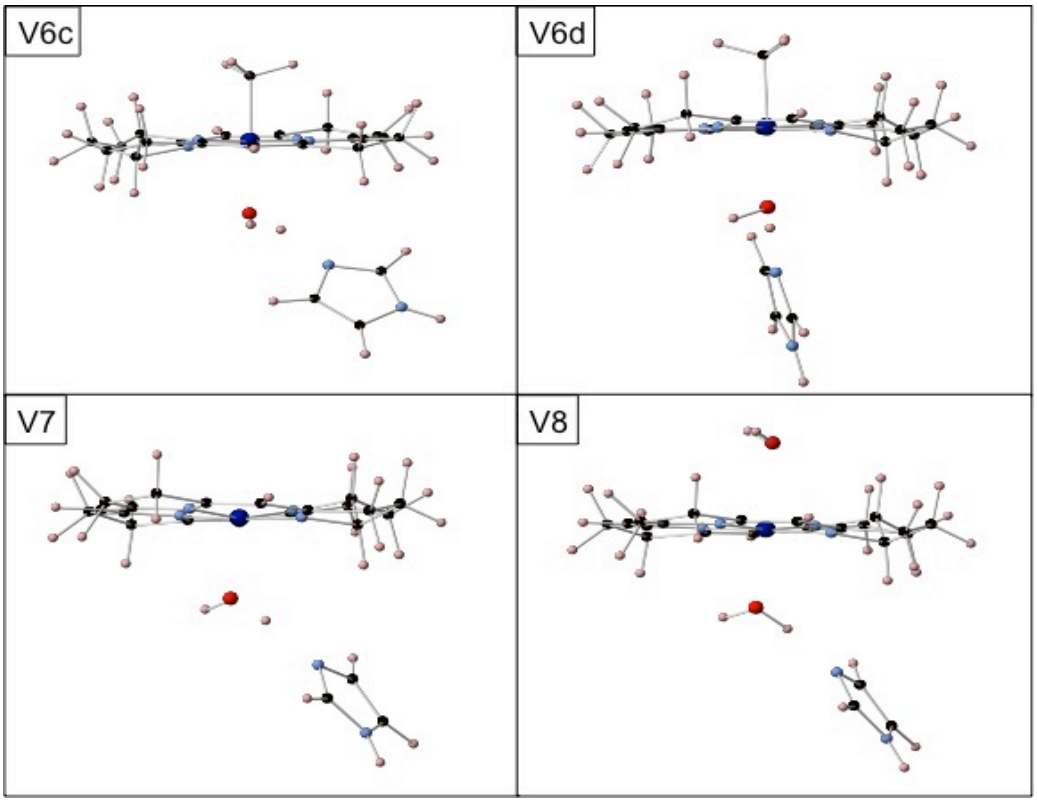


Fig. 5.40 c: DFT optimized MeCbl model structures. V6c and V6d are the same as V6a but with the axial nitrogen in conformations - 3 and 4 respectively. V7 and V8 are intermediate/product states with upper-axial carbon detached (V7), but substituted with water (V8) and the lower axial nitrogen ligated to the cobalt through a water bridge.

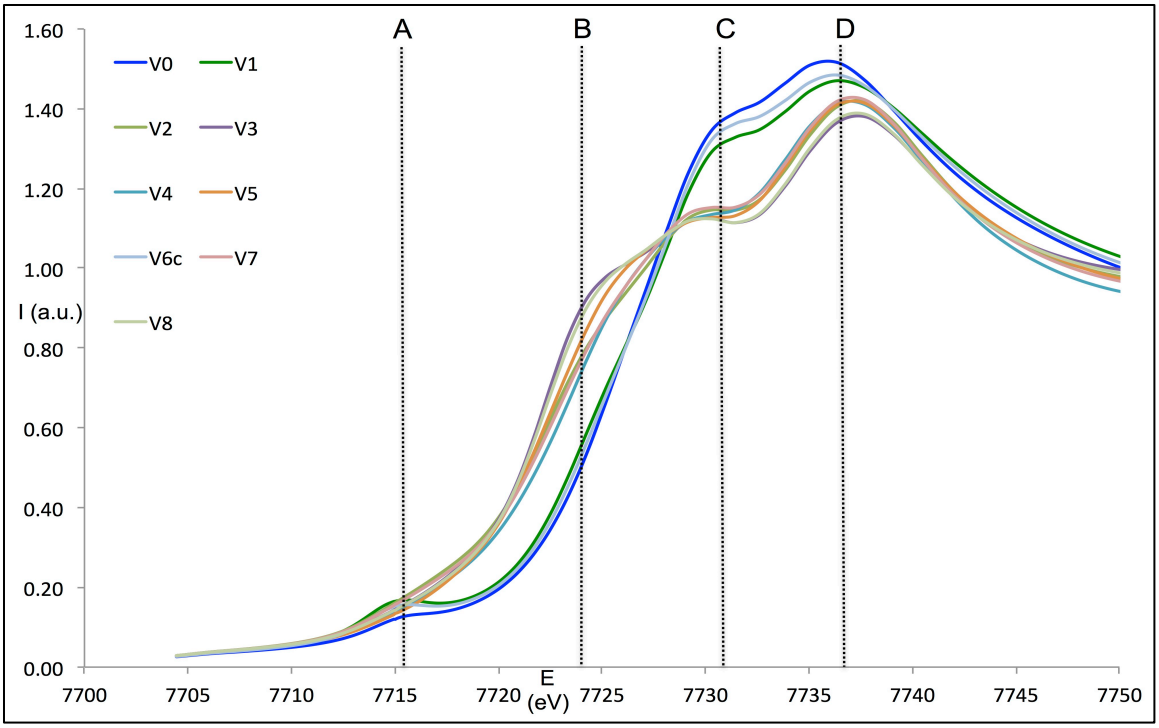


Fig. 5.41: XANES simulations using FEFF of all the DFT optimized model structures.

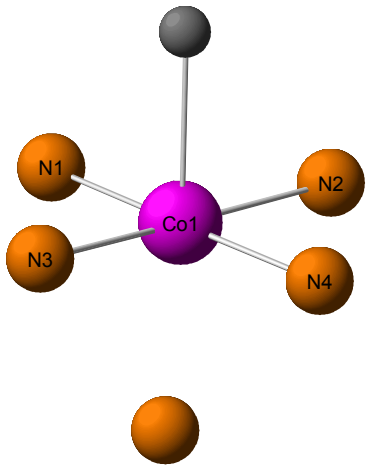


Fig. 5.42: V0 reference model structure for 1st shell atom labels and nomenclature.

	Co-N2	Co-N4	Co-N3	Co-N1	Co-L (upper axial)	Co-L (lower axial)
V0	1.87	1.87	1.93	1.93	1.97	2.14
V0-solv	1.87	1.87	1.93	1.93	1.97	2.12
V0-solv-S1	1.87	1.87	1.97	1.97	1.98	2.02
V1	1.86	1.86	1.92	1.92	1.95	2.27
V2	1.86	1.86	1.92	1.92		2.22
V3	1.86	1.86	1.92	1.92	2.34	2.34
V4	1.86	1.86	1.93	1.92		2.10
V5-OH	1.87	1.88	1.94	1.94	1.87	2.00
V5-OH-OH	1.87	1.87	1.95	1.94	1.93	1.93
V5	1.86	1.86	1.93	1.92	2.58	2.16
V6a	1.86	1.86	1.92	1.93	1.95	2.18
V6b	1.86	1.87	1.92	1.92	1.96	2.18
V6c	1.86	1.86	1.93	1.92	1.96	2.18
V6d	1.86	1.86	1.93	1.93	1.95	2.19
V7	1.86	1.86	1.92	1.92		2.15
V8	1.86	1.86	1.92	1.92	2.22	2.43

Table 5.1: Bond-distances (Å) for the different models for the equatorial and axial ligands in 1st shell. Color code: Black – Carbon, Brown – Nitrogen, Aqua blue – Water, Blueberry – Water bridged imidazole, Midnight – Hydroxyl.

4.4 Analysis of Ground State Reactants:

In analyzing the data, first comparisons will be made with ground state structures. Experimental and simulated spectra of the ground state reactant MeCbl pH7 and pH2 are presented together in fig. 5.43. From the two spectra, it is visible that the qualitative features of the experimental spectra are well represented in the theoretical spectrum. The energy scale has been shifted to match the simulated spectra to the experimental spectra. It is important to specify here that the V1 and V6c models, both seem to suit the pH2 MeCbl experimental spectra. However, it clearly makes a difference to have an explicit water ligation to the central cobalt atom

rather than an empty lower axial ligand or adding the effect of solvation. To look into further detail of the simulated spectra, the ldos for the three cases of V0, V1 and V6c are presented next.

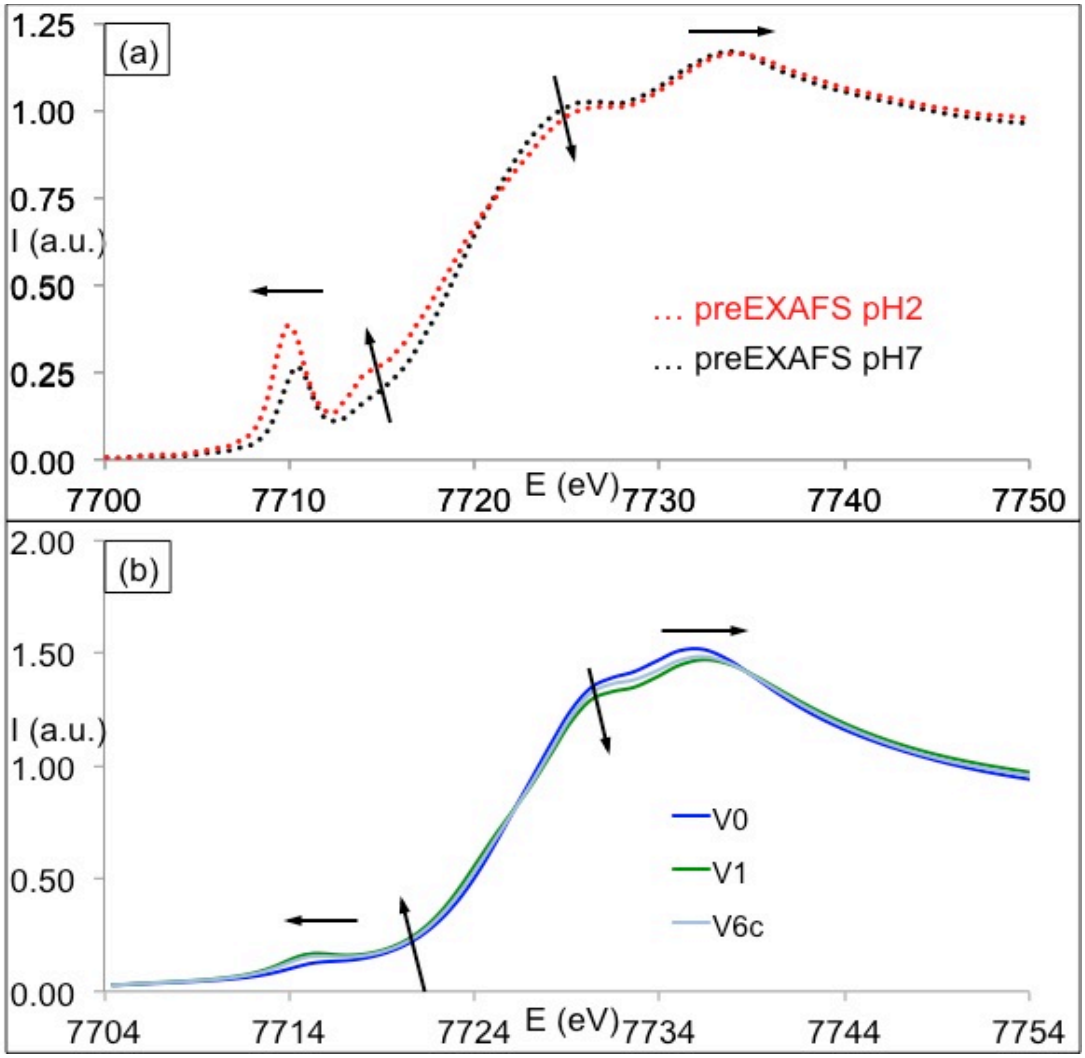


Fig. 5.43: MeCbl XANES spectrum for ground state pH7 and pH2 (a) experimental and (b) FEFF simulations of DFT models (V0 – pH7 and V1 and V6c – pH2.)

A point to note here is that the FEFF code predicts for the shift in the Fermi Energy during the ldos calculations. However, shifting the energy axis by this value is known to not be sufficient. In the case of the FEFF simulations here, the energy needed to be adjusted by about 6-7 eV post the FEFF correction. This was chosen by visual inspection rather than by using a rigorous procedure since the aim of the ldos calculation was only to obtain a qualitative reference

for the changes noticed in the XANES simulations across the different models. In all further discussions involving the lDOS, the energy axis of the XANES (in bold) will be taken as reference to note positions of peaks.

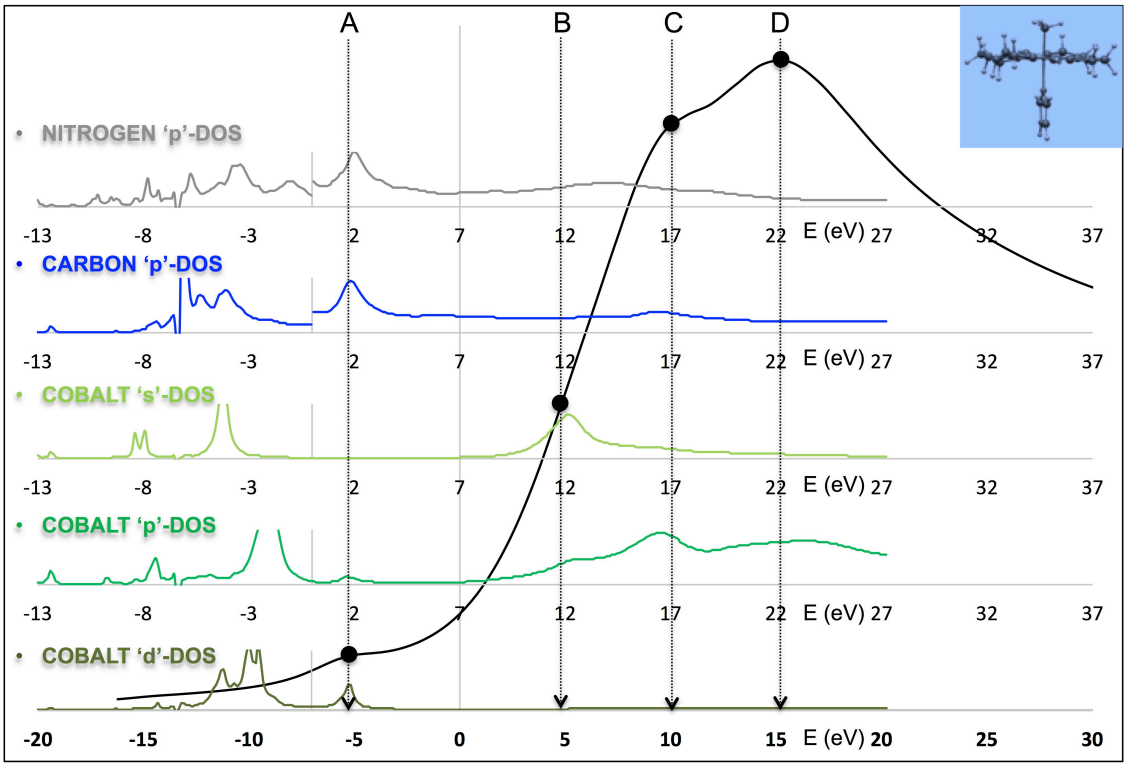


Fig. 5.44: Simulated XANES (black solid line) and ldos for V0. Inset shows the structure of V0 (ground state pH7 MeCbl with upper axial Carbon and lower axial Imidazole)

Fig. 5.44 displays the XANES spectrums of V0 (pH7 ground state of MeCbl) along with the individual ldos components. The lowest energy axis is the XANES spectrum and the rest correspond to the energy axes of the DOS calculations that are shifted to align with the XANES. From initial inspection it is evident that the predominant contribution to the XANES comes from the cobalt pDOS. This is expected since the excitation is of the K edge, which implies the electron is excited from the 1s orbital. The final state is expected to be dominated by a 4p character since the dipole contribution (in the multipole expansion) is the strongest contributor to the transition. All the three features on the absorption edge (B, C and D) are clearly seen to appear from the cobalt pDOS (fig. 5.44).

The pre-edge region (peak A) seems to have a contribution from the pDOS and dDOS along with the carbon and nitrogen pDOS. This can be interpreted as two types of transitions – (i) a direct 1s-4p transition as suggested by the cobalt pDOS and (ii) a 1s-3d(2p) transition that involves the mixing of the 3d metal orbital with the 2p ligand orbital giving the other quadrupolar 1s-3d transition some dipole character with the ligand p mixing.

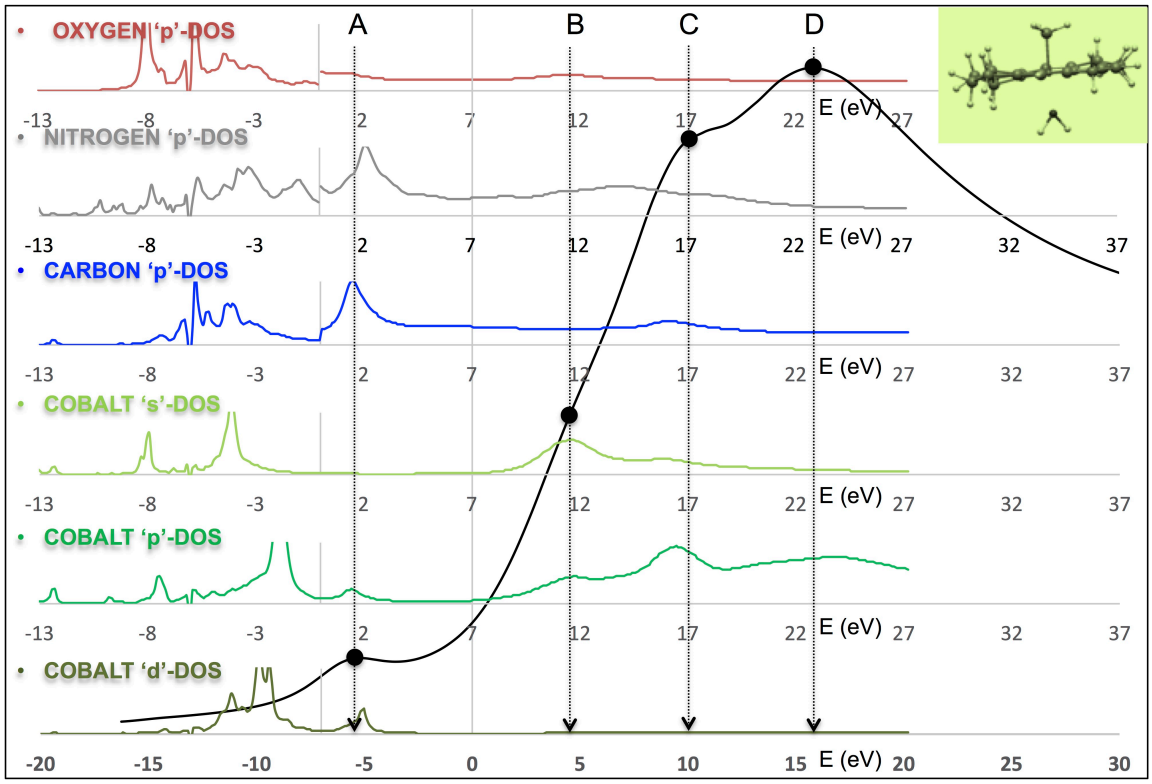


Fig. 5.45: Simulated XANES (black solid line) and ldos for V1. Inset shows the structure of V1 (ground state pH2 MeCbl with upper axial Carbon and lower axial water)

Performing the same type of calculations for V1 it can be seen (fig. 5.45) that the predominant contributions to the XANES is from the cobalt pDOS as with V0. The same explanations as for V0 holds with the pre-edge peak A. However, as was shown in fig. 5.43 b, there are some subtle differences in the peak features and intensities that differentiate V0 from V1. To understand the possible reasons for the difference, it would warrant comparing a few partial DOS for both these models together as in fig. 5.46.

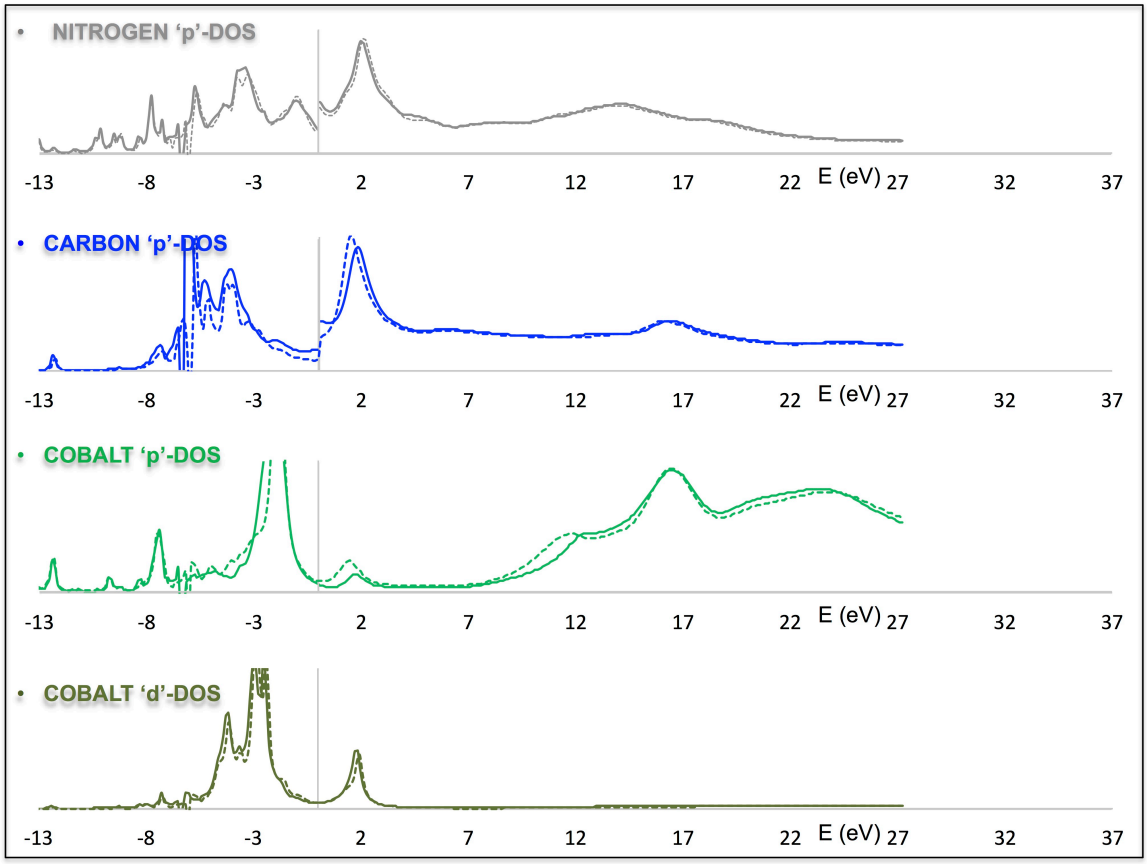


Fig. 5.46: Simulated ldos for V0 (solid line) and V1 (dotted line).

Selected partial DOS are compared for V0 and V1 in fig. 5.46. Starting from cobalt pDOS, there are three important inferences. First, the ldos about peak A position (pre-edge) of V1 is red-shifted and possesses more vacant orbital density (seen as the increased intensity). This could already explain the red-shifted and stronger intense pre-edge feature in fig. 5.43 b. In addition, the metal 3d ligand 2p mixing depends on the energy and vacancies of both the metal and the ligand 2p orbitals. For the case of V0 and V1, while the metal 3d orbitals are pretty similarly occupied and positioned in energy, there is an observable difference in the position and intensity of the ligand 2p orbitals. The carbon 2p orbitals of the V1 model are red-shifted at the pre-edge region and also show a higher unoccupied DOS. This is then in effect the same type of contribution as the metal 4p. In going from V0 to V1 there are two major changes –

(i) lower axial nitrogen is absent: This means that from 5 nitrogen ligands in V0 (4 from the corrin ring and one lower axial) the shift to V1 happens with only the equatorial corrin ring nitrogens as ligand. However, a glance at fig. 46 for the nitrogen pDOS reveals that there is almost no visible difference in the pDOS for the two models. At this juncture it is important to understand what FEFF means when it provides a pDOS output. FEFF's partial DOS is the angular momentum projected DOS for the element. This means, as for the element nitrogen, it calculates the p-type DOS for all the nitrogens in the molecule (a cumulative DOS for the element nitrogen). This is different from DFT or MO based methods that can project the orbitals of every atom (not element) onto the central atom (in this case cobalt) to obtain the projected density of states from a specific atom. This being the case, the way to interpret the comparison between V0 and V1 pDOS of nitrogen is that, starting from V0 and moving to V1 the total number of nitrogen atoms in the molecule is reduced by one (the lower axial nitrogen is absent in V1). If the nitrogen atom absent had been from a further shell, then its influence on the cobalt atom would be meager. However, the lower axial nitrogen is of the first coordination shell so the impact of its presence or absence on cobalt is crucial. But, as mentioned above, there is hardly any difference in the pDOS of nitrogen in moving from V0 to V1. The only explanation to this is that the V0 lower axial nitrogen bound through an imidazole is a long bond (bond length = 2.13 Å) and hence it is weak and qualitatively has no significance in affecting the cobalt DOS.

(ii) Oxygen (of water) ligation to cobalt: V1 has the only oxygen in the entire molecule ligated directly to the cobalt center. This is entirely different from V0 where there are no oxygen atoms in the system at all. However, fig. 5.45 shows the impact of oxygen in the pre-edge region is not substantial. There is no strong density of unoccupied states in the region. This could make sense under the assumption that the water ligation is not a strong bond and is weak and long (DFT optimized V1 has a Co-O bond distance of 2.269 Å).

Moving to the absorption edge, there are three changes noticed (fig. 5.43 b) in V1 in comparison with V0. Peak B is more intense and red-shifted. This is clearly evident from the cobalt pDOS in fig. 5.45 where the dotted lines (V1) are red shifted and marginally more intense than the solid lines (V0). The next change noticed is the intensity of peak C that is reduced in V1

as compared to V0. This is something that the DOS do not directly make an analysis case for. The cobalt pDOS at peak C is about exactly the same for both V0 and V1 as per models. The only difference could be in the multiple scattering or the oscillator strengths of the respective transitions, which could be a small yet significant change not accounted by using just the DOS. Finally the peak D intensity in fig. 5.43 b is blue shifted and marginally lower in V1 than in V0. The pDOS of cobalt (fig. 46) show the exact same trend. A quick remark is to be made on the only other model that represents the ground state reactant V6c (fig. 5.47). V6c is almost identical to V1 in the XANES (fig. 43b) and LDOS. There is almost no difference in the partial DOS components of most elements with the exception of oxygen. Even with oxygen, the qualitative features are same and changes seems to be only with intensity of the unoccupied DOS at a given energy. Hence either of V1 or V6c could represent the ground state pH2 base-off structure.

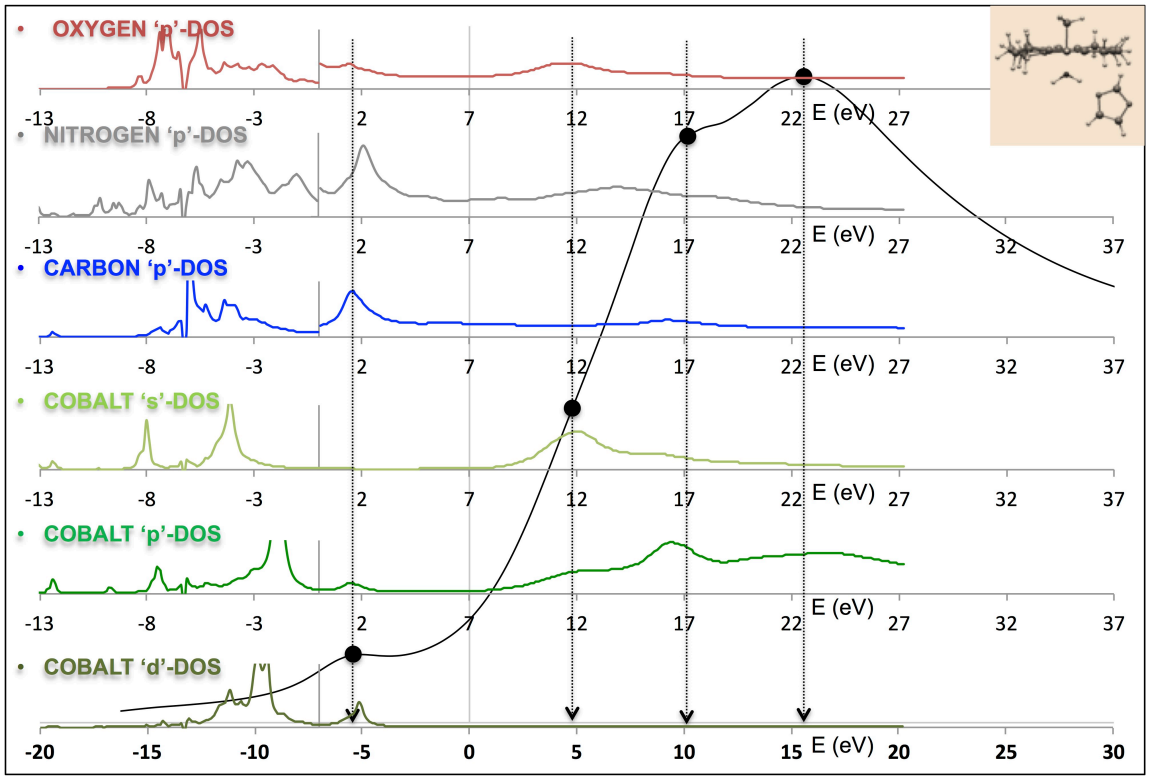


Fig. 5.47: Simulated XANES (black solid line) and ldos for V6c. Inset shows the structure of V6c (ground state pH2 MeCbl with upper axial Carbon and lower axial water bridged imidazole).

In the analysis of the ground state reactants, the models V0 and V1/V6c represent to a very good extent, the sample structure and chemistry.

4.5 Analysis of Ground State Products:

The post-EXAFS spectra of the ground state products are presented in fig. 5.48 along with the pre-EXAFS spectra for both pH7 and pH2.

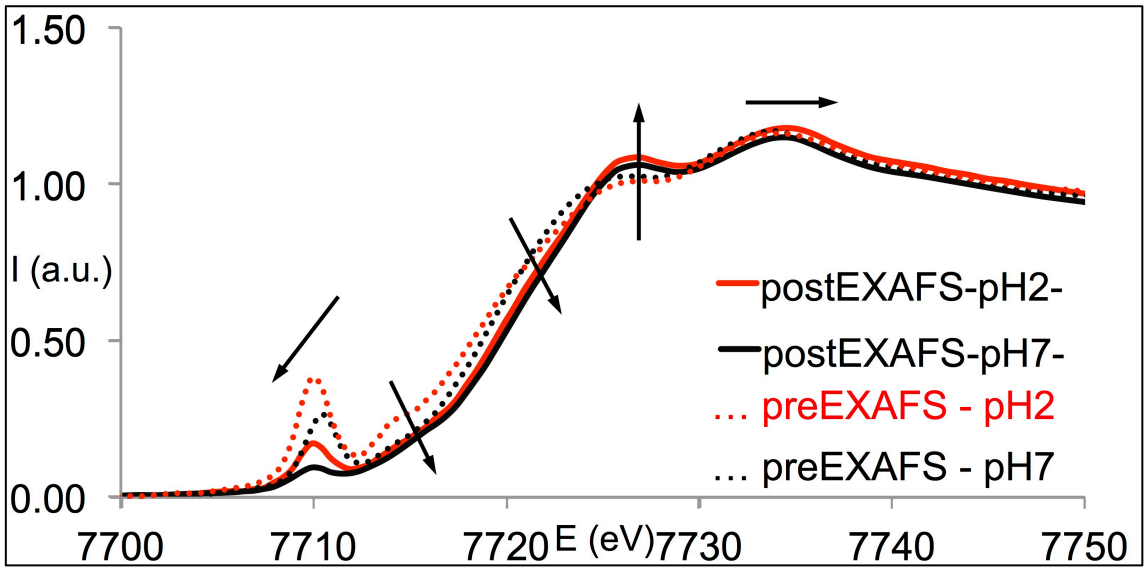


Fig. 5.48: MeCbl XANES experimental spectrum for ground state experimental pH7 and pH2 before (dotted lines) and after (solid lines) laser exposure.

It is very interesting to note that the post EXAFS spectra for both pH2 and pH7 are very similar except for the pre-edge peak intensity. As mentioned with reference to fig. 5.36, the post-EXAFS spectra are strongly indicative of the formation of aquo-cobalamin. The variation in the pre-edge intensity could arise from the different (pump-laser) exposure times for the pH2 (1 hour) as compared to the pH7 (3 hours), which would consequently determine the amount of products formed. In such a case, since the sample would not have completely converted to the products, the final spectrum in both cases would be a linear superposition of the ground state reactant spectrum (base-on / base-off MeCbl) and the pure product spectrum (aquocobalamin).

4.6 Analysis of Intermediate States:

Before the analysis extends to the intermediate state spectra, there is an important concern to be clarified. In the process of the X-ray measurement, there is a possibility of measuring spectra from damaged species or other unwanted products formed unintentionally. As the X-ray probes the liquid jet isotropically without any selection, the spectra would contain signal from any cobalt that is present. So in order to confirm that the X-ray spectra contain the difference signal of the excited state with the reactant, pump-probe transient optical absorption measurements were performed (at ASU) on these samples post X-ray exposure at the APS.

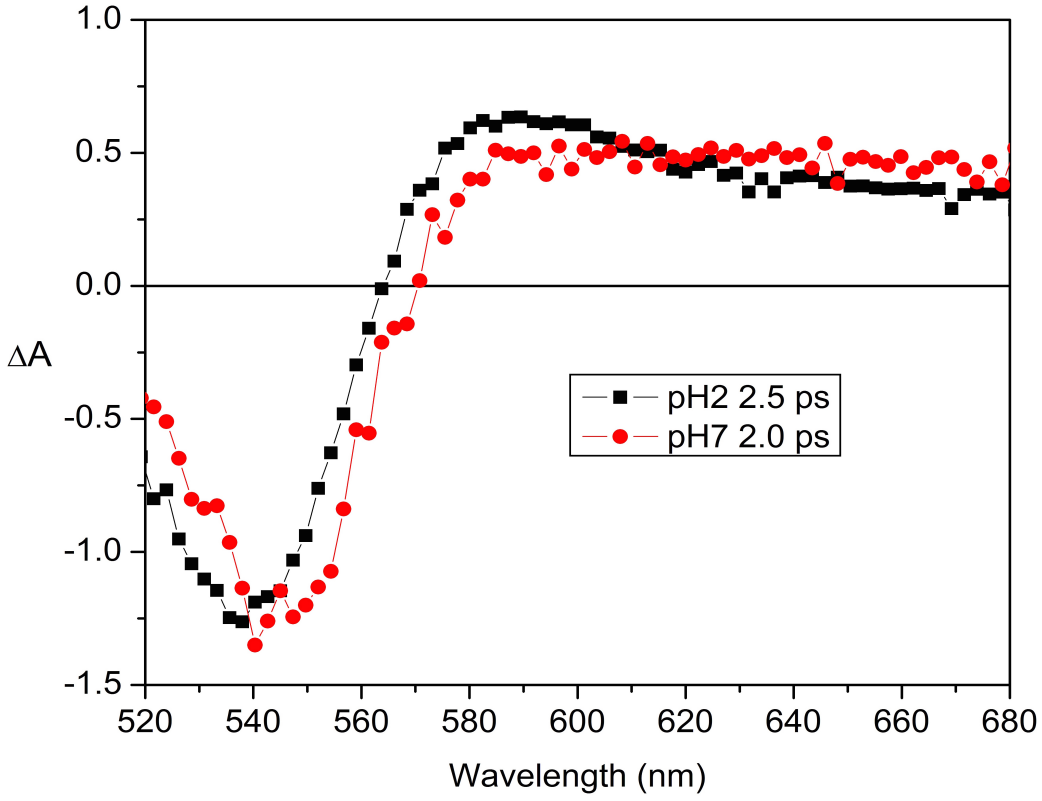


Fig. 5.49: Pump-probe transient optical spectra of the TR-XAS exposed samples for both pH7 and pH2 (containing products).

Fig. 5.49 shows the difference absorption spectra for pH7 and pH2 of the samples that have been laser-pumped and X-ray probed at APS. Interestingly, there is only one time constant at about 2 ps for both cases. This is indicative of two factors – (a) with the time constants being

so short it can be safely assumed that no excited states are contributed by the products formed over time such that deconvoluting the excited state XAS spectrum before calculating the difference with the ground state is necessary. Hence all the data presented in fig. 5.38 are valid as the difference spectra obtained from exciting the ground state MeCbl (pH7 / pH2) and the corresponding excited state. (b) From prior literature [64,65] it can be seen that this is consistent with the presence of a hydroxyl (aqua) axial ligand (CN-Cbl also has very similar time constants. However, there is no scope for the formation of cyanide in this experiment and hence can be safely ruled out). Hence this is yet another confirmation of the presence of an aqua/hydroxyl ligand in the product.

Having established that the difference spectra in fig. 5.38 do not have any contributions from excited states of the products, it is now apt to analyze fig. 5.41 in detail in comparison. Fig. 5.50 is the 200 ps light vs. dark X-ray absorption spectrum showing the changes at the different positions A-D (as highlighted by the inset). We chose this specific spectrum because it has the best and sufficient signal above noise to evaluate the spectral features without discrepancies.

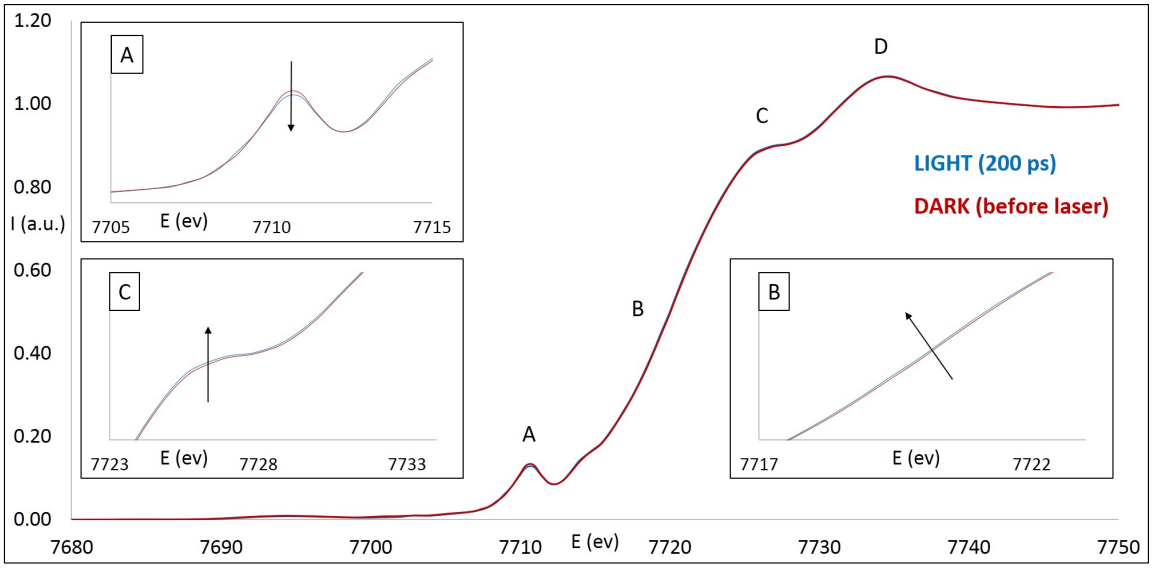


Fig. 5.50: X-ray absorption spectra of the dark (before laser) and light (200 ps after laser) pH7 sample. Insets highlight the regions corresponding to peak A, B and C.

This is the same as fig. 38 b except that it shows the complete absorption spectrum rather than the difference making it easier to compare with fig. 5.41. In fig. 5.41, all the intermediate/product spectra (V2, V3, V4, V5, V7 and V8) have a broad and weak pre-edge feature A, a strong red shift of the absorption edge (B) and reduced intensity at peak C. The experimental spectra on the other hand, the intermediate (fig. 5.50) and product spectra (fig. 5.48), both show a reduced pre-edge intensity and a stronger (more intense than the ground state) peak C. However, at the absorption edge rise (B); while for the intermediate spectrum the absorption edge is red-shifted, the product spectrum is blue-shifted as compared to the ground state. The simulated spectrum matches (and qualitatively) only in the pre-edge region where the peak intensity is reduced and broadened.

A strong discrepancy of the simulated spectra from the models and the experimental (both intermediate and product) is at the position of peak C. The simulated spectra (as explained before) show that the peak intensity is weaker at C for the intermediate/product models than the ground state, opposite to the experimental spectra.

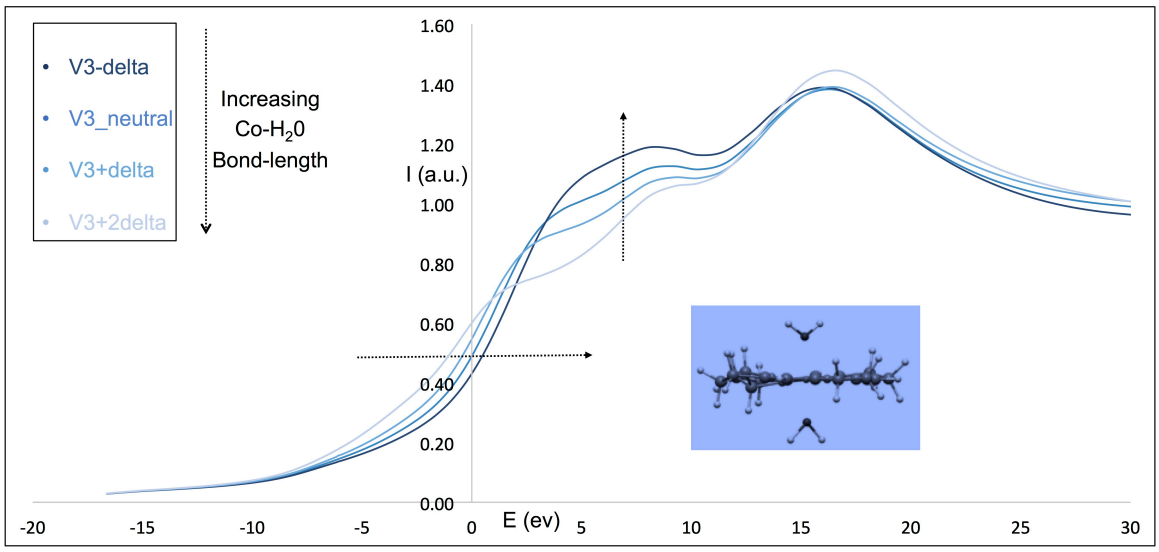


Fig. 5.51: V3 simulated XAS spectrum. Increasing bond-lengths in steps of 0.1 Å. V3_neutral is the DFT optimized distance for the axial aqua ligands to cobalt. V3-delta is the bond distance reduced by 0.1 Å. V3+delta is the bond distance increased by 0.1 Å. V3+2delta is the bond distance increased by 0.2 Å.

To analyze this further, the V3 model was studied in more detail. Taking the DFT optimized structure, the bond distance of the aqua ligands (both upper and lower axial) to cobalt is modified (+delta meaning an increase in bond length and –delta vice versa, in steps of 0.1 Å. This is an impractical however educative example of the effect of the axial bond distance on the absorption spectrum assuming that the rest of the system remains unchanged. Fig. 5.51 shows that in such an analysis, the closer the axial bonds (decreasing bond length) are to the cobalt, the stronger the peak C intensity is. This is consistent with prior claims of similar nature [59] in spectroelectrochemical studies on the cobalamin and related systems.

It is also observed that the shorter bond blue shifts the edge rise (fig. 5.51). This is important to note since although the intermediate is red-shifted (compared to ground state) at B, the product is blue shifted. None of the simulations reflect this, the reason for which needs to be identified. Nevertheless, the above signature is noteworthy and to understand the effect of bond, the V5 model system was analyzed next. In this case, the upper axial aqua ligand was first substituted by a hydroxyl ligand (V5-OH) and then subsequently, the lower axial imidazole was also substituted by a hydroxyl (V5-OH-OH) as shown in fig. 5.52. (Each of these modified models were DFT optimized unlike in the case of fig. 5.51)

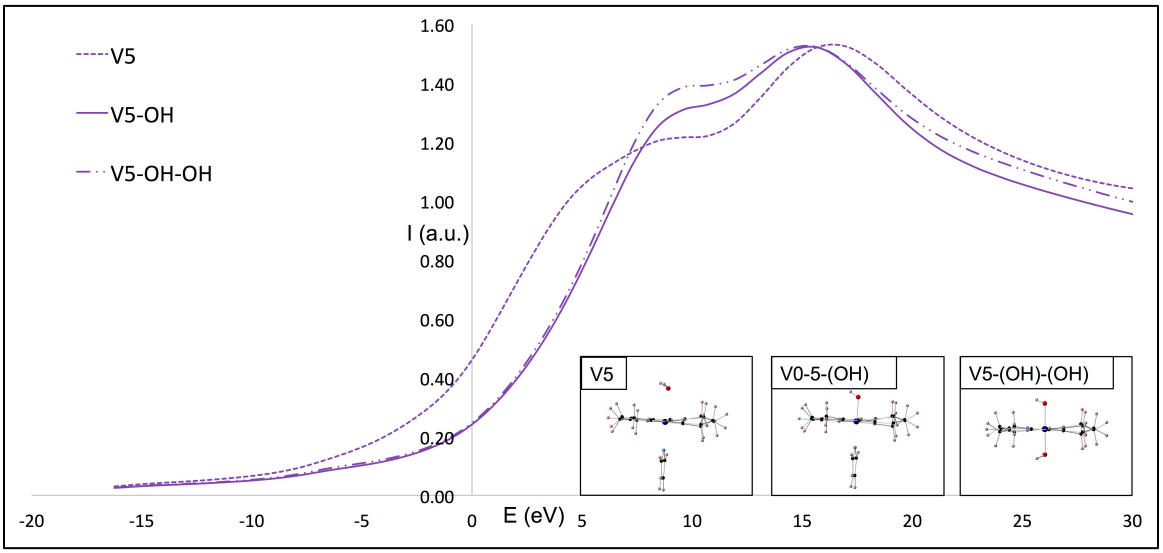


Fig. 5.52: Simulated XAS with different nature of ligands for the V5 (upper aqua and lower imidazole), V5-OH (upper hydroxyl and lower imidazole) and V5-OH-OH (both hydroxyl).

Similar to the case in V3 (fig. 5.51), here as well it can be noticed that the modified models show a much stronger intensity at the peak C and strongly blue shifted edge rise at B. However, here the change is the nature of the ligand. The water (aqua) ligand in the original V5 model is linked to the cobalt through a coordinate bond while the hydroxyl forms a stronger and shorter covalent bond (table 5.1). Interestingly, in going from the single substituted V5-OH to the doubly substituted V5-OH-OH, although the peak C intensity increases, the edge-rise has a slight red- shift. This poses very interesting scenario with the axial ligands.

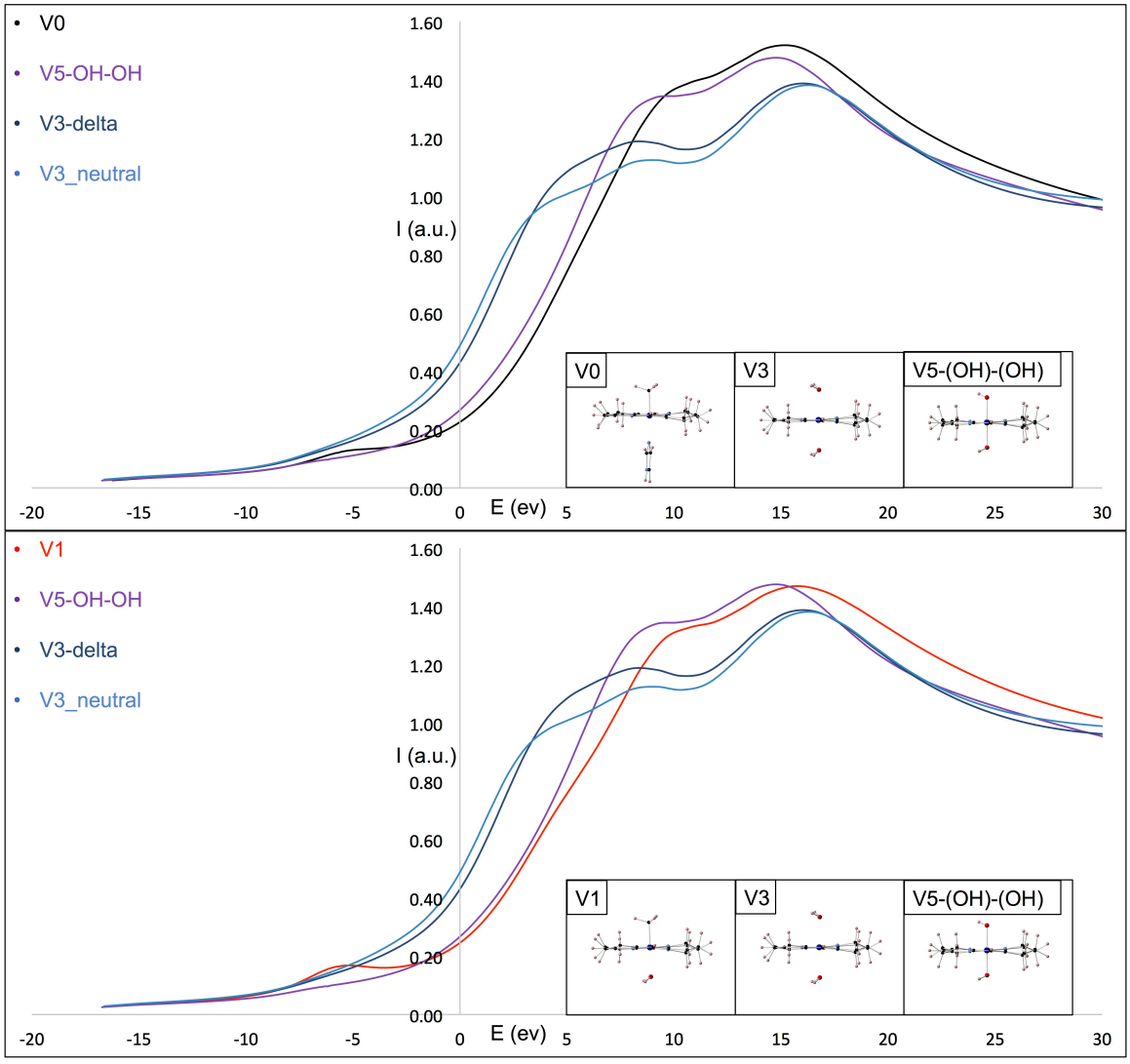


Fig. 5.53: Simulated XAS comparison of V3, V3-delta and V5-OH-OH models with ground state. TOP: Comparison with pH7 ground state (V0). BOTTOM: Comparison with pH2 ground state (V1)

In the next figure (fig. 5.53) the modified best case scenarios of the V3 and V5 models are compared with the ground state models (V0 and V1). It is very clear from this that compared to the original models for intermediate/product (V3_neutral), the modified models V3-delta and V5-OH-OH are both better suited to match the experimental spectra for the intermediate/products. However, it is very interesting to note that thus far V5-OH-OH is the closest to the experimental intermediate/product spectra and it still falls short of predicting the increase in the peak C intensity as is seen in fig. 5.48 and fig. 5.50. It is also noteworthy that though the peak C intensity of V5-OH-OH is marginally higher than V1, it is smaller than V0.

5 Conclusions

This chapter elucidates the possibility of performing frequency domain measurements of dynamics of biomolecules using X-ray as a probe. The specific case of interest is Methylcobalamin. With a substantial amount of prior information on this system studied using various different techniques already present in literature, certain fundamental questions regarding the nature of the intermediate state in the photolysis of MeCbl still remained unanswered. For the first time, time-resolved pump-probe X-ray absorption spectra were measured on the system in the ps-ns time regime. Models were created, which are truncated versions of the entire MeCbl molecule, optimized with DFT and TD-DFT whose structures were then fed to the FEFF code to simulate the XAS.

The simulations of XAS match well with the ground state pH7 and pH2 MeCbl qualitatively matching the energy shifts and the relative intensities of the peaks. It is noticed here that for the case of pH2 MeCbl, there is a need for an explicit water molecule (stand alone or acting as a bridge to link cobalt to the imidazole) as a ligand to simulate the XA spectrum well. For the postEXAFS corresponding to the product state, the experimental spectra measured at APS match qualitatively with prior experimental data. This coupled with the post X-ray optical steady state and transient measurements suggest that the final products are ligated with water/hydroxyl groups. The transient spectra show difference, which are mostly similar to the difference spectra of products vs. the reactants. They both show a reduced pre-edge intensity,

and a stronger (shoulder) peak C. However, the intermediate difference spectrum shows a red shifted absorption edge-rise as opposed to the case of the product spectrum, which shows a blue shift in the same. The DFT+FEFF simulations have not been able to convincingly reproduce either of the intermediate/product spectra. However, there is a major thrust towards the role (nature and distance) of the axial ligand in the absorption spectral features.

The next stage in the research would be to pursue the role of the axial ligand in an elaborate and systematical manner to identify and isolate the model(s) that match the experimental spectra. Subsequently, more experiments shall be performed at the other time delays to obtain a broader and complete picture of the evolution of the intermediate state with time. Finally, it would be ideal for the experiments to be repeated at the LCLS, which supports X-ray pulses that are 3-4 orders of magnitude shorter than at the APS to study the electronic transitions at the sub, and few ps timescales before nuclear rearrangement sets in. This would then complete the understanding of the photolysis of MeCbl across the timescales from fs to ms.

6 References:

1. Mccaddon, A., Hudson, P. R. & Mccaddon, B. A. Spectrums : Treatment of Alzheimer ' s Disease-Related Cognitive Decline L-METHYLFOLATE , METHYLCOBALAMIN , AND A / - ACETYLCYSTEINE IN THE TREATMENT OF ALZHEIMER'S DISEASE-RELATED COGNITIVE DECLINE. 2–5 (2015). doi:10.1017/S1092852900027589

2. Solomon, L. R. Disorders of cobalamin (vitamin B12) metabolism: emerging concepts in pathophysiology, diagnosis and treatment. *Blood Rev.* **21,** 113–30 (2007).

3. Smith, A. D. The worldwide challenge of the dementias: A role for B vitamins and homocysteine? in *Food and Nutrition Bulletin* **29,** (2008).

4. Schalinske, K. L. & Smazal, A. L. Homocysteine Imbalance: a Pathological Metabolic Marker. *Adv. Nutr. An Int. Rev. J.* **3,** 755–762 (2012).

5. Scott, J. M. Folate and vitamin B 12 . *Proc. Nutr. Soc.* **58,** 441–448 (1999).

6. Miller, A., Korem, M., Almog, R. & Galboiz, Y. Vitamin B12, demyelination, remyelination and repair in multiple sclerosis. in *Journal of the Neurological Sciences* **233,** 93–97 (2005).

7. Okada, K. *et al.* Methylcobalamin increases Erk1/2 and Akt activities through the methylation cycle and promotes nerve regeneration in a rat sciatic nerve injury model. *Exp. Neurol.* **222,** 191–203 (2010).

8. Munro, O. Q. *et al.* Molecular mechanics study of the ruffling of metalloporphyrins. *J. Am. Chem. Soc.* **114,** 7218–7230 (1992).

9. Hay, B. P. & Finke, R. G. Thermolysis of the Co-C Bond in Adenosylcorrins. III. Quantification of the Axial Base Effect in Adenosylcobalamin by the Synthesis and Thermolysis of axial base-free adenosylcobinamide. Insights into the energetics of Enzyme-assisted cobalt-carbon bond h. *J. Am. Chem. Soc.* **109,** 8012–8018 (1987).

10. Jarrett, J. T. *et al.* Mutations in the B12-binding region of methionine synthase: how the protein controls methylcobalamin reactivity. *Biochemistry* **35,** 2464–2475 (1996).

11. Grate, J. H. & Schrauzer, G. N. Chemistry of cobalamins and related compounds. 48. Sterically induced, spontaneous dealkylation of secondary alkylcobalamins due to axial base coordination and conformational changes of the corrin ligand. *J. Am. Chem. Soc.* **101,** 4601–4611 (1979).

12. Marzilli, Luigi G., Toscano, P. J., Randaccio, L., Bresciani-Pahor, N. & Calligaris, M. An unusually long cobalt-carbon bond. Molecular structure of trans-bis(dimethylglyoximato)(isopropyl)(pyridine)cobalt(III). Implications with regard to the conformational trigger mechanism of cobalt-carbon bond cleavage in coenzyme B12. *J. Am. Chem. Soc.* **101,** 6754–6756 (1979).

13. Ng, F. T. T. & Rempel, G. L. Ligand Effects on Transition-Metal-Alkyl Bond Dissociation Energies. *J. Am. Chem. Soc.* **104,** 621–623 (1982).

14. Dong, S., Padmakumar, R., Banerjee, R. & Spiro, T. G. Resonance raman co-C stretching frequencies reflect bond strength changes in alkyl cobalamins, but are unaffected by trans ligand substitution. *J. Am. Chem. Soc.* **118,** 9182–9183 (1996).

15. Puckett, J. M. J., Mitchell, M. B., Hirota, S. & Marzilli, L. G. Near-IR FT-Raman Spectroscopy of Methyl-B12 and Other Cobalamins and of Imidazole and Imidazolate Methylcobinamide Derivatives in Aqueous Solution. *Inorg. Chem.* **35,** 4656–4662 (1996).

16. Dong, S., Padmakumar, R., Maiti, N. & May, R. V. Resonance Raman Spectra Show That Coenzyme B 12 Binding to Methylmalonyl-Coenzyme A Mutase Changes the Corrin Ring Conformation but Leaves the Co - C Bond Essentially Unaffected. **3,** 9947–9948 (1998).

17. Hogenkamp, H. P. C. The Photolysis of Methylcobalamin. *Biochemistry* **5,** 417–422 (1966).

18. Schrauzer, G. N., Sibert, J. W. & J., W. R. Photochemical and thermal cobalt-carbon bond cleavage in alkylcobalamins and related organometallic compounds. Comparative study. *Org. Biol. Chem.* **590,** 6681–6688 (1968).

19. Taylor, R. T., Smucker, L., Hanna, L. M. & Gill, J. Aerobic Photolysis of Alkylcobalamins : Quantum Yields and light-action spectra. *Arch. Biochem. Biophys.* **156,** (1973).

20. Endicott, J. F. & Netzel, T. L. Early events and transient chemistry in the photolysis of alkylcobalamins. *J. Am. Chem. Soc.* **101,** 4000–4002 (1979).

21. Chen, E. & Chance, M. R. Continuous- Wave Quantum Yields of Various Cobalamins Are Influenced by Competition between Geminate Recombination and Cage Escape ? 1480–1487 (1993).

22. Grissom, C. B. & Chagovetz, A. M. MAGNETIC-FIELD EFFECTS IN MODEL B(12) ENZYMATIC-REACTIONS - THE PHOTOLYSIS OF METHYLCOB(III)ALAMIN. *Zeitschrift fur Phys. Chemie* **182,** 181–188 (1993).

23. Natarajan, E. & Grissom, C. B. The Origin of Magnetic Field dependent recombination in alkylcobalamin radical pairs. *Photochem. Photobiol.* **64,** 286–295 (1996).

24. Lott, W. B., Chagovetz, A. M. & Grissom, C. B. Alkyl Radical Geometry Controls Geminate Cage Recombination in Alkylcobalamins. *J. Am. Chem. Soc.* **117,** 12194–12201 (1995).

25. Walker, L. A. *et al.* Time-Resolved Spectroscopic Studies of B 12 Coenzymes : The Identification of a Metastable Cob (III) alamin Photoproduct in the Photolysis of Methylcobalamin. **7863,** 3597–3603 (1998).

26. Shiang, J. J., Ii, L. A. W., Anderson, N. A., Cole, A. G. & Sension, R. J. Time-Resolved Spectroscopic Studies of B 12 Coenzymes : The Photolysis of Methylcobalamin Is Wavelength Dependent. 10532–10539 (1999).

27. Sension, R. J., Harris, D. A. & Cole, A. G. Time-resolved spectroscopic studies of B 12 coenzymes: Comparison of the influence of solvent on the primary photolysis mechanism and geminate recombination of methyl-, ethyl-, n-propyl-, and 5???-deoxyadenosylcobalamin. *J. Phys. Chem. B* **109,** 21954–21962 (2005).

28. Harris, D. A., Stickrath, A. B., Carroll, E. C. & Sension, R. J. Influence of Environment on the Electronic Structure of Cob (III) alamins : Time-Resolved Absorption Studies of the S 1 State Spectrum and Dynamics. *J. Am. Chem. Soc.* **129,** 7578–7585 (2007).

29. Peng, J. *et al.* Ultrafast Excited-State Dynamics and Photolysis in Base-Off B 12 Coenzymes and Analogues : Absence of the trans-Nitrogenous Ligand Opens a Channel for Rapid Nonradiative Decay. 12398–12405 (2010).

30. Jones, A. R. *et al.* Ultrafast Infrared Spectral Fingerprints of Vitamin B 12 and Related Cobalamins. *J. Phys. Chem. A* **116,** 5586–5594 (2012).

31. Jaworska, M. & Lodowski, P. Electronic spectrum of Co-corrin calculated with the TDDFT method. *J. Mol. Struct. THEOCHEM* **631,** 209–223 (2003).

32. Kozlowski, P. M., Kuta, J. & Galezowski, W. Reductive Cleavage Mechanism of Methylcobalamin : Elementary Steps of Co - C Bond Breaking. 7638–7645 (2007).

33. Jaworska, M., Lodowski, P., Andrunio, T. & Kozlowski, P. M. Photolysis of Methylcobalamin : Identification of the Relevant Excited States Involved in Co - C Bond Scission. 2419–2422 (2007).

34. Lodowski, P., Jaworska, M., Andrunio, T., Kumar, M. & Louis, V. Photodissociation of Co - C Bond in Methyl- and Ethylcobalamin : An Insight from TD-DFT Calculations. 6898–6909 (2009).

35. Solheim, H., Kornobis, K., Ruud, K. & Kozlowski, P. M. Electronically Excited States of Vitamin B 12 and Methylcobalamin : Theoretical Analysis of Absorption , CD , and MCD Data. *J. Phys. Chem. B* **115,** 737–748 (2011).

36. Kornobis, K. *et al.* Electronic structure of the S1 state in methylcobalamin: insight from CASSCF/MC-XQDPT2, EOM-CCSD, and TD-DFT calculations. *J. Comput. Chem.* **34,** 987–1004 (2013).

37. Kornobis, K. *et al.* Electronically Excited States of Vitamin B 12 : Benchmark Calculations Including Time-Dependent Density Functional Theory and Correlated ab Initio Methods. 1280–1292 (2011).

38. Lodowski, P., Jaworska, M., Kornobis, K., Andruni, T. & Kozlowski, P. M. Electronic and Structural Properties of Low-lying Excited States of Vitamin B 12. 13304–13319 (2011).
39. Lodowski, P., Jaworska, M., Andruniów, T., Garabato, B. D. & Kozlowski, P. M. Mechanism of the S 1 excited state internal conversion in vitamin B 12. *Phys. Chem. Chem. Phys.* **16,** 18675 (2014).

40. Wiley, T. E., Arruda, B. C., Miller, N. A., Lenard, M. & Sension, R. J. Excited electronic states and internal conversion in cyanocobalamin. *Chinese Chem. Lett.* **26,** 439–443 (2015).

41. Stich, T. A., Buan, N. R. & Brunold, T. C. Spectroscopic and Computational Studies of Co 2 + Corrinoids : Spectral and Electronic Properties of the Biologically Relevant Base-On and Base-Off Forms of Co 2 + Cobalamin. 4820–4829 (2004).

42. Brooks, A. J., Vlasie, M., Banerjee, R. & Brunold, T. C. Spectroscopic and Computational Studies on the Adenosylcobalamin-Dependent Methylmalonyl-CoA Mutase : Evaluation of Enzymatic Contributions to Co - C Bond Activation in the Co 3 + Ground State. 8167–8180 (2004).

43. Conrad, K. S. & Brunold, T. C. Spectroscopic and Computational Studies of Glutathionylcobalamin : Nature of Co À S Bonding and Comparison to Co À C Bonding in. 8755–8766 (2011).

44. Brunold, T. C., Conrad, K. S., Liptak, M. D. & Park, K. Spectroscopically validated density functional theory studies of the B12 cofactors and their interactions with enzyme active sites. *Coord. Chem. Rev.* **253,** 779–794 (2009).

45. Park, K. & Brunold, T. C. Combined Spectroscopic and Computational Analysis of the Vibrational Properties of Vitamin B 12 in its Co 3 + , Co 2 + , and Co 1 + Oxidation States. (2013).

46. Lodowski, P., Jaworska, M., Andrunio, T., Garabato, B. D. & Kozlowski, P. M. Mechanism of Co − C Bond Photolysis in the Base-On Form of Methylcobalamin. (2014).

47. Sagi, I., Wirt, M. D., Chen, E., Frisbie, S. M. & Chance, M. R. Structure of an Intermediate of Coenzyme B12 Catalysis by EXAFS: Cobalt(II)B12. **112,** (1990).

48. Sagi, I. & Chance, M. R. Extent of Trans Effects in (Nonalky1)cobalamins: Steric Effects Control the Co-N Distance to 5,6-Dimet h ylbenzimidazole. *J. Am. Chem. Soc.* **114,** 8061–8066 (1992).

49. Wirt, M. D., Kumar, M., Ragdale, S. W. & Chance, M. R. X-ray Absorption Spectroscopy of the Corrinoid/Iron-Sulfur Protein Involved in Acetyl Coenzyme A Synthesis by Clostridium thermoaceticum. *J. Am. Chem. Soc.* **115,** 2146–2150 (1993).

50. Wirt, M. D. *et al.* Structural and Electronic Factors in Heterolytic Cleavage: Formation of the Co(I) Intermediate in the CorrinoidDron- Sulfur Protein from Clostridium thermoaceticum. *Biochemistry* **34,** 5269–5273 (1995).

51. Scheuring, E. M. *et al.* Time-Resolved X-ray Absorption Spectroscopy of Photoreduced Base-off Cob (II) alamin Compared to the Co (II) Species in Clostridium thermoaceticum. *J. Phys. Chem.* **100,** 3344–3348 (1996).

52. Champloy, F., Gruber, K., Jogl, G. & Kratky, C. XAS spectroscopy reveals X-ray-induced photoreduction of free and protein-bound B12 cofactors. *J. Synchrotron Radiat.* **7,** 267–73 (2000).

53. Van Stokkum, I. H. M., Larsen, D. S. & Van Grondelle, R. Global and target analysis of time-resolved spectra. *Biochimica et Biophysica Acta - Bioenergetics* **1657,** 82–104 (2004).

54. Ravel, B. & Newville, M. ATHENA, ARTEMIS, HEPHAESTUS: Data analysis for X-ray absorption spectroscopy using IFEFFIT. in *Journal of Synchrotron Radiation* **12,** 537–541 (2005).

55. Ankudinov, A. L., Ravel, B., Rehr, J. J. & Conradson, S. D. Real-space multiple-scattering calculation and interpretation of x-ray-absorption near-edge structure. **58,** 7565–7576 (1998).

56. Stöhr, J. *NEXAFS Spectroscopy*. *Book* **25,** (1992).

57. Firth, R., Hill, H. A. O. & Pratt, J. M. The Circular Dichroism and Absorption Spectra of Some Vitamin B12 Derivatives. *Biochemistry* **6,** 2178–2189 (1967).

58. Kozlowski, P. M., Garabato, B. D., Lodowski, P. & Jaworska, M. Photolytic properties of cobalamins: a theoretical perspective. *Dalton Trans.* **45,** 4457–70 (2016).

59. Giorgetti, M. *et al.* In situ X-ray absorption spectroelectrochemical study of hydroxocobalamin. *J. Biol. Inorg. Chem.* **5,** 156–66 (2000).

60. Meier, M. & Van Eldik, R. Ligand-substitution reactions of aquacobalamin (vitamin B12a) revisited. Conclusive evidence for the operation of a dissociative interchange mechanism. *Inorg. Chem.* **32,** 2635–2639 (1993).

61. Ahmad, I., Husaain, W. & Fareedi, A. L. Photolysis of cyanocobalamin. *J. Pharm. Biomed. Anal.* **10,** 9–15 (1992).

62. Wiley, T. E. *et al.* Photostability of Hydroxocobalamin: Ultrafast Excited State Dynamics and Computational Studies. *J. Phys. Chem. Lett.* **7,** 143–147 (2016).

63. Frisch, M. J. *et al.* Gaussian 09 Revision E.01. (2009).

64. Shiang, J. J. *et al.* Ultrafast excited-state dynamics in vitamin B12 and related cob(III)alamins. *J. Am. Chem. Soc.* **128,** 801–808 (2006).

65. Rury, A. S., Wiley, T. E. & Sension, R. J. Energy Cascades, Excited State Dynamics, and Photochemistry in Cob(III)alamins and Ferric Porphyrins. *Acc. Chem. Res.* **48,** 860–867 (2015).

CONCLUSIONS

The research presented in this dissertation was devoted to the study the use of electrons and X-rays in elucidating the biomolecular structure and dynamics. Among the multiple ways of studying biomolecular structure function relation, a couple of standout techniques have been crystallography and spectroscopy. Especially with the question of dynamics, there is a strong push towards the use of ultra-short and ultrafast probes. With proteins, the issue of radiation induced damage and outrunning the damage offers an additional reason to push towards the ultrafast probing methods.

In the use of electron crystallography (transmission electron diffraction), the crucial factor is to maintain the diffraction within the kinematic limit. It is identified that this poses an upper limit on the protein crystal thickness. Within the limitations of the simulations and a reasonable error in the experimental diffracted intensities, the maximum permissible thickness of protein crystal for transmission electron diffraction based crystallography measurements is about 100 nm. Additionally, the analysis on molecular replacement as a technique to provide phases to reconstruct the electron density for proteins shows that it is heavily biased on the phases from the model. This is noticed to the extent that it takes about 100% error in diffracted intensities before the hen-egg white lysozyme sample modeled on turkey-egg white lysozyme actually shows an error in the reconstructed electron density map.

A solution to work around this fundamental problem pertaining to crystallography is to adopt the single particle imaging technique. Being consistent with the theme of working with electrons as a probe, a new scheme of microscopy that is reciprocal to ADF-STEM termed (hollow-cone) HC-TEM is proposed as the alternative. Analogous to the X-ray field, the electron as a probe also introduces damage in the sample and the mechanisms and timescales of such processes (different from that of X-rays) are discussed and point to the inevitability of using an ultrafast and ultra-short electron source (femtosecond pulsed-electron beam) similar to the XFELs. It is assessed that under the current experimentally realized conditions of electron beam focusing parameters, a single particle imaging experiment would fail to produce even 1 electron

per image per shot. However, in the scenario that the probe can be converged to the size of the particle (few 10s to 100 nm), the electron count statistics would be good enough to establish the technique as a practically viable method.

Complementary to the crystallography and imaging technique is the energy-loss spectroscopy technique EELS. It is strongly related to the core and low loss spectroscopy, which are electronic transitions within a material. However, the development of the monochromated and aberration corrected NiON UltraSTEM enabled, for the first time, to measure energy-loss spectra at 10 meV resolution in energy and angstrom resolution in space. This then provides the possibility to measure phonon modes in a TEM at high spatial resolution. Using established Fourier techniques, the spectra can then be transformed to the time-domain to provide the dynamics of phonon excitation modes in materials that are not just global, but also localized in space. This provides, for the first time, an opportunity to study chemical and biochemical reactions (such as bond formation/dissociation/rearrangement) in the frequency domain using a continuous probe beam rather than using crystallography or other such techniques in the time domain. The example data of protein and water is analyzed using the Kramers-Kronig and compared with the Drude-Lorentz time bottom-up model to assign meaning to the spectral parameters. The proof of principle with the NiON UltraSTEM is provided using BaF_2 low-loss spectra, which is also analyzed using the Drude-Lorentz model to identify the time-dependent decay of the electronic excitation representing color centers.

The revolution in X-ray science is the advent of the hard X-ray free electron lasers that enable measuring the diffracted signals before the sample is damaged, thanks to the femtosecond duration and extreme flux of 10^{12} per pulse. To reap benefits of this invention however, a major upgrade of the means of delivering the sample to the X-ray beam would be necessary. This chapter briefly discusses the GDVN and LCP sample injector systems, their characteristics and a few striking scientific outcomes that exhibit the success and importance of this development.

Finally, similar to the case with the electron probe, the X-ray probe as well can be used to study the biomolecular structure and dynamics through the spectroscopic means. The case in question was to study the photolysis of Methylcobalamin, a cofactor of vitamin B_{12} that acts as a catalytic precursor to methyl-transfer reactions. For the first time, ps time resolved X-ray absorption data has been recorded at the APS. Using DFT and TD-DFT to model the reactants, products and intermediates and the FEFF code to simulate the XANES spectra for the same. The DFT+FEFF show that the pH2 MeCbl has the lower axial imidazole bond detached and either substituted by water or bridged by water. Additionally, it was also identified that the product has strong qualitative similarity to the XAS spectra of Aq/OH-Cbl for both pH2 and pH7 MeCbl. This indicates that over a period of laser exposure (with the upper axial methyl cleaved and the lower axial DBI weakened/detached), one or possibly two water molecules can ligate the axial sites to substitute the methyl and/or DBI. The time-resolved intermediate spectrum (200 ps) shows features of increased centrosymmetry with the pre-edge intensity substantially reduced, the absorption edge red shifted and the shoulder peak C more intense in comparison to the ground state spectrum. For this case though, the DFT models fail to predict the structure to provide a qualitative match to the experimental spectrum. Nevertheless, the models suggest a strong role played by the axial ligand nature and distance from the cobalt center. It is worth mentioning here that in other proteins bearing tetra-dentate ligands (such as prophyrins), the role of the axial ligands in the catalytic activity is pretty well established and studied and thus it would not be surprising to notice a similar instance with the corrin macrocycles as well. Thus, to achieve a completion of this project would imply a rigorous analysis of the impact of the nature and distance of the axial ligand in various combinations to shortlist and select the strongest contending models that would represent the intermediate spectrum. Such a structure would then be the first ever intermediate structure to be proposed with direct experimental evidence of the photolytic bond cleavage occurring in MeCbl.

FUTURE PERSPECTIVES

This dissertation elaborated on some examples on the study of structure and dynamics of biomolecules using different probes and different techniques. But, two recent technological advancements have catapulted the structural biology engine into a rocket launch mode. The first is the advent of the X-ray free electron laser (as described in chapter 4). With the ability to outrun radiation damage and obtain structures of proteins and biomolecules hitherto unknown, the XFEL has revolutionized this field. However, the current state of experiments reflects an important (if not the most important) threshold in successfully solving the structures in action – crystal quality. Although single particle imaging efforts have been running in parallel to provide a complementary route to structure identification without the need to crystallize these samples, they are yet to breakthrough a critical resolution barrier that would get the biological community interested in working this option. An alternative option that has yet to be taken advantage of is the use of X-ray spectroscopy. X-ray spectroscopy has the advantage of not needing to crystallize samples compromising nevertheless on the fact that the concentration of the central (excited) atom is nominally about 0.1 – 0001/atom. The way to counter this struggle is to use a combination of techniques – X-ray diffraction and spectroscopy. A fine example of demonstrating this idea is the work by Kern et al. (*Science* 340.6131 (2013): 491-495.). X-ray spectroscopy can yield high-resolution information on the local coordination and chemistry and this coupled with the global atom coordinates and structure that diffraction provides can be a very powerful methodology in studying the biomolecular dynamics. An added advantage in such a study is the ability to identify electronic damage (that occur at faster timescales and can be identified through spectroscopy) as independent from nuclear damage (that occur at relatively longer timescales and are reflected in diffraction). In this regard, one of the future aims is to take the MeCbl system to the XFELs to perform simultaneous X-ray diffraction and spectroscopy to track the intermediate states from the few fs timescales to the slower timescales that are accessible at the synchrotrons. In this way, XFELs provide the opportunity to study electronic transitions (that are not accessible with the time

resolution offered at synchrotrons) that occur at ultrashort times and hence provide completion to the understanding of the evolution of the excited state in the photolytic process.

The second technological advancement is in the cryo-EM community where with the advent of the new generation of detectors, single particle imaging using the electron microscope has enabled to provide structures of molecules at 3 Å or less. This method has rejuvenated the EM community to study single particle biomolecular structure at near atomic resolution thereby providing a stiff competition to the X-ray work. However, there are some important limitations in this technique, the most significant of them being the restriction in measuring dynamics. Pulsed electron probes have not yet been realized in imaging biological molecules as successfully as with the X-rays (reasons mentioned in chapter 2). Nevertheless, the ability to measure a high-resolution structure with a few thousand particles is a great development as compared to the case of XFELs where tens to hundreds of thousands of crystals need to be shot to obtain an equivalent structure clearly positions single particle cryo-EM imaging as a technique to sustain.

In fact the next decade will pose an interesting duel, one that will benefit the scientific community and the society at large irrespective of the outcome, between the electron probe and X-ray probe in being better abled in elucidating the relation between protein structure and function. With high energy and gravitational physics making substantial strides in the understanding and experimental verification of the big, will structural biology uncover truths of the small such that during our lifetime we could be participants and witnesses to a quantum leap in answering the fundamental question ‘What is life?’ …

COMPLETE LIST OF REFERENCES

INTRODUCTION

1. Nurse SP. The great ideas of biology. 2003;3(6).

2. Kessel A, Ben-Tal N. *Introduction to Proteins: Structure, Function, and Motion*.; 2012.

3. Fischer E. Einfluss der Configuration auf die Wirkung der Enzyme. *Ber Dtsch Chem Ges*. 1894;27:2985-2993.

4. Mirsky AE, Paulin L. On the structure of native, denatured, and coagulated proteins. In: *Proceedings of the National Academy of Sciences of the United States of America*. Vol 22. ; 1936:439-447.

5. Anfinsen CB. Principles that Govern the Folding of Protein Chains. *Science (80-)*. 1973;181(4096):223-230.

6. Drenth J, Mesters J. *Principles of Protein X-Ray Crystallography: Third Edition*.; 2007.

7. Spence JCH, Weierstall U, Chapman HN. X-ray lasers for structural and dynamic biology. *Rep Prog Phys*. 2012;75(10):102601.

8. Teng TY, Moffat K. Primary radiation damage of protein crystals by an intense synchrotron X-ray beam. *J Synchrotron Radiat*. 2000;7(Pt 5):313-317.
9. Weik M, Ravelli RBG, Kryger G, et al. Specific chemical and structural damage to proteins produced by synchrotron radiation. *Proc Natl Acad Sci*. 2000;97(2):623-628.

10. Tenboer J, Basu S, Zatsepin N, et al. Time-resolved serial crystallography captures high-resolution intermediates of photoactive yellow protein. *Science*. 2014;346(6214):1242-1246.

11. Zhang H, Unal H, Gati C, et al. Structure of the angiotensin receptor revealed by serial femtosecond crystallography. *Cell*. 2015;161(4):833-844.

12. Li X, Mooney P, Zheng S, et al. Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. *Nat Methods*. 2013;10(6):584-590.

13. Henderson R. The potential and limitations of neutrons, electrons and X-rays for atomic resolution microscopy of unstained biological molecules. *Q Rev Biophys*. 1995;28:171-193.

14. Glaeser RM, Taylor KA. Radiation damage relative to transmission electron microscopy of biological specimens at low temperature: a review. *J Microsc*. 1978;112(1):127-138.

15. Brenner S, Horne RW. A negative staining method for high resolution electron microscopy of viruses. *Biochim Biophys Acta*. 1959;34:103-110.

16. Kühlbrandt W. Cryo-EM enters a new era. *Elife*. 2014;3:e03678.

17. Callaway E. The revolution will not be crystallized: a new method sweeps through structural biology. *Nat News*. 2015;525(7568):172.

18. Cavalli A, Salvatella X, Dobson CM, Vendruscolo M. Protein structure determination from NMR chemical shifts. *Proc Natl Acad Sci U S A*. 2007;104:9615-9620.

19. Bax A. Two-dimensional NMR and protein structure. *AnnRevBiochem*. 1989;58:223-256.

20. Tugarinov V, Kanelis V, Kay LE. Isotope labeling strategies for the study of high-molecular-weight proteins by solution NMR spectroscopy. *Nat Protoc*. 2006;1(2):749-754.

21. Downing AK. *Protein NMR Techniques*. Vol 278.; 2004.

22. Corker J, Lefebvre F, Evans J, et al. Catalytic Cleavage of the C-H and C-C Bonds of Alkanes by Surface Organometallic Chemistry: An EXAFS and IR Characterization of a Zr-H Catalyst. *Science (80-)*. 1996;271(5251):966-969.

23. Penner-Hahn JE. X-ray Absorption Spectroscopy. *Compr Coord Chem II*. 2003:159-186.

24. Penner-Hahn JE. X-ray absorption spectroscopy in coordination chemistry. *Coord Chem Rev*. 1999;190-192:1101-1123.

25. Debeer S. X-ray absorption spectroscopy. *Methods Mol Biol (Clifton, NJ)*. 2011;766:165-176.

26. Bergmann U, Glatzel P. X-ray emission spectroscopy. *Photosynth Res*. 2009;102(2-3):255-266.

27. P. G, A. J, Glatzel P, Juhin A. X-ray Absorption and Emission Spectroscopy. *Local Struct Charact*. 2014:89-171.

CHAPTER 1

1. Henderson, R. *et al.* Model for the structure of bacteriorhodopsin based on high-resolution electron cryo-microscopy. *J. Mol. Biol.* **213,** 899–929 (1990).

2. Kühlbrandt, W., Wang, D. N. & Fujiyoshi, Y. Atomic model of plant light-harvesting complex by electron crystallography. *Nature* **367,** 614–621 (1994).

3. Shi, D., Hsiung, H. H., Pace, R. C. & Stokes, D. L. Preparation and analysis of large, flat crystals of Ca(2+)-ATPase for electron crystallography. *Biophys. J.* **68,** 1152–1162 (1995).

4. Taylor, K. A. & Varga, S. Similarity of Three-dimensional Microcrystals of from Pig Kidney and from Skeletal Muscle Sarcoplasmic Reticulum *. *J. Biol. Chem* **269,** 10107–10111 (1994).
5. Jap, B. K. High-resolution electron diffraction of reconstituted PhoE porin. *J. Mol. Biol.* **199,** 229–231 (1988).

6. Fujiyoshi, Y. *et al.* Structure and function of water channels. *Current Opinion in Structural Biology* **12,** 509–515 (2002).

7. Löwe, J., Li, H., Downing, K. . & Nogales, E. Refined structure of αβ-tubulin at 3.5 Å resolution. *J. Mol. Biol.* **313,** 1045–1057 (2001).

8. Chang, S., Head-Gordon, T., Glaeser, R. M. & Downing, K. H. Chemical bonding effects in the determination of protein structures by electron crystallography. *Acta Crystallogr A* **55 (Pt 2 ,** 305–313 (1999).

9. Glauber, R. & Schomaker, V. The Theory of Electron Diffraction. *Phys. Rev.* **89,** 667–671 (1953).

10. Son, J., Millward, G. R. & Thomas, J. M. The Role of Multiple Scattering in the Study of Lattice Images of Graphitie CarbOns. *Acta Crystallogr. Sect. A Cryst. Physics, Diffraction, Theor. Gen. Crystallogr.* **32,** (1976).
11. De Rosier, D. J. & Klug, a. Reconstruction of Three Dimensional Structures from Electron Micrographs. *Nature* **217,** 130–134 (1968).

12. Glaeser, R. M. & Downing, K. H. High-resolution electron crystallography of protein molecules. *Ultramicroscopy* **52,** 478–486 (1993).

13. Anstis, G. R., Lynch, D. F., Moodie, a. F. & O'Keefe, M. a. n -Beam lattice images. III. Upper limits of ionicity in W 4 Nb 26 O 77. *Acta Crystallogr. Sect. A* **29,** 138–147 (1973).

14. McKeown, J. T. & Spence, J. C. H. The kinematic convergent-beam electron diffraction method for nanocrystal structure determination. *J. Appl. Phys.* **106,** 074309 (2009).

15. Gjønnes, J. & Moodie, a. F. Extinction conditions in the dynamic theory of electron diffraction. *Acta Crystallogr.* **19,** 65–67 (1965).

16. Huang, W. J., Zuo, J. M., Jiang, B., Kwon, K. W. & Shim, M. Sub-ångström-resolution diffractive imaging of single nanocrystals. *Nat. Phys.* **5,** 129–133 (2009).

17. Nederlof, I., Li, Y. W., Van Heel, M. & Abrahams, J. P. Imaging protein three-dimensional nanocrystals with cryo-EM. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **69,** 852–859 (2013).

18. Glaeser, R. M. Review: electron crystallography: present excitement, a nod to the past, anticipating the future. *J. Struct. Biol.* **128,** 3–14 (1999).

19. Dorset, D. L. *Structural electron crystallography*. (Springer Science & Business Media, 2013).

20. Read, R. J. [7] Model phases: Probabilities and bias. *Methods Enzymol.* **227,** 110 (1997).

21. Lattman, E. & DeRosier, D. Why phase errors affect the electron function more than amplitude errors. *Acta Crystallogr. A.* **64,** 341–4 (2008).

22. Scapin, G. Molecular replacement then and now. in *Acta Crystallographica Section D: Biological Crystallography* **69,** 2266–2275 (2013).

23. Rossmann, M. G. The molecular replacement method. *Acta Crystallogr. Sect. A* **46,** 73–82 (1990).
24. Brünger, A. T. Extension of molecular replacement: a new search strategy based on Patterson correlation refinement. *Acta Crystallogr. Sect. A* **46,** 46–57 (1990).

25. Rossmann, M. G. Molecular replacement - Historical background. *Acta Crystallogr. - Sect. D Biol. Crystallogr.* **57,** 1360–1366 (2001).

26. Abergel, C. Molecular replacement: Tricks and treats. in *Acta Crystallographica Section D: Biological Crystallography* **69,** 2167–2173 (2013).

27. Evans, P. & McCoy, A. An introduction to molecular replacement. *Acta Crystallogr. D. Biol. Crystallogr.* **64,** 1–10 (2008).

28. Doyle, P. A. & Turner, P. S. Relativistic Hartree-Fock X-ray and electron scattering factors. *Acta Cryst (1968). A24, 390-397 [doi10.1107/S0567739468000756]* **3,** 1–8 (1968).

29. Zuo, J. M. & Spence, J. C. H. *Electron microdiffraction*. (Springer Science & Business Media, 2013).

30. Meyer, J. C. *et al.* Experimental analysis of charge redistribution due to chemical bonding by high-resolution transmission electron microscopy. *Nat. Mater.* **10,** 209–215 (2011).

31. Spence, J. C. H. *High-resolution electron microscopy*. (OUP Oxford, 2013).

32. Zuo, J. M. Measurements of electron densities in solids: a real-space view of electronic structure and bonding in inorganic crystals. *Reports Prog. Phys.* **67,** 2053–2103 (2004).

33. Cowley, J. M. *Diffraction Physics*. (Elsevier, 1995).

34. Vincent, R. & Midgley, P. A. Double conical beam-rocking system for measurement of integrated electron diffraction intensities. *Ultramicroscopy* **53,** 271–282 (1994).

35. Reimer, L. & Kohl, H. *Transmission Electron Microscopy*. *Book* (2008). doi:10.1007/978-0-387-34758-5

36. Boutet, S. *et al.* High-resolution protein structure determination by serial femtosecond crystallography. *Science (80-.).* **337,** 362–364 (2012).

37. Wilson, A. J. C. Largest likely values for the reliability index. *Acta Crystallogr.* **3,** 397–398 (1950).

38. Marks, L. D. & Plass, R. Atomic structure of Si (111)-(5x 2)-Au from high resolution electron microscopy and heavy-atom holography. *Phys. Rev. Lett.* **75,** 2172 (1995).

39. Own, C. S.-Y. System Design and Verification of the Precession Electron Diffraction Technique. (Northwestern University, 2005).

40. Kirian, R. A. *et al.* Femtosecond protein nanocrystallography—data analysis methods. *Opt. Express* **18,** 5713 (2010).

41. Dorset, D. L. Electron diffraction intensities from bent molecular organic crystals. *Acta Crystallogr. Sect. A Cryst. Physics, Diffraction, Theor. Gen. Crystallogr.* **36,** 592 – 600 (1980).

42. Nannenga, B. L., Shi, D., Leslie, A. G. W. & Gonen, T. High-resolution structure determination by continuous-rotation data collection in MicroED. *Nat. Methods* **11,** 927–30 (2014).

43. Liu, H. *et al.* Atomic structure of human adenovirus by cryo-EM reveals interactions among protein networks. *Science* **329,** 1038–43 (2010).

44. Kissick, D., Wanapun, D. & Simpson, G. Second-order nonlinear optical imaging of chiral crystals. *Annu. Rev. Anal. …* **4,** 419–437 (2011).

45. Abdallah, B. G., Kupitz, C., Fromme, P. & Ros, A. Crystallization of the large membrane protein complex photosystem i in a microfluidic channel. *ACS Nano* **7,** 10534–10543 (2013).

CHAPTER 2

1. Chapman, H. N. *et al.* Femtosecond diffractive imaging with a soft-X-ray free-electron laser. *Nat. Phys.* **2,** 839–843 (2006).

2. Spence, J. C. H., Weierstall, U. & Chapman, H. N. X-ray lasers for structural and dynamic biology. *Rep. Prog. Phys.* **75,** 102601 (2012).

3. Zhang, J. *et al.* Agonist-bound structure of the human P2Y12 receptor. *Nature* **509,** 119–22 (2014).

4. Tenboer, J. *et al.* Time-resolved serial crystallography captures high-resolution intermediates of photoactive yellow protein. *Science* **346,** 1242–6 (2014).

5. van der Schot, G. *et al.* Imaging single cells in a beam of live cyanobacteria with an X-ray laser. *Nat Commun* **6,** (2015).

6. Henderson, R. The potential and limitations of neutrons, electrons and X-rays for atomic resolution microscopy of unstained biological molecules. *Q. Rev. Biophys.* **28,** 171–193 (1995).

7. Spence, J. C. H. *High-resolution electron microscopy*. (OUP Oxford, 2013).

8. Rez, P. Comparison of phase contrast transmission electron microscopy with optimized scanning transmission annular dark field imaging for protein imaging. *Ultramicroscopy* **96,** 117–24 (2003).

9. Subramanian, G., Basu, S., Liu, H., Zuo, J. M. & Spence, J. C. H. Solving protein nanocrystals by cryo-EM diffraction: Multiple scattering artifacts. *Ultramicroscopy* **148,** 87–93 (2015).

10. Barty, A. *et al.* Self-terminating diffraction gates femtosecond X-ray nanocrystallography measurements. *Nat. Photonics* **6,** 35–40 (2012).
11. Muro'oka, Y. *et al.* Transmission-electron diffraction by MeV electron pulses. *Appl. Phys. Lett.* **98,** 251903 (2011).

12. Musumeci, P., Moody, J. T. & Scoby, C. M. Relativistic electron diffraction at the UCLA Pegasus photoinjector laboratory. *Ultramicroscopy* **108,** 1450–1453 (2008).

13. Huldt, G., Szoke, A. & Hajdu, J. Diffraction imaging of single particles and biomolecules. *J. Struct. Biol.* **144,** 219–227 (2003).

14. Seibert, M. M. *et al.* Single mimivirus particles intercepted and imaged with an X-ray laser. *Nature* **470,** 78–81 (2011).

15. Dashti, A. *et al.* Trajectories of the ribosome as a Brownian nanomachine. *Proc. Natl. Acad. Sci. U. S. A.* **111,** 17492–7 (2014).

16. Lahme, S., Kealhofer, C., Krausz, F. & Baum, P. Femtosecond single-electron diffraction. *Struct. Dyn.* **1,** (2014).

17. LaGrange, T. *et al.* Nanosecond time-resolved investigations using the in situ of dynamic transmission electron microscope (DTEM). *Ultramicroscopy* **108,** 1441–1449 (2008).

18. Pennycook, S. J. & Nellist, P. D. *Scanning Transmission Electron Microscopy: Imaging and Analysis*. (Springer Science & Business Media, 2011).

19. Saxton, W. O. TEM observations using bright field hollow cone illumination. *Optik (Stuttg).* **49,** 505–510 (1978).

20. Helmholtz, H. *Handbuch der physiologischen*. *Optik* **9,** (1867).

21. Cowley, J. M. Image Contrast in a Transmission Scanning Electron Microscope. *Appl. Phys. Lett.* **15,** 58 (1969).

22. Zeitler, E. & Thomson, M. G. R. Scanning transmission electron microscopy. *Optik (Stuttg).* **31,** 258 (1970).

23. Pogany, a. P. & Turner, P. S. Reciprocity in electron diffraction and microscopy. *Acta Crystallogr. Sect. A Cryst. Physics, Diffraction, Theor. Gen. Crystallogr.* **24,** 103–109 (1968).

24. Spence, J. C. H., Vecchione, T. & Weierstall, U. A coherent photofield electron source for fast diffractive and point-projection imaging. *Philos. Mag.* **90,** 4691–4702 (2010).

25. Hernandez Garcia, C. & Brau, C. A. Pulsed photoelectric field emission from needle cathodes. in *Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **483,** 273–276 (2002).

26. Musumeci, P. *et al.* Capturing ultrafast structural evolutions with a single pulse of MeV electrons: Radio frequency streak camera based electron diffraction. *J. Appl. Phys.* **108,** 1–6 (2010).

27. Bazarov, I. V., Dunham, B. M. & Sinclair, C. K. Maximum achievable beam brightness from photoinjectors. *Phys. Rev. Lett.* **102,** (2009).

28. Li, R. K. & Musumeci, P. Single-Shot MeV Transmission Electron Microscopy with Picosecond Temporal Resolution. *Phys. Rev. Appl.* **2,** 024003 (2014).

29. Chatelain, R. P. *et al.* Space-charge effects in ultrafast electron diffraction patterns from single crystals. *Ultramicroscopy* **116,** 86–94 (2012).

CHAPTER 3

1. Han, P. Y. *et al.* A direct comparison between terahertz time-domain spectroscopy and far-infrared Fourier transform spectroscopy. *Journal of Applied Physics* **89,** 2357–2359 (2001).

2. Zhang, W., Azad, A. K. & Grischkowsky, D. Terahertz studies of carrier dynamics and dielectric response of n-type, freestanding epitaxial GaN. *Appl. Phys. Lett.* **82,** 2841–2843 (2003).

3. George, P. A. *et al.* Ultrafast optical-pump terahertz-probe spectroscopy of the carrier relaxation and recombination dynamics in epitaxial graphene. *Nano Lett.* **8,** 4248–4251 (2008).
4. Ulbricht, R., Hendry, E., Shan, J., Heinz, T. F. & Bonn, M. Carrier dynamics in semiconductors studied with time-resolved terahertz spectroscopy. *Rev. Mod. Phys.* **83,** 543–586 (2011).

5. Fischer, B. M., Walther, M. & Uhd Jepsen, P. Far-infrared vibrational modes of DNA components studied by terahertz time-domain spectroscopy. *Phys. Med. Biol.* **47,** 3807–3814 (2002).

6. Castro-Camus, E. & Johnston, M. B. Conformational changes of photoactive yellow protein monitored by terahertz spectroscopy. *Chem. Phys. Lett.* **455,** 289–292 (2008).

7. Xu, J., Plaxco, K. W. & Allen, S. J. Collective dynamics of lysozyme in water: Terahertz absorption spectroscopy and comparison with theory. *J. Phys. Chem. B* **110,** 24255–24259 (2006).

8. Schmuttenmaer, C. A. Exploring dynamics in the far-infrared with terahertz spectroscopy. *Chem. Rev.* **104,** 1759–1779 (2004).

9. Ashkenov, N. *et al.* Infrared dielectric functions and phonon modes of high-quality ZnO films. *J. Appl. Phys.* **93,** 126–133 (2003).

10. Schubert, M., Tiwald, T. E. & Herzinger, C. M. Infrared dielectric anisotropy and phonon modes of sapphire. *J. Appl. Phys.* **61,** (2000).

11. Li, Z. Q. *et al.* Dirac charge dynamics in graphene by infrared spectroscopy. *Nat. Phys.* **4,** 6–9 (2008).

12. Zhang, S. *et al.* Experimental demonstration of near-infrared negative-index metamaterials. *Phys. Rev. Lett.* **95,** (2005).

13. Silaghi, S. D. *et al.* Dielectric functions of DNA base films from near-infrared to ultra-violet. in *Physica Status Solidi (B) Basic Research* **242,** 3047–3052 (2005).

14. Fecko, C. J., Eaves, J. D., Loparo, J. J., Tokmakoff, a & Geissler, P. L. Ultrafast hydrogen-bond dynamics in the infrared spectroscopy of water. *Science* **301,** 1698–1702 (2003).

15. Jones, A. R. *et al.* Ultrafast Infrared Spectral Fingerprints of Vitamin B 12 and Related Cobalamins. *J. Phys. Chem. A* **116,** 5586–5594 (2012).

16. Aspnes, D. E. & Studna, A. A. Dielectric functions and optical parameters of Si, Ge, GaP, GaAs, GaSb, InP, InAs, and InSb from 1.5 to 6.0 eV. *Phys. Rev. B* **27,** 985–1009 (1983).

17. Lauret, J.-S. *et al.* Ultrafast carrier dynamics in single-wall carbon nanotubes. *Phys. Rev. Lett.* **90,** 057404 (2003).

18. Lim, S. G. *et al.* Dielectric functions and optical bandgaps of high- K dielectrics for metal-oxide-semiconductor field-effect transistors by far ultraviolet spectroscopic ellipsometry. *J. Appl. Phys.* **91,** 4500–4505 (2002).

19. Nelson, F. J. *et al.* Electronic excitations in graphene in the 1-50 eV range: The π and π + σ peaks are not plasmons. *Nano Lett.* **14,** 3827–3831 (2014).

20. Johari, P. & Shenoy, V. B. Tunable dielectric properties of transition metal dichalcogenides. in *ACS Nano* **5,** 5903–5908 (2011).

21. Roth, F., Bauer, J., Mahns, B., Büchner, B. & Knupfer, M. Electronic structure of undoped and potassium-doped coronene investigated by electron energy-loss spectroscopy. *Phys. Rev. B - Condens. Matter Mater. Phys.* **85,** (2012).

22. Roth, F., Mahns, B., Büchner, B. & Knupfer, M. Dynamic response and electronic structure of potassium doped picene investigated by electron energy-loss spectroscopy. *arXiv* 13 (2011). doi:10.1103/PhysRevB.83.144501

23. Roth, F. & Knupfer, M. Electronic excitation spectrum of doped organic thin films investigated using electron energy-loss spectroscopy. *J. Electron Spectros. Relat. Phenomena* **204,** 23–28 (2015).

24. Mahns, B. *et al.* Electronic properties of spiro compounds for organic electronics. *J. Chem. Phys.* **136,** 124702 (2012).

25. Isaacson, M. Interaction of 25 keV Electrons with the Nucleic Acid Bases, Adenine, Thymine, and Uracil. II. Inner Shell Excitation and Inelastic Scattering Cross Sections. *J. Chem. Phys.* **56,** 1813 (1972).

26. Isaacson, M. Interaction of 25 keV Electrons with the Nucleic Acid Bases, Adenine, Thymine, and Uracil. I. Outer Shell Excitation*. *J. Chem. Phys.* **56,** 1803 (1972).

27. Hainfeld, J. & Isaacson, M. The use of electron energy loss spectroscopy for studying membrane architecture: A preliminary report. *Ultramicroscopy* **3,** 87–95 (1978).

28. Garvie, L. A. J., Rez, P., Alvarez, J. R. & Buseck, P. R. Interband transitions of crystalline and amorphous SiO2: An electron energy-loss spectroscopy (EELS) study of the low-loss region. *Solid State Commun.* **106,** 303–307 (1998).

29. Berger, S. D., McKenzie, D. R. & Martin, P. J. EELS analysis of vacuum arc-deposited diamond-like films. *Philos. Mag. Lett.* **57,** 285–290 (1988).

30. Krivanek, O. L., Lovejoy, T. C., Dellby, N. & Carpenter, R. W. Monochromated STEM with a 30 meV-wide, atom-sized electron probe. *J. Electron Microsc. (Tokyo).* **62,** 3–21 (2013).

31. Krivanek, O. L. *et al.* Vibrational spectroscopy in the electron microscope. *Nature* **514,** 209–12 (2014).

32. Crozier, P., Aoki, T. & Liu, Q. Detection of Water and Its Derivatives on Individual Nanoparticles using Vibrational Electron Energy-Loss Spectroscopy. *Ultramicroscopy* **169,** 30–36 (2016).

33. Rez, P. *et al.* Damage-free vibrational spectroscopy of biological materials in the electron microscope. *Nat. Commun.* **7,** 10945 (2016).

CHAPTER 4

1. Mccaddon, A., Hudson, P. R. & Mccaddon, B. A. Spectrums : Treatment of Alzheimer ' s Disease-Related Cognitive Decline L-METHYLFOLATE , METHYLCOBALAMIN , AND A / - ACETYLCYSTEINE IN THE TREATMENT OF ALZHEIMER'S DISEASE-RELATED COGNITIVE DECLINE. 2–5 (2015). doi:10.1017/S1092852900027589

2. Solomon, L. R. Disorders of cobalamin (vitamin B12) metabolism: emerging concepts in pathophysiology, diagnosis and treatment. *Blood Rev.* **21,** 113–30 (2007).

3. Smith, A. D. The worldwide challenge of the dementias: A role for B vitamins and homocysteine? in *Food and Nutrition Bulletin* **29,** (2008).

4. Schalinske, K. L. & Smazal, A. L. Homocysteine Imbalance: a Pathological Metabolic Marker. *Adv. Nutr. An Int. Rev. J.* **3,** 755–762 (2012).

5. Scott, J. M. Folate and vitamin B 12 . *Proc. Nutr. Soc.* **58,** 441–448 (1999).

6. Miller, A., Korem, M., Almog, R. & Galboiz, Y. Vitamin B12, demyelination, remyelination and repair in multiple sclerosis. in *Journal of the Neurological Sciences* **233,** 93–97 (2005).

7. Okada, K. *et al.* Methylcobalamin increases Erk1/2 and Akt activities through the methylation cycle and promotes nerve regeneration in a rat sciatic nerve injury model. *Exp. Neurol.* **222,** 191–203 (2010).

8. Munro, O. Q. *et al.* Molecular mechanics study of the ruffling of metalloporphyrins. *J. Am. Chem. Soc.* **114,** 7218–7230 (1992).

9. Hay, B. P. & Finke, R. G. Thermolysis of the Co-C Bond in Adenosylcorrins. III. Quantification of the Axial Base Effect in Adenosylcobalamin by the Synthesis and Thermolysis of axial base-free adenosylcobinamide. Insights into the energetics of Enzyme-assisted cobalt-carbon bond h. *J. Am. Chem. Soc.* **109,** 8012–8018 (1987).

10. Jarrett, J. T. *et al.* Mutations in the B12-binding region of methionine synthase: how the protein controls methylcobalamin reactivity. *Biochemistry* **35,** 2464–2475 (1996).

11. Grate, J. H. & Schrauzer, G. N. Chemistry of cobalamins and related compounds. 48. Sterically induced, spontaneous dealkylation of secondary alkylcobalamins due to axial base coordination and conformational changes of the corrin ligand. *J. Am. Chem. Soc.* **101,** 4601–4611 (1979).

12. Marzilli, Luigi G., Toscano, P. J., Randaccio, L., Bresciani-Pahor, N. & Calligaris, M. An unusually long cobalt-carbon bond. Molecular structure of trans-bis(dimethylglyoximato)(isopropyl)(pyridine)cobalt(III). Implications with regard to the conformational trigger mechanism of cobalt-carbon bond cleavage in coenzyme B12. *J. Am. Chem. Soc.* **101,** 6754–6756 (1979).

13. Ng, F. T. T. & Rempel, G. L. Ligand Effects on Transition-Metal-Alkyl Bond Dissociation Energies. *J. Am. Chem. Soc.* **104,** 621–623 (1982).

14. Dong, S., Padmakumar, R., Banerjee, R. & Spiro, T. G. Resonance raman co-C stretching frequencies reflect bond strength changes in alkyl cobalamins, but are unaffected by trans ligand substitution. *J. Am. Chem. Soc.* **118,** 9182–9183 (1996).

15. Puckett, J. M. J., Mitchell, M. B., Hirota, S. & Marzilli, L. G. Near-IR FT-Raman Spectroscopy of Methyl-B12 and Other Cobalamins and of Imidazole and Imidazolate Methylcobinamide Derivatives in Aqueous Solution. *Inorg. Chem.* **35,** 4656–4662 (1996).

16. Dong, S., Padmakumar, R., Maiti, N. & May, R. V. Resonance Raman Spectra Show That Coenzyme B 12 Binding to Methylmalonyl-Coenzyme A Mutase Changes the Corrin Ring Conformation but Leaves the Co - C Bond Essentially Unaffected. **3,** 9947–9948 (1998).

17. Hogenkamp, H. P. C. The Photolysis of Methylcobalamin. *Biochemistry* **5,** 417–422 (1966).

18. Schrauzer, G. N., Sibert, J. W. & J., W. R. Photochemical and thermal cobalt-carbon bond cleavage in alkylcobalamins and related organometallic compounds. Comparative study. *Org. Biol. Chem.* **590,** 6681–6688 (1968).

19. Taylor, R. T., Smucker, L., Hanna, L. M. & Gill, J. Aerobic Photolysis of Alkylcobalamins : Quantum Yields and light-action spectra. *Arch. Biochem. Biophys.* **156,** (1973).

20. Endicott, J. F. & Netzel, T. L. Early events and transient chemistry in the photolysis of alkylcobalamins. *J. Am. Chem. Soc.* **101,** 4000–4002 (1979).

21. Chen, E. & Chance, M. R. Continuous- Wave Quantum Yields of Various Cobalamins Are Influenced by Competition between Geminate Recombination and Cage Escape ? 1480–1487 (1993).

22. Grissom, C. B. & Chagovetz, A. M. MAGNETIC-FIELD EFFECTS IN MODEL B(12) ENZYMATIC-REACTIONS - THE PHOTOLYSIS OF METHYLCOB(III)ALAMIN. *Zeitschrift fur Phys. Chemie* **182,** 181–188 (1993).

23. Natarajan, E. & Grissom, C. B. The Origin of Magnetic Field dependent recombination in alkylcobalamin radical pairs. *Photochem. Photobiol.* **64,** 286–295 (1996).

24. Lott, W. B., Chagovetz, A. M. & Grissom, C. B. Alkyl Radical Geometry Controls Geminate Cage Recombination in Alkylcobalamins. *J. Am. Chem. Soc.* **117,** 12194–12201 (1995).

25. Walker, L. A. *et al.* Time-Resolved Spectroscopic Studies of B 12 Coenzymes : The Identification of a Metastable Cob (III) alamin Photoproduct in the Photolysis of Methylcobalamin. **7863,** 3597–3603 (1998).

26. Shiang, J. J., Ii, L. A. W., Anderson, N. A., Cole, A. G. & Sension, R. J. Time-Resolved Spectroscopic Studies of B 12 Coenzymes : The Photolysis of Methylcobalamin Is Wavelength Dependent. 10532–10539 (1999).

27. Sension, R. J., Harris, D. A. & Cole, A. G. Time-resolved spectroscopic studies of B 12 coenzymes: Comparison of the influence of solvent on the primary photolysis mechanism and geminate recombination of methyl-, ethyl-, n-propyl-, and 5???-deoxyadenosylcobalamin. *J. Phys. Chem. B* **109,** 21954–21962 (2005).

28. Harris, D. A., Stickrath, A. B., Carroll, E. C. & Sension, R. J. Influence of Environment on the Electronic Structure of Cob (III) alamins : Time-Resolved Absorption Studies of the S 1 State Spectrum and Dynamics. *J. Am. Chem. Soc.* **129,** 7578–7585 (2007).

29. Peng, J. *et al.* Ultrafast Excited-State Dynamics and Photolysis in Base-Off B 12 Coenzymes and Analogues : Absence of the trans-Nitrogenous Ligand Opens a Channel for Rapid Nonradiative Decay. 12398–12405 (2010).

30. Jones, A. R. *et al.* Ultrafast Infrared Spectral Fingerprints of Vitamin B 12 and Related Cobalamins. *J. Phys. Chem. A* **116,** 5586–5594 (2012).

31. Jaworska, M. & Lodowski, P. Electronic spectrum of Co-corrin calculated with the TDDFT method. *J. Mol. Struct. THEOCHEM* **631,** 209–223 (2003).

32. Kozlowski, P. M., Kuta, J. & Galezowski, W. Reductive Cleavage Mechanism of Methylcobalamin : Elementary Steps of Co - C Bond Breaking. 7638–7645 (2007).

33. Jaworska, M., Lodowski, P., Andrunio, T. & Kozlowski, P. M. Photolysis of Methylcobalamin : Identification of the Relevant Excited States Involved in Co - C Bond Scission. 2419–2422 (2007).

34. Lodowski, P., Jaworska, M., Andrunio, T., Kumar, M. & Louis, V. Photodissociation of Co - C Bond in Methyl- and Ethylcobalamin : An Insight from TD-DFT Calculations. 6898–6909 (2009).

35. Solheim, H., Kornobis, K., Ruud, K. & Kozlowski, P. M. Electronically Excited States of Vitamin B 12 and Methylcobalamin : Theoretical Analysis of Absorption , CD , and MCD Data. *J. Phys. Chem. B* **115,** 737–748 (2011).

36. Kornobis, K. *et al.* Electronic structure of the S1 state in methylcobalamin: insight from CASSCF/MC-XQDPT2, EOM-CCSD, and TD-DFT calculations. *J. Comput. Chem.* **34,** 987–1004 (2013).

37. Kornobis, K. *et al.* Electronically Excited States of Vitamin B 12 : Benchmark Calculations Including Time-Dependent Density Functional Theory and Correlated ab Initio Methods. 1280–1292 (2011).

38. Lodowski, P., Jaworska, M., Kornobis, K., Andruni, T. & Kozlowski, P. M. Electronic and Structural Properties of Low-lying Excited States of Vitamin B 12. 13304–13319 (2011).
39. Lodowski, P., Jaworska, M., Andruniów, T., Garabato, B. D. & Kozlowski, P. M. Mechanism of the S 1 excited state internal conversion in vitamin B 12. *Phys. Chem. Chem. Phys.* **16,** 18675 (2014).

40. Wiley, T. E., Arruda, B. C., Miller, N. A., Lenard, M. & Sension, R. J. Excited electronic states and internal conversion in cyanocobalamin. *Chinese Chem. Lett.* **26,** 439–443 (2015).

41. Stich, T. A., Buan, N. R. & Brunold, T. C. Spectroscopic and Computational Studies of Co 2 + Corrinoids : Spectral and Electronic Properties of the Biologically Relevant Base-On and Base-Off Forms of Co 2 + Cobalamin. 4820–4829 (2004).

42. Brooks, A. J., Vlasie, M., Banerjee, R. & Brunold, T. C. Spectroscopic and Computational Studies on the Adenosylcobalamin-Dependent Methylmalonyl-CoA Mutase : Evaluation of Enzymatic Contributions to Co - C Bond Activation in the Co 3 + Ground State. 8167–8180 (2004).

43. Conrad, K. S. & Brunold, T. C. Spectroscopic and Computational Studies of Glutathionylcobalamin : Nature of Co À S Bonding and Comparison to Co À C Bonding in. 8755–8766 (2011).

44. Brunold, T. C., Conrad, K. S., Liptak, M. D. & Park, K. Spectroscopically validated density functional theory studies of the B12 cofactors and their interactions with enzyme active sites. *Coord. Chem. Rev.* **253,** 779–794 (2009).

45. Park, K. & Brunold, T. C. Combined Spectroscopic and Computational Analysis of the Vibrational Properties of Vitamin B 12 in its Co 3 + , Co 2 + , and Co 1 + Oxidation States. (2013).

46. Lodowski, P., Jaworska, M., Andrunio, T., Garabato, B. D. & Kozlowski, P. M. Mechanism of Co − C Bond Photolysis in the Base-On Form of Methylcobalamin. (2014).

47. Sagi, I., Wirt, M. D., Chen, E., Frisbie, S. M. & Chance, M. R. Structure of an Intermediate of Coenzyme B12 Catalysis by EXAFS: Cobalt(II)B12. **112,** (1990).

48. Sagi, I. & Chance, M. R. Extent of Trans Effects in (Nonalky1)cobalamins: Steric Effects Control the Co-N Distance to 5,6-Dimet h ylbenzimidazole. *J. Am. Chem. Soc.* **114,** 8061–8066 (1992).

49. Wirt, M. D., Kumar, M., Ragdale, S. W. & Chance, M. R. X-ray Absorption Spectroscopy of the Corrinoid/Iron-Sulfur Protein Involved in Acetyl Coenzyme A Synthesis by Clostridium thermoaceticum. *J. Am. Chem. Soc.* **115,** 2146–2150 (1993).

50. Wirt, M. D. *et al.* Structural and Electronic Factors in Heterolytic Cleavage: Formation of the Co(I) Intermediate in the CorrinoidDron- Sulfur Protein from Clostridium thermoaceticum. *Biochemistry* **34,** 5269–5273 (1995).

51. Scheuring, E. M. *et al.* Time-Resolved X-ray Absorption Spectroscopy of Photoreduced Base-off Cob (II) alamin Compared to the Co (II) Species in Clostridium thermoaceticum. *J. Phys. Chem.* **100,** 3344–3348 (1996).

52. Champloy, F., Gruber, K., Jogl, G. & Kratky, C. XAS spectroscopy reveals X-ray-induced photoreduction of free and protein-bound B12 cofactors. *J. Synchrotron Radiat.* **7,** 267–73 (2000).

53. Van Stokkum, I. H. M., Larsen, D. S. & Van Grondelle, R. Global and target analysis of time-resolved spectra. *Biochimica et Biophysica Acta - Bioenergetics* **1657,** 82–104 (2004).

54. Ravel, B. & Newville, M. ATHENA, ARTEMIS, HEPHAESTUS: Data analysis for X-ray absorption spectroscopy using IFEFFIT. in *Journal of Synchrotron Radiation* **12,** 537–541 (2005).

55. Ankudinov, A. L., Ravel, B., Rehr, J. J. & Conradson, S. D. Real-space multiple-scattering calculation and interpretation of x-ray-absorption near-edge structure. **58,** 7565–7576 (1998).

56. Stöhr, J. *NEXAFS Spectroscopy*. *Book* **25,** (1992).

57. Firth, R., Hill, H. A. O. & Pratt, J. M. The Circular Dichroism and Absorption Spectra of Some Vitamin B12 Derivatives. *Biochemistry* **6,** 2178–2189 (1967).

58. Kozlowski, P. M., Garabato, B. D., Lodowski, P. & Jaworska, M. Photolytic properties of cobalamins: a theoretical perspective. *Dalton Trans.* **45,** 4457–70 (2016).

59. Giorgetti, M. *et al.* In situ X-ray absorption spectroelectrochemical study of hydroxocobalamin. *J. Biol. Inorg. Chem.* **5,** 156–66 (2000).

60. Meier, M. & Van Eldik, R. Ligand-substitution reactions of aquacobalamin (vitamin B12a) revisited. Conclusive evidence for the operation of a dissociative interchange mechanism. *Inorg. Chem.* **32,** 2635–2639 (1993).

61. Ahmad, I., Husaain, W. & Fareedi, A. L. Photolysis of cyanocobalamin. *J. Pharm. Biomed. Anal.* **10,** 9–15 (1992).

62. Wiley, T. E. *et al.* Photostability of Hydroxocobalamin: Ultrafast Excited State Dynamics and Computational Studies. *J. Phys. Chem. Lett.* **7,** 143–147 (2016).

63. Frisch, M. J. *et al.* Gaussian 09 Revision E.01. (2009).

64. Shiang, J. J. *et al.* Ultrafast excited-state dynamics in vitamin B12 and related cob(III)alamins. *J. Am. Chem. Soc.* **128,** 801–808 (2006).

65. Rury, A. S., Wiley, T. E. & Sension, R. J. Energy Cascades, Excited State Dynamics, and Photochemistry in Cob(III)alamins and Ferric Porphyrins. *Acc. Chem. Res.* **48,** 860–867 (2015).