Characterization of Metalloproteins and Biomaterials by X-ray Absorption Spectroscopy and X-ray Diffraction

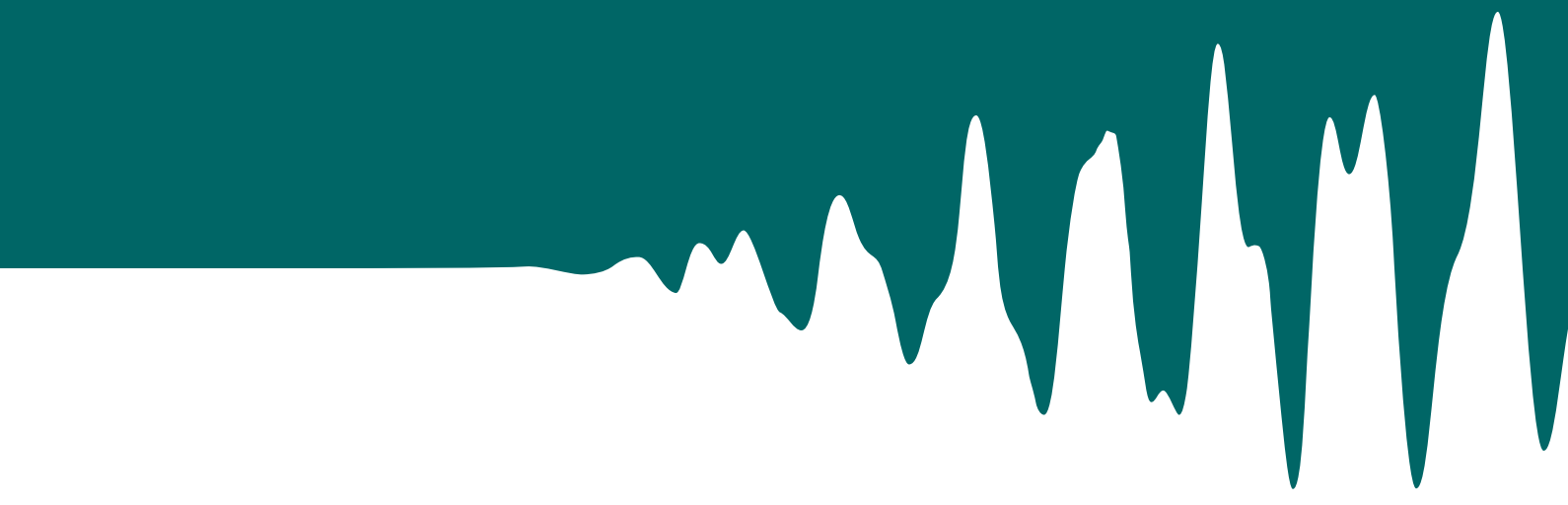
Frankær, Christian Grundahl; Ståhl, Kenny; Harris, Pernille

Publication date:
2012

Document Version
Publisher's PDF, also known as Version of record

Characterization of Metalloproteins and Biomaterials by X-ray Absorption Spectroscopy and X-ray Diffraction

Christian Grundahl Frankær



PhD Thesis
Department of Chemistry
Technical University of Denmark
2012

Characterization of Metalloproteins and Biomaterials by X-ray Absorption Spectroscopy and X-ray Diffraction

PhD Thesis

Christian Grundahl Frankær

Physical Chemistry Group

Department of Chemistry

Technical University of Denmark

March 2012

Preface and acknowledgements

The present dissertation is submitted to the Technical University of Denmark and partially fulfils the requirements for the PhD degree. The work has been carried out at the Department of Chemistry in the Physical Chemistry Group with Associate Professor Kenny Ståhl as main supervisor and Associate Professor Pernille Harris as co-supervisor. The PhD scholarship was funded by the Technical University of Denmark.

First and foremost, I would like to thank my supervisors Kenny Ståhl and Pernille Harris for their help and guidance throughout the project and for always being available for questions and inspiring scientific discussions. I shall also thank them for their company and encouragement, particularly during the long days and nights without sleep at our synchrotron trips, struggling at the beam line.

I shall thank Astrid Schøneberg and Lise Lotte Berring for practical assistance, when it was needed the most. Special thanks to Astrid Schøneberg for assistance, persistence and patience with searching and mounting useful crystals of strontium ranelate among tiny fragments. I will miss the pleasant atmosphere in the X-ray Group, which is now a part of the Physical Chemistry Group, at Department of Chemistry and particularly our Friday morning breakfasts.

I owe a special thank to Marianne Vad Knudsen for optimizing the crystal growth of different insulin conformations during her BSc project. I would also like to thank Associate Professor Irene Shim for helping me with *ab initio* Hartree-Fock calculations on copper clusters. I thank Professor Annette Henriksen at Carlsberg Laboratories for testing powder diffraction from proteins on an in-house micro-source X-ray diffractometer.

I highly acknowledge employees at MAX-lab for helpful support and special thank to Katarina Norén for introducing me to the data treatment and analysis of X-ray absorption spectroscopy data and Thomas Ursby for estimation of radiation damage. I also acknowledge Danscatt for funding of my synchrotron trips. I would like to thank Elena Nazarenko

from University of Gothenburg for introducing me to quantitative interpretation of XANES spectra, which was a completely new topic to me. The acquaintance has resulted in a fruitful collaboration. I had the opportunity to attend the HERCULES course in Grenoble, which provided me with unsurpassed knowledge and new acquaintances for which I am very grateful.

From the Analytical Chemistry Group I would like to thank Associate Professor Jens E. T. Andersen for placing the bone samples for my disposal and PhD Anders C. Raffalt for performing the quantitative analyses and for a constructive collaboration. I also acknowledge Stephan Christgau and Osteorologix A/S for providing the bone samples.

I would like to thank Monika N. Løvgreen for cheerful company and fruitful scientific discussions during almost three years in the same office. Finally, I would like to thank my friends and family for their support and I dedicate my special thanks to my loving wife Sarah Maria and my daughter Elisabeth for their patience and love.

Kgs. Lyngby, March 17th 2012

Christian Grundahl Frankær

Abstract

This thesis presents the work on combining complementary X-rays techniques for studying the structures of proteins and other biomaterials, and consists of three different projects: (*i*) Characterization of protein powders with X-ray powder diffraction (XRPD). (*ii*) The combination of X-ray crystallography and X-ray absorption spectroscopy (XAS) applied to studying different hexameric insulin conformations. (*iii*) The structures of polymorphs of strontium ranelate and the distribution of strontium in bone tissue.

A procedure for fast identification and verification of protein powders using XRPD was developed and tested on micro-crystals of lysozyme and insulin. Different protein crystal forms were identified by comparing experimental powder diffraction patterns with patterns calculated from PDB coordinates. The key factor to bring the calculated patterns in agreement with the observed patterns was correction for disordered bulk-solvent, but also correction for background and optimization of unit cell parameters have to be taken into account. A sample holder was designed for collecting powder diffraction data on a standard laboratory X-ray powder diffractometer. The background was reduced by use of pinholes integrated in the sample holder design, and crystal forms of well diffracting protein powders were safely identified from only 7 μl of sample. The sample holder was furthermore adaptable for XAS experiments.

The crystal structures of three conformations of hexameric bovine zinc insulin (T_6, T_3R_3 and R_6) were solved by single crystal X-ray diffraction (XRD) to 1.40 Å, 1.30 Å and 1.80 Å resolution, respectively. The zinc coordination in each conformation was studied by XAS including both extended X-ray absorption fine structure (EXAFS) spectroscopy and X-ray absorption near edge structure (XANES) spectroscopy. High quality XAS spectra were collected at the Zn K-edge on insulin micro-crystals with EXAFS signal up to 14 $Å^{-1}$ in k-space. Regarding the hexamer as a dimer of trimers where each trimer contains a Zn^{2+} ion, zinc adopted tetrahedral coordination in all R_3-trimers, and octahedral coordination

in T_3-trimers. The coordination distances to zinc were refined from EXAFS with standard deviations below 0.01 Å for the first coordination sphere. In contrast to the distances determined from the medium resolution crystal structures, the XAS results were in good agreement with distances found in small molecules with similar coordination geometries, as well as in other high resolution insulin structures. As the radiation doses for the XRD experiments were significantly higher compared to that of XAS experiments, the single crystals were exposed to a higher degree of radiation damage affecting the zinc coordination in the T_3-sites, in particular. Furthermore, XANES spectra for the zinc sites in T_6 and R_6 insulin were successfully calculated using finite difference methods, and the bond distances and angles were optimized from a quantitative XANES analysis.

T_6 insulin was furthermore crystallized with iron, nickel and copper, and their XRD crystal structures were solved to 1.90 Å, 1.50 Å and 1.45 Å resolution, respectively. As the affinity to iron is low, iron insulin crystals were grown in presence of small amounts of zinc. The two metal sites in the XRD structure thus contained respectively one Fe^{2+} and one Zn^{2+} ion, with respectively tetrahedral and octahedral coordination geometry. The metal sites in nickel and copper insulin were studied by XAS. Coordination distances were refined from EXAFS showing a very regular octahedral coordination of Ni^{2+}, which was further verified by calculated XANES spectra. The coordination geometry of copper was, however, ambiguously determined in the XRD structure of copper insulin. The XANES spectrum revealed the presence of copper in oxidation state +I, which was caused by photoreduction of Cu^{2+} during the XAS experiments. Photoreduction was not observed for nickel and zinc insulin. EXAFS analysis showed a tetrahedral coordination of copper, which was further supported by *ab initio* Hartree-Fock calculations, as well as calculated XANES spectra.

Strontium ranelate, a pharmaceutical used in the treatment of osteoporosis, was synthesized and crystallized and the crystal structure of the nonahydrate was solved. The crystals were highly sensitive to humidity and temperature and easily deteriorate upon dehydration. Four other hydrates were observed by *in situ* XRPD upon dehydration. These were identified as heptahydrate, pentahydrate, trihydrate and dihydrate by estimation of the water content by thermogravimetric analysis.

Bone tissue from dogs treated with strontium malonate was studied using XAS. A new approach for analysing the X-ray absorption spectra resulted in a compositional model, from which the relative distribution of strontium in the different bone composites was estimated. Approximately 35–45 % of the strontium present was incorporated into calcium hydroxyapatite by substitution of some of the calcium ions, and at least 30 % was located at sites with a high structural disorder similar to Sr^{2+} in solution. The remaining strontium was absorbed in collagen. The relative distribution of strontium among the different composites seemed to be independent of treatment period and dose level.

Dansk resumé

Nærværende ph.d.-afhandling omfatter arbejdet med at kombinere komplementære røntgenmetoder til at undersøge strukturen af proteiner og biomaterialer og indeholder tre forskellige projekter: (*i*) Karakterisering af proteinpulvre med røntgenpulverdiffraktion (XRPD). (*ii*) Kombinationen af røntgenkrystallografi og røntgenabsorptionsspektroskopi (XAS) anvendt på et studium af forskellige konformationer af insulinhexamere. (*iii*) Strukturerne af strontiumranelat polymorfer samt fordelingen af strontium i knoglevæv.

En procedure blev udviklet til hurtigt at kunne identificere og verificere proteinpulvre med XRPD, og blev afprøvet på mikrokrystaller af lysozym og insulin. Forskellige proteinkrystalformer blev identificeret ved at sammenligne eksperimentelt opnåede pulverdiffraktogrammer med diffraktogrammer beregnet ud fra PDB-koordinater. Korrektion for uordnet solvent blev identificeret som den vigtigste faktor til at opnå god overensstemmelse mellem beregnede og eksperimentelle diffraktogrammer, men også baggrundskorrektion samt optimering af enhedcelleparametre anses for vigtige. En prøveholder blev designet til opsamling af pulverdiffraktionsdata på standard pulverdiffraktometre. Pinholes integreret i prøveholderdesignet bevirkede reduktion af baggrunden, og krystalformer af veldiffrakterende proteinpulvre blev identificeret fra blot 7 µl prøve. Prøveholderen var designet til også at kunne anvendes til XAS-målinger.

Krystalstrukturerne af tre konformationer af bovin insulinhexamere med zink (T_6, T_3R_3 og R_6) blev løst med røntgendiffraktion (XRD) fra énkrystaller til opløsninger på hhv. 1.40 Å, 1.30 Å og 1.80 Å. Koordinationen af zink i de enkelte konformationer blev undersøgt med XAS, hvilket både inkluderede EXAFS (extended X-ray absorption fine structure) og XANES (X-ray absorption near edge structure). XAS-spektre blev målt på Zn K-kanten på insulinmikrokrystaller med EXAFS-signal ud til $k = 14$ Å^{-1}. Ved at betragte hexameren som en dimer af trimere, hvor hver trimer indeholder en Zn^{2+}-ion, antog zink tetraedrisk koordinationsgeometri i alle R_3-trimere, og oktaedrisk koordination i T_3-trimerene. Koor-

dinationsafstande til zink blev forfinet vha. EXAFS med standardafvigelser under 0.01 Å for den inderste koordinationssfære. I modsætning til de zinkafstande, der forekommer i krystalstrukturerne til middel opløsning, stemte XAS resultaterne bedre overens med zinkafstande observeret blandt småmolekylstrukturer med lignende koordinationsgeometrier, såvel som blandt andre insulinstrukturer til høj opløsning. Eftersom strålingsdoserne for XRD-forsøgene var signifikant højere i forhold til doserne for XAS-forsøgene, var énkrystallerne udsatte for større strålingsskade, hvilket påvirkede særligt koordinationen af zink i T_3-trimerene. Endvidere blev XANES-spektre af zinkcentrene i T_6 og R_6 insulin beregnet vha. FDM (finite difference methods), og bindingsafstande samt -vinkler blev optimeret vha. en kvantitativ analyse af XANES-spektrene.

T_6 insulin blev endvidere krystalliseret med jern, nikkel og kobber, og deres krystalstrukturer blev løst til opløsninger på hhv. 1.90 Å, 1.50 Å and 1.45 Å. Eftersom affiniteten overfor jern er lav, blev jerninsulin krystalliseret under tilstedeværelse af små mængder af zink. Dermed optog Fe^{2+} og Zn^{2+} hhv. hver af de to metalcentre i krystalstrukturen, og med hhv. tetraedrisk og oktaedrisk koordinationsgeometri. Metalcentrene in nikkel- og kobberinsulin blev undersøgt med XAS. Koordinationsafstande blev forfinet vha. EXAFS, som viste en meget regulær oktaedrisk koordination af Ni^{2+}, hvilket yderligere blev bekræftet vha. beregnede XANES-spektre. Kobbers koordinationsgeometri var flertydigt bestemt i krystalstrukturen af kobberinsulin. XANES-spektret påviste tilstedeværelsen af kobber i oxidationstrin +I, hvilket var forårsaget af en fotoreduktion af Cu^{2+} i løbet af XAS-målingerne. Fotoreduktion blev ikke observeret for nikkel- og zinkinsulin. En efterfølgende EXAFS-analyse påviste tetraedrisk koordination af kobber, hvilket yderligere blev bekræftet vha. *ab initio* Hartree-Fock beregninger samt beregnede XANES-spektre.

Strontium ranelat, som er et lægemiddel brugt til behandling af knogleskørhed, blev syntetiseret og krystalliseret, hvorefter krystalstrukturen af nonahydratet blev løst. Krystallerne var yderst følsomme overfor luftfugtighed og temperatur, og gik let i stykker under dehydrering. Fire andre hydrater blev observeret under dehydreringen vha. *in situ* XRPD. Disse hydrater blev identificeret som heptahydrat, pentahydrat, trihydrat samt dihydrat, ved at estimere vandindholdet ud fra en termogravimetrisk analyse.

Knoglevæv fra hunde behandlet med strontiummalonat blev undersøgt med XAS. En ny måde at gribe dataanalysen af absorptionsspektrene an på, resulterede i en model, hvor fra den relative fordeling af strontium i de forskellige knoglebestanddele kunne estimeres. Cirka 35–45 % af det tilstedeværende strontium var inkorporeret i calciumhydroxyapatit under udskiftning af nogle af calciumionerne med strontium, og mindst 30 % optog uordnede sites som strukturelt ligner dem for Sr^{2+}-ioner i vandig opløsning. Den resterende mængde strontium var absorberet i kollagen. Den relative fordeling af strontium i de forskellige knoglebestanddele forekom at være uafhængig af behandlingstid samt dosis.

Publications and conference contributions

Publications

Paper I In-house characterization of protein powder. Christian Grundahl Hartmann, Ole Faurskov Nielsen, Kenny Ståhl & Pernille Harris. *J. Appl. Cryst.* (2010). **43,** 876−882. (Appendix A.1)

Paper II In-house characterization of protein powder. Christian Grundahl Hartmann, Pernille Harris & Kenny Ståhl. *Z. Kristallogr. Proc.* (2011). **1,** 163−168. (Appendix A.2)

Paper III A sample holder for in-house X-ray powder diffraction studies of protein powders. Christian Grundahl Frankær, Pernille Harris & Kenny Ståhl. *J. Appl. Cryst.* (2011). **44,** 1288−1290. (Appendix A.3)

Paper IV Polymeric strontium ranelate nonahydrate. Kenny Ståhl, Christian Grundahl Frankær, Anders Christer Raffalt, Søren Roed Sørensen & Jens Enevold Taulov Andersen. *Acta Cryst.* (2011). E**67,** m471−m472. (Appendix A.4)

Paper V The structures of T_6, T_3R_3 and R_6 bovine insulin: Combining X-ray diffraction and absorption spectroscopy. Christian Grundahl Frankær, Marianne Vad Knudsen, Katarina Norén, Elena Nazarenko, Kenny Ståhl and Pernille Harris. *Manuscript has been prepared based on Chapter 7 and submitted to Acta Cryst. D.*

Paper VI On the polymorphs on strontium ranelate. Kenny Ståhl, Christian Grundahl Frankær, Anders Christer Raffalt & Jens Enevold Taulov Andersen. *Manuscript has been prepared partly based on Chapter 10.*

Paper VII Studies of strontium localisation in bones by X-ray absorption spectroscopy. Christian Grundahl Frankær, Anders Christer Raffalt, Jens Enevold Thaulov Andersen & Kenny Ståhl. *Manuscript has been prepared based on Chapter 11.*

Conference contributions

I Locating strontium ions in bone tissue using EXAFS. Christian Grundahl Hartmann, Jens Enevold Thaulov Andersen & Kenny Ståhl. Poster at Danish Crystallographers Meeting (DKM39), Technical University of Denmark, 2009.

II The structure and phase transitions of strontium ranelate Kenny Ståhl, Christian G. Hartmann, Carsten B. Prag, Søren R. Sørensen, Anders C. Raffalt, Stephen Christgau & Jens E. T. Andersen. Poster at Danish Crystallographers Meeting (DKM39), Technical University of Denmark, 2009.

III The crystal structures of polymorphs of strontium ranelate. Kenny Ståhl, Christian G. Hartmann & Jens E. T. Andersen. Poster at European Crystallography Meeting, (ECM25), Istanbul, Turkey, 2009.

IV Exploring alternative methods for characterization of proteins: Testing XRPD and EXAFS on hexameric insulin. Christian G. Hartmann, Pernille Harris, Katarina Norén & Kenny Ståhl. Poster at HERCULES 20th anniversary conference, Grenoble, France 2010.

V Protein characterization by in-house X-ray powder diffraction. Christian G. Hartmann, Kenny Ståhl & Pernille Harris. Oral presentation at Danish Crystallographers Meeting (DKM40), University of Copenhagen, 2010

VI In-house X-ray characterization of protein powder Christian G. Hartmann, Kenny Ståhl & Pernille Harris. Oral presentation and poster at European Powder Diffraction Conference (EPDIC12), Darmstadt, Germany, 2010.

VII In-house characterization of proteins using X-ray powder diffraction. Christian G. Frankær, Pernille Harris & Kenny Ståhl. Poster at DTU.protein 4th work shop, Technical University of Denmark, 2010.

VIII In-house characterization of protein powder. Christian Grundahl Frankær. Poster at DTU Chemistry Graduate School Symposium I, Schäffergården, Gentofte, 2010.

IX Bovine insulin conformations studied by single crystal X-ray diffraction, X-ray powder diffraction and X-ray absorption spectroscopy. Christian Grundahl Frankær, Kenny Ståhl & Pernille Harris. Oral presentation at Danish Crystallographers Meeting (DKM41), Aarhus University, 2011

X The complementarity of X-ray techniques for structure determination of proteins: Combining single crystal X-ray diffraction, X-ray powder diffraction and X-ray absorption spectroscopy. Christian Grundahl Frankær, Kenny Ståhl and Pernille Harris. Poster at DTU.protein 5th work shop, Technical University of Denmark, 2011.

XI Conformational studies of bovine insulin. Christian Grundahl Frankær, Kenny Ståhl & Pernille Harris. Poster at the Congress and General Assembly of the International Union of Crystallography (IUCrXXII), Madrid, Spain, 2011.

XII Conformational studies of bovine insulin. Christian Grundahl Frankær, Kenny Ståhl & Pernille Harris. Poster at CoLuAa XX Crystallographers Meeting, University of Copenhagen, 2011.

XIII Bovine insulin conformations studied by complementary X-ray techniques Christian Grundahl Frankær. Oral presentation at DTU Chemistry Graduate School Symposium II, Rungstedgaard, Rungsted Kyst, 2011.

List of abbreviations

AFM	Atomic Force Microscopy
ASU	Asymmetric Unit
CaHA	Calcium Hydroxyapatite
CCD	Charge Coupled Device
CD	Circular Dichroism
CSD	Cambridge Structure Database
CuNiR	Copper Nitrite Reductase
DLS	Dynamic Light Scattering
EM	Electron Microscopy
EPR	Electron Paramagnetic Resonance
ESRF	European Synchrotron Radiation Facility (Grenoble, France)
EXAFS	Extended X-ray Absorption Fine Structure
FDM	Finite Difference Methods
FT	Fourier Transform
FWHM	Full Width at Half Maximum
HA	Hydroxyapatite
HF	Hartree-Fock
ICP-MS	Inductively Coupled Plasma Mass Spectrometry
IR	Infrared light
MAX-lab	National electron accelerator laboratory for synchrotron radiation research, nuclear physics and accelerator physics (Lund University, Sweden)
MOF	Metal Organic Framework
MS	Multiple Scattering
MT	Muffin-Tin
NEXAFS	Near Edge X-ray Absorption Fine Structure
NMR	Nuclear Magnetic Resonance
PDB	Protein Data Bank
PIPS	Passivated Implanted Planar Silicon
PSII	Photosystem II

R	Relaxed insulin conformation
RMS	Root Mean Square
RT	Room Temperature
SANS	Small Angle Neutron Scattering
SAXS	Small Angle X-ray Scattering
SEM	Scanning Electron Microscopy
SLS	Static Light Scattering
SOD	Superoxide Dismutase
SOLEIL	French national synchrotron facility, (Saint-Aubin, France)
SrHA	Strontium Hydroxyapatite
SrR	Strontium Ranelate
SS	Single Scattering
STM	Scanning Tunnelling Microscopy
T	Tensed insulin conformation
TGA	Thermogravimetric Analysis
THF	Tetrahydrofuran
TLS	Translation Libration Screw-motion
T2DNiR	Type II Depleted Nitrate Reductase
UV	Ultraviolet light
UV-vis	Ultraviolet-visible light
WHO	World Health Organization
XAFS	X-ray Absorption Fine Structure
XAS	X-ray Absorption Spectroscopy
XANES	X-ray Absorption Near Edge Structure
XRD	X-ray Diffraction
XRPD	X-ray Powder Diffraction
XRF	X-ray Fluorescence

Software and programs used are written in italic throughout the thesis.

The displacement factors refined by EXAFS are given as σ^2 or as $2\sigma^2$ depending on the program used. The reported displacement factor type used in each case is always specified in the tables.

Furthermore, "strontium", "strontium ions" "Sr" and "Sr^{2+}" are used interchangeably. It is implied that strontium always exists in oxidation state +II in aqueous solutions and in biological systems. The same applies to zinc, copper, nickel and iron, which exist in oxidation state +II in hexameric insulin, unless another oxidation state is stated.

Contents

Chapter 1

Introduction

This chapter gives an overview of different techniques used to study the structure of proteins and biological matter, with emphasis on X-ray diffraction and X-ray absorption spectroscopy, which are the main techniques used in this work. A brief introduction to the above mentioned X-ray techniques are given, focussing on the combination and synergy between them, which is further described through selected examples of applications. Finally, an outline of the dissertation is presented.

1.1 Techniques used to study the structure of biological matter

Structures of proteins and biological soft materials are studied using many different techniques. The more commonly used are summarized in Figure 1.1. The diversity among the techniques is high and so are the types of structural information which can be obtained from them. The appropriateness of each technique depends on the system, and many techniques highly complement each other. One of the best examples of combining the strengths of the different methods is the structure determination of the ribosome, which awarded Ada E. Yonath, Venkatraman Ramakrishnan and Thomas A. Steinz the Nobel Prize in Chemistry in 2009. Using a combination of high resolution methods to determine the three-dimensional atomic structure of the subunits and low resolution techniques to locate these subunits relative to each other resulted in a complete structure. A review on the solution of the ribosome structure is presented by Moore (1998).

The three-dimensional structure of macromolecules is best determined by diffraction techniques and nuclear magnetic resonance (NMR). A look in the protein data bank (PDB)

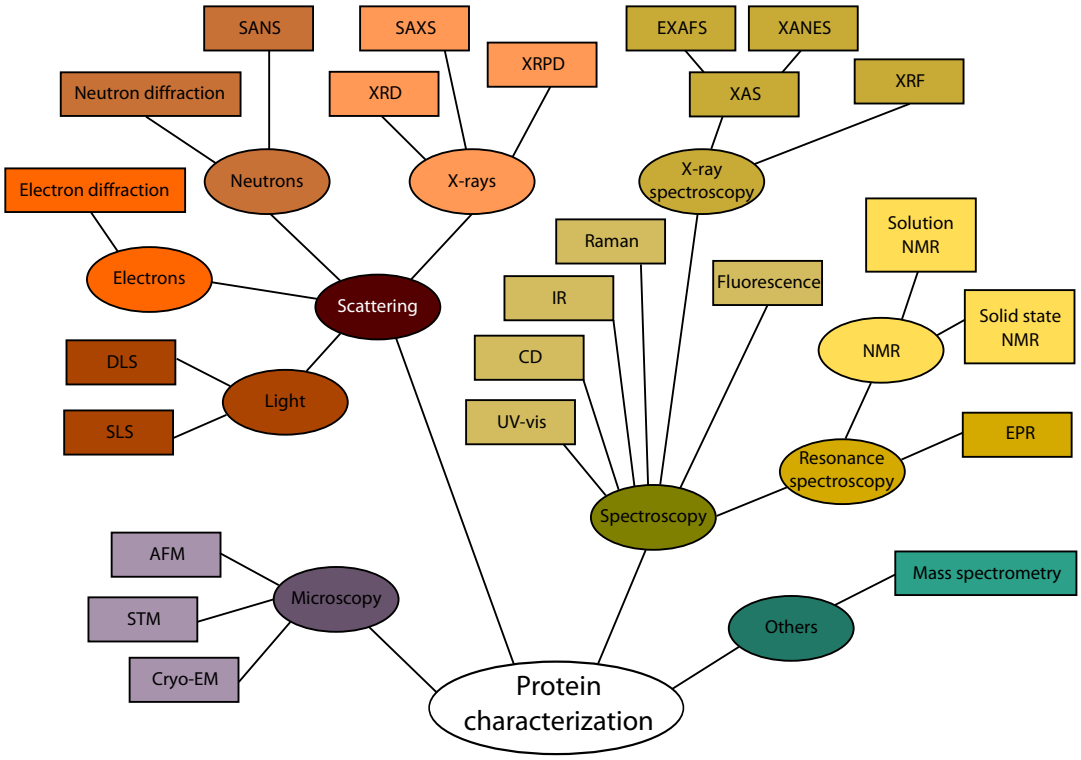


Figure 1.1: Overview of the more commonly used techniques to provide structural information of proteins and biomaterials.

reveals that 87.6 % of about 80,000 deposited structures are determined from X-ray diffraction (XRD), whereas NMR counts for 11.7 %. The importance of XRD and NMR structures is primarily due to the high level of resolution provided by these techniques. XRD and NMR are somewhat complementary, as NMR gives a picture of the protein in solution, which better resembles its natural biological environment, in contrast to the ordered array of molecules found in protein crystals. Another advantage of NMR is that no crystals are needed. However, NMR is in practice only applicable for small proteins (< 30 kDa).

Similar resolution levels are obtained from neutron diffraction, which is complementary to X-rays as the scattering properties of neutrons and X-rays are different. Neutron experiments are however time-consuming and expensive to perform and require very large crystals, wherefore only about 50 neutron diffraction structures are found in PDB.

Lower resolution methods count small angle X-ray scattering (SAXS) and small angle neutron scattering (SANS) in which the overall shape and size of the macromolecule is determined. These techniques can be performed on dilute protein solutions and are strong tools for studying assemblies and fibrils in combination with XRD structures of the separate domains. Electron diffraction provides structural information of typically $20 - 3$ Å resolution from two-dimensional crystals, which may be the only obtainable crys-

tals for membrane proteins. X-ray powder diffraction (XRPD) from micro crystals can be obtained to resolutions of about 3 Å (von Dreele, 2005) and the structures can be refined from Rietveld refinements, see Chapter 4 for further details. Microscopy covering cryo-electron microscopy (cryo-EM), scanning tunnelling microscopy (STM) and atomic force microscopy (AFM) is well-suited for studying larger proteins (≥ 100 kDa) and quarternary structures of protein complexes, as the resolution typically is above 5 Å.

Spectroscopic methods generally provide information about the electronic structure. Infrared (IR) and Raman spectroscopy probe the vibrations of the chemical bond and are thereby used to fingerprint specific bond types. In proteins, α-helices and β-strands can be distinguished and the distribution of secondary structure elements can be estimated from these methods. Similar information are obtained from circular dichroism (CD), which is used to classify the protein folding. As conjugated double bonds absorb electromagnetic radiation in the ultraviolet (UV) and visible regions of the electromagnetic spectrum, UV-vis spectroscopy provides information about aromatic residues (Tyr, Phe and Trp), and is widely used for determination of protein concentration. Furthermore UV-vis is used to obtain electronic information from metalloproteins as many *d-d* electron transitions have energies in the visible region. Fluorescence spectroscopy is mostly used for imaging and mapping. Distributions of proteins in cells and tissue is mapped by attachment of fluorophores to the peptides and the distribution of specific elements can be mapped with X-ray fluorescence (XRF).

Structural information to very high resolution may be obtained by X-ray absorption spectroscopy (XAS), which includes both extended X-ray absorption fine structure (EXAFS) and X-ray absorption near edge structure (XANES). XAS is element specific and probes the three-dimensional structure within a radius of 5–6 Å around the specific element, which makes the method well-suited for studying the local structure around the metal centres in metalloproteins. Electron paramagnetic resonance (EPR) is another method providing information on metal sites, but only for paramagnetic ions, whereas XAS is applicable for any metal.

1.2 Combination of XRD and XAS for metalloprotein studies

Metals play a central role in biological processes and over 30 % of the proteins known are metalloproteins. To fully understand their role and the biological processes they are involved in, it is important to know the three-dimensional structure, particularly in the metal sites in which the reactions take place. Many of these reactions involves redox chemistry where structural changes are very small (approx. 0.1 Å). Therefore it is important to obtain high resolution structural information to study these subtle differences (Strange *et al.*, 2005; Strange & Feiters, 2008).

X-ray diffraction and X-ray absorption spectroscopy belong to the techniques from which three-dimensional structure can be determined at highest resolution. The entire three-dimensional structure is obtained from XRD. However, only 2 % of the XRD structures in PDB are solved to resolution better than 1.2 Å, as crystals with high diffraction power and high stability towards radiation damage and photoreduction during the data collection are required (Ascone *et al.*, 2005).

In contrast to XRD, XAS is applicable to metalloproteins only, and the structural information is limited to a sphere of radius 5–6 Å around the absorbing metal atom. Hereby the coordination of the metal (i.e. number and type of the neighbouring atoms) and the atomic distances are determined with accuracies and precision comparable to those achieved from small-molecule crystallography. XAS also provides information about the electronic structure of the absorbing metal (i.e. oxidation state and spin state), and thereby offers a method to characterize spectroscopically silent metals (Penner-Hahn, 2005). Another advantage of XAS is that data can be collected on both aqueous and crystalline samples.

A major drawback of the method is that almost any model can be fitted to XAS data. Therefore *a priori* knowledge about the system is highly required. A realistic starting model may be provided from crystallography. Analysis of XAS data must be carefully performed to avoid over determinacy in the refined models. Only parameters which significantly contribute to the XAS signal must be included in the refinements.

Combination of medium resolution XRD and XAS may improve a structure to atomic resolution accuracy in the metal centres as demonstrated by Hasnain & Hodgson (1999) and Hasnain & Strange (2003). Likewise, XAS can be combined with NMR to determine the metal coordination, which is not obtained from NMR (Banci *et al.*, 2005). The output from the XAS analysis can be directly used as geometrical restraints in the final refinement of the entire protein structure from low and medium resolution XRD data. Simultaneous collection of both XAS and XRD data provides an excellent opportunity to implement this strategy, and is available at more beam lines, e.g. Proxima 1 at the SOLEIL synchrotron radiation facility in France (Ascone *et al.*, 2007).

1.3 Examples of applications of XRD and XAS

The majority of the XAS done on biological systems involves the $3d$ transition metals, since they often constitute or are part of the active centre in metalloproteins, but metal centres which have structural properties, as calcium, have also been studied. Since the development of the modern XAS theory in the 1970'es, XAS has been found well suited for studying the metal centres in metalloproteins (Sayers *et al.*, 1975). A small selection of metal centres studied by XAS is shown in Figure 1.2.

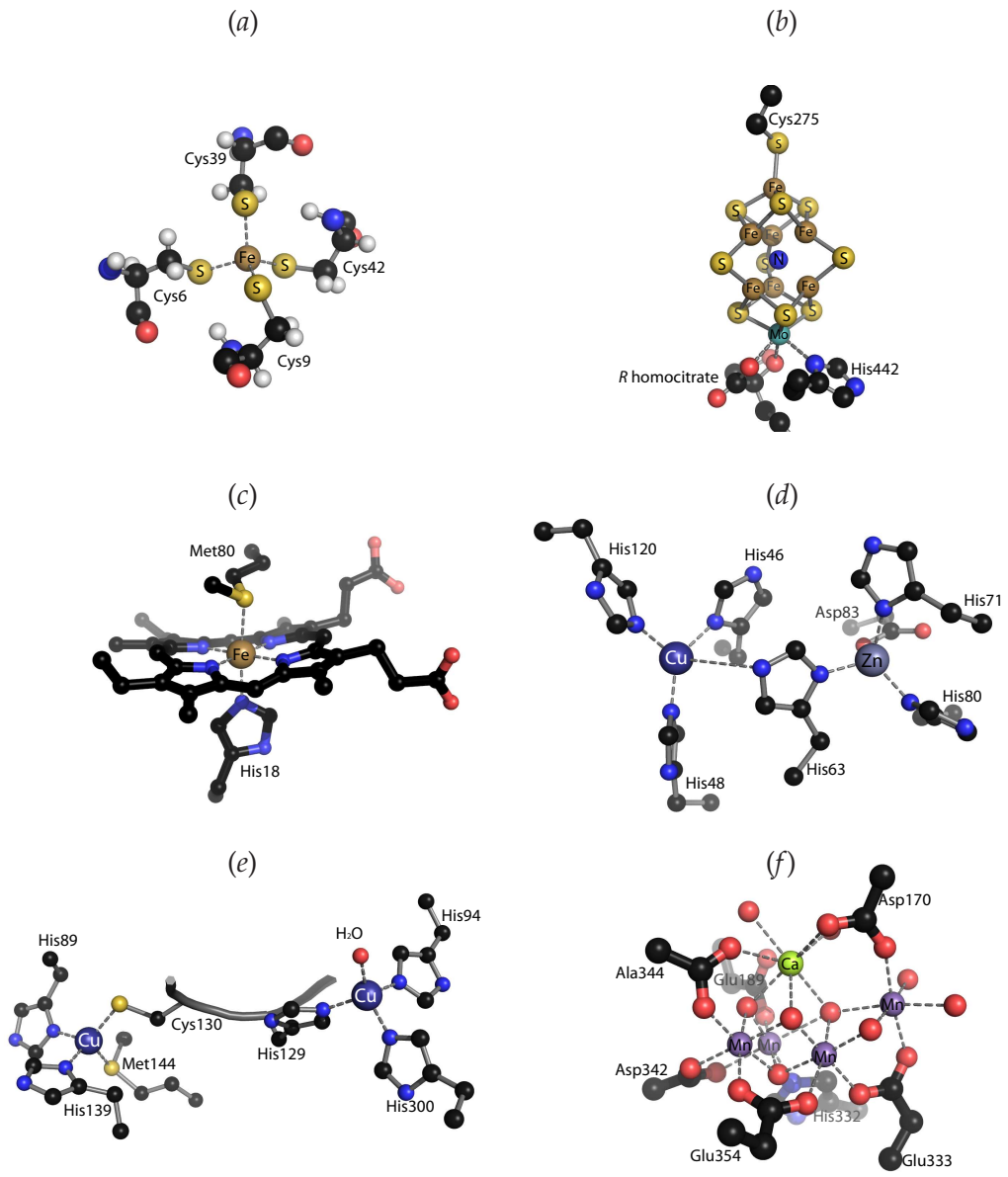


Figure 1.2: Selected examples of metal centres in metalloproteins studied with X-ray absorption spectroscopy. The figures are based on the PDB-coordinates stated. (*a*) The FeS_4-cluster in rubredoxin, PDB-entry 5RXN (Watenpaugh *et al.*, 1980). (*b*) The Fe-Mo cofactor $MoFe_7S_9X$, in nitrogenase, PDB-entry 2MIN (Einsle *et al.*, 2002). (*c*) The haem-protein cytochrome *c* (PDB-entry 1HRC) (Bushnell *et al.*, 1990). (*d*) Binuclear Cu–Zn superoxide dismutase, PDB-entry 2C9V (Strange *et al.*, 2006). (*e*) Cu-nitrite reductase, PDB-entry 1OE1 (Ellis *et al.*, 2003). (*f*) Mn_4Ca cluster in the oxygen-evolving photosystem II, PDB-entry 3ARC (Umena *et al.*, 2011). Carbon atoms are shown in black, nitrogen in blue, oxygen in red, hydrogen in white, sulfur in yellow, calcium in light green, manganese in purple, iron in brown, copper in dark blue, zinc in light blue and molybdenum in turquoise.

Rubredoxin was the first reported metalloprotein studied by EXAFS. Shulman *et al.* (1975) studied the iron-sulfur cluster in rubredoxin, Figure 1.2(*a*), and showed that the four Fe–S distances were equal in length in contrast to the XRD-structure available at that time. Other more complex iron-sulfur clusters have been studied with XAS, for instance the FeMo-cofactor ($MoFe_7S_9X$, where X is either C, N or O) in nitrogenase (Chen *et al.*, 1993; Corbett *et al.*, 2005; George *et al.*, 2008), see Figure 1.2(*b*). These XAS studies on the Fe K-edge and the Mo K and L-edges highly complement the crystal structures, and have provided supplementary electronic information about these clusters. Also haem-proteins e.g. cytochrome *c*, Figure 1.2(*c*) have been studied with XAS (Cheng *et al.*, 1999). The XAS spectra of these porphyrin systems are highly influenced by multiple scattering (MS), see Section 2.4.2. MS can be used to obtain information of not only the bond distances, but also the angles between the ligands coordinated to the central iron atom. This is illustrated in a recent study of different heam-complexes using both EXAFS and XANES data (D´Angelo *et al.*, 2008).

Superoxide dismutase (SOD) is a binuclear metalloprotein containing copper and zinc, Figure 1.2(*d*). Oxidized and the reduced SOD have been studied by XAS performed on both the Cu and Zn K-edges (Blackburn *et al.*, 1984; Murphy *et al.*, 1997), revealing a change in coordination geometry around the copper ion upon reduction. This was not observed in the first crystal structure of reduced SOD (Rypniewski *et al.*, 1995), which illustrates the difficulty in obtaining a particular oxidation state during XRD experiments. The XAS studies thereby provides complementary information about oxidation state and detailed coordination geometry.

Copper nitrite reductase (CuNiR) is another binuclear metalloprotein containing two copper sites, respectively type I and type II copper, Figure 1.2(*e*). In general, proteins containing more clusters with the same element, but with different coordination hereof, highly increases the complexity of the XAS experiments. As XAS probes the average structure around the absorbing metal, the XAS spectrum of CuNiR is an average of copper site I and site II. To separate the different copper sites, Strange *et al.* (1999) recorded the EXAFS spectrum of a type II depleted nitrite reductase (T2DNiR) containing less than 10 % type II copper, whereafter the type I copper EXAFS spectrum was achieved from the difference EXAFS between T2DNiR and CuNiR.

Following a reaction while it occurs is the ultimate goal to understand the function of a protein and the biological system it is involved in. As the reactions may take place within microseconds *in situ* XAS are challenging. Different methods to study biological reactions by XAS are summarized by Haumann *et al.* (2005) and Westenhoff *et al.* (2010). Freeze-quenching, in which the reaction is quenched at different stages, has been used to slow down the reaction or trap intermediates, for instance in an XAS study of the ferritin reaction between Fe^{2+} and O_2 (Hwang *et al.*, 2000). Rapid-scan XAS is another approach,

in which the time resolution on the spectrometer is increased (Richwin *et al.*, 2001). This is required for the pump-and-probe experiments, where a reaction is initiated by photolysis, as in a recent study of different intermediates of the Mn_4Ca-cluster in the light sensitive complex photosystem II (PSII), Figure 1.2(*f*) (Haumann *et al.*, 2005; Yano *et al.*, 2006; Yano & Yachandra, 2008).

1.4 Dissertation outline

The first part of this dissertation consists of two chapters introducing the theory and experimental principles behind XAS. Chapter 2 provides an introduction to the theory of X-ray absorption spectroscopy, and defines the central terms used in XAS. The chapter covers absorption of X-rays and theory of EXAFS. Chapter 3 summarizes the most important issues when designing XAS experiments and provides a brief introduction to the experimental setup and acquisition of XAS data. The chapter is a supplement to the experimental descriptions of XAS data collection procedures in the following chapters. The data treatment and analysis procedures are described and different strategies used to analyse the XAS data in this project are presented. Afterwards the results from the experimental work carried out during this PhD study are presented. The work covers three areas:

I: Characterization of protein powder by XRPD

Chapter 4 summarizes the results from an XRPD study of protein powders. The focus is on calculation of powder diffraction patterns where the contribution from the disordered solvent are taken into account. The patterns are useful for identification and verification of different crystal forms. The entire study is published in *J. Appl. Cryst.*, (Paper I) and *Z. Kristallogr. Proc.* (Paper II), which are included in Appendices A.1 and A.2, respectively. Chapter 5 describes the experimental procedure by which protein XRPD-patterns can be collected on in-house equipment at reasonable time scales. For this purpose different sample holders have been developed, which are described. A descriptive note of the final sample holder design, which is also adaptable for XAS experiments, has been published in *J. Appl. Cryst.*, (Paper III) included in Appendix A.3.

II: XRD and XAS studies on hexameric insulin conformations

Chapter 6 contains an introduction to the structure of insulin with focus on the hexamer conformations and their metal binding sites. Chapter 7 describes a study of three different conformations (T_6, T_3R_3 and R_6) of hexameric zinc insulin from bovine pancreas. All three structures are solved by single crystal XRD and the coordination of zinc in each of the conformers is studied by XAS, which include analysis of both EXAFS and XANES data. Based

on this work, which illustrate the complementarity between XRD and XAS, a manuscript has been prepared and submitted to *Acta Cryst. D*. Chapter 8 describes another XRD and XAS study of T_6 insulin hexamers containing iron, nickel and copper. Focus is particularly on the metal coordination. A paper is currently in preparation.

III: Strontium for osteoporosis treatment and XAS studies on bones

Chapter 9 contains an introduction to bone physiology, osteoporosis and briefly summarizes the mechanisms of renewal of bone tissue, which are involved in the uptake of strontium in the bones. An overview of other XAS studies of bone and calcified tissue is furthermore presented. Chapter 10 presents the structure of polymorphs of strontium ranelate, which is a medicament against osteoporosis. The content of water bound in the structure is very dependent on temperature and relative humidity. The crystal structure of fully hydrated strontium ranelate is published in *Acta Cryst. E*, (Paper IV). A reprint is included Appendix A.4. The other polymorphs, which occur upon dehydration, are studied by XRPD and TGA. A manuscript has been prepared and will be submitted to *Powder Diffraction*. Chapter 11 describes an XAS-study of the localisation of strontium incorporated in bone tissue on femur (thigh bones) and calvariae (skullcaps) from dogs treated with strontium malonate. A linear combination of XAS spectra from reference samples are fitted to the XAS spectra allowing estimation of the relative distribution of strontium in each composite. Based on this work a manuscript has been prepared and will be submitted to *J. Bone Miner. Res.*

Final conclusions about the techniques used as well as general conclusions are given in Chapter 12. Furthermore, suggestions for future work are outlined.

Four appendices are included: Appendix A contains reprints of the scientific papers published during this PhD study. Appendix B describe the structures of two insulin hexamers crystallized with monovalent cations (rubidium and caesium). Appendix C describes the EXAFS analysis on selected strontium salts used for reference compounds for the XAS study of bone from dogs treated with strontium. Appendix D contains some initial results from an EXAFS pre-study of strontium containing bone samples performed at room temperature.

Throughout the thesis numerous figures are presenting structures and results. All protein structures are created using *PyMOL* (DeLano, 2006), inorganic structures using *ATOMS* (Shape Software, 2006) and all data curves are plotted using *MATLAB* (The Mathworks, 2008).

Chapter 2

Introduction to X-ray absorption spectroscopy

Since the discovery of X-rays in 1895 by W. C. Röntgen (Röntgen, 1896), X-rays have been used for obtaining information about the structure of matter. X-rays are high energy electromagnetic radiation with energies ranging from approximately 0.1 to 100 keV, with corresponding wavelengths between 100 and 0.1 Å. There are two fundamental physical phenomena when the oscillating electric field of the X-rays interfere with matter: Photoelectric absorption and scattering (inelastic and elastic). An introduction to X-ray absorption spectroscopy (XAS) is given in this chapter.

2.1 Historical background

X-ray absorption fine structure (XAFS) observed at energies above an absorption edge was first discovered by Fricke (1920) and Hertz (1920). An early attempt to interpret the edge region were done by Kossel (1920) and the first theories were suggested by Kronig (1931, 1932), but it took 40 years before the phenomenon was correctly explained. The realization that the fine structure could be used to obtain structural information led to the derivation of the modern theory by Lytle, Stern and Sayers, (Sayers *et al.*, 1971) in the beginning of the 1970'es. As synchrotron radiation sources have developed during the last 40 years, the XAS theory has been further improved resulting in the curved-wave multiple scattering theory, (Lee & Pendry, 1975; Ashley & Doniach, 1975; Gurman *et al.*, 1984; Rehr & Albers, 1990), which is implemented in modern XAS software. An in-depth review of the XAS theory

and its development is produced by Rehr & Albers (2000). The principles of absorption of X-rays and XAS are described in the following (Scott, 1985; Koningsberger *et al.*, 1988; Jalilehvand, 2000; Newville, 2004; Glatzel *et al.*, 2011).

2.2 Absorption of X-rays in general

When X-rays pass through a sample of thickness x the intensity of transmitted photons, I, is related to the incoming photon intensity, I_0 by the inverse exponential power. This is commonly referred to as the Beer-Lambert law:

$$I = I_0 \ \exp(-\mu \ x) \tag{2.1}$$

where μ is the X-ray linear absorption coefficient, and describes how much of the incoming photon energy is absorbed by matter. μ depends on the energy of the incoming photons E and the elemental composition of the sample.

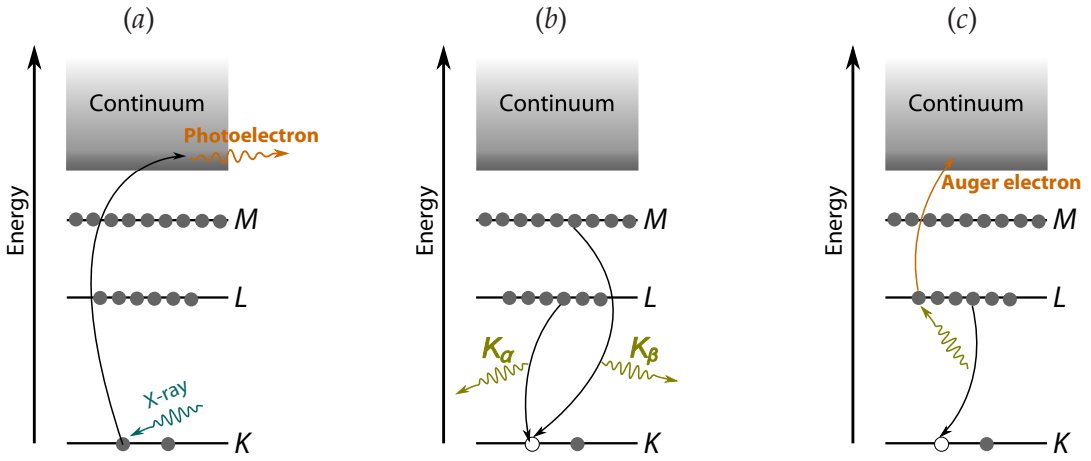


Figure 2.1: Electronic processes taking place during absorption of X-rays. (*a*) Excitation of core electron under ejection of photoelectron. (*b*) Fluorescence where the core hole is closed by an electron in a neighbouring shell. (*c*) Emission of an Auger electron which is promoted to the continuum from another core-level. Revised figure from Newville (2004).

During the absorption process the energy from the X-ray photons is transferred to the sample where it is used for excitation processes, under ejection of photoelectrons, Figure 2.1(*a*). This photoexcitation is by far the most dominant mechanism for absorption of hard X-rays. Any excess energy from the incoming radiation is given to the ejected photoelectron. The energies required to eject deep core electrons from the K, L_{I}, L_{II} and L_{III}-shells to a continuum fall in the X-ray regime and create corresponding core holes in the $1\,s$, $2\,s$, $2\,p_{1/2}$ and $2\,p_{3/2}$ orbitals, respectively. The atoms in excited states decays via competitive processes: (*i*) Fluorescence, where an electron from a neighbour shell closes the core hole under emission of fluorescence, Figure 2.1(*b*). (*ii*) Secondary processes via the formation of Auger

electrons, which are most distinct for lighter elements, Figure 2.1(*c*). For radiation in the hard X-ray regime ($E > 2$ keV) fluorescence is the dominating decaying process. Furthermore, scattering processes including both elastic (diffraction) and inelastic (Compton) scattering also occur.

2.3 X-ray absorption spectroscopy

X-ray absorption spectroscopy constitutes the technique of measuring the variations of μ as a function of energy, E, and the corresponding absorption spectrum is a plot of $\mu(E)$, Figure 2.2(*a*). The energy at which a sharp increase in absorption is observed is referred to as the absorption edge and corresponds to the energy required to eject a core electron to the continuum, which increases with increasing atomic number, Z, for a given edge, Figure 2.2(*b*). This transition is associated with the absorption threshold and characterized by

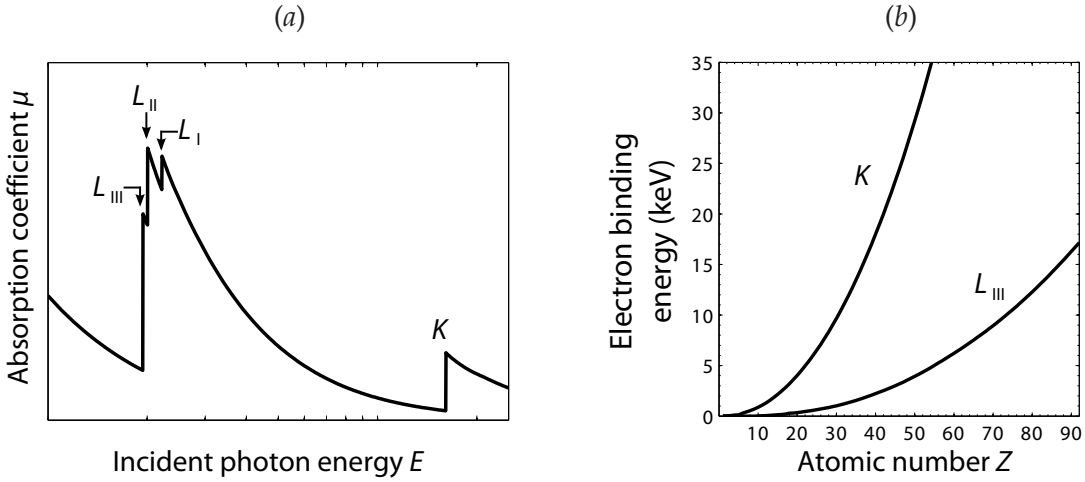


Figure 2.2: (*a*) Absorption coefficient μ as a function of the incoming photon energy E. The position of the K-edge and the L-edges are indicated. Revised figure from Rehr & Albers (2000). (*b*) K and L_{III} absorption edge energies as a function of atomic number, Z.

the energy E_0, which theoretically indicates the energy available for photoelectric absorption via an open continuum channel. Experimentally, E_0 is set as the energy corresponding to the first inflection point of the first derivative below the edge. The exact determination of E_0 is however not possible because its position is influenced by the band-width of the X-ray source, pre-edge structures and instrumental resolution. At energies above the absorption edge photoelectrons are ejected and the kinetic energies of these photoelectrons, E_k, correspond to the excess energy $E_k = E - E_0$.

Practically, the absorption coefficient $\mu(E)$ can be measured in two ways: (*i*) by transmission, where the absorption is measured directly by measuring the number of photons transmitted through the sample or (*ii*) by fluorescence, where the intensity of the emitted

fluorescent X-rays, I_f, is measured, see Section 3.3. Since I_f is directly proportional to the amount of absorbed X-rays in the sample, the fluorescence yield (I_f/I_0) is proportional to the absorption, μ.

The X-ray absorption fine structure (XAFS) is the modulation of μ at and above the absorption edge. The XAFS spectrum can be divided into two different regions: X-ray absorption near edge structure (XANES) and extended X-ray absorption fine structure (EXAFS), Figure 2.3(*a*), which contain related, but slightly different information about the local coordination and chemical state of an element.

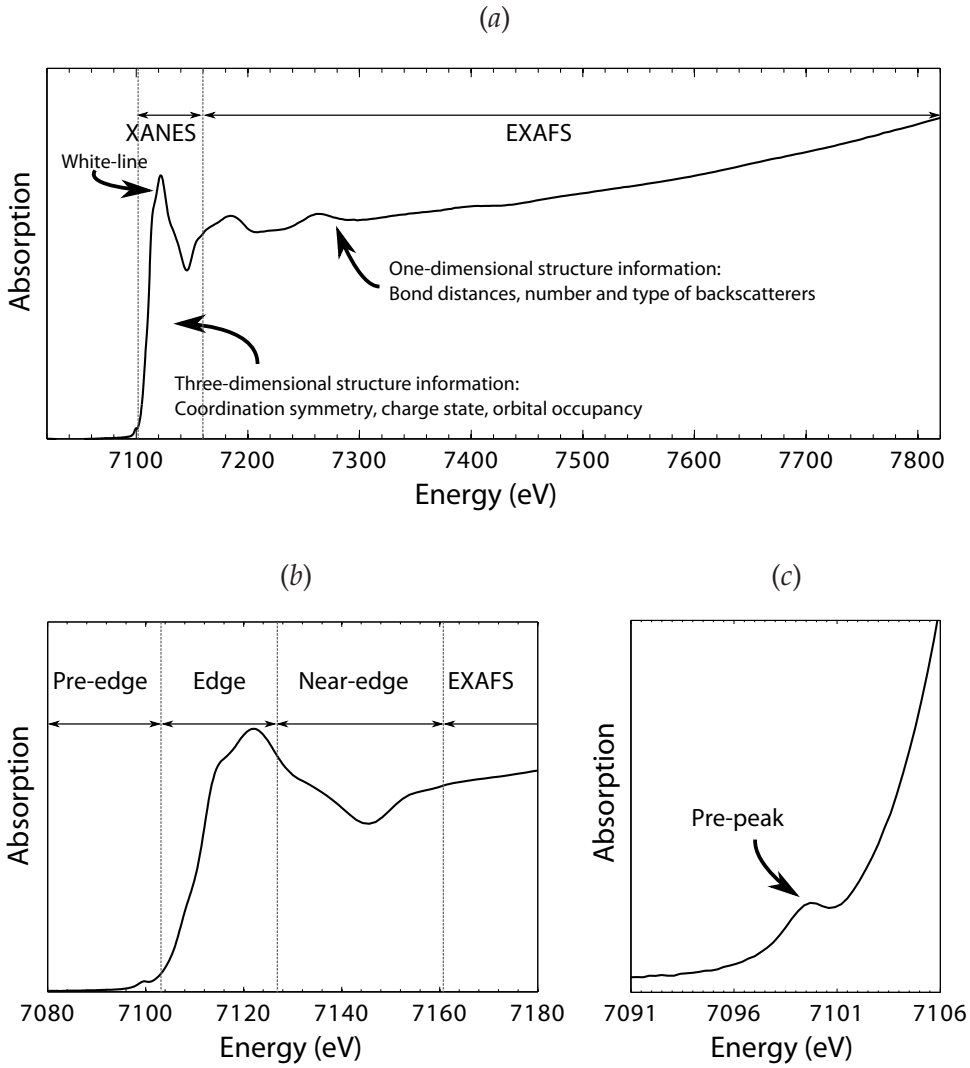


Figure 2.3: Fe K-edge X-ray absorption spectra of reduced horse heart cytochrome *c*, (Frankær, 2011). (*a*) Definition of the XANES and EXAFS regions. (*b*) Zoom of the regions close to the edge: pre-edge, edge and near-edge. (*c*) Zoom of the pre-peak which occurs just below the edge.

2.3.1 X-ray absorption near edge structure

The XANES includes the pre-edge, the edge region and the near-edge region from approx. 10 eV below E_0 to 50 eV above, and contains informations about the local site symmetry, charge state and orbital occupancy, Figure 2.3(*b*).
The pre-edge region is dominated by electronic transitions from the core levels to unoccupied or partially occupied energy levels, and the spectral features are observed as small so-called pre-peaks before the edge in the absorption spectrum, Figure 2.3(*c*). The energy levels of the valence shells are determined by the geometric arrangement of atoms around the absorbing atom. This means that the pre-edge region contains information about the coordination geometry.

At the edge the energy matches the energy required for excitation of an electron to the continuum, which is seen as an increase in absorption. The position of the edge depends on the oxidation state of the absorbing atom and may be shifted by several eV per unit change in oxidation state. On top of the absorption edge a so-called white-line is frequently observed. The intensity of this sharp and intense peak depends on the electronic structure and coordination geometry of the absorbing atom and originates from the fact that the probability of absorption increases at the often closely spaced energy levels near the continuum, mainly $2\,p$ to $n\,d$ transitions, (Koningsberger *et al.*, 1988).

The near-edge region includes the energy range from a few eV to 50 eV above the absorption edge. The term near edge X-ray absorption fine structure, (NEXAFS), is sometimes used to describe this region, but is most common for soft X-ray absorption spectroscopy (edge energies < 1 keV). The absorption features in this region originates from the scattering of photoelectrons with relatively low kinetic energy. This includes contributions from multiple scattering paths between the absorbing atom and its first shell and sometimes also tightly bound second shell neighbour atoms.

Qualitatively, XANES can be used as a fingerprint for the structural arrangement of atoms around the absorbing atom by comparing with model compounds. For a more quantitative interpretation of the XANES recent developments by including full multiple scattering theory (e.g. *FEFF* (Ankudinov *et al.*, 1998)), and using finite difference methods (e.g. *FDMNES* (Joly, 2001)) have allowed calculation of more realistic XANES spectra.

2.3.2 Extended X-ray absorption fine structure

The EXAFS extends from approx. 50 to 1000 eV above E_0. The region is characterized by the scattering of a photoelectron ejected from the absorbing atom by the photoelectric effect, and is mostly dominated by single scattering events. The EXAFS contains information about the local structure, i.e. bond distances, number and type of neighbour atoms. The modulations in the absorption arise from the interference between the outgo-

ing and the backscattered photoelectron and depends on the kinetic energy of the photoelectron. Recent development of data treatment software taking into account the contribution from multiple scattering has permitted inclusion of the near-edge region data and thereby 'stretched' the EXAFS region into the XANES. The EXAFS is further described in the following section.

2.4 Extended X-ray absorption fine structure theory

The fine structure originates from the interference between the photoelectron ejected from the absorbing atom and the backscattered photoelectron from the surrounding atoms. The photoelectrons are thereby very sensitive probes to the distribution of charge and the arrangement of the neighbouring atoms. The interference is correlated to the probability that the backscattered photoelectron returns to the absorbing atom in its excited state, i.e. before the core hole is closed. As the released photoelectron only travels as long as the core hole exists, the limitation for how far from the absorbing atom the structure can be determined by EXAFS depends on the life time of the core hole.

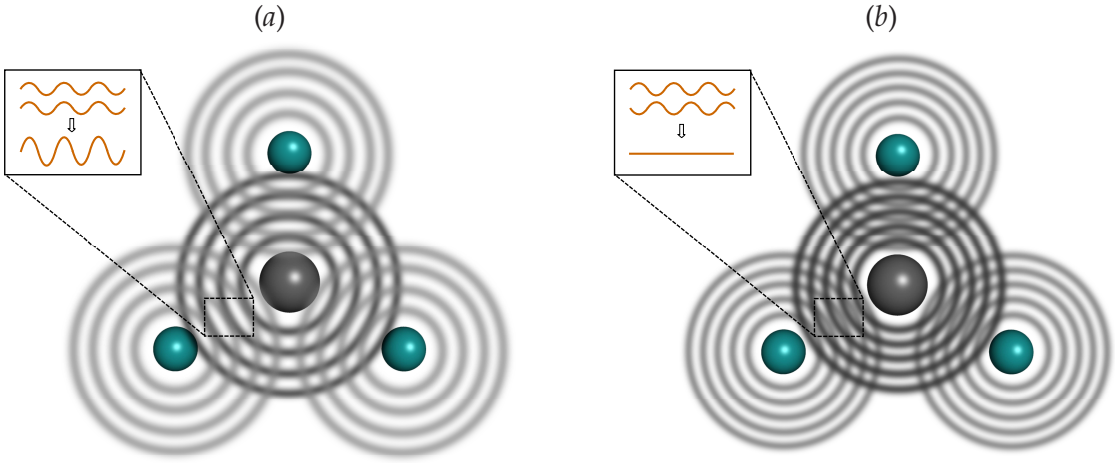


Figure 2.4: Interference between outgoing (dark grey) and backscattered photoelectron (light gray), which is represented by waves, and can either be constructive (*a*) or destructive (*b*). The interference depends on distance between atoms and wavelength (i.e. energy) of the photoelectron.

Different ratios between the distance to the neighbour atom R and the wavelength of the photoelectron λ leads to constructive and destructive interference between the outgoing and backscattered photoelectron, Figure 2.4. Constructive and destructive interferences respectively results in denser and lighter electron densities, which are seen as maxima and minima in the absorption spectrum, respectively. The interference pattern, and thereby the modulations in the absorption spectrum, contains information about the arrangement of atoms around the absorbing atom.

The fine structure modulation in absorbance as a function of energy, $\chi(E)$, is extracted by subtracting the absorption for the isolated atom in the field of its neighbours, without specific interactions, $\mu_0(E)$, from the measured absorption, $\mu(E)$, and then normalizing to one absorption event by dividing with $\mu_0(E_0)$:

$$\chi(E) = \frac{\mu(E) - \mu_0(E)}{\mu_0(E_0)} \tag{2.2}$$

Since EXAFS is an interference effect between the photoelectric waves, the absorption is better represented as a function of photoelectron wave number k, rather than X-ray energy, E. The conversion from E to k is:

$$k = \sqrt{\frac{2\ m_e\ (E - E_0)}{\hbar^2}} \tag{2.3}$$

where m_e is the electron mass. The practical aspects of extraction of χ is further described in Section 3.4

2.4.1 The EXAFS equation

The EXAFS contribution from each backscattering coordination shell j, including atoms of a particular type and similar distances from the origin of the initial photoelectron, can be approximated by a damped sine wave with amplitude, A, and phase, Ψ, characteristic for the associated scattering path. Contracting all contributions into a sum of sine waves results in the EXAFS equation (Sayers *et al.*, 1971), by which the fine structure modulations, $\chi(k)$, can be modelled:

$$\chi(k) = \sum_j A_j(k)\ \sin(\Psi_{ij}) \tag{2.4}$$

$$\chi(k) = \sum_j \frac{N_j\ S_0^2(k)}{k\ R_j^2} \cdot \left|f_j^{\text{eff}}(k)\right| \cdot \exp\left(-2\ k^2\sigma_j^2\right) \cdot \exp\left(\frac{-2\ R}{\Lambda(k)}\right) \cdot \sin\left(2\ k\ R_j + \phi_{ij}(k)\right) \tag{2.5}$$

N_j is the number of backscatterers in the j^{th} shell.
R_j is the distance from the absorbing atom i to the backscatterer in the j^{th} shell.
$f_j^{\text{eff}}(k)$ is the effective amplitude function.
$S_0^2(k)$ is the amplitude reduction factor.
$2\sigma_j^2$ is the Debye-Waller term representing the mean-square disorder (static and thermal) in the j^{th} shell.
$\Lambda(k)$ is the mean free path of the photoelectron.
$\phi_{ij}(k)$ is the phase shift arising from the Coulomb potential between atom i and j.

The amplitude of the EXAFS signal is proportional to the number of backscatters N around the central atom Figure 2.5(*a*), whereas the phase contains information about the bond distances from the central atom to the neighbouring atoms, R. As the photoelectric wave must travel from the absorbing atom to the scatterer and back, the frequency of each EXAFS wave is related to R, in a way that high frequencies corresponds to long distances and *vice versa*, Figure 2.5(*b*).

The amplitude of each EXAFS wave further depends on the number and the backscattering power of the neighbouring atoms. The backscattering power of an atom is described by the effective amplitude function, $f_j^{\text{eff}}(k)$, which only depends on the nature of the backscattering atom. The backscattering power is generally higher for heavier backscatters, and the amplitude function $f_j^{\text{eff}}(k)$ takes maximum values at low k-values (3–4 Å^{-1}) for light backscatterers (C, N, O etc.) and at higher k-values for heavier backscatterers, Figure 2.5(*c*). In modern curved wave theory, (in which the effect arising from the curvature of the atoms is taken into account) $f_j^{\text{eff}}(k)$ also slightly depends on the distance to the central atom, R. On the way the photoelectron will interfere with the Coulombic interactions of the absorber, i, and the backscatterer, j. This results in a total phase shift, $\phi_{ij}(k)$, which is a sum of the k-dependent phase shifts the photoelectron experiences during the scattering process: two from the absorber, $2\delta_i$, (both outgoing and incoming) and one from the scattering δ_j. The k-dependence is shown in Figure 2.5(*d*). Calculations of both phase shifts and amplitude functions are implemented in modern algorithms for *ab initio* calculations of XAFS spectra, which use the methods described in Section 2.5.

The amplitude reduction factor, $S_0^2(k)$ mostly describes the reduction of signal due to inelastic losses within the absorbing atom (intrinsic). Outer-shell electrons of the absorbing atom may either be ejected to empty orbitals or to the continuum using the kinetic energy of the photoelectron, (shake-up and shake-off processes, respectively). This energy loss results in photoelectrons which EXAFS contributions are shifted in energy and phase, and thereby reduces the total EXAFS amplitude. Also experimental factors influence on $S_0^2(k)$, e.g. self absorption of highly absorbing samples where the fluorescence emitted is absorbed by the sample itself, which for fluorescence measurements result in a reduced intensity emitted from the sample. The contributions to $S_0^2(k)$ are difficult to quantify, and $S_0^2(k)$ is in practise estimated from a standard sample of a well known structure. The fact that the EXAFS phases are better determined than the amplitudes generally makes the EXAFS-technique more accurate for bond distances than for coordination numbers.

The Debye-Waller term, $\exp(-2\sigma_j^2\ k^2)$, describes both the static and thermal disorder effects. σ represents the half width at half maximum in a Gaussian distribution of distances. The distribution of distances is normally assumed to be symmetric around the mean distance, but asymmetric distributions can be modelled by using expanded Debye-Waller terms.

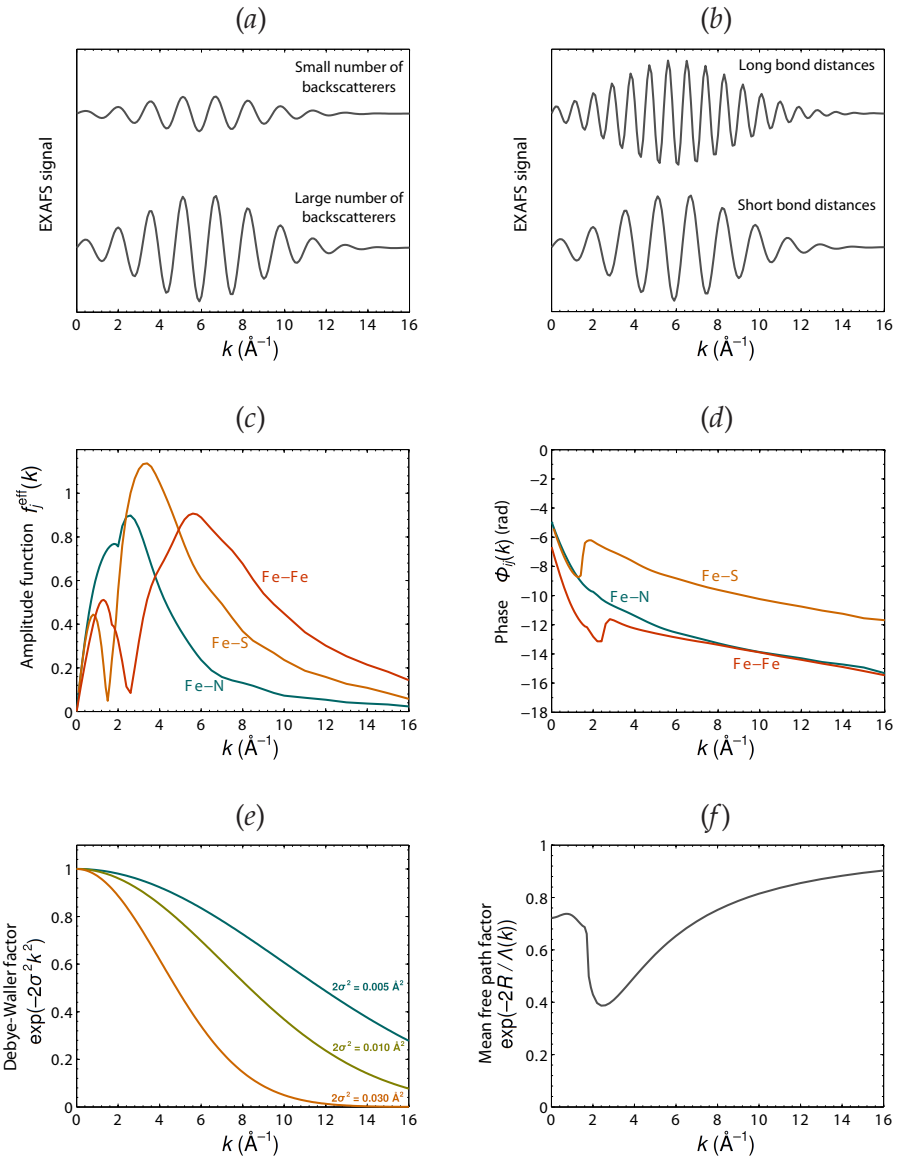


Figure 2.5: The k-dependence of selected factors which are included in the EXAFS equation. (*a*) The influence on the EXAFS signal of the number of backscatterers N_j within the j^{th} shell. The signal is amplified by more backscattering atoms. (*b*) The influence on the EXAFS signal of the bond distance R_j to the j^{th} shell. Long distances results in high frequencies and *vice versa*. (*c*) The effective amplitude function, $f_j^{\text{eff}}(k)$ and (*d*) phase shifts ϕ_{ij} calculated using Fe as absorber atom for different backscatters at different distances: N at 1.99 Å, S at 2.30 Å and Fe at 2.46 Å. The position of the maximum of $f_j^{\text{eff}}(k)$ in k-space increases with increasing atom number. The calculations are done using *FEFF8* (Ankudinov *et al.*, 1998). (*e*) The Debye-Waller term $\exp(-2\ \sigma^2\ k^2)$ for different values of $2\ \sigma^2$. The damping effect increases with increasing $2\ \sigma^2$-values. (*f*) The mean free path term $\exp(-2\ R_j/\Lambda(k))$ calculated using 2.46 Å Fe–Fe scattering path.

Regarding the photoelectron scattering from an atom located at a mean distance R and with a distribution of distances σ from the central atom, the backscattered contributions will be slightly out of phase and thereby damping the EXAFS signal, in particular at high k-values, Figure 2.5(*e*).

Inelastic scattering of the photoelectron involving ejection of valence electrons from the neighbouring atoms (extrinsic) may also occur. These extrinsic losses will not contribute to the interference between the outgoing photoelectric wave with backscattered wave, and will thereby damp the EXAFS signal. Further the limiting life time of the core hole will also damp the signal, as the backscattered photoelectron must return to the core hole before the excited state decays. To describe these losses a damped spherical wave which includes the mean free path of the photoelectron, Λ, is used. The damping factor is plotted in Figure 2.5(*f*). Λ describes the probability of how long the photoelectron is able to travel in average and generally increases with increasing photoelectron energy. The mean free path factor together with the $1/R^2$ dependence makes EXAFS a local atomic probe.

2.4.2 Multiple scattering

The sum over contributions in the EXAFS equation includes single scattering (SS) paths from many shells of atoms, in which the photoelectron is travelling from the central atom to the backscatterer and back again, Figure 2.6(*a*). But the contributions may also include so-called multiple scattering (MS) paths in which the photoelectron is scattered by more than one atom before it returns to the central atom, Figure 2.6(*b*).

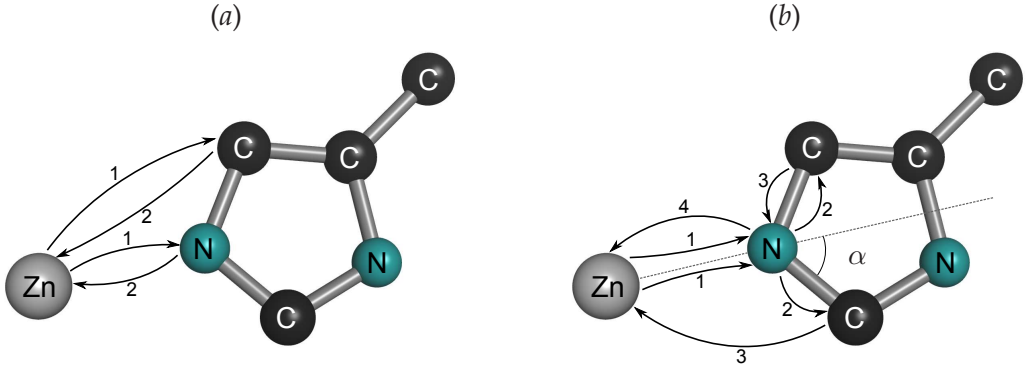


Figure 2.6: The definition of scattering pathways. (*a*) Single scattering events where the photoelectron is only scattered once, $n_{\text{leg}} = 2$. (*b*) Multiple scattering pathways where the photoelectron is scattered by more atoms before returning to the absorbing atom for a three-leg pathway $n_{\text{leg}} = 3$, and a four-leg pathway $n_{\text{leg}} = 4$. The angle α describing the deviation from linearity is indicated.

An MS pathway is characterized by its number of legs, n_{leg}, which corresponds to the number of distances the photoelectron travels before it returns to the absorber. Hence for

single scattering pathways $n_{\text{leg}} = 2$, whereas multiple scattering pathways have $n_{\text{leg}} \geq 3$. MS pathways are always longer than SS pathways from the inner coordination shell, and their amplitudes and phases are modified. MS becomes more important at low energies (XANES region), which explains why the XANES regime is sensitive to the spatial arrangement of atoms around the absorbing one relative to each other, and not only their radial distances. This is illustrated in Figure 2.7. The effect from an MS pathway depends on the angle α, describing the deviation from linearity of the atoms involved in the MS path, as defined in Figure 2.6(*b*). The contribution is generally enhanced for linear systems ($\alpha = 0°$), since the forward scattering from the intermediate atom is strong (nitrogen atom in Figure 2.6(*b*)). This is known as the focussing effect. The contribution will decrease with increasing α.

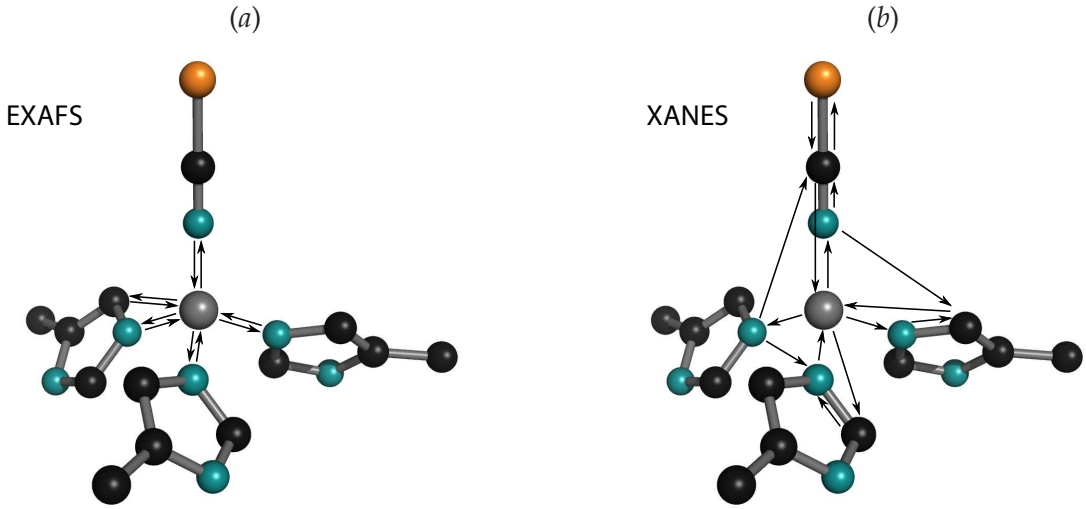


Figure 2.7: (*a*) In the EXAFS part of an absorption spectrum, where the photoelectron has high energy, single scattering dominates. (*b*) The XANES region is more sensitive to multiple scattering events, as these occurs at lower photoelectron energies. Thereby EXAFS is a one-dimensional probe determining bond distances R very accurately, whereas XANES is also sensitive to bond angles.

2.4.3 Fourier transformation

A Fourier transform of the EXAFS function in k-space, $\chi(k)$, results in a pseudo radial distribution function, describing the electron density around the absorbing atom in a one dimensional R-space.

For simplicity, the phase shift terms, ϕ_{ij}, are normally omitted in the preliminary Fourier transforms. As a consequence of not including these terms, the peak positions giving the average absorber-backscatterer distances, R_j, are typically shifted by approx. 0.5 Å be-

low their true values. The phases are however calculated in the fitting procedure to obtain accurate values of R_j.[1]

The spatial resolution, δR, depends on the k-range used for the Fourier transform, Δk. δR can be estimated by the following equation:

$$\delta R = \frac{\pi}{2\,\Delta k} \tag{2.6}$$

Small k-intervals give broader peaks and thereby lower resolution. Often atomic shells of similar distances can not be resolved and have to be modelled as one shell. In order to avoid truncation effects $\chi(k)$ must take values close to zero in the end points. Convolution of $\chi(k)$ with a window function, which is constant over the selected k-range, except at the end point where it smoothly approaches zero is often used to obtain clean peaks in the Fourier transform. During fitting analysis individual peaks in the Fourier transform can be cut out and back-Fourier transformed in to k-space. This technique known as Fourier filtering may help to identify different pathways, as they can be analysed individually.

2.5 Concepts on XANES calculation

For a quantitative interpretation of the XANES part of an absorption spectrum it is necessary to calculate realistic spectra, which is often accompanied by heavy quantum mechanical calculations. Development of new methods providing short-cuts to these heavy calculations has improved the quantitative interpretation of XANES.

In the one-electron approximation, in which only one electron is assumed to be ejected during the excitation of the atom, the X-ray absorption coefficient μ is proportional to the transition rate, which is given by Fermi's golden rule:

$$\mu(\omega) \propto \sum_f \left|\langle \psi_f \left|\mathbf{p}\cdot\mathbf{A}(\mathbf{r})\right|\psi_i\rangle\right|^2 \delta\left(E_f - E_i - \hbar\omega\right) \tag{2.7}$$

where $\hbar\omega$ is the energy and $\mathbf{A}(\mathbf{r})$ the vector potential of the electromagnetic field of the incident light. $\mathbf{p}$ is the momentum operator and ψ_i and ψ_f are the wave functions corresponding to the initial and final eigenstates with energies E_i and E_f, respectively. In the one-electron approach ψ_i is a $1\,s$ wave function. It is calculated by solving the Schrödinger equation for the initial state, which is relatively easy as the atom is in its ground state. Solving the Schrödinger equation for the final state, however, is very complicated as all electrons in principle are involved, leading to many-body processes. The final state is strongly dependent on the environment. To overcome calculation of ψ_f several approximations and calculation strategies is used, which will be briefly described in the following:

[1] Plotting Fourier transforms without phase correction is common procedure in many programs (e.g. *WinXAS*, (Ressler, 1998)).

2.5.1 Muffin-tin approximation

As the ejected photoelectron propagates in the sample it moves in a potential. This potential determines the scattering strength at each site and must be as realistic as possible for performing accurate calculations of spectra. The muffin-tin (MT) approximation, is called so, as its profile resemble the shape of a muffin tin, and assumes non-overlapping spherical scattering potentials centred on each atom and a constant value in the interstitial regions. The calculations are eased by using this approximation, and it describes the actual potential very well near the centre of each atom. As the high energy photoelectrons are mostly scattered from these inner regions, the MT-potential is highly adequate for most EXAFS calculations. The shape of the potential in the outer edges of the atoms as well as the interatomic potential however becomes more important in the near edge region, where the photoelectron energies are lower. A more accurate description is often needed taking into account the fact that the atomic potentials may overlap and that the charge distribution, which is determined by the bonding properties in the material, is anisotropical.

2.5.2 Green's formalism

In the Fermi golden rule the wave functions ψ can be substituted with Green's functions, which eases the calculations of the final states. The Green formalism has been essential for the development of the multiple scattering theory and is implemented in many programs, e.g. *FEFF* and *FDMNES*.

2.5.3 Finite difference methods

Finite difference methods, (Kimball & Shortley, 1934) is a general mathematical method to solve differential equations by discretize them over a grid of points, and has been implemented in *FDMNES*, (Joly, 2001). The Schrödinger equation has to be solved at each point within a spherical volume which is centred on the absorbing atom. The radius of this sphere defines the cluster size. The calculation time is strongly dependent on the cluster size. The FDM is today the most precise approach to calculate XANES spectra (Jacquamet *et al.*, 2009a) as the method offers an approach to solve the Schrödinger equation in a free shape potential, and thereby goes further than the muffin-tin approximation.

Chapter 3

Experimental X-ray absorption spectroscopy

This chapter gives a brief introduction to experimental XAS including specifications of the samples, the experimental setup, acquisition and treatment of XAS data, and is a supplement to the experimental descriptions of XAS data collection procedures in the following chapters. The two final sections respectively present an overview of the different strategies used in this project to analyse the EXAFS data as well as quantitative analysis of XANES data.

3.1 Availability of absorption edges

XAS experiments can be performed for almost any type of absorber atom in the periodic table. For biological XAS studies hard X-rays ($E > 2$ keV) is mostly used, and the energy available at synchrotron XAS beam lines typically range form 2 to 40 keV. In principal there is no upper limit of Z, since elements with K-edges above the available photon energy range, will have L-edges which fall within the available range, see Figure 2.2(b). The lower level is around 2.5 keV, which corresponds to absorbers with atomic number $Z \geq 16$, (sulfur K-edge is 2472 eV). For energies $E < 2$ keV absorption of the X-rays by air becomes an important issue even with experiments performed in He-atmosphere.

3.2 Sample specifications

XAS can be obtained from samples in any state, e.g. crystals, micro crystals, amorphous solids, liquids, frozen solutions or solutions at room temperature or even directly from tissue and cells. The technique is very sensitive wherefore it can be applied to study trace metals and dilute samples. Obtaining useful XAS spectra highly depend on the sample and the preparation hereof. Important issues about sample preparation are summarized in the following points:

- The sample must be homogeneous. This is primarily an issue for solid material, e.g. bone tissue, which must be carefully ground.
- The amount of absorbing atoms must be optimized to the experiment to obtain the optimal signal-to-noise ratio. This is done by optimizing the sample thickness x and/or the concentration for solutions. For fluorescence experiments on metalloproteins, the samples typically contain metal concentrations around $1 - 5$ mM (Ascone *et al.*, 2003), but using specially designed setups high quality XAS data can be collected on samples with metal concentrations down to 100 μM, (Ranieri-Raggi *et al.*, 2003).
- The sample must be stable during the entire measurement, meaning that photo-reduction and degradation from radiation damage must be avoided. This is done by cooling the samples in a liquid nitrogen or helium cryostat. Some systems are more sensitive to photo-reduction. By performing faster scans and thereby collecting more spectra allows for monitoring an eventual evolution caused by degradation. Moving the beam position on the sample to an unexposed spot or changing the sample with a freshly prepared one are strategies used to overcome this problem.
- For metalloproteins, excess amounts of the absorbing metal must be eliminated, (i.e. metal atoms which do not coordinate to the protein, but can be present in the solution).

3.3 Experimental setup and data acquisition

A typical setup for an X-ray absorption spectroscopy experiment on a synchrotron facility is shown in Figure 3.1.

As the absorption fine structure oscillations generally are weak signals (1–10 % of the absorption coefficient), synchrotron radiation sources are required for performing XAS experiments. Synchrotron radiation is generated in a high vacuum storage ring in which electrons, accelerated to the speed of light, are injected. The storage ring is in principle a

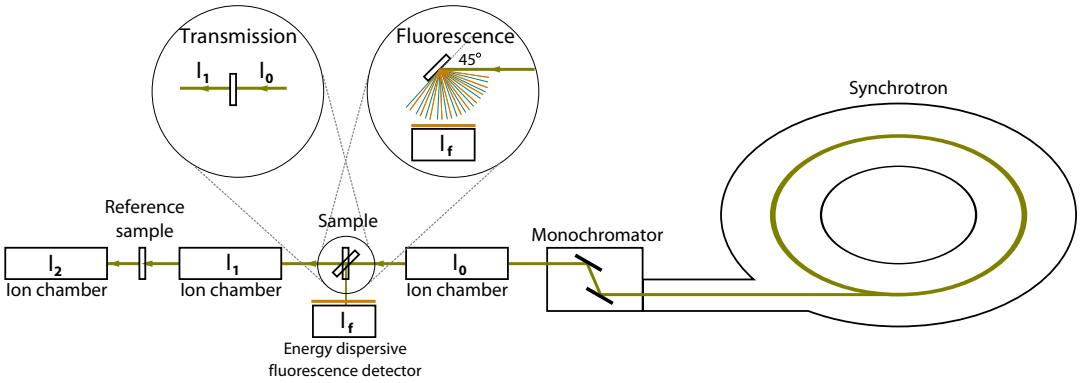


Figure 3.1: Schematic representation of an experimental arrangement for XAS. Synchrotron radiation is generated from accelerated electrons in the storage ring. The monochromator provides a monochromatic beam which can be tuned in energy. The intensities I_0, I_1 and I_2 are measured in the ion chambers before the sample, after the sample and after having passed a reference sample, respectively. For transmission experiments the sample is aligned perpendicular to the beam. For fluorescence experiments, the sample is rotated with an angle of 45° to both the incident beam and the energy dispersive fluorescence detector, by which the intensity of the fluorescence is measured, I_f.

polygon consisting of curved and straight sections which define the trajectory of the electrons. The electrons are forced to follow the curvature of the ring by magnetic fields from strong magnets. When the electrons hit the curved sections (bending magnets) they loose energy which is emitted as electromagnetic radiation tangential to the ring. The energy of the electrons can be tuned and modified in the straight sections by undulators and wigglers providing radiation with the wished characteristics. The emitted radiation is further focussed and cut in the guide line leading the radiation from the ring to the experimental hutch, by mirrors and slits.

The normal procedure for collecting XAS spectra is by collecting data points at discrete photon energies, which is achieved by a scanning high-resolution monochromator. The monochromator contains a pair of monochromator crystals, in which Bragg reflections from a certain diffraction plane is used to monochromatize the radiation. At beam line 811 at MAX-lab interchangeable pairs of Si(111) and Si(311) crystals are used in the monochromator to cover an energy range from 2.3 to 20 keV. By varying the angle between the incident X-rays and the Bragg planes different energies can be sorted out.

XAS data can be collected in two different modes: Transmission and fluorescence. Transmission experiments are conducted on concentrated samples, whereas fluorescence experiments are more adequate for dilute samples used for biological applications where the concentration of absorbing element typically is below 1 %, (Ascone *et al.*, 2003).

In transmission experiments the sample is aligned perpendicular to the beam, and the detector is placed behind. The absorption is obtained in analogy to optical spectroscopy by measuring the intensities of the incident beam, I_0, and the transmitted intensity, I_1.

The absorbance μ is calculated as $\log(I_0/I_1)$. The intensities I_0, I_1 and I_2 are measured in ion chambers. The ion chambers are filled with mixtures of inert gases (He, N_2, Ne, Ar or Kr), and a voltage is applied over the chambers. When the beam transverses through the ion chamber, the gas is ionized, which is converted into an electrical signal. An optimal composition of the gas ionization chambers should provide an absorption of 20 % of the incident beam in I_0, and 20–40 % of the incident beam in I_1. The gas mixture in I_2 absorbs the majority of the remaining beam. As the absorption depends on the concentration of absorbing element and the thickness of the sample, the sample preparation must be optimized with respect to sample thickness and concentration to obtain spectra with optimal signal-to-noise ratios. Solid samples containing high concentrations of the absorbing element are typically diluted with boron nitride.

In fluorescence experiments the optimal orientation of the sample is with an angle of 45° to both the incident beam and the fluorescence detector. The fluorescence emitted from the sample I_f is detected by an energy-dispersive detector placed parallel to the beam, see Figure 3.1. The fluorescence yield, which is proportional to the absorbance μ is calculated as (I_f/I_0).

A concurrent transmission experiment is conducted on a reference sample for internal calibration of the spectra. The reference sample is typically a foil of pure metal of the same type as the absorbing atom, and placed between ion chambers I_1 and I_2. The absorbance μ is thus calculated as $\log(I_1/I_2)$.

3.4 Data treatment

The data treatment procedure of an experimental XAS spectrum is divided into several steps. For extraction of XANES spectra only the first three steps are required, whereas all steps are necessary for extracting the EXAFS oscillations $\chi(E)$. In this work the data treatment was carried out in *WinXAS*, (Ressler, 1998).

Firstly, an experimental spectrum must corrected for background, which originates from absorption by other atoms present in the sample, instrumental effects and absorption by sample holder and surroundings. For transmission data the background is decreasing, Figure 3.2(*a*), since the absorption of X-rays decreases with increasing energies, whereas the opposite is observed for fluorescence experiments, Figure 3.2(*b*), as the fluorescence increases with increasing energy. The background is empirically modelled by a smooth polynomial (typically first order) which is fitted to the pre-edge of the spectrum and subtracted.

Secondly, the experimental spectrum is normalized by scaling the height of the absorption step, μ_{step} to unity, Figure 3.2(*c*). The concentration of absorbing atoms is proportional

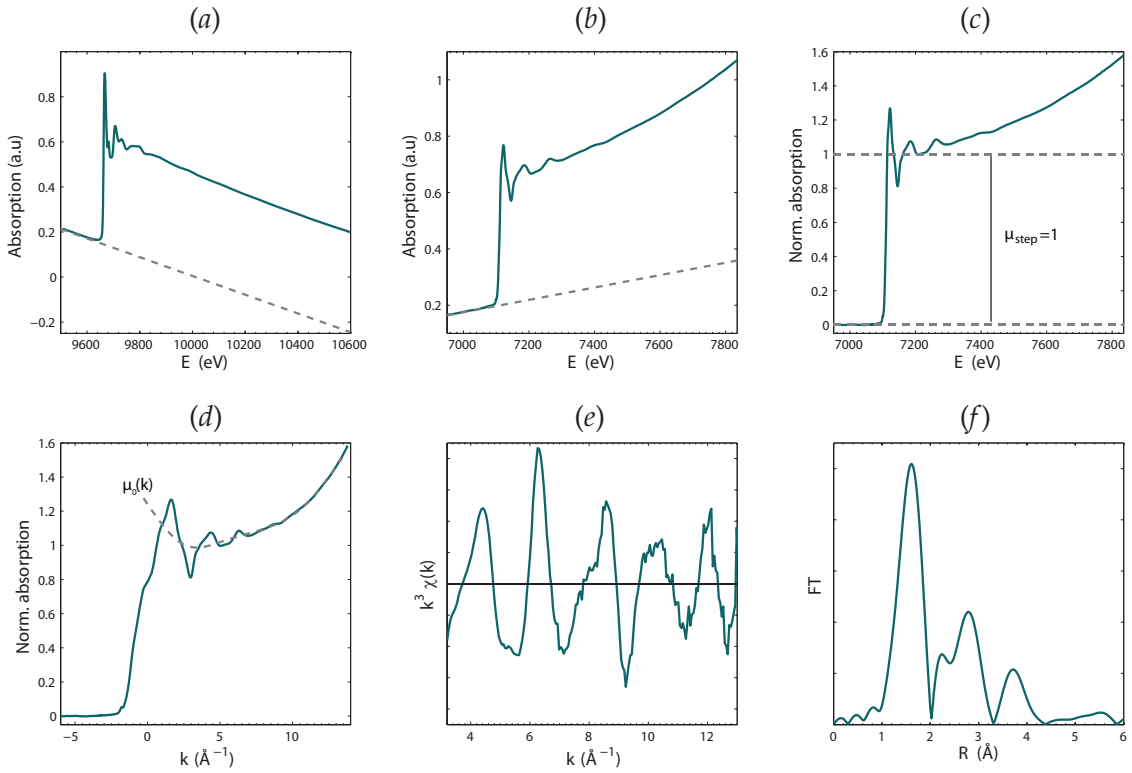


Figure 3.2: Extraction of the EXAFS oscillations $\chi(k)$ from raw spectra. (*a*) Transmission spectrum of a zinc imidazole complex (Frankær, 2011) and (*b*) fluorescence spectrum of reduced horse heart cytochrome *c* (Frankær, 2011) in which the backgrounds are modelled by a first order polynomial fitted to the pre-edge (dashed lines) before subtraction. The background is generally decreasing in transmission spectra (*a*) and increasing in fluorescence spectra (*b*). (*c*) Normalization is carried out by scaling the absorption step μ_{step} to unity. (*d*) After setting the threshold energy E_0, the spectrum is converted to k-space and an empirically modelled spline function, $\mu_0(k)$ (dashed line), imitating the absorption of the isolated atom in the field of its neighbours without specific interactions, is subtracted. The EXAFS function $\chi(k)$ is obtained after normalization by $\mu_0(k)$. (*e*) The k^3-weighted EXAFS function $\chi(k)$, amplifying the features at higher k-values. (*f*) Fourier transformation of $\chi(k)$ results in the radial distribution function.

to this step height and varies in different samples. A normalization allows EXAFS spectra from samples with different concentrations of absorbing atom to be compared.

Thirdly, the spectrum must be energy calibrated by setting the threshold energy E_0. There is no simple way to determine E_0 experimentally, wherefore the threshold energy offset ΔE_0 is refined upon model fitting to an EXAFS spectrum. However, when more spectra are compared it is of utmost importance that the spectra are energy calibrated relatively to each other. Therefore the spectra are calibrated by using a concurrently collected spectrum from a well-characterized reference sample. A careful energy calibration is also important when more experimental spectra are averaged.

For extraction of the EXAFS oscillations, it is convenient to convert the spectrum $\mu(E)$ into a function of wave number $\mu(k)$ using Equation (2.3). The fine structure $\chi(k)$, is ex-

tracted by subtraction of a spline $\mu_0(k)$, which is an empirically determined background function imitating the absorption for the isolated atom in the field of its neighbours, without specific interactions, Figure 3.2(*d*). Division by $\mu_0(k)$ normalizes the spectrum to a per atom basis.

$$\chi(k) = \frac{\mu(k) - \mu_0(k)}{\mu_0(k)} \tag{3.1}$$

The EXAFS function $\chi(k)$ can be k^n-weighted in order to enhance the features at higher k-values, Figure 3.2(*e*). For biological samples which primarily contain light backscatterers n is typically 3. Finally, the EXAFS function is Fourier transformed resulting in the radial distribution function, Figure 3.2(*f*), as described in Section 2.4.3.

3.5 Model fitting to EXAFS spectra

The data analysis always involves some form of curve fitting of a model function to the experimental EXAFS spectra, thereby extracting structural and chemical information of the samples. The theoretical model function is constructed by a sum of individual contributions as given in Equation (2.5). Backscattering amplitudes and atomic phase shifts are obtained from *ab initio* calculations for selected pathways from the theoretical model. The values of the parameters (distances R_j, number of backscatterers N_j, Debye-Waller factors σ_j^2 and threshold energy offset ΔE_0) are then adjusted and refined by a least-squares fitting procedure. The final model includes all pathways which significantly contribute to the experimental spectrum.

There are strong correlations between some of the parameters. As both N_j, and σ_j^2 influence on the amplitude of $\chi(k)$ these parameters usually correlates, whereas R_j and ΔE_0 influence on the phase of $\chi(k)$ and correlate.

As the three-dimensional structure information obtained by EXAFS is decomposed into a one-dimensional spectrum, the number of observables are limited and depends on the data quality. The number of independent parameters N_p can be estimated by:

$$N_p = \frac{2\,\Delta k\,\Delta R}{\pi} + 2 \tag{3.2}$$

in which Δk and ΔR and the k-range and Fourier filtering R-range, respectively. The number of parameters to be refined can be limited by fixing the number of backscatterers N_j, as they can be determined from other experiments (e.g. X-ray crystallography), and by grouping similar distances R_j and Debye-Waller factors, σ_j^2. Three different approaches to fit models to the experimental EXAFS spectra are illustrated in the work described in the following chapters.

- The cluster approach, in which atoms within the same molecule are grouped. The groups can be constrained by which each group is assumed to be rigid, or the groups

can be imposed to geometrical restraints, which allows some movement of the atoms within the molecule (Binsted *et al.*, 1992; Gurman, 1995). This approach was applied in the study of hexameric insulin described in Chapter 7 and 8 using *EXCURVE* (Gurman *et al.*, 1984, 1986; Binsted *et al.*, 1991), which perform full multiple scattering calculations. Thereby three-dimensional information is indirectly obtained from the EXAFS. Using initial PDB-coordinates from crystal structures this approach, outlined as 3D-EXAFS by Strange *et al.* (2005), is probably the best suited for metalloprotein studies.

- The shell approach, by which atoms at equal distances from the absorbing atom are grouped in shells. By the separate inclusion of pathways (both SS and MS pathways) this approach is not the best suited for metalloproteins, as metal clusters in proteins often contains many shells and thereby results in too many parameters. This approach was applied for strontium reference compounds in the study of bone tissue using *WinXAS*, (Ressler, 1998). The EXAFS results of the reference compounds are presented in Appendix C.
- A third approach, in which a linear combination of experimental EXAFS spectra from model compounds are fitted to the experimental EXAFS spectrum, may be more adequate for highly disordered structures. This approach was applied in the study of strontium localisation in bone tissue described in Chapter 11 using the linear combination fit procedure included in *ATHENA* (Newville, 2001; Ravel & Newville, 2005).

All refinement evaluate the fit by a residual. The residuals are differently defined and presented in the respective chapters.

3.6 Quantitative model fitting to XANES spectra

Advanced fitting algorithms against XANES spectra have recently been developed and improved during the last decade (Benfatto *et al.*, 2001; Smolentsev & Soldatov, 2006; Sarangi *et al.*, 2008; Jacquamet *et al.*, 2009b). The main advantage of quantitative XANES analysis is the possibility to obtain the information about three-dimensional environment of absorbing atom. Unlike the EXAFS shell approach, which provides only the radial distribution of atoms, the XANES enables to withdraw the bond lengths and angles.

For quantitative fitting of XANES spectra an initial model is required, where the coordinates can be provided from a preliminary EXAFS analysis or an XRD structure. A few structural parameters (typically two to four) can be varied and optimized by multidimensional interpolation of calculated XANES spectra. The XANES fitting was applied in the study of hexameric zinc insulin described in Chapter 7, using *FDMNES* (Joly, 2001) and *FitIt* (Smolentsev & Soldatov, 2007).

Chapter 4

Characterization of protein powder by XRPD

X-ray diffraction on single crystals is the preferable method to solve and characterize protein structures. The results are usually unambiguously clear, when the structure is revealed. The need to find the right conditions for growing single-crystals of sufficient size and quality is one of the major bottlenecks in this process, and being able to screen early precipitates for crystals would save large amounts of time and efforts. Furthermore, as structure based drug design is a growing field and drug candidates are more often proteins themselves, the need for quick characterization and verification of protein polymorphs and substrate-protein complexes is growing. X-ray powder diffraction (XRPD) offers a convenient and quick way to characterize crystalline protein precipitate in the laboratory and could thereby become a valuable tool for product characterization in the pharmaceutical industry, as well as in the enzyme production industry.

For reliable identification of crystal forms, calculation of realistic protein powder patterns based on for example Protein Data Bank (PDB) coordinate files are required. A procedure which includes straightforward corrections for background, unit cell parameters, disordered bulk-solvent and geometric (Lorentz) factors has been developed and tested with good agreements between measured and calculated patterns. A description of the procedure has been summarized in this chapter and published in Paper I (Hartmann *et al.*, 2010) Appendix A.1 and Paper II (Hartmann *et al.*, 2011) Appendix A.2.

The calculation procedure have been tested with tetragonal lysozyme and trigonal and cubic insulin. In this chapter only the results from lysozyme are summarized. Further details and the results from the two insulin crystal forms are found in Paper I, (Hartmann *et al.*, 2010) Appendix A.1.

4.1 Introduction

Metmyoglobin was the first protein structure to be refined from powder diffraction data (von Dreele, 1999) by combining high-resolution diffraction data with restrained Rietveld refinements. During the last decade X-ray powder diffraction has proven its ability for solving and refining small protein structures such as insulin (von Dreele *et al.*, 2000), lysozyme (von Dreele, 2001, 2005; Basso *et al.*, 2005) and SH3-ponsin domain (Margiolaki *et al.*, 2007b) from high resolution powder data using synchrotron radiation (von Dreele *et al.*, 2006; Margiolaki & Wright, 2008; Wright *et al.*, 2008; von Dreele, 2009; Doebbler & von Dreele, 2009).

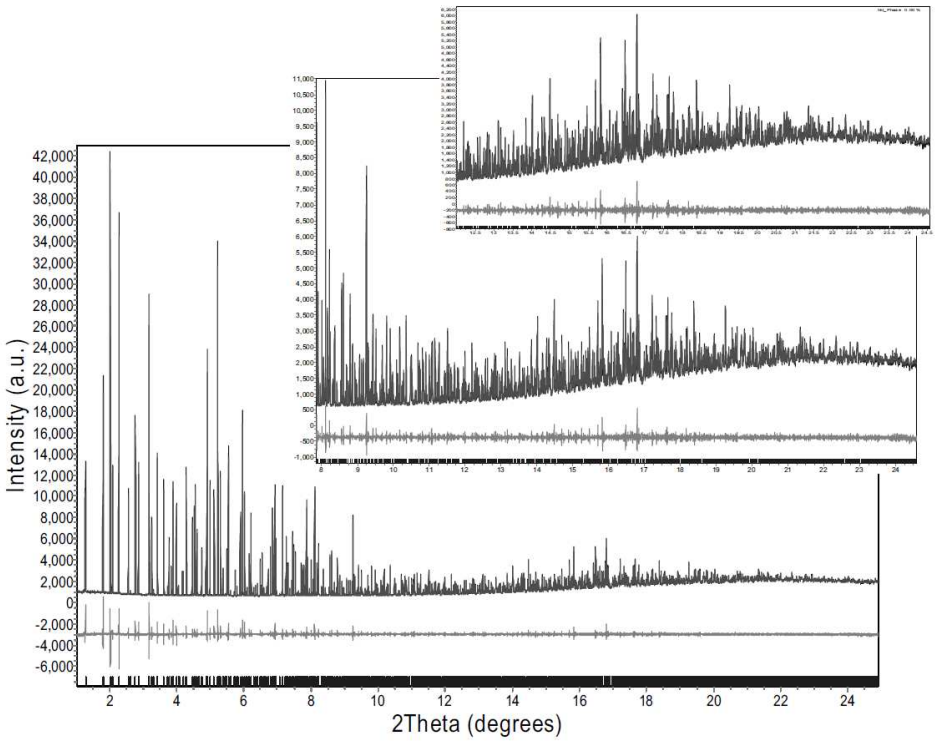


Figure 4.1: High resolution powder diffraction patterns for tetragonal lysozyme recorded on beam line ID31 at ESRF, Grenoble, France. Magnification of the high angle regions are shown in the insets. This beam line provides data with extremely high angular resolution (FWHM of around 0.003° in 2θ) and is thus well-suited for protein powder diffraction experiments. Adapted from (Margiolaki *et al.*, 2007a) by permission from Oldenbourg Wissenschaftsverlag GmbH, ©2007.

The critical point in refining protein structures from powder data is to obtain data with resolution high enough to separate the diffraction peaks which otherwise overlap. However, even at specially designed beam lines using analyser crystals some peaks may still be overlapping. A strategy in which diffraction peaks can be separated includes combination of more data sets collected at different conditions. In a study by (Basso *et al.*, 2005) an alteration of pH and temperature induced small systematic changes of the tetragonal unit cell of lysozyme, which shifted the diffraction peaks. A high resolution powder pattern for lysozyme is shown in Figure 4.1 Powder diffraction patterns with well resolved peaks have also been collected on highly optimized laboratory equipment, (Margiolaki *et al.*, 2007a), but at the cost of longer data collection times.

Going from high resolution to medium resolution XRPD, a recent study using synchrotron radiation demonstrated the ability to quickly and effectively distinguish between different crystal forms of insulin in combination with multivariate data analysis (Norrman *et al.*, 2006). Based on these previous experiences the aim was to study whether protein crystal forms could be sufficiently characterized and fingerprinted by using standard laboratory diffractometers.

Collecting XRPD data with sufficient signal-to-noise ratios and calculation of realistic powder patterns require special procedures compared to those used for small molecules. Some challenges concerning sample handling and data treatment have been identified:

- The large unit cells result in Bragg peaks occurring at very low angles and at higher angles the reflections are suffering from a severe peak overlap, which limit the data resolution to about 4 Å (2θ =2–20° using Cu $K\alpha_1$-radiation).
- The measured intensities fit poorly with those calculated from known structures. Protein structures available from the PDB, do not include a description of the contribution from the disordered solvent in the solvent channels, which is the major source for this discrepancy.
- The measured peak positions sometimes deviate from those derived from the literature. In many cases this is due to a difference in data collection temperature, but small differences in the amount of solvent and the exact composition of the solvent may also slightly affect the unit cell parameters.
- Due to the large molecules and the high solvent content, protein crystals are soft and in general poor scatterers. Combined with the non-Bragg scattering from the surrounding mother liquor and the sample holder, it results in large background, which must be subtracted manually.
- Working on laboratory sources relatively large amount of sample and/or measuring time is needed.

4.2 Experimental

4.2.1 Growing of lysozyme crystals

Tetragonal lysozyme crystals were grown using the vapour diffusion technique. Lyophilized lysozyme from chicken egg white, *Gallus gallus* (95 % Sigma Aldrich L-6876), was dissolved in 0.02 M sodium acetate and the protein concentration was determined to 13.8 mg ml^{-1} using ε_{280} of 37470 M^{-1}cm^{-1}. The protein was equilibrated in a sitting drop against a reservoir of 0.1 M HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid) and 0.8 M potassium sodium tartrate ($KNaC_4H_4O_6$) at pH 7.5 (cf. Crystal Screen HR-2-110 no. 29, Hampton Research). The drops were formed of 200 µl protein solution mixed with 200 µl reservoir solution. The crystals were grown during three weeks, Figure 4.2

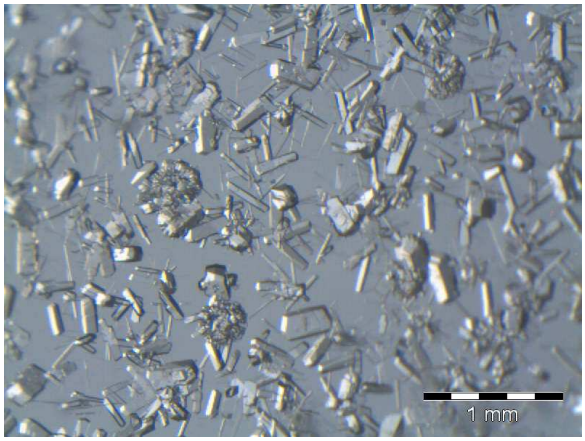


Figure 4.2: Crystals of tetragonal lysozyme.

4.2.2 X-ray powder diffraction

After crystallization excess mother liquor was removed and the wetted crystals were gently crushed to a concentrated suspension of protein powder. The diffraction experiments have been carried out at an in-house Huber G670 diffractometer using Cu $K\alpha_1$ radiation ($\lambda = 1.54059$ Å). The samples were mounted in a specially designed sample holder (description of the design is reported in Section 5.2.1), which contains appr. 150 µl of the powder suspension (corresponding to appr. 5 mg protein). The measurements were performed on rotating samples at room temperature. Data were collected at room temperature for 15×4 h. No indication of time decay was observed. The background was subtracted in the program *PROTPOW* [1] by manually picking between 10 and 20 background points, fit-

[1] Available from http://www.xray.kemi.dtu.dk/English/Computer%20Programs/PROTPOW.aspx

ting to a spline function and subtracting. The manual procedure is necessary to distinguish background from overlapping signal.

4.3 Calculation of XRPD patterns from PDB coordinates

A powder diffraction pattern of tetragonal lysozyme was calculated from the PDB-coordinates of a lysozyme structure, which was refined from a crystal originating from the same batch as the lysozyme powder (Hartmann, 2008; Hartmann *et al.*, 2010). The resulting structure factors, before and after corrections, were entered into *PROTPOW* for powder diffraction pattern calculations. All corrections were carried out on the calculated patterns.

4.3.1 Optimization of unit cell parameters

It was necessary to optimize the unit cell parameters to get fully comparable powder patterns. This could be due to differences in data collection temperatures (lysozyme powder data were collected at room temperature while the single crystal data set was collected at 100 K) or variations in solvent concentrations. Optimization of the unit cell parameters was done by a full-pattern profile fit. Optimized parameters describing the peak width (FWHM) and peak shape of a pseudo-Voigt function are included in the pattern fit by *PROTPOW*. A powder pattern was calculated in *PROTPOW* from the PDB coordinate file using the optimized unit cell and peak profile parameters, Figure 4.3. In general the peak positions fit well, but without taking into account the disordered bulk-solvent, the intensities fit poorly with the measured intensities, in particular at low 2θ.

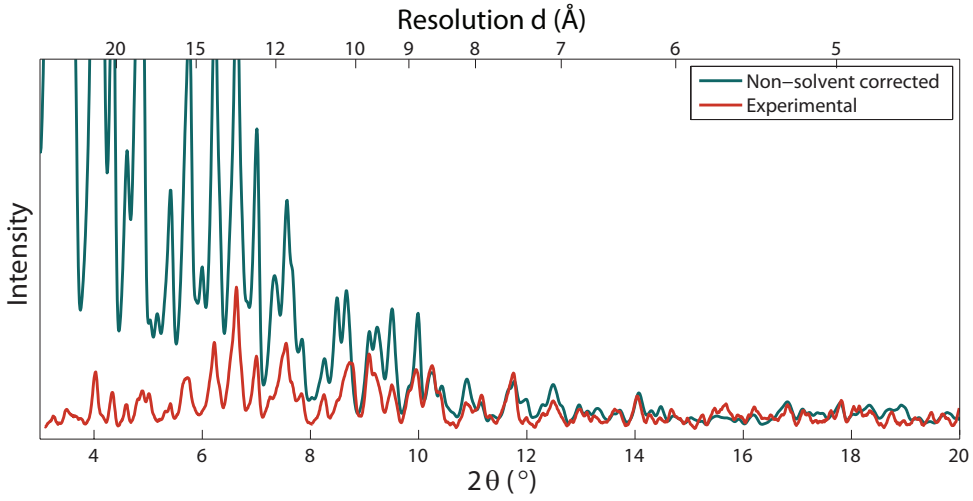


Figure 4.3: Calculated powder diffraction pattern for lysozyme before solvent correction (blue). An experimental powder pattern (red) is included for comparison.

4.3.2 Correction of disordered solvent

The contribution from the solvent was calculated in *PHENIX* (Adams *et al.*, 2010) using the flat bulk-solvent model (Phillips, 1980) where a constant level of electron density is assumed in the voids between the macromolecule. The bulk-solvent model takes two parameters k_{sol} defining the level of electron density and B_{sol} defining the steepness of the border between the solvent and the macromolecular regions. The parameters are refined in today's single-crystal X-ray structure refinement software, but in case no other information is available following average values can be used $k_{\mathrm{sol}} = 0.35$ e Å^{-3} and $B_{\mathrm{sol}} = 46$ Å^2 (Fokine & Urzhumtsev, 2002). The contribution from the solvent is added to the calculated structure factors. The improved fit from the powder pattern based on calculated structure factors is seen on Figure 4.4.

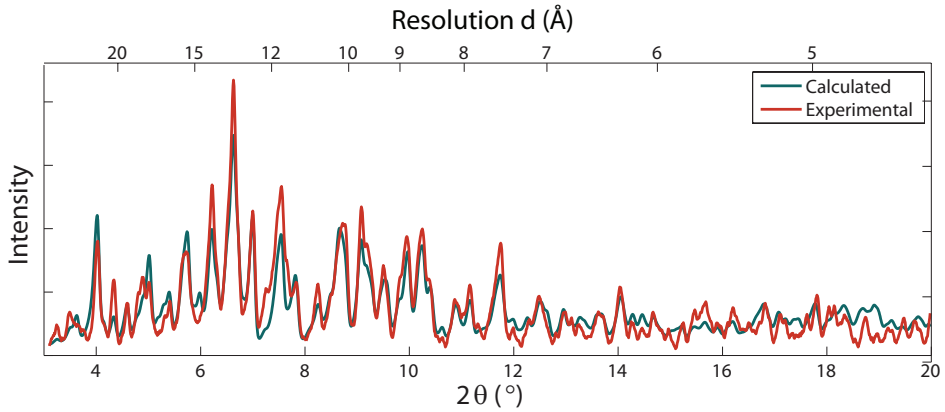


Figure 4.4: Calculated (blue) and experimental (red) powder diffraction patterns for lysozyme. The calculated pattern is corrected for bulk-solvent.

4.3.3 Other factors

As the conventional Lorentz correction tends to infinity when approaching $2\theta = 0^\circ$ a revised geometrical intensity correction factor has been derived. Applying this improved geometrical factor gave a minor, but, significant improvement to the fit in the low angle region.

Protein structures available from the PDB do not generally contain hydrogen positions. Including calculated hydrogen positions did not improve the overall fit and was abandoned. Further details are reported in Paper I (Hartmann *et al.*, 2010).

4.4 Discussion

The combined effects from applying background subtraction, optimizing the unit cell parameters and bulk-solvent correction are seen in Figure 4.4. The agreement between the experimental and the corrected calculated patterns are generally very good. Even in the region where $2\theta = 12 - 20^{\circ}$ the agreement is good, meaning that for well diffracting crystals it is possible to detect the detailed diffraction pattern down to resolutions of 4 Å on in-house equipment. However, there still remain minor disagreements between the calculated and experimental patterns, which have not been accounted for. Some of the factors which may influence on the experimental data are:

- Preferred crystal orientation which causes a non-random distribution of reflecting planes.
- Insufficient or excess background correction. The overlap problem may result in a background subtraction which also includes some of the signal.
- Despite solvent corrections the structural model may still be deficient.

4.5 Concluding remarks

Using a standard in-house X-ray powder diffractometer, powder patterns up to 4 Å resolution were collected from well-diffracting powder samples of lysozyme. Calculated powder patterns were generated from PDB coordinate data. By performing corrections for background, unit cell parameters, disordered bulk-solvent and geometrical factors the agreement between the experimental and the calculated patterns is generally very good. The key factor to bring the calculated patterns in agreement with the observed patterns is the bulk-solvent correction.

Chapter 5

Practical aspects and sample handling for in-house XRPD studies of protein

A procedure by which protein XRPD-patterns can be collected on in-house equipment at reasonable time scales has been developed. For this purpose two sample holders have been designed and described in this chapter. The final design furthermore supports the sample holder to be adaptable for XAS experiments, and has been used for the insulin study described in Chapter 7 and 8. A descriptive note of the final sample holder design has been published in Paper III (Frankær *et al.*, 2011) Appendix A.3.

5.1 Introduction

Obtaining powder diffraction patterns with sufficient signal-to-noise ratios (S/N) from protein powders on in-house X-ray sources requires careful design of the sample holders and collimators. Some general issues concerning the design are mentioned here:

- As protein crystals contain 30–80 % of disordered solvent they must be prevented from drying out by keeping them in their the mother liquors in sealed sample chambers.

- The geometry of the sample chamber: To expose as many crystallites to the X-ray beam as possible a large surface area is required. However, the sample volumen must be minimized in order to minimize the sample amount.

- The sample should be easy to mount in the holder.
- Rotation should be applied to avoid sedimentation on the windows and walls of the sample chamber and so avoid multiple layers of powder.
- The background should be minimized by optimizing the collimation of the radiation using pinholes and slits.
- Cryo-cooling can improve the signal, but will require a thorough optimization of the cryo-protectant. Furthermore, successful cryo-coolong depends on the geometry of the sample holder in order to maintain a laminar flow of cold nitrogen.

5.2 Sample holder designs

Two sample holders designs were developed for the Huber G670 diffractometer equipped with a flat powder specimen holder, which allows the samples to be rotated around an axis χ orthogonal to the detector. To collect as much signal as possible on the in-house X-ray sources, the general principle behind the designs is to support a thin film of protein crystallites with a relatively large surface area. Drawings and photographs of the sample holders are seen in Figure 5.1. These designs, however, obstruct cooling of the samples as turbulence from the cryo-stream will form when hitting the flat holder.

For capillary measurements a lead shield was designed to offer a simple and efficient way to ensure reproducible collimation and significant background reduction. A photograph is seen in Figure 5.2. For an effective sample cooling a laminar cryo-flow can be maintained using capillaries.

Approaching very small sample volumes a loop, similar to those used for single crystal protein crystallography, has been tested with 1.0 µl of protein powder suspension on a micro-source diffractometer. Dimensions and sample volumes of all sample holders are listed in Table 5.1.

Table 5.1: Dimensions of the sample chambers of the different sample holders. All sample chambers are cylindrical (except the loop which is assumed spherical), hence the dimensions are given by a height and a diameter.

	Model 1	Model 2a	Model 2b	Capillary	Loop
Diameter (mm)	10	4.5	3.0	0.5	1
Height/thickness (mm)	1.5–2.0	1.0	1.0	25	–
Sample chamber volume (µl)	120–150	16	7	5	1
Powder suspension needed (µl)	120–150	16	7	15	1
Rotation around	χ	χ	χ	φ	φ
Cooling	Ineffective	Ineffective	Ineffective	Possible	Required

5.2.1 XRPD sample holder, model 1

The first sample holder model, encases a thin layer of protein powder wetted by its mother liquor. The protein powder film is approximately 0.5–1.0 mm thick allowing the diffractometer to work in transmission mode. It consists of a rubber O-ring between two iron discs with circular windows (6.5 mm in diameter), Figure 5.1(*a*) and 5.1(*b*). Due to the Guinier geometry of the powder diffractometer the profile of the inner edge has been cut in a 45° angle to prevent shielding of the incoming X-rays. The discs are plated with a thin layer of gold for increased corrosion resistance. Two pieces of mylar foil are placed between each disc and the O-ring so they cover the windows. The mylar windows have low absorption of X-rays, the scattering from the material gives an amorphous and relatively low contribution to the background and the strong and non-permeable material is able to endure the pressure that occurs when the sample holder is assembled. Three screws and appertaining nuts keep the device assembled. Each disc is supplied with an engraved circular furrow matching the diameter of the O-ring and when assembled the O-ring presses against the foil and provides for a tight fit.

5.2.2 XRPD sample holder, model 2

The improved sample holder consists of a 1.0 mm thick polycarbonate plate with a hole, (either 4.5 mm or 3.0 mm in diameter, creating sample chamber volumes of 16 and 7 µl, and referred to as model 2a and 2b, respectively). Further dimensions are specified in Figure 5.1(*c*) and 5.1(*d*). Metal-free polycarbonate was chosen to avoid fluorescence signal from the sample holder which for XAS experiments would conflict with the fluorescence signal from metalloproteins. Small holes have been drilled so the holder fits the cryostat (Oxford Instruments, ITC503) available at the 811 XAS beam line at the synchrotron at MAX-lab, Lund, Sweden. The sample holder takes a small cylindrical volume of protein powder wetted by its mother liquor, and with a sample thickness of 1.0 mm this holder still allows the diffractometer to work in transmission mode. The sample is sealed by two pieces of mylar foil which are glued to the polycarbonate plate with either epoxy glue or double-sided adhesive tape. For powder diffraction a lead plate with a pinhole matching the diameter of the sample chamber is glued on top of the assembled holder and serves for efficient shielding of the background scattering. The inner edge profile of the lead plate has been cut in a 45° angle to prevent shielding of the incoming X-ray beam. The pinhole plate fits exactly into the flat powder specimen holder at the Huber G670 diffractometer, allowing the sample to be centred in the beam. The relatively compact shape of the sample chamber, compared to the first model, eases mounting of the sample.

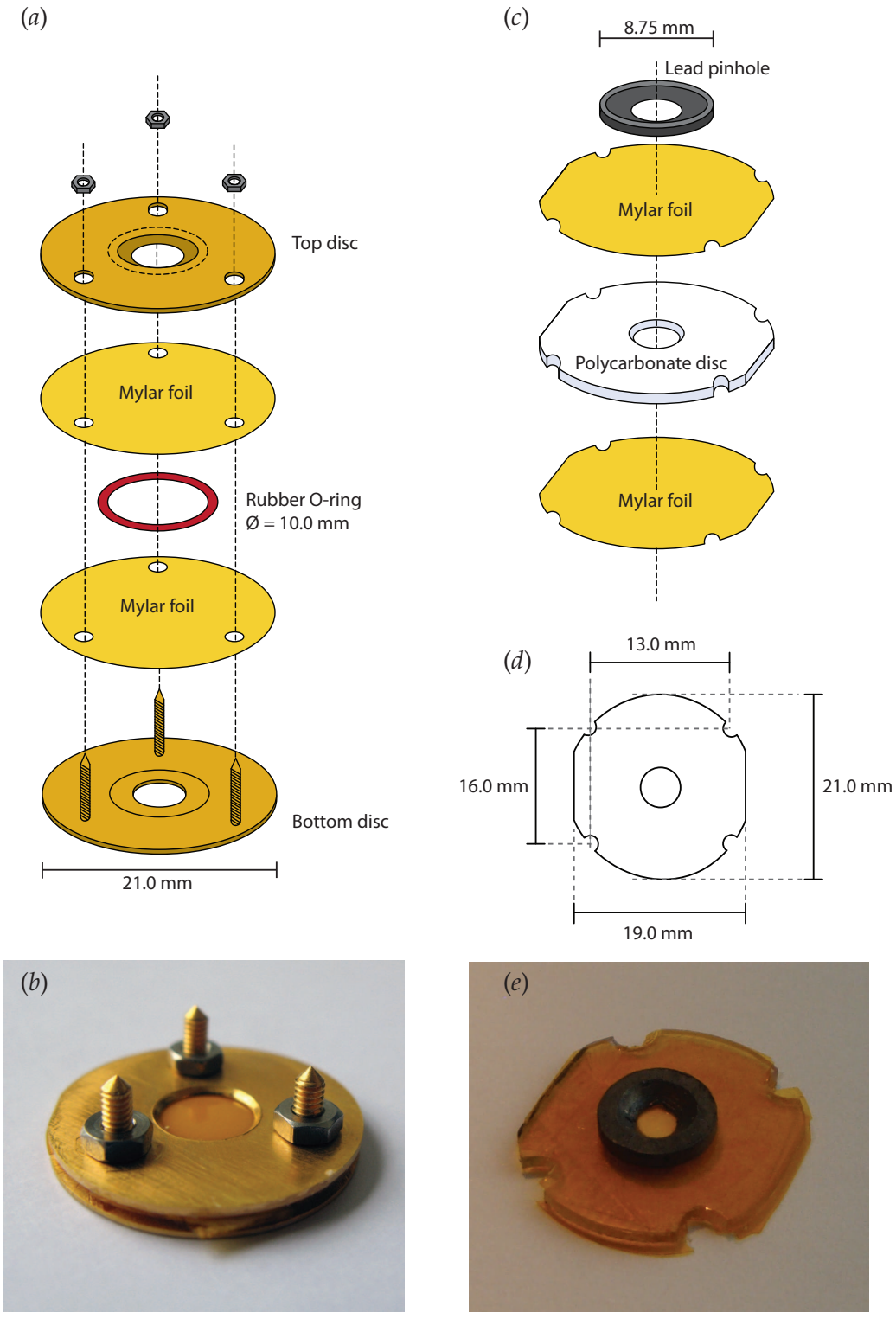


Figure 5.1: Drawings and photographs XRPD sample holders for Huber G670 diffractometer. The dimensions are specified on the drawings. (*a*) and (*b*) model 1. (*c*), (*d*) and (*e*) model 2.

5.2.3 Lead shield for capillaries

The lead shield was made from a 1.0 mm thick lead plate (16 mm $\times$ 30 mm) with a slit (1 mm $\times$ 11 mm) and was designed to be mounted on the capillary sample holder at the Huber G670 diffractometer, Figure 5.2. When mounted on the sample holder the slit is centred 5 mm from of the capillary and efficiently shields the scattering from the background.

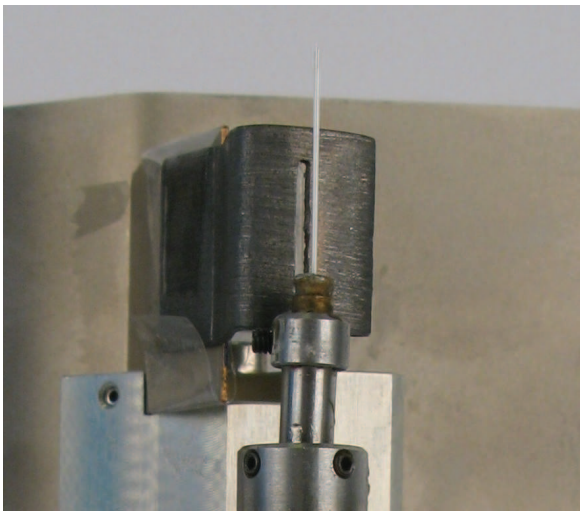


Figure 5.2: Capillary mounted on a goniometer with the lead shield seen in the background.

5.3 Experimental

The sample holders were tested with powder of tetragonal lysozyme crystals which were grown according to the procedure described by Section 4.2.1. After crystallization excess mother liquor was removed and the wetted crystals were gently crushed to a concentrated suspension of protein powder.

5.3.1 Sample preparation and data collection in sample holders

The sample holders were loaded by filling the partly assembled holder with protein powder suspensions. Amounts of concentrated protein suspension are listed in Table 5.1. The top parts of the sample holders were carefully put in place in order to avoid trapping air bubbles in the sample chambers. The sample holders were mounted on the flat powder specimen holder at the Huber G670 diffractometer and rotation around χ was applied to the samples. Data were collected for 4 h at room temperature using Cu $K\alpha_1$ radiation (λ=1.5406 Å).

5.3.2 Sample preparation and data collection in capillaries

Glass capillaries with an inner diameter of 0.5 mm were packed with 15 µl suspension of protein powder in its mother liquor by centrifugation (4×3 min at 7500 rpm). The packing efficiency after centrifugation increased 2–3 times resulting in sample volumes of 3–5 µl densely packed powder. For low temperature experiments with cryo-protection glycerol was added to the suspensions resulting in final concentrations of 15–30 vol.%. After sedimentation the capillaries contained 15–25 mm high columns of densely packed crystals. The capillaries were then sealed with epoxy glue and mounted on a goniometer head fitting the capillary powder sample holder at the diffractometer. The sample holder was further equipped with the lead pinhole for capillaries. Rotation around an axis φ perpendicular to the beam and parallel to the capillary was applied, and data were collected for 4 h using Cu $K\alpha_1$ radiation both at room temperature and at 120 K using a cryo-stream of nitrogen (Oxford Cryosystems).

5.3.3 Sample preparation and data collection in loops

A commercial loop (Hampton Research) with a diameter of 1.0 mm was loaded with approximately 1.0 µl of protein powder suspension. The loop was mounted on a diffractometer with micro-source beam: GeniX beam delivery system 50 kV, 50 W generating Cu $K\alpha$-radiation (λ=1.5418 Å). The beam is focussed on the sample with a spot size of 230 µm. The samples were cooled to 130 K to prevent them from drying out and data were collected for 15 and 45 minutes on a Rigaku R-AXIS IV++ 2D-detector in a single frame covering $\Delta\varphi$ of 90°. Distance to detector was 400 mm. The 2-dimensional powder patterns, (Figure 5.3) were integrated to one dimension using *FIT2D* ((Hammersley, 1998; Hammersley *et al.*, 1996). The beam stop was masked out and the tilting of the detector and the position of the direct beam was optimized using the signal from the hexagonal ice as an internal standard.

5.4 Results

5.4.1 Performance of the sample holders

The background intensity levels and the signal-to-noise ratios, which are somehow interrelated, were estimated by looking at the raw lysozyme powder patterns presented in Figure 5.4. The patterns for all X-ray tube source experiments are directly comparable, as the data collection times are similar, whereas for the micro-source, the integrated signal was simply scaled to the model 2a sample holder pattern. Note that the peaks for the cold samples are slightly shifted towards higher 2θ-values as a result of the contracted unit cell. Separating background from diffraction is however not straightforward due to the heavy

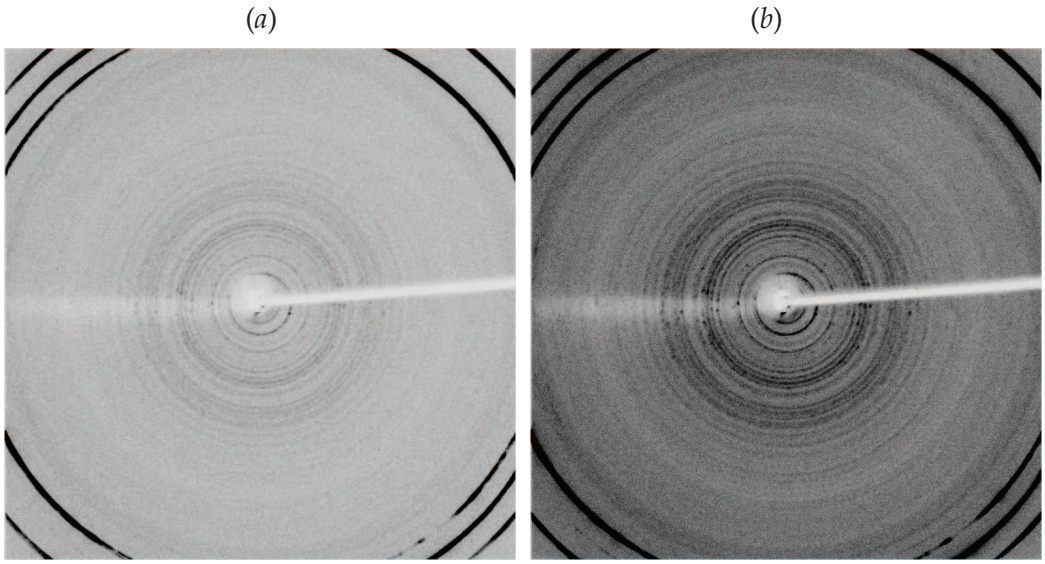


Figure 5.3: X-ray powder diffraction rings collected at 130 K on an in-house X-ray micro-source diffractometer for 1 µl samples of lysozyme powder. Data collection times of (*a*) 15 min and (*b*) 45 min. The intensities are scaled relatively to each other.

peak overlap. An estimation of the background intensity was therefore done by scaling a calculated pattern to the experimental. Prior to the calculation of a powder pattern unit cell and peak shape parameters were optimized using the procedure described in Section 4.3. The background level, BG, was evaluated as a sum of intensities I in the 2θ-range 2–20° by the following equation:

$$BG = \frac{\sum I_{\text{Background}}}{\sum \left(I_{\text{Bragg}} + I_{\text{Background}}\right)} = \frac{\sum \left(I_{\text{Total}} + I_{\text{Bragg}}\right)}{\sum I_{\text{Total}}} \tag{5.1}$$

where I_{Bragg} was evaluated from a calculated pattern that was scaled to the experimental using *PROTPOW*. The background levels are presented in Table 5.2.

Table 5.2: The performance of the different sample holders evaluated as the background level, BG.

	Temperature (K)	BG (%)
Model 1	293	90.2
Model 2a	293	87.4
Model 2b	293	92.3
Capillary	293	88.3
Capillary	120	89.3
Loop (15 min)	130	90.5
Loop (45 min)	130	82.2

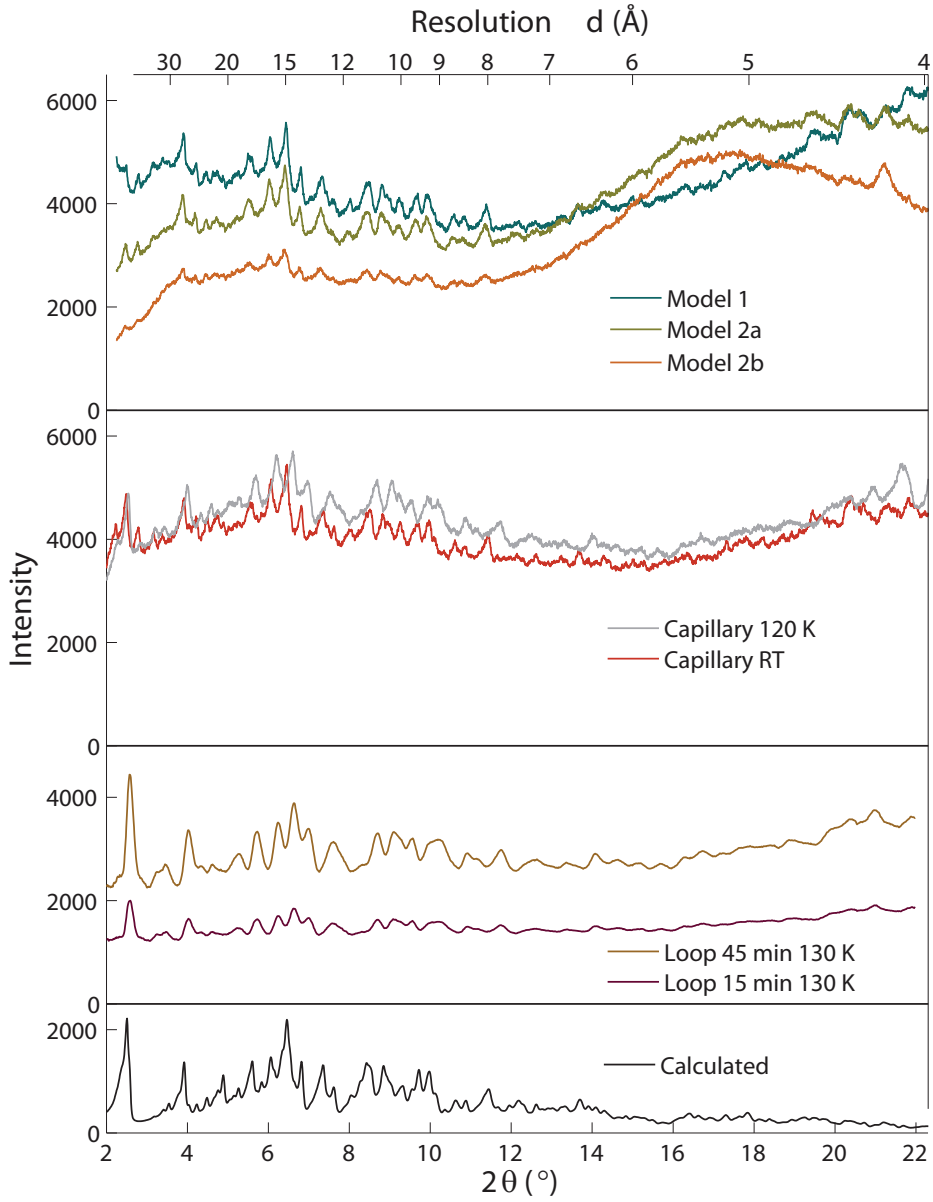


Figure 5.4: XRPD patterns of tetragonal crystals of hen egg white lysozyme prior to background subtraction. The data collection time was 4 h for sample holder model 1, 2a, 2b and capillaries, and 15 or 45 min for the loop. The powder patterns for the loop has been converted to Cu $K\alpha_1$ radiation ($2\theta = 1.5406$ Å) and the intensities scaled relative to the other patterns. A pattern has been calculated using the unit cell and peak shape parameters, optimized from the model 2a powder pattern, and is shown for comparison (bottom).

With the improved sample holder (model 2) the minimum amount of sample has been reduced by an order of magnitude resulting in 7–16 µl, which is similar to the amounts required for capillary samples. Furthermore the new design eased the sample preparation. Rotation of the sample was shown to be important for these relatively thick samples. Without rotation the crystallites may sediment on the windows and walls of the sample chamber and result in multiple layers of powder.

The S/N-ratios and the background levels depend on the collimation of the beam. Pinholes, which are integrated in the flat specimen sample holder designs, ensure that the beam only hits the homogeneous part of the sample and reduce the background, in particular at low 2θ-angles. After an optimization of the beam stop position, it is possible to measure reflections down to 2.0° on the Huber G670 diffractometer. For capillaries the lead slit seen in Figure 5.2 has reduced the background level by a factor of two, resulting in S/N-ratios similar to the flat specimen holders. In general the background level, which primarily originates from the mother liquor, seems to have reached a constant level around 90 % independent on which sample holder is used, and may lead to the conclusion that the upper limit of what can be measured on an in-house powder diffractometer using a standard X-ray tube as source has been reached, provided that other post-crystallization treatments such as dehydration of the crystals as described by (Heras & Martin, 2005) are not taken into account.

5.4.2 Comments on micro-source

A micro-source diffractometer, generating a smaller beam with a higher photon flux (2×10^8) ph/s), permits that powder patterns with S/N comparable with the previous results can be obtained in only 15 minutes for 1.0 µl droplets of samples supported by a loop and flash-cooled. The high collimation of the micro-source beam obviates the need for pinholes and a lower background was observed (82.2 %). With the 2D-detector reflections occurring at 2θ-angles down to 1.0° was measured. Collecting the entire Debye-Scherrer rings gave a better average and smoother patterns. At the micro-source diffractometer the beam is focused on the sample, which results in less resolved peaks compared to the dedicated powder diffractometer, where the beam is focused on the detector.

5.4.3 Comments on cryo-protecting protein powder

For this system, no radiation damage was observed for the room-temperature samples. Cooling the crystals would, in principle, reduce the thermal motions and thereby increase the crystal order and improve the S/N. However, the cooling process may also induce stress and disorder in the crystal, making it diffucult to predict whether the net diffraction quality will increase or decrease (Juers & Matthews, 2004). Comparing the two cap-

illary patterns a slightly higher background was observed for the cold sample. It is possible to protect the crystallites with a cryo-protectant (Jenner *et al.*, 2007), wherefore a pre-experiment was carried out where different concentrations of glycerol was added to the powder suspension. The crystallites were partly dissolved wherefore more sample was required, and a decreased S/N with increasing glycerol concentrations was observed. Most likely the glycerol slightly etches the surface of crystals, which is not noticeable when treating large single crystals. However, for a powder sample the crystallite size is significantly reduced or even dissolved, requiring that the cryo-protection condition are even better optimized for powder samples than for single crystals.

5.5 Concluding remarks

Optimizing the collimation by pinholes the sample holder design has been improved to operate with sample volumes down to 7 µl which is comparable with capillaries, but easier to prepare. The improved sample holder design has been successfully applied to samples used for both powder diffraction and XAS experiments. Capillary samples are generally more cumbersome to prepare, but may with suitable collimation produce useful diffraction patterns for tests prior to synchrotron experiments. Using an in-house micro-source diffractometer an acceptable powder pattern can be collected in 15 minutes from 1.0 µl of powder suspension. Working with in-house X-ray sources cooling is not necessary to avoid radiation damage and it does not significantly improve the scattered intensity at the low angles of interest. Adding cryo-protectant requires thorough optimization, and may complicate the experiments unnecessarily. However, cooling is needed for samples mounted in loops to prevent them from drying out.

Chapter 6

Introduction to the insulin structure and its conformations

A general introduction to insulin with emphasis on its structure is given in this chapter. Divalent cations play a special role in the formation and stabilization of insulin hexamers, which are believed to be involved in the storage and release mechanisms of insulin into the blood stream. The insulin structure is therefore described with special emphasis on the hexamer assembly and its metal ion binding sites, which are studied with XRD and XAS and reported in the following two chapters.

6.1 Introduction

Since the biological function of insulin was discovered by Banting & Best (1922), insulin has been extensively studied. The protein is mainly known for its regulation of the sugar level in the blood and is given to patients suffering from diabetes, which is a still increasing global public health problem. The World Health Organization (WHO) estimates that 346 million people worldwide have diabetes, and that the number of diabetes death will double between 2005 and 2030 (World Health Organization, 2012).

6.2 Biosynthesis of insulin

Insulin is a protein hormone produced in the endocrine β-cells located in the islets of Langerhans in the pancreas. Insulin is first synthesized as pre-proinsulin containing three

domains (chain B, C and A) and an N-terminal signal peptide of 24 amino acid residues (see Figure 6.1). The signal peptide is required for transporting the peptide across the membrane of the endoplasmatic reticulum (Steiner *et al.*, 1967), and is later removed by a post-translational process. The resulting proinsulin is transported into the Golgi apparatus, where it assembles into hexamers in presence of zinc (Emdin *et al.*, 1980; Huang & Arvan, 1995), and stored in vesicles to protect the protein from enzymatic degradation. The β-cells contain high concentrations of zinc, which indicates the importance of this element for the storage and secretion of insulin. Two endopeptidases are responsible for a proteolytic cleavage, which converts proinsulin into insulin by removing the C-chain (Davidson *et al.*, 1988). As the solubility of insulin is lower compared to the proinsulin, the insulin hexamers start to form micro-crystals in the β-granules. This compact organization allows the insulin to be tightly packed and furthermore serves for protection against proteolytic degradation (Halban *et al.*, 1987). When insulin is secreted into the blood stream vessels by exocytosis (Howell, 1984) the zinc concentration is lowered and the micro-crystals experience a jump in pH from approx. 5.5 inside the storage vesicles to 7.4 in the blood. This have a destabilizing effect on the hexamer which rapidly disintegrates into monomers, see Section 6.6.

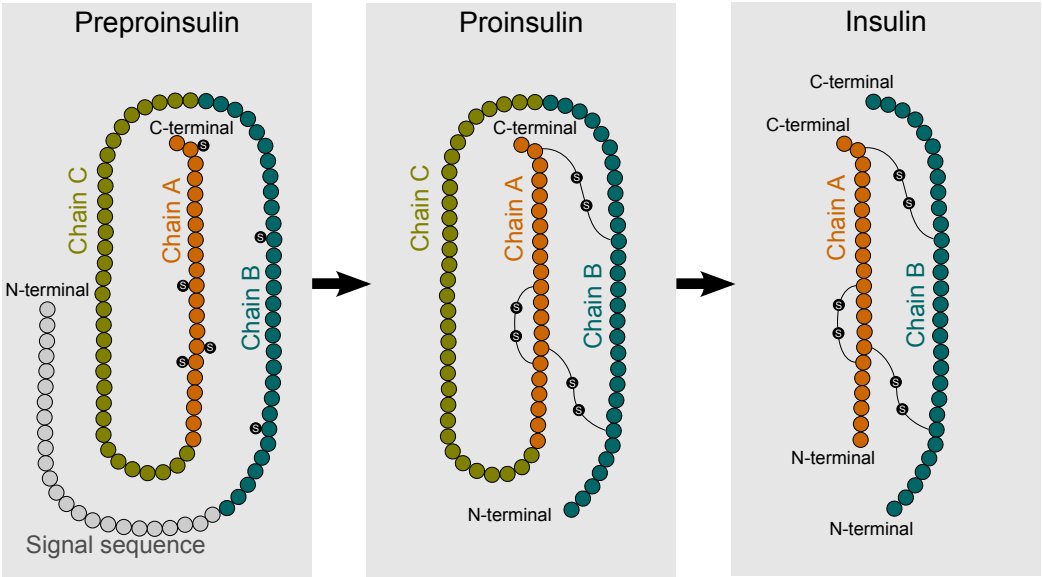


Figure 6.1: Schematic overview of the biosynthesis of insulin. Proinsulin is formed from preproinsulin by cleavage of the N-terminal signal peptide. Insulin is formed from proinsulin by proteolytic cleavage of the C-chain. The residues are represented by circles, and the disulfur bridges which connect the insulin A- and B-chain are included.

6.3 The insulin structure

After the insulin amino acid sequence was determined (Ryle *et al.*, 1955), the first three-dimensional structure of porcine insulin was solved from X-ray diffraction to 2.8 Å resolution by Adams *et al.* (1969). Since then, the number of insulin structures has increased and today more than 150 insulin structures from XRD and NMR are deposited in the Protein Data Bank originating from different source organisms, mutants and in different conformations.

6.3.1 The insulin monomer

The insulin molecule is built up of a 21-residue A-chain and a 30-residue B-chain which are connected by two interchain disulfide bonds (CysA7–CysB7 and CysA20–CysB19) and an intrachain disulfide bond in the A-chain (CysA6–CysA11) (Blundell *et al.*, 1972). The conservation of the amino acid sequence is high for mammals. The sequence of porcine insulin differs from bovine insulin at two residues A8 and A10 (respectively Thr and Ile in porcine → Ala and Val in bovine) and from human insulin by one residue B30 (Ala in porcine → Thr in human). The sequence of bovine insulin is shown in Figure 6.2(*a*).

The tertiary structure of the molecule is maintained by a hydrophobic interior and the overall folding is shown in Figure 6.2(*b*). The folding of the A-chain is shown in Figure 6.3(*a*) and consists of two antiparallel α-helical segments (A1–A8 and A13–A19) con-

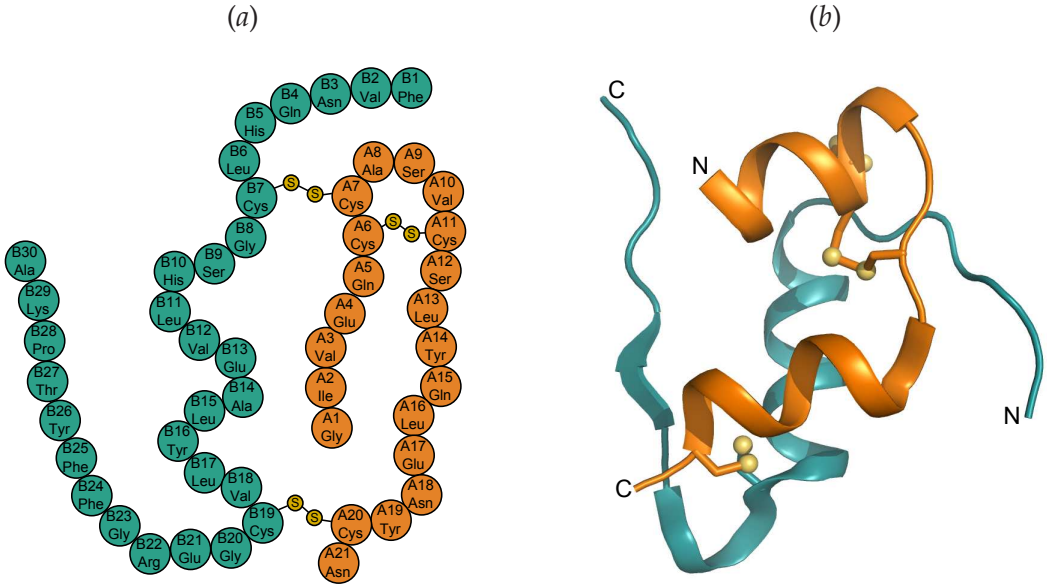


Figure 6.2: (*a*) Schematic drawing including the sequence of bovine insulin. (*b*) The tertiary folding of insulin including the disulfide bondings (yellow spheres). The figure is based on the atomic coordinates from the Zn T_6 insulin presented in Chapter 7. The A-chain is shown in orange and the B-chain in blue.

nected by a loop (A9–A12). The intrachain disulfide bond A6–A11 stabilizes the loop. The N- and C-terminal are brought together and further stabilized by van der Waals contact between IleA2 and TyrA19. The B-chain consists of a central α-helical segment (B9-B19) flanked by two extended regions, see Figure 6.3(*b*). The two interchain disulfide bonds linking to the A-chain are located at each end of the central α-helix. In presence of phenolic derivates and high lyotropic anion concentrations the N-terminal extended part (B1–B8) can adopt different conformations where it is partly or entirely curled up in an α-helix resulting in the formation of a 19-residues long helix. These conformations are further described in Section 6.5.

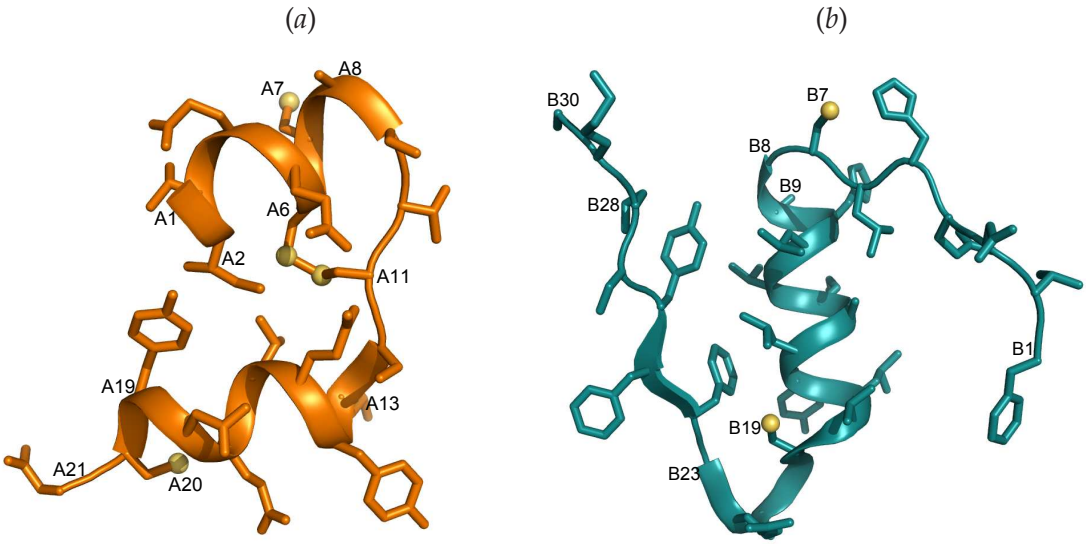


Figure 6.3: The tertiary folding of the insulin A-chain (*a*) and B-chain (*b*). The figure is based on the atomic coordinates from the Zn T_6 insulin presented in Chapter 7.

6.3.2 The insulin dimer

The surface of the insulin monomer has two extensive non-polar regions, which play an essential role for the contacts formed under aggregation of the monomers into both dimers and hexamers. The two insulin molecules in the dimer are related by a local approximate two-fold rotation axis. The strongest monomer-monomer contact involves residues from the central B-chain α-helix and from the B-chain C-terminal, as seen in Figure 6.4. The dimer is further stabilized by a hydrogen bonding network, where an antiparallel β-pleated sheet structure is formed involving the residues B23–B28 from each molecule. The two PheB25 residues in the centre of the β-sheet are located directly opposite to each other. In order to avoid clash these residues are forced to adopt an asymmetric conformation (Figure 6.4), which breaks down ideal two-fold symmetry. The smallest possible asymmetric unit is thus constituted by the dimer. Generally, the region of contacts between

the monomers has a high population of aromatic residues (PheB24, PheB25 and TyrB26), which stabilize the dimer further by hydrophobic interactions and forms a non-polar core in the centre of the dimer (Blundell *et al.*, 1972).

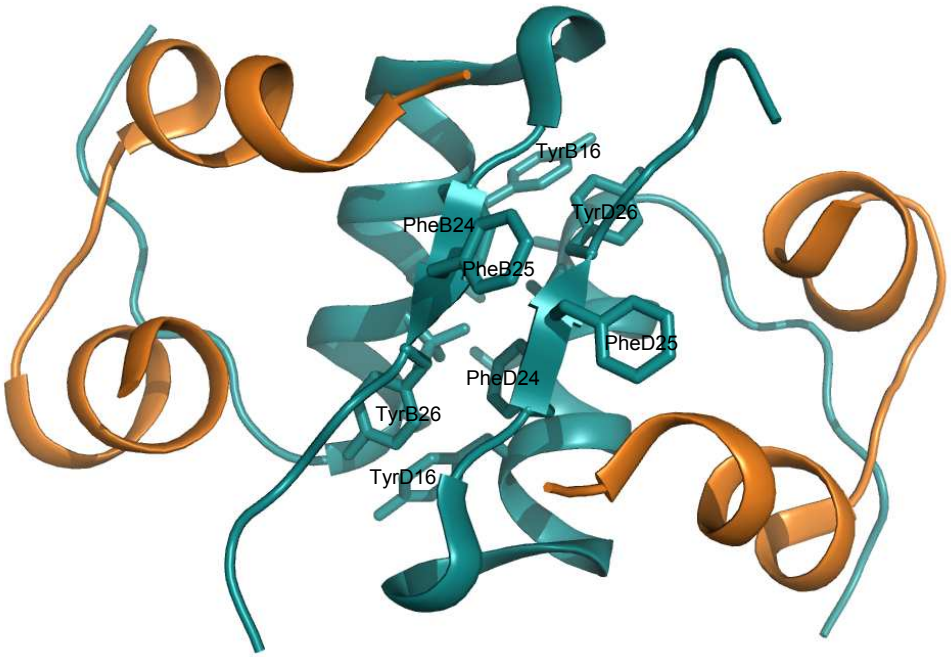


Figure 6.4: The dimer and its interactions between the aromatic residues in the antiparallel β-sheet and around a local approximate two-fold axis. The figure is based on the atomic coordinates from the Zn T_6 insulin presented in Chapter 7.

6.3.3 The insulin hexamer

Insulin forms hexamers in presence of salt and divalent metal ions (typically Zn^{2+}). The six insulin molecules compose a hexamer with an overall torus shape of approx. 35×50 Å. A schematic representation of the hexamer is seen in Figure 6.5.

The hexamer can be regarded as a trimer of dimers where the 3-fold axis (Q) generates the hexamer from the dimers. In rhombohedral crystals the three-fold axis is coinciding with the crystallografic c-axis. A pseudo two-fold axis P relates the two monomers of the dimer, as described in Section 6.3.2. As a consequence of the three-fold symmetry, another pseudo two-fold rotation axis P' is generated perpendicular to Q and 60° to P. This axis relates one insulin molecule to another one in its neighbour dimer. The contact surface between two dimers at P', includes non-polar residues from the central B-chain α-helix, the B-chain N-terminal and the A-chain interhelical region, as seen in Figure 6.6.

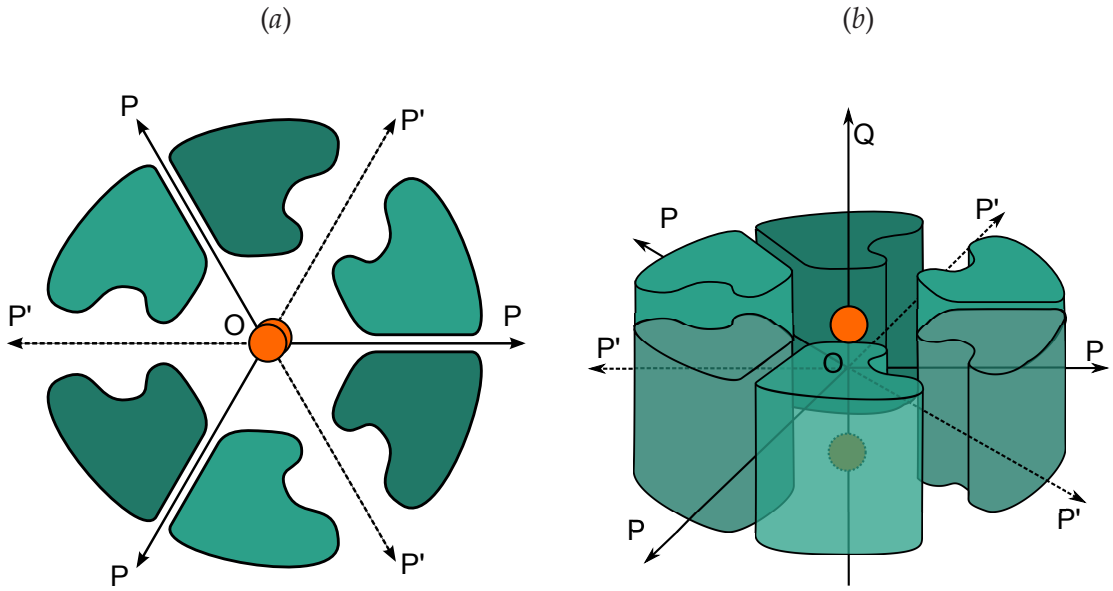


Figure 6.5: Schematic representation of the insulin hexamer, (*a*) seen perpendicular to the three-fold axis, (*b*) seen from the side. The pseudo two-fold axis P relates the two monomers of the dimer and P′ relates two dimers to each other. The two metal sites located on the three-fold axis Q are separated by approx. 15.8 Å and shown in orange.

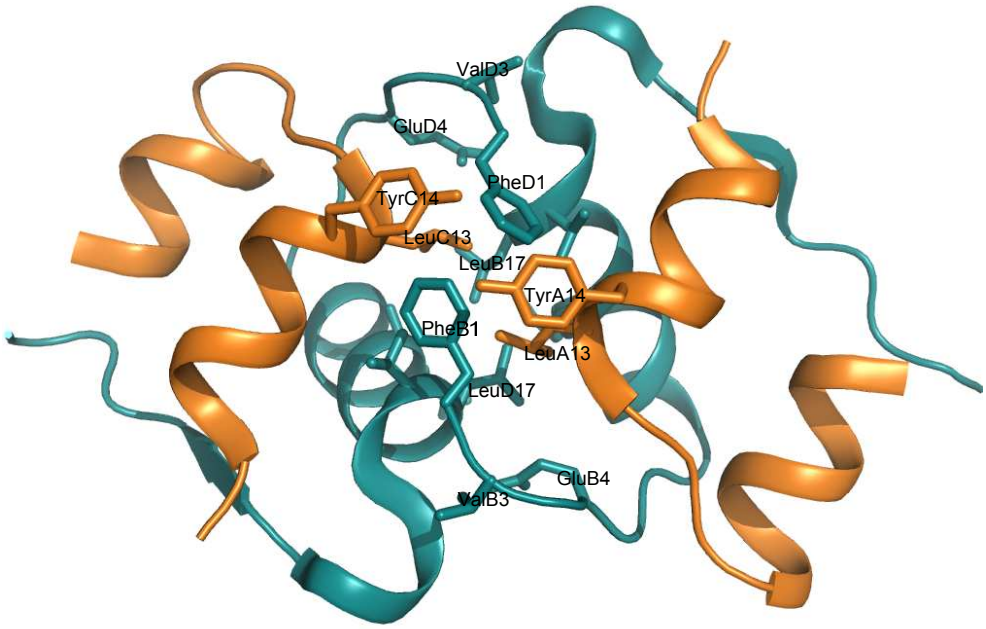


Figure 6.6: The surface interactions between two adjacent dimers (i.e. around the local two-fold axis P′). The figure is based on the atomic coordinates from the Zn T_6 insulin presented in Chapter 7.

6.4 Metal ion binding in the hexamer

The hexamer is a zinc insulin complex, where each hexamer binds two Zn^{2+}-ions. The zinc ions are located on the three-fold rotation axis approx. 8 Å above and below the centre of the hexamer, Baker *et al.* (1988). Each zinc ion is coordinated to the $N^{\epsilon 2}$ atoms of three symmetry related HisB10 residues, and can be either octahedrally- or tetrahedrally coordinated, Figure 6.7. Octahedral coordination is fulfilled by further coordination of three water molecules, and tetrahedral coordination is fulfilled by coordination of one lyotropic anion (typically chloride), which is also located on the three-fold symmetry axis.

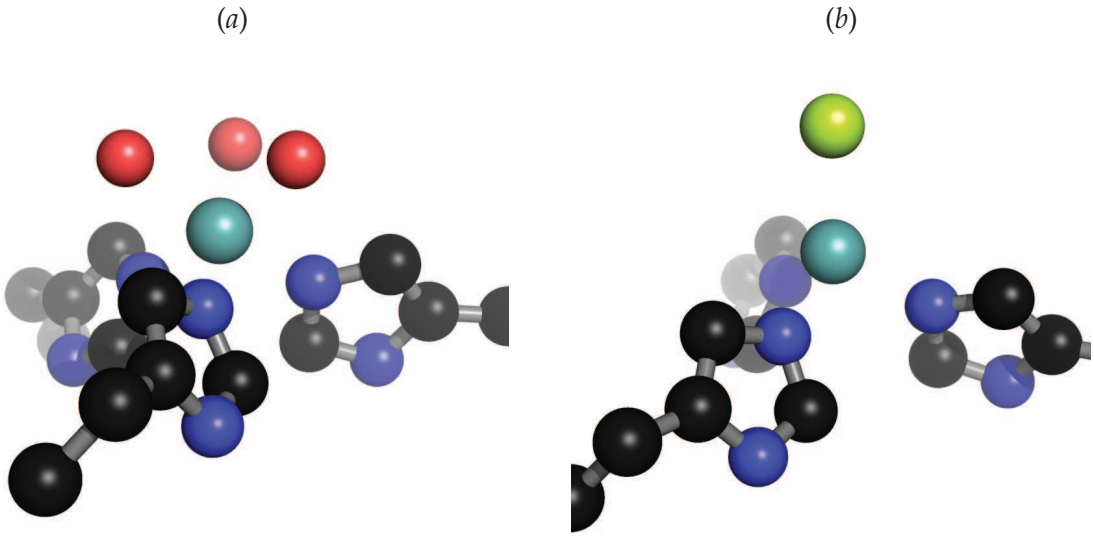


Figure 6.7: Possible coordination geometries of the hexameric insulin zinc sites. Both ions are located on a three-fold symmetry axis going through the hexamer and coordinate to three symmetry-related histidine $N^{\epsilon 2}$ atoms. (*a*) Octahedral coordination is fulfilled by further coordination of three water molecules. (*b*) Tetrahedral coordination is fulfilled at both Zn-sites by coordination of one lyotropic anion, (e.g. chloride), which is also located on the three-fold symmetry axis. The atoms are coloured according to their type, carbon is black, nitrogen is blue, oxygen is red, chlorine is green and zinc in light blue.

Hexamers containing more than two zinc ions have also been observed (Emdin *et al.*, 1980). At higher zinc concentrations additional off-axial binding sites are formed in which zinc is tetrahedrally coordinated to $N^{\epsilon 2}$ atoms of a HisB10 residue in another conformation, a HisB5 residue from the neighbour dimer and two anions or water molecules (Bentley *et al.*, 1976; Smith *et al.*, 1984).

The presence of divalent metal has for a long time been known to be crucial for the hexamer formation (Scott, 1934; Scott & Fisher, 1935). Zn^{2+} has the best affinity to stabilize the hexamer, but hexamers can be well stabilized by other divalent metal ions, e.g. Mn^{2+}, Fe^{2+}, Co^{2+}, Ni^{2+}, Cu^{2+} and Cd^{2+} (Schlichtkrull, 1956). These metals have unoccupied 3 *d*-orbitals and are thus better suited for spectroscopy, which have been used to study confor-

mational changes of hexameric insulin (Roy *et al.*, 1989; Brader *et al.*, 1992; Bloom *et al.*, 1997; Brader *et al.*, 1997; Kadima, 1999). The crystal structure of human insulin has recently been solved for hexamer complexes containing manganese (PDB-entry 2R34) (Sreekanth *et al.*, 2009), cobalt (PDB-entry 1M5A) (Nicholson *et al.*, 2006), nickel (PDB-entries 3EXX, 2R36 and 3INC) (Prugovečki *et al.*, 2009; Sreekanth *et al.*, 2009), copper (PDB-entries 3TT8 and 3IR0) and strontium (PDB-entry 3ILG).

6.5 Hexamer conformations

In its hexameric form insulin is an allosteric complex and exists in different conformations. The conformations are distinguished by the secondary structure of the first eight residues of the B-chain, and are presented in Figure 6.8. The residues B1 to B8 can either adopt an extended conformation, which is denoted as tensed (T), Figure 6.8(*a*), or adopt a relaxed conformation (R), Figure 6.8(*c*), where the central B-chain α-helix is continuous from B1 to B19 (Kaarsholm *et al.*, 1989). A third conformation in which the continuous α-helix is limited to residues B4 to B19, while the residues B1 to B3 adopt extended conformation has been reported as the frayed R-conformation, denoted R^f (Ciszak *et al.*, 1995), see Figure 6.8(*b*). A fourth conformation, where the PheB1 residue is extended thereby forming a conformation between pure R- and R^f-conformation, has been observed for one out of six molecules in a hexamer in monoclinic R_6 insulin structures (Smith *et al.*, 2000). This conformation will be denoted R^g in the following.

The insulin hexamer has been observed in three different conformations, as seen in Figure 6.9. Using the T/R-nomenclature introduced by Kaarsholm *et al.* (1989), where a

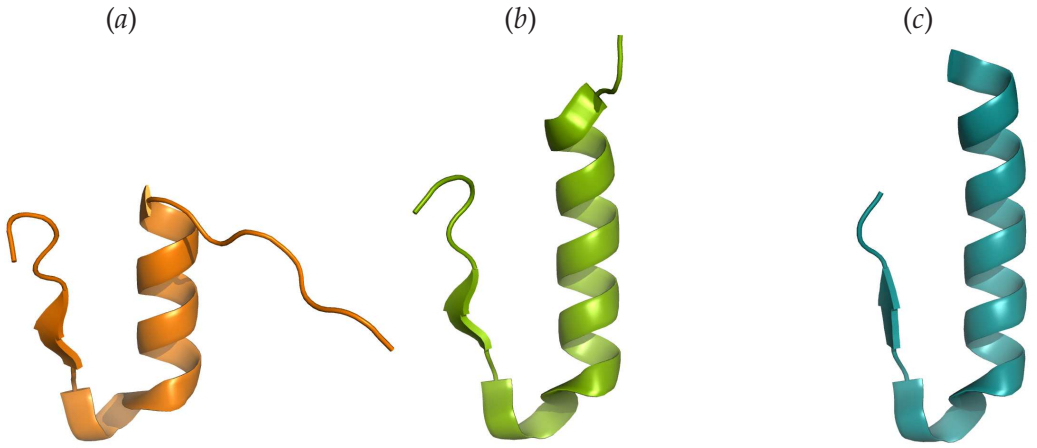


Figure 6.8: Different conformations of the insulin B-chain. (*a*) Tensed state (T), where the residues B1–B8 adopt an extended structure. (*b*) Relaxed and frayed (R^f), where the residues B1–B3 are extended and B4–B19 form an α-helix. (*c*) Relaxed (R), where the α-helical structure is adopted throughout the entire N-terminal (B1–B19).

number in the subscript refers to the number of monomers present in each conformation, the T_6 denotes the conformation where all six monomers adopt a T-conformation. This conformation was earlier reported as 2Zn-insulin due to the two zinc ions bound in the structure (Adams *et al.*, 1969; Blundell *et al.*, 1971; Dodson *et al.*, 1979; Baker *et al.*, 1988).

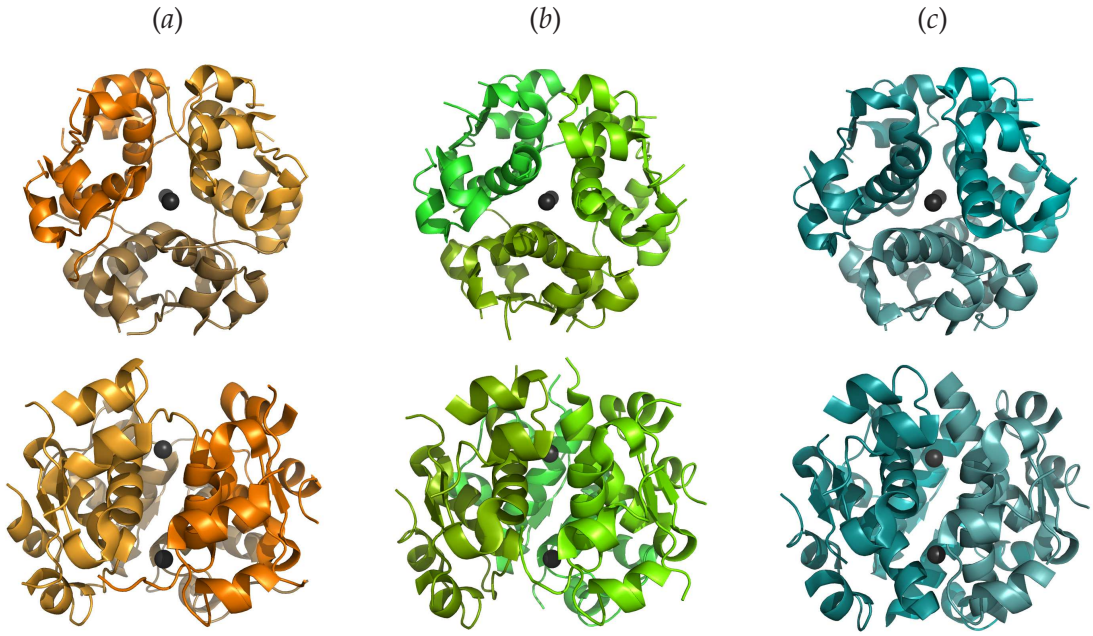


Figure 6.9: Different conformations of the insulin hexamer viewed along the 3-fold axis (top) and perpendicular to the 3-fold axis (bottom). (*a*) T_6, where both molecules in the dimer adopt T-conformation. (*b*) $T_3R_3^f$, where one insulin molecule in the dimer adopts T-conformation and the other R^f-conformation. (*c*) R_6, where both molecules in the dimer adopt R-conformation. The figure is based on atomic coordinates from PDB-entries 1MSO (Smith *et al.*, 2003), 1TRZ (Ciszak & Smith, 1994) and 1EV3 (Smith *et al.*, 2000), respectively.

It is often useful to consider the hexamer as a dimer of trimers, where each trimer contain one zinc ion. In the T_3R_3 conformation one of the two trimers will adopt the helical R-conformation while the other will remain tensed. The conformation of B1 to B3 in the three monomers constituting in the R_3 trimer were later concluded to be frayed, why this conformation was renamed $T_3R_3^f$ (Ciszak *et al.*, 1995). However, the T_3R_3 notation can still be used when the conformation of the B1–B3 segment is unspecified. The $T_3R_3^f$ conformation was first seen in the crystal structure of 4Zn-insulin (Bentley *et al.*, 1976), which had been known from chemical analyses to contain four zinc ions per hexamer (Schlichtkrull, 1958). The $T_3R_3^f$ conformation has later been observed in hexamers which only contain two zinc ions (Ciszak & Smith, 1994), wherefore it seemed reasonable to name the hexamer species by their conformation rather than their zinc content.

The R_6 conformation contains six insulin monomers all adopting the R-conformation, and was first observed in the structure determined from monoclinic crystal form (Dere-

wenda *et al.*, 1989). The structure has later been solved from rhombohedral crystal of R_6 insulin (Smith & Dodson, 1992; Smith *et al.*, 2000).

Spectroscopic and structural studies of the allosteric property of the insulin hexamer, demonstrate that it exists in a state of dynamic equilibrium, following the reaction scheme:

$$T_6 \rightleftharpoons T_3R_3 \rightleftharpoons R_6$$

The T_6 state is obtained at low concentration of anions and in the absence of phenol derivates (Baker *et al.*, 1988; Smith *et al.*, 2003). Shifting the equilibrium from the T_6 state to the T_3R_3 state is done by addition of inorganic lyotropic anions such as chloride or thiocyanide in high concentrations (Whittingham *et al.*, 1995). These anions bind at the zinc site and stabilize helix formation of the residues B4 to B8 in the R_3^f-trimer (Ciszak *et al.*, 1995). The increased lyotropic anion concentration can only stabilize helix formation in one of the two trimers.

Besides the presence of lyotropic anions, shifting the equilibrium to the R_6 state requires addition of phenol or phenol derivates. Addition of phenol-like molecules creates two hydrophobic phenol binding sites, located at each dimer-dimer interface, and fully stabilizes the helix formation of the first eight N-terminal residues of the B-chains in both trimers (Derewenda *et al.*, 1989; Smith & Dodson, 1992). The phenolic compounds bind in these hydrophobic pockets, and form hydrogen bonds between the hydroxyl group of the phenol and the carbonyl oxygen of CysA6 and the backbone nitrogen atom of CysA11.

Rhombohedral crystals were first obtained by Abel (1926), and were found to belong to the space group $R3$ by X-ray crystallographic measurements (Crowfoot, 1938). Today rhombohedral crystals have been obtained from all three conformations and crystal structures belonging to the space group $R3$ have been solved. As the trimers undergo a conformation from T_3 to R_3, the length of the hexamer slightly increases along the three-fold symmetry axis which is coincident with the crystallographic c-axis. A systematic increase in the unit cell c-axis from 34.0 Å in T_6, to 37.6 Å in $T_3R_3^f$ and 40.4 Å in R_6 has been observed and has earlier been used for characterization of 2Zn-insulin and 4Zn-insulin (Harding *et al.*, 1966; Bentley *et al.*, 1978).

6.6 Hexamer assembly and stability

The formation of hexamers has been proposed to occur by the formation of a dimer from two monomers, which then interacts with a second dimer forming a tetramer and finally completed by interaction with a third dimer. As the associative forces between the dimers are not as strong as between the two monomers of the dimer, divalent metal ions are required to stabilize the hexamer assembly. The stability formed by binding Zn^{2+} in these

sites may overcome the repulsion forces of the close lying GluB13 residues in the hexamer centre, and complete the hexamer formation. (Hill *et al.*, 1991)

The different allosteric states are thought to be involved in the zinc uptake and release, which respectively stabilizes and destabilizes the hexamer, and thereby play an important role in the storage and release of insulin.

In the T_6 state, the extended conformation of the N-terminal B-chains forms a wide cleft leading into the zinc ions and exposes them to the surrounding solvent, Figure 6.9(*a*). The T_3 sites provides room for an octahedral coordination geometry where three water molecules fulfil the coordination. The zinc ions exchange on a time scale of seconds, and may thereby easily diffuse from the hexamer. As the N-terminal B-chain helices are stabilized the channel leading into the zinc site is closed and the zinc ions become shielded from its surroundings. Also the coordination of zinc is affected in these R_3 sites. R_3 zinc sites are more compact, compared to those found in T_3 sites due to the three helical N-termini are closer to each other. Thereby R_3 zinc sites only leaves room for tetrahedral coordination. The stability of the hexamer conformations has been verified from spectroscopic studies (Rahuel-Clermont *et al.*, 1997), which showed that R_6 is the most stable of the three conformations and that T_3R_3 is more stable than T_6. This high stability of R_6 led to increased interest in co-crystallizing insulin with phenol derivates, e.g. *m*-cresol, resorcinol (Smith *et al.*, 2000), 4'-hydroxyacetanilide (Smith & Ciszak, 1994), 4-hydroxybenzamid (Smith *et al.*, 1996), methylparaben (Whittingham *et al.*, 1995) to produce R_6 crystals for long term acting insulin formulations.

The stability of the hexamer is also pH dependent. A stabilizing hydrogen bonding network involving six glutamate residues (GluB13) is observed in the centre cavity of the hexamer. At pH around 6.0 the glutamate residues are hydrogen bonded to each other. When pH exceeds 7.0 the carboxylic acids are deprotonated, and the hexamer assembly becomes destabilized by repulsive forces, from the six negative charges in the core of the hexamer. At the pH of blood 7.4, these interactions are increasingly disfavoured and coordination of zinc ions are essential to overcome their repulsive forces. When insulin are secreted into the blood stream, the dilution of zinc ions in the serum means that the zinc ion coordination can not be maintained, and the hexamer rapidly disintegrates into monomers.

The important role of GluB13 was studied by mutating the glutamate to glutamine. This recombinant insulin mutant was shown to form stable hexamers in absence of Zn-ions at pH 7.0 (Bentley *et al.*, 1992), and the hexamer from this mutant would presumably not be able to disintegrate into monomers upon secretion of insulin from the storage vesicles into the blood stream.

Chapter 7

XRD and XAS studies of T_6, T_3R_3 and R_6 bovine insulin with Zn

The work presented in this chapter illustrates the complementarity between X-ray diffraction and X-ray absorption spectroscopy. These techniques have been used to study hexameric bovine insulin in the three different conformations T_6, T_3R_3 and R_6, with emphasis on the zinc ion binding sites. Based on this work a manuscript has been written and submitted.

7.1 Introduction

In hexameric insulin crystallized in the rhombohedral space group $R3$, unstructured electron densities around the zinc ions bound in T_3 binding sites are often observed among the structures deposited in the Protein Data Bank, in particular for the low resolution structures (Harding *et al.*, 2010). These electron densities may be modelled with atoms at chemically unreasonable bond distances. For instance a tetrahedrally coordinated Zn–O distance of 1.18 Å is observed in a T_3 binding site in the structure of an insulin analogue to 2.0 Å resolution (PDB-entry 6INS) (Derewenda *et al.*, 1991). A bovine insulin structure to 2.56 Å resolution (PDB-entry 2ZP6) contains T_3-sites where both tetrahedral and octahedral coordination geometry of zinc are observed, and Zn–O distances up to 2.63 Å.

Possible explanations for these unstructured densities include radiation damage, which is a general problem for single crystal XRD, due to the high energy load per volume of sample. Metal containing proteins, in particular, suffer from radiation damage as pro-

tein crystals with heavy atoms generally absorb more radiation and thereby have shorter lifetimes (Holton, 2009). Furthermore, artefacts from data truncation (so-called Fourier ripples) occurring around heavy scatters or artefacts generated from the three-fold symmetry axis, on which the zinc ions are located, may affect the electron densities at these zinc sites. Such problems are avoided by X-ray absorption spectroscopy, which also provides a lower energy load per sample. This technique may thus be better suited for studying the zinc coordination in hexameric insulin.

The structures of all three conformations of bovine insulin have been solved by single crystal XRD, and corresponding samples have been studied by XAS. The single crystal structures of T_3R_3 and R_6 are new, as for bovine insulin only the structure of T_6-conformation has hitherto been solved (Smith *et al.*, 2005). The developed procedure for verification of crystal forms by XRPD has been applied to characterize the insulin samples before the XAS experiments. Despite the fact that insulin is a well studied protein only one EXAFS study of the zinc sites in a single conformation of hexameric insulin has been reported (Bordas *et al.*, 1983), and modelling performed using only single scattering paths. This study has been updated using both XANES and EXAFS, including full multiple scattering. Finally, the XAS results are compared with what has been observed in small molecules with similar coordination geometries, as well as in other insulin structures to evaluate the accuracy of XAS.

7.2 Experimental

Lyophilized insulin from bovine pancreas, *Bos taurus*, was purchased from Sigma Aldrich (I-5500). Other reagents used were stock chemicals, analytical grade.

7.2.1 Growth of single-insulin crystals

After several crystallization trials it was found that full control of the concentration of Zn^{2+} was crucial for the growths of single crystals large enough for XRD analysis. As the insulin contains approximately 0.5 %(w/w) zinc, the zinc ions were removed from the insulin before use. Metal-free insulin was prepared by chelating using the method of Coffman & Dunn (1988). 20 mg insulin was suspended in 1 mL ultra-pure water (18.2 MΩ·cm, milli-Q, Millipore) and dissolved by raising the pH to 10.5 with 1 M NaOH. 360 mg Chelex100 resin was added and pH adjusted to 8.7 with 1 M HCl and the suspension was stirred for 2 h while maintaining the pH at 8.7. In additional 150 mg Chelex100 resin was added to ensure complete metal chelation and the resin was separated from the insulin solution by syringe filtration (0.22 μm). Following the solution was diluted with milli-Q water. All concentrations were determined by UV-vis spectrophotometry using an extinction coeffi-

cient at 280 nm, ε_{280}, of 5960 $M^{-1}cm^{-1}$, which was estimated by *ProtParam* (Gasteiger *et al.*, 1995).

For the growth of T_6 and R_6 insulin crystals, the vapour diffusion technique was used. 2 µl of a solution containing 5 mg ml^{-1} insulin adjusted to pH 6.8 using HCl(aq) was mixed with 2 µl reservoir solution and equilibrated in a hanging drop against 1 ml reservoir with a composition given in Table 7.1. After 7 days 200–500 µm large crystals were observed, Figure 7.1.

Table 7.1: Reservoir composition for single crystal growth of T_6 and R_6 insulin using the vapour diffusion technique.

	T_6 insulin	R_6 insulin
Sodium citrate (M)	0.05	0.12
Zinc acetate (mM)	0.5	17
Acetone (vol.%)	15	22
m-cresol[†] (vol.%)	–	0.5
NaCl (M)	–	1.67
pH	7.2	8.4

[†] *m*-cresol was dissolved in the acetone before addition.

T_3R_3 insulin was crystallized using the batch crystallization technique. The crystallization procedure is described in Section 7.2.2.

Single crystals with dimensions of 300–500 µm were soaked in cryo-protective solution containing 15 vol.% PEG 400, 25 vol.% glycerol and 60 vol.% mother liquor (Smith *et al.*, 2001, 2003) and stored in liquid nitrogen until diffraction analysis.

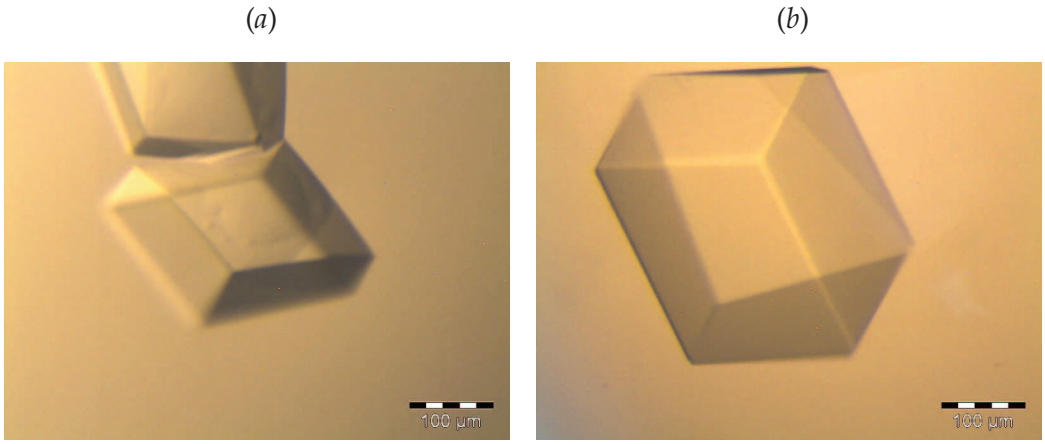


Figure 7.1: Single crystals of (*a*) T_6 and (*b*) R_6 insulin.

7.2.2 Preparation of insulin micro-crystals

Micro-crystal samples of the three conformations were used for X-ray absorption spectroscopy measurements. All three conformations were prepared using the batch crystallization technique. The compositions of the crystallization solutions are given in Table 7.2 as the concentration for each compound in the final crystallization solution. The crystallization compounds were added in the order given in the table.

Table 7.2: Crystallization conditions for preparation of micro-crystals of T_6 T_3R_3 and R_6 insulin by batch crystallization. The compositions are given as the concentration for each compound in the final crystallization solution.

	T_6 insulin	T_3R_3 insulin	R_6 insulin
Insulin (mg ml^{-1})	5.0	6.5	5.0
HCl (M)	0.01	0.01	0.01
Zinc acetate (mM)	5.0	7.0	7.5
Sodium citrate (M)	0.02	0.05	0.05
Acetone (vol.%)	15	15	20
KSCN (M)	–	0.2	–
m-cresol[†] (vol.%)	–	–	0.5
NaCl (M)	–	–	1.0
pH	6.0	6.4	8.2

[†] *m*-cresol was dissolved in the acetone before addition.

T_6 insulin microcrystals were obtained using a slightly modified procedure of the method by Schlichtkrull (1956). The crystallization solution was warmed to 323 K and pH was adjusted to 6.0 with HCl. The solution was allowed to cool to room temperature and microcrystals were precipitated during 24 h under agitation. The micro-crystals were saved by filtration using an Ultra-free-MC centrifuge filter (Amicon/Millipore), with a pore size of 0.22 μm. Excess zinc was removed by re-crystallization under analogue conditions, but with exclusion of zinc acetate, resulting in stochiometric ratios between insulin and zinc.

T_3R_3 insulin crystals were grown with thiocyanate using a slightly modified procedure of the method by Whittingham *et al.* (1995). After the crystallization solution (Table 7.2) was prepared pH was subsequently raised to 8.0 with 1 M NaOH. The solution was warmed to 323 K to ensure complete solution and maintained at approximately 323 K while pH was lowered to 6.4 with 1 M HCl. The solution was placed in an incubator at 323 K and cooled to room temperature over a period of four days. The crystallization resulted in a few 500 μm large crystals and several smaller rhomboid shaped crystals. The crystals were recovered by filtration on a centrifuge filter. A large crystal was saved for single crystal XRD while the rest were gently smashed to a powder. The powder was washed three times with a solution of 0.04 M sodium citrate, 12 vol.% acetone and 0.2 M KSCN, pH = 6.0 to eliminate excess zinc in the sample.

R_6 insulin was co-crystallized with the phenol derivative *m*-cresol using a crystallization solution with an analogue composition to that from the procedure by Smith *et al.* (2000), see Table 7.2. pH was raised to 9.4 with 1 M NaOH to ensure complete solution. After adjusting the final pH to 8.2 the solution was placed in an incubator at 277 K and after two days micro-crystals were observed. The crystals were separated by filtration on a centrifuge filter and excess zinc was eliminated by washing three times with a zinc free solution containing 0.035 M sodium citrate, 23 vol.% acetone, 0.3 % *m*-cresol and 1 M NaCl.

For all three samples the correct crystal form was verified with XRPD.

7.2.3 Single crystal diffraction

Single crystal diffraction data were collected at MAX-lab, Lund, Sweden, on beam line 911-2, MAX-II using a MarResearch CCD detector. Data collection and data processing statistics for all three crystals are summarized in Table 7.3. The data were processed and scaled using the programs *XDS* and *XSCALE* (Kabsch, 1993). As twinning previously has been observed for hexameric insulin crystals (Knudsen, 2011), the processed data were tested for twinning using the on-line merohedral twin detector algorithm by Padilla & Yeates (2003). The structures were refined using the programs *REFMAC5* (Murshudov *et al.*, 1997) included in the *CCP4* suite (Collaborative Computational Project, Number 4, 1994) and *PHENIX* (Adams *et al.*, 2010). The model building and editing were carried out using *WinCoot* (Emsley & Cowtan, 2004), whereas ordered solvent was modelled using the auto insertion procedure in *PHENIX*, followed by a manual inspection to remove insignificant water molecules. For each data set a random 5 % subset of all reflections was reserved for validation. The structures were validated using *PROCHECK* (Laskowski *et al.*, 1993), *WHAT_CHECK* (Hooft *et al.*, 1996) and the *STructure ANalysis* server *STAN* (Kleywegt & Jones, 1996). Refinement and validation statistics are listed in Table 7.4.

The residue numbers are assigned according to the nomenclature used by Smith *et al.* (2000, 2001), in which the decimal portion refers to the monomer number in the dimers. In the deposited coordinates A, C, E, G, I, K, M, O, Q, S, U, W, Y, 1, 3 and 5 refer to A-chains and B, D, F, H, J, L, N, P, R, T, V, X, Z, 2, 4 and 6 to B-chains. For instance, PheB1.3 designates the first phenylalanine residue in the B-chain of insulin monomer 3, which corresponds to PheF1 in the deposited PDB-file.

T_6 insulin

The structure was refined using the peptide chain from the 1.0 Å structure of human T_6 insulin, PDB-entry 1MSO (Smith *et al.*, 2003). Two zinc atoms were inserted, and the residues specific for bovine insulin were mutated. The N-terminal of the B.2-chain (PheB1.2 to GlnB4.2) was remodelled and the side chains of residues ValB12.1, LeuB17.1, CysA11.2 and

Table 7.3: Data collection parameters and processing statistics. All data were collected at beam line 911-2, MAX II, MAX-lab, Lund, Sweden.

	T_6 insulin	T_3R_3 insulin	R_6 insulin
Data collection			
Wavelength (Å)	1.04002	1.03914	1.04002
Temperature (K)	100	100	100
No. of frames	180	200	180
Oscillation range (°)	1	0.5	1
Exposure time per frame (s)	8	10	10
Detector distance (mm)	80	70	120
Resolution collected (Å)	1.30	1.21	1.70
Data processing			
Resolution (Å)	20.78–1.40 (1.50–1.40)†	21.27–1.23 (1.30–1.23)†	28.88–1.80 (2.00–1.80) †
No. of reflections	86987 (16137)†	70690 (7380)†	365308 (96841)†
No. of unique reflections	15790 (2963)†	23527 (3060)†	64897 (17655)†
Redundancy	5.51 (5.45)†	3.00 (2.41)†	5.63 (5.49)†
Completeness (%)	98.1 (99.8)†	93.0 (79.0)†	97.5 (97.9)†
R_{sym} (%)‡	4.2 (60.8)†	4.6 (41.4)†	7.1 (50.1)†
$\langle I/\sigma(I)\rangle$	19.28 (2.62)†	14.68 (3.52)†	11.65 (2.95)†
Space group	$R3$	$R3$	$R3$
No. of molecules per asu	2	2	16
a (Å)	80.98	79.20	156.24
c (Å)	33.49	37.22	78.88
Average mosaicity (°)	0.72	0.37	0.42
Solvent content§ (%)	33.2	37.2	39.1

† Values in parentheses are for the outermost resolution shell.
‡ R_{sym} is defined as $\sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| \,/\, \sum_{hkl}\sum_i I_i(hkl)$, where $\langle I(hkl)\rangle$ is the mean intensity of a set of equivalent reflections.
§ Estimated by the program *Matthews* (Kantarjieff & Rupp, 2003).

LeuB17.2 were modelled in two alternating conformations. A total of 84 water molecules were inserted. Restrained refinement was carried out in *PHENIX* and hydrogen atoms were included. The two zinc ions were modelled by anisotropic refinement and the peptide chain by a combination of TLS refinement and isotropic refinement. The TLS domains were chosen in a way that each secondary structure element constitute a TLS domain resulting in seven domains in total: Residues 1–8 and 13–19 in A-chains, 9–18 in the B-chains and a group containing the residues 23–27 of two adjacent B-chains. Other were refined isotropically. Validation showed that only one residue SerA9.1 fell into the outlier region in the Ramachandran plot using the regions defined by Kleywegt & Jones (1996).

T_3R_3 insulin

The structure was refined using the peptide chain from the structure of human T_3R_3 insulin co-crystallized with 4-hydroxybenzamide, PDB-entry 1BEN (Smith *et al.*, 1996). Two zinc ions and one molecule of thiocyanate were included. The refinement procedure is

Table 7.4: Data-refinement and validation statistics.

	T_6 insulin	T_3R_3 insulin	R_6 insulin
Resolution cut (Å)	1.40	1.30	1.80
No. of atoms in the model			
Total non H-atoms	903	897	6816
Total H-atoms	787	750	0
Total protein atoms	1604	1544	6218
Total ordered water molecules	84	98	454
Total ligand atoms	0	3	136
Total Zn^{2+}-ions	2	2	8
***B*-factors**[§]			
Overall (Å^2)	32	27	41
Main chain (Å^2)	26	20	36
Side chains and water molecules (Å^2)	34	29	45
M^{2+}-ions (Å^2)	16	11	24
RMS deviation from ideal			
Bonds (Å)	0.013	0.008	0.007
Angles (°)	1.449	1.038	0.982
Ramachandran plot[‡]			
In core regions (%)	98.9	98.8	98.4
Outliers (%)	1.1	1.2	1.6
***R*-factors**[†]			
R	0.1938	0.1439	0.2088
R_{free}	0.2285	0.1794	0.2717

[§] The analysis of B-factors was done using *BAVERAGE* included in *CCP4* (Collaborative Computational Project, Number 4, 1994)

[†] R and $R_{\text{free}} = \sum ||F_{\text{obs}}| - |F_{\text{calc}}|| \,/\, \sum |F_{\text{obs}}|$, where F_{obs} and F_{calc} are the observed and calculated structure-factor amplitudes, respectively. R_{free} was calculated with a random 5 % subset of all reflections excluded from the refinement.

[‡] The definition of the Ramachandran plot regions according to Kleywegt & Jones (1996).

analogue to the one for T_6 insulin. The residues PheB1.2, LysB29.2 and AlaB30.2 were disordered and could not be modelled. The final structure contained 98 water molecules and the side chains of residues GlnA5.1, LeuA16.1, GluB13.1, and SerB9.2 were modelled with alternating conformations. Hydrogen atoms were included and anisotropic refinement of the atomic displacement factors was applied. Validation showed that only one residue SerA9.1 fell into the outlier region in the Ramachandran plot.

R_6 insulin

The structure was solved by molecular replacement using an entire insulin hexamer generated from the structure from human R_6 insulin, PDB-entry 1EV3 (Smith et al., 2000), as template. Molecular replacement was carried out using *Phaser* (McCoy *et al.*, 2007) which

found two hexamers in the asymmetric unit. According to the doubling of cell parameters when compared to the other conformations, 16 insulin molecules were expected to constitute the asymmetric unit. Another molecular replacement was performed on the difference map using an insulin dimer as template. This revealed another two dimers located close to the three fold rotation axis generating two hexamers with their symmetry-related dimers. The asymmetric unit was thus found to contain a total of 16 insulin molecules distributed as two hexamers (insulin momomers 1–6 and 7–12) and two dimers (monomers 13–14 and 15–16). A schematic overview of all insulin molecules in the R_6 conformation is shown in Figure 7.2(*a*).

The electron difference density maps clearly revealed the zinc sites and showed tetrahedral coordination, wherefore eight chloride ions were further included in the model. The residues specific for bovine insulin were changed and 16 molecules of m-cresol were inserted in the binding pocket located next to the CysA6 backbone oxygens. The two C-terminal residues LysB29 and AlaB30 in all monomers were disordered and were not modelled. Special attention was made to model the phenylalanine residues at the N-terminals of all B-chains, as they were found to adopt different conformations in the different chains, see Figure 7.2. Alternate conformations were assigned to the residues AsnB3.1, ValB18.1, ValB18.2, GlnB4.3, ValB18.4, ValB18.6, LeuB17.7, AsnB3.10, SerA9.12, ValB18.12, AsnB3.14, LeuB17.16 and ValB18.16 with occupancies refined in *PHENIX* (the minimum occupancy was found to be 0.33). 454 water molecules were included in the structure and the atomic displacement factors were refined by a combination of TLS refinement and isotropic refinement using seven TLS domains per dimer in analogy to the T_6-structure, resulting in 56 domains in total. All other atoms were refined isotropically. Validation showed that 11 ValB2 residues in monomers 3, 4, 5, 7, 8, 9, 11, 12, 13, 15 and 16 fell into the outlier region in the Ramachandran plot.

7.2.4 X-ray powder diffraction

The micro-crystal samples for XAFS analysis were analysed with X-ray powder diffraction (XRPD) before and after the XAFS experiment to verify the conformation and to monitor eventual degradations of the samples. Small amounts (5–7 µl) of wetted insulin microcrystals samples were mounted in a 1 mm thick sample holder for XRPD as well as XAS (Frankær *et al.*, 2011), as described in Chapter 5. Using the procedure to calculate XRPD patterns from PDB coordinates described in Chapter 4, the conformations were verified in 30 minutes on a Huber G670 diffractometer using Cu $K\alpha_1$ radiation ($\lambda = 1.5406$ Å). XRPD data was furthermore collected for R_6 insulin at 120 K and RT for 64 and 16 h, respectively. These R_6-samples were mounted in capillaries in order to apply a laminar flow of nitrogen for cooling (Oxford Cryosystems).

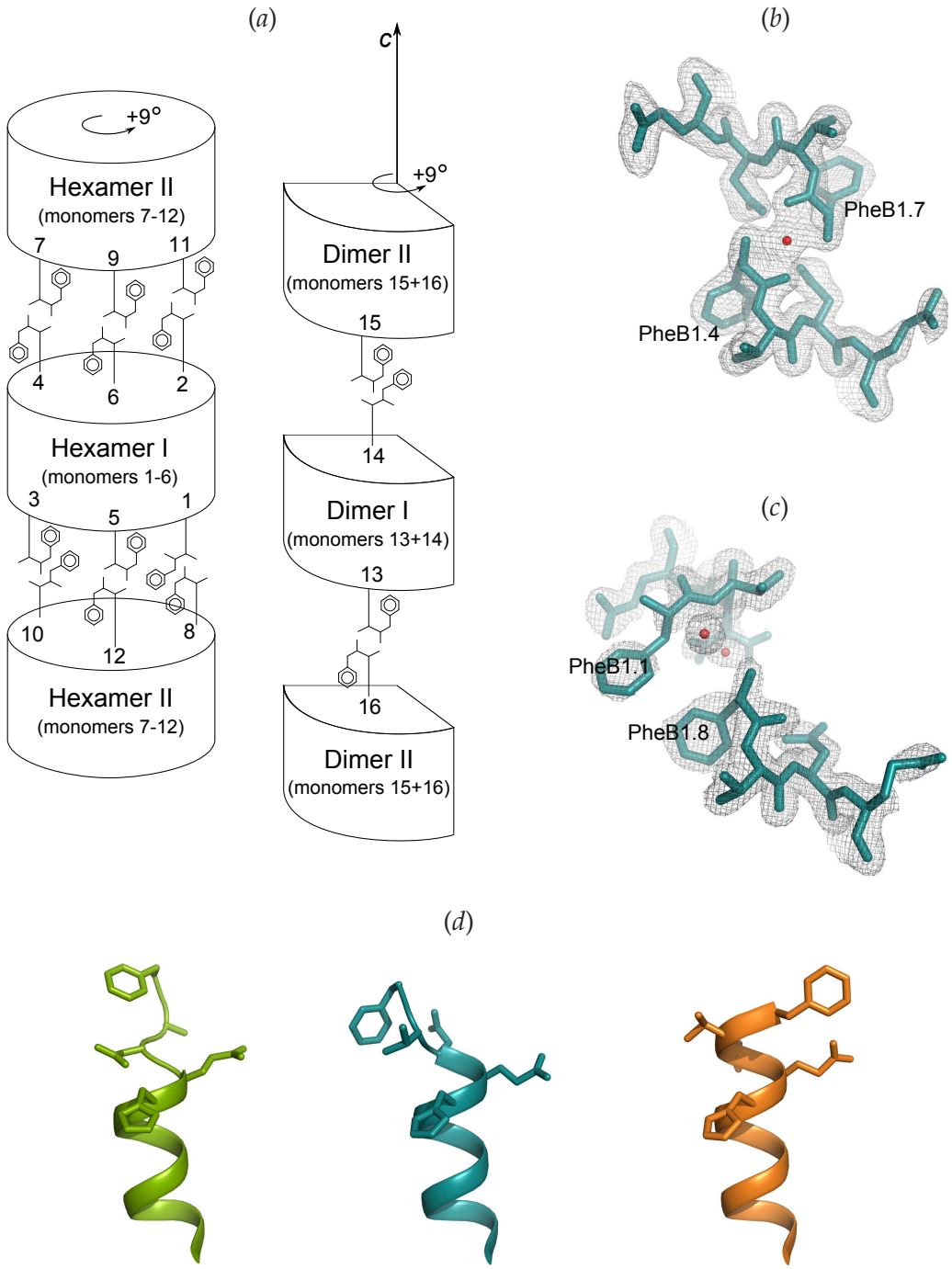


Figure 7.2: (*a*) Schematic representation of the over all packing of the eight dimers constituting the asymmetric unit, with the different conformations of PheB1 schematically drawn. The N-terminal residues B1–B4 are shown in reality: (*b*) for the site where monomers 4 and 7 meet and (*c*) for the site where monomers 1 and 8 meet. Ordered water molecules are shown as red spheres. (*d*) The B-chain N-terminal for R^f-conformation (green) seen in R_3-sites of T_3R_3 insulin (e.g. PDB-entry 1BEN), the new conformation (blue), seen in monomers 2–9, 11–13, 15 and 16 and the pure R-conformation (orange), seen in monomers 1, 10 and 14.

7.2.5 X-ray absorption spectroscopy

Data collection

XAFS measurements at the Zn K-edge were carried out on beam line 811 at the synchrotron at MAX-lab, Lund, Sweden using a Si(111) double-crystal monochromator. The storage ring was operating with ring currents between 150 and 300 mA, and harmonics were suppressed by detuning the peak intensity by 60 % at 10660 eV.

XAS data for the insulin samples as prepared for XRPD were collected in fluorescence mode. A photograph of a mounted insulin sample is shown in Figure 7.3. XAS data were also collected for two model compounds representing respectively octahedrally and tetrahedrally coordinated Zn: The six-coordinated zinc in hexakisimidazolezinc(II) chloride (Sandmark & Brändén, 1967) and the four-coordinated zinc in bisimidazolezinc(II) chloride (Edsall *et al.*, 1954). The model compounds were diluted with boron nitride in the ratio 1:6 to avoid self absorption and measured in transmission mode.

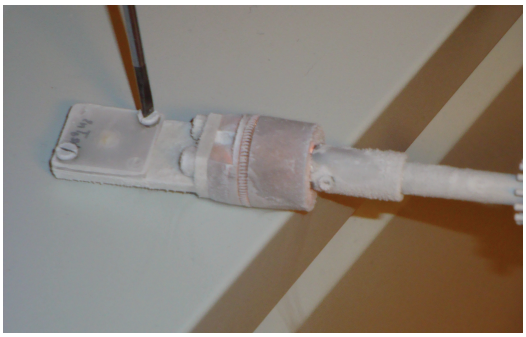


Figure 7.3: XAS sample of insulin mounted at the cryostat holder on beam line 811, MAX-lab, Lund, Sweden.

The samples were kept at 100 K in a liquid nitrogen cryostat. Transmission data were collected using ion chambers and fluorescence data were collected with a PIPS (Passivated Implanted Planar Silicon) detector equipped with solar slits and a copper filter to surpress non-Zn fluorescence scattering. Data were collected in the region 9510–10660 eV in following intervals: Pre-edge data (150–30 eV before the edge) were collected in steps of 5 eV for 1 s, the edge (from 30 eV before to 30 eV above the edge) in steps of 0.3 eV for 1 s, and the EXAFS (30–1000 eV above the edge, corresponding to 16 $Å^{-1}$ in k-space) in steps of 0.05 $Å^{-1}$ for 1–15 s. Three to four spectra were collected for each sample to ensure reproducibility and that no radiation damage of the sample had taken place.

XAS data reduction

The spectra were deglitched, averaged, energy calibrated, background corrected and the EXAFS function $\chi(k)$ extracted using *WinXAS* (Ressler, 1998), as described in Section 3.4.

The energy calibration was carried out using a simultaneous XAFS transmission spectrum of a Zn reference foil. The first inflection point of the foil spectrum was assigned to the Zn K-edge (9659 eV).

EXAFS analysis

Simulations of the calibrated, averaged and background corrected EXAFS, $\chi(k)$, were carried out using *EXCURVE* (Gurman *et al.*, 1984, 1986; Binsted *et al.*, 1991). Potentials were calculated using the relaxed approximation and phase shift by the Hedin-Lundqvist method. The starting models, consisting of atoms within a radius of 5.6 Å from the zinc atoms, (which thereby includes the C^{β} atom of the histidine residues) were taken from the crystal structures described in Section 7.2.3.

Multiple scattering, which occurs in the imidazole and the linear thiocyanate molecule, in particular, was included throughout all refinements. The amplitude reduction factor, S_0^2, was set to 0.95 due to a relatively low self absorption caused by the low zinc concentrations. The models were refined by first a constrained and later a restrained refinement (Binsted *et al.*, 1992). In the constrained refinement E_0, distances and Debye-Waller factors were refined iteratively where each histidine is assumed to be rigid and refined by only one distance and one angular parameter, (Binsted *et al.*, 1992). The number of refineable parameters was kept low by grouping Debye-Waller factors of atoms with similar distances to the central atom (e.g. the $C^{\varepsilon 1}$ and $C^{\delta 2}$ in the histidine residue). During refinement, the ordered water molecules included in the crystal structures in the outer coordination shells were stepwise included in the model, and justified only if the fit-index, ϵ_v^2, decreased by more than 5 % (Joyner *et al.*, 1987). The fit index taking account of the degree of over determinacy in the system is given as:

$$\epsilon_v^2 = \frac{1}{N_{\text{ind}} - N_{\text{p}}} \frac{N_{\text{ind}}}{N} \sum_i^N \frac{1}{\sigma_i^2} \left(\chi_{\text{exp}}(k_i) - \chi_{\text{teo}}(k_i)\right)^2 \tag{7.1}$$

in which N is the number of data points, N_{ind} is the number of independent data points and N_{p} is the number of parameters. $1/\sigma_i$ is a normallization term used for simplifying the numerical methods employed in the algorithm and is further defined in Binsted *et al.* (1992). The fit was evaluated by a residual, R_{exafs}, defined as (Binsted *et al.*, 1992):

$$R_{\text{exafs}} = \sum_i^N \frac{1}{\sigma_i} \left(\left|\chi_{\text{exp}}(k_i) - \chi_{\text{teo}}(k_i)\right|\right) \cdot 100\% \tag{7.2}$$

Hereafter restrained refinement was performed in which the restraints were set up for the histidine residue restraining the imidazole and the C^{β} atom as defined by Engh & Huber (1991) and shown in Figure 7.4. The geometric restraints and the EXAFS were weighted by 50 % each.

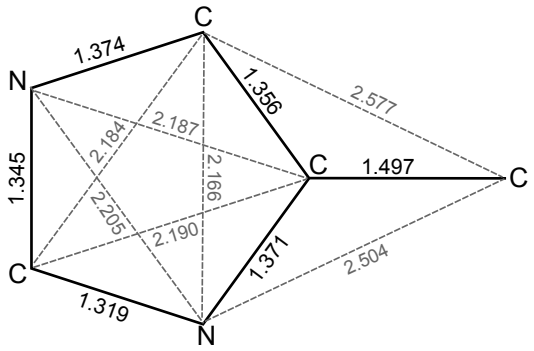


Figure 7.4: Ideal bond lengths of the histidine residue used in restrained refinement (Engh & Huber, 1991). Dashed lines indicates 'secondary' bond distances which are used to define the angles.

XANES analysis

For XANES analysis the calibrated, averaged and background corrected experimental spectra, given as a function of energy, $\mu(E)$, are compared with other experimental XANES spectra (qualitative analysis) and calculated XANES spectra (quantitative analysis). Calculation of XANES spectra were performed using the *FDMNES* code (Joly, 2001), which perform full potential calculations using finite difference methods (FDM).

For quantitative fitting of XANES spectra the coordinates optimized from the EXAFS analysis were used as starting models. Full multiple scattering calculations were performed at the Zn K-edge using the *FDMNES*-code (Joly, 2001) testing both the muffin-tin (MT) approximation and the FDM approach. A cluster with a radius of 4.5 Å from the Zn-ion was found suitable to reproduce the features of the XANES spectra still minimizing the calculation time. The calculated spectra were convoluted with an energy-dependent broadening using an arctangent function to account for inelastic processes, and with a Gaussian function to simulate the experimental energy resolution of the instrument. For the fitting procedure geometrical transformations were decomposed into parameters and selected by the criterion that each parameter must have a significant influence on the XANES spectra. Four structural parameters (two ligand-zinc distances and two angles) have been chosen for T_6 and R_6 insulin conformers. All geometry transformations and FDM calculations were carried out and the structural parameters were optimized by multidimensional interpolation of calculated XANES spectra by *FitIt* (Smolentsev & Soldatov, 2007). The best fit was found by finding a suitable interpolation polynomial and a XANES spectrum of the optimized geometry was calculated for verification. The fit between calculated and experimental spectra is evaluated with an R-factor, given as:

$$R_{\text{xanes}} = \sum_{i}^{N} \frac{\left|\mu_{\text{exp}}(E_i) - \mu_{\text{calc}}(E_i)\right|}{\mu_{\text{exp}}(E_i)} \cdot 100\% \tag{7.3}$$

7.2.6 Comparison analysis

An analysis of Zn-coordination in other insulin structures was carried out by searching the Protein Data Bank (PDB). 67 X-ray structures of hexameric insulin (including mutant proteins) containing Zn were found. Structures solved at non-cryogenic temperatures were discarded and only the structures with a resolution of 2.00 Å or better were included in the analysis since the geometrical restraints are tighter for low resolution structures. This resulted in a total of 31 structures. The Zn sites in these structures were classified as either a T_3-site or an R_3-site and the bond distances from Zn to its inner coordination sphere ligands were extracted.

Zn-coordination in small molecules were analysed by searching the Cambridge Structure Database (CSD) among structures solved at temperatures below 200 K containing zinc which coordinates to three monodentate ligands through sp^2-hybridized nitrogen atoms. For the T_3-site and the R_3-site, respectively, only structures containing Zn coordinated to six atoms (3 nitrogen ligands and 3 other atoms) and Zn coordinated to four atoms (3 nitrogen ligands and 1 chloride) were included in the analysis. Among the structures containing hexa-coordinated zinc only Zn–O-distances were included in the analysis besides the three Zn–N-distances.

7.3 Results

7.3.1 Single-crystal X-ray diffraction structures

General conformation

The conformations of the B-chain N-terminals, which determine the overall insulin conformation, have been verified in all three structures: Both insulin molecules in the T_6-structure adopt tensed conformation in which the backbone of residues B1–B8 is extended. In the T_3R_3-structure the insulin molecule present in the T_3-trimer also adopts extended conformation, whereas for the molecule involved in the R_3-trimer an R^f conformation was observed in analogy to what has been observed for human and porcine T_3R_3 insulin (Smith *et al.*, 2001; Whittingham *et al.*, 1995). In the R_6 insulin structure pure R-conformation, where all residues in the range B1–B18 take part of an α-helix, was only observed in monomers 1, 10 and 14 (see Figure 7.2(*a*)–(*c*). The three first residues B1–B3 in the rest of the 16 insulin molecules were neither observed to adopt a pure R- nor an R^f-conformation, but something in between, where the PheB1 residue is extended, Figure 7.2(*d*). This conformation will be designated R^g in the following. As a consequence the ψ and φ-torsion angles of ValB2 take values around $-89°$ and $+50°$, respectively which are found to be located at the border to the allowed regions in the Ramachandran plot, thereby justifying the outliers of the ValB2 residues. The PheB1 residues are located on the protein surface and

are thereby involved in the hexamer packing. The backbone orientation of the two hexamers and the two dimers, respectively, were observed to differ by a rotation of 9° about an axis parallel to the c-axis, see Figure 7.2(*a*).

Geometry of Zn-sites

The coordination of the zinc sites are presented in Figure 7.5 and 7.6. For the T_6-conformation, octahedral coordination geometry was verified at both Zn-sites, fulfilled by three symmetry related water molecules at 2.29 Å on average from the Zn-atom. As the Zn-sites in T_6 insulin are exposed to the solvent, a large electron density was observed above the Zn-ion, Figure 7.5(*a*), which was best modelled by water molecules connected by a hydrogen bonding network. The Zn-sites in R_6 insulin, Figure 7.5(*b*), are not exposed to the solvent. They are encased by the helix formation of the N-terminal of the B-chains, and thereby only allowed room for a tetrahedral coordination with a chloride ion completing the coordination. The Zn–Cl distances ranged from 2.07 to 2.40 Å, throughout the eight Zn-sites, with an average of 2.21±0.10 Å.

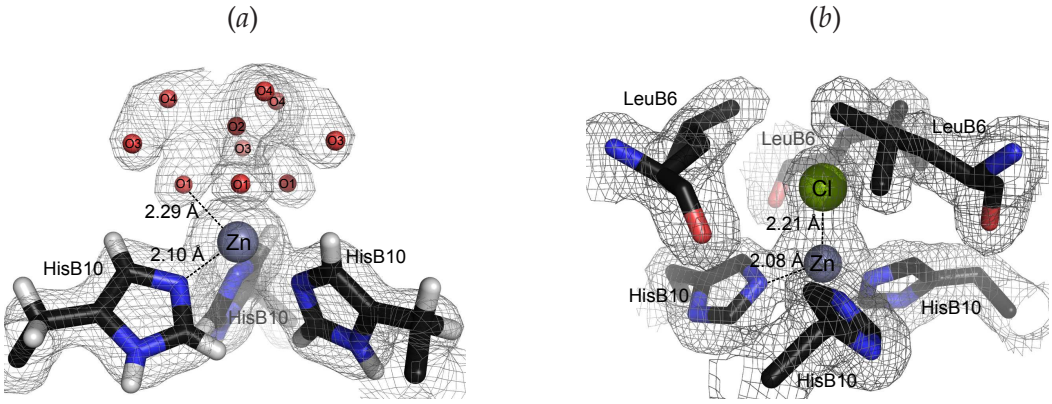


Figure 7.5: (*a*) Octahedrally coordinated Zn in the 1.4 Å T_6 insulin structure, where the electron density has been modelled with four symmetry related water molecules, of which one is located on the three-fold axis (O2). (*b*) Tetrahedrally coordinated Zn in the 1.8 Å R_6 insulin structure, in which the LeuB6 residues encase the metal site. Distances to the first coordination sphere as determined by single crystal X-ray diffraction are shown. As the T_6 and R_6 insulin respectively contain two and eight Zn per asymmetric unit, the distances shown are the average of the different Zn-sites in each structure. The σ_A-weighted $2F_o - F_c$ maps have been contoured at 1.0 σ.

In the T_3R_3-structure both Zn-site geometries were observed, Figure 7.6. One thiocyanate ion fulfils the tetrahedral Zn-site coordination, Figure 7.6(*a*), as observed earlier for the porcine T_3R_3-structure (Whittingham *et al.*, 1995). A clear octahedral Zn coordination was observed at the other Zn-site, and the electron density was modelled by water molecules at a distance 2.47 Å from the Zn-atom, Figure 7.6(*b*).

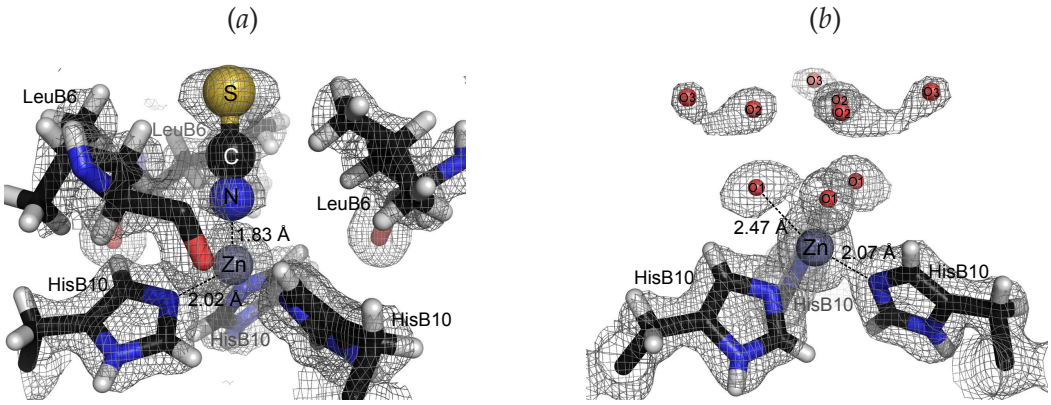


Figure 7.6: Zn coordination in the 1.3 Å T_3R_3 insulin structure. (*a*) A thiocyanate fulfils the tetrahedral coordination and the site is further encased by LeuB6 residues in analogy to the Zn-sites in R_6 insulin. (*b*) Octahedrally coordinated Zn where the electron density has been modelled with three symmetry related water molecules. Distances to the first coordination sphere as determined by single crystal X-ray diffraction are shown. The σ_A-weighted $2F_o - F_c$ maps have been contoured at 1.0 σ.

7.3.2 X-ray powder diffraction

Using the procedures described in Chapter 4 the crystal forms were verified from the powder patterns. The unit cell parameters were determined from a full pattern profile fit and listed in Table 7.5. The unit cell parameters are different for the different conformations, e.g. the unit cell parameter c is increased by approximately 3 Å for each trimer undergoing a T to R conformational change. This systematic variation of unit cell parameters is thereby the primary factor for the identification of the conformations. Experimental and calculated patterns of all three insulin conformations are seen in Figure 7.7.

Table 7.5: Unit cell parameters determined by XRPD from full pattern profile fits.

	Identification			Phase analysis	
	T_6 insulin	T_3R_3 insulin	R_6 insulin	R_6 insulin	R_6 insulin
Temperature (K)	RT	RT	RT	RT	120
Data collection time (h)	0.5	0.5	0.5	16	64
a (Å)	82.86	80.99	80.39	80.02	156.49
c (Å)	33.96	37.45	40.27	40.11	78.53
R_p (%)	0.83	1.09	0.94	0.79	0.22

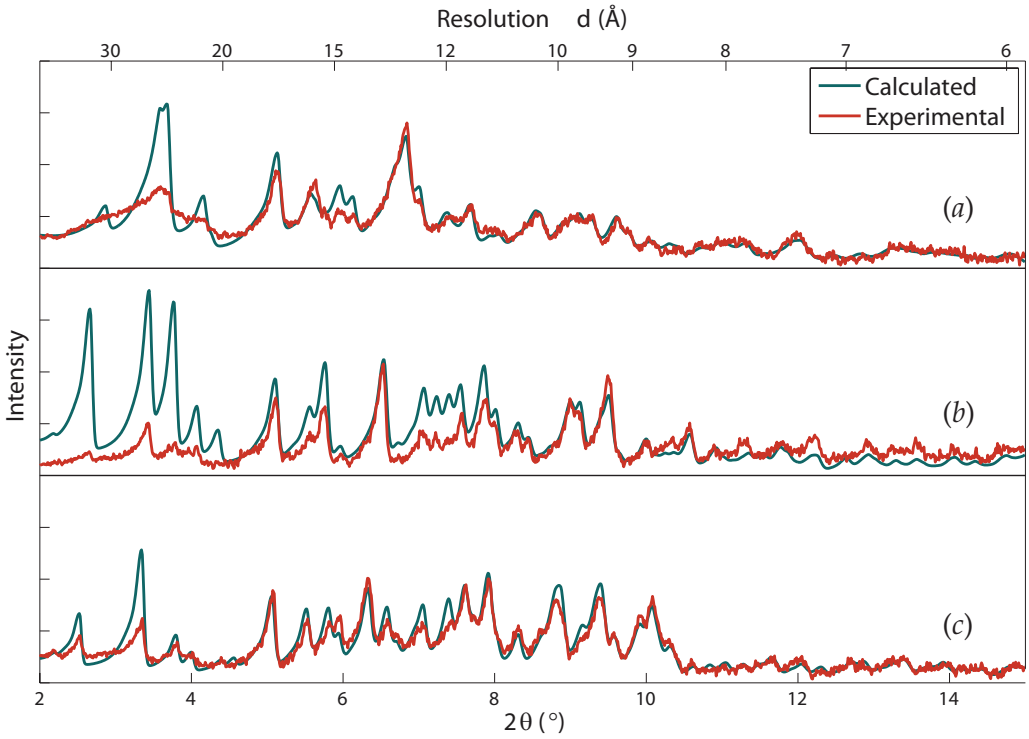


Figure 7.7: Experimental (red) and calculated (blue) XRPD patterns of (*a*) T_6 , (*b*) T_3R_3 and (*c*) R_6. The verification of the crystal forms was done from experimental data collected in 30 minutes. Unit-cell and peak shape parameters were optimized from full pattern profile fits, and solvent correction using average values of $k_{sol} = 0.35$ e Å^{-3} and $B_{sol} = 46$ Å^2 was applied prior to calculation of the powder patterns.

In order to detect whether the R_6 insulin undergoes a phase transition when cryo-cooled, powder patterns with longer data collection times were collected at RT and at 120 K. The unit cell doubling observed in the single crystal experiment was detected from the XRPD patterns, as well, by looking at the low angle region ($2\theta = 3 - 5°$) in the powder patterns, where the peaks are better resolved. In the cold pattern, reflections which are not in agreement with the non-doubled unit cell are observed, for instance the reflection at 3.25°, which is pointed out on Figure 7.8(*a*) can only originate from a large unit cell. Analogously, from the powder pattern collected at RT all reflections are in accordance with the non-doubled unit cell, Figure 7.8(*b*).

7.3.3 Qualitative XANES spectroscopy

In Figure 7.9 XANES spectra of the three insulin conformations T_6, R_6 and T_3R_3, having respectively six-coordinated zinc, four-coordinated zinc and a mixture hereof, are compared with the two model compounds. The octahedrally coordinated zinc complexes with six ligands give strong white lines (> 1.5) with normalized edge steps, whereas tetrahedrally

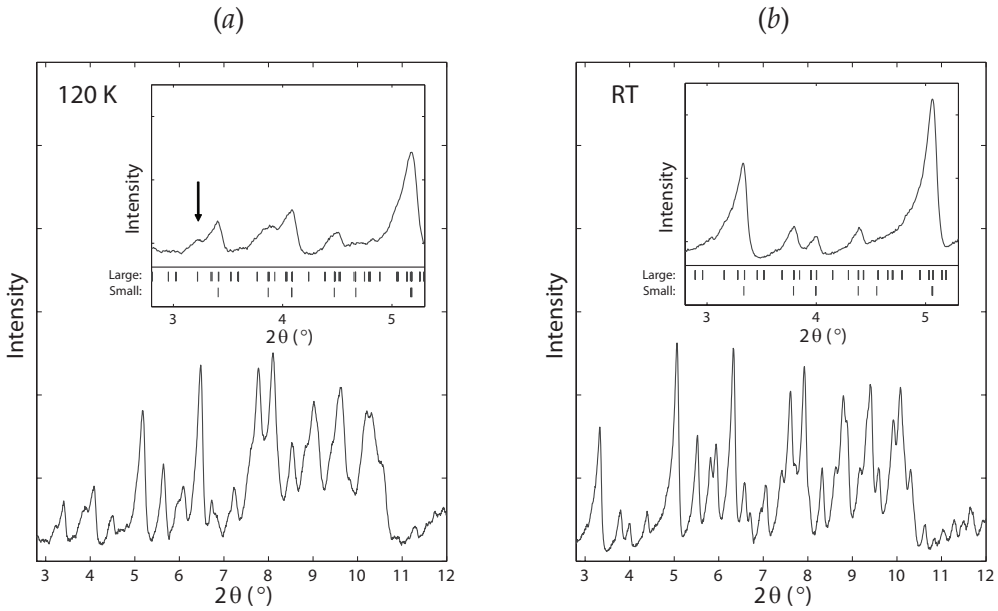


Figure 7.8: XRPD patterns of R_6 insulin collected at 120 K (*a*) and RT (*b*). Insets: magnification of the low angle region ($2\theta = 3 - 5°$) and the positions of the Bragg peaks for the large ($a \approx 160$ Å, $c \approx 80$ Å) and small ($a \approx 80$ Å, $c \approx 40$ Å) unit cell, respectively. The peak pointed out at 3.25° in the cold spectrum indicates that the cold phase must have doubled unit cell parameters.

coordinated zinc complexes give white lines (< 1.5) with normalized edge steps (Feiters *et al.*, 2003). The medium intensity of the white line for the T_3R_3-conformation indicates coordination of both octahedral and tetrahedral character, which is in agreement with the dual coordination observed in the crystal structure.

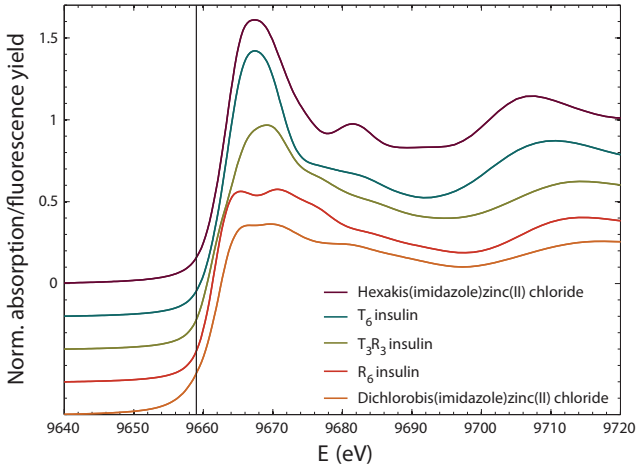


Figure 7.9: Energy calibrated XANES spectra of the three protein samples and the two reference compounds: Hexakisimidazolezinc(II) chloride, T_6 insulin, T_3R_3 insulin, R_6 insulin, bisimidazolezinc(II) chloride. The location of the K-edge of metallic zinc (9659 eV) is pointed out.

7.3.4 EXAFS spectroscopy

The extracted k^3-weighted EXAFS spectra and the modulus of the phase-corrected Fourier transforms of the three insulin conformations are presented in Figure 7.10. The major single peak in the radial distribution function is observed around 2.0 Å and corresponds to the inner coordination shell which is mostly dominated by the three imidazole nitrogen atoms and common for all three conformations. The shells occurring around 3 and 4 Å verifies the imidazole coordination of histidine in all conformations. The imidazole coordination can also be directly identified from the 'camel-back' feature around 4–5 $Å^{-1}$ which originates from the interference between the backscattered electron waves from the first, second and third shell atoms of the imidazole ring (Bordas *et al.*, 1983).

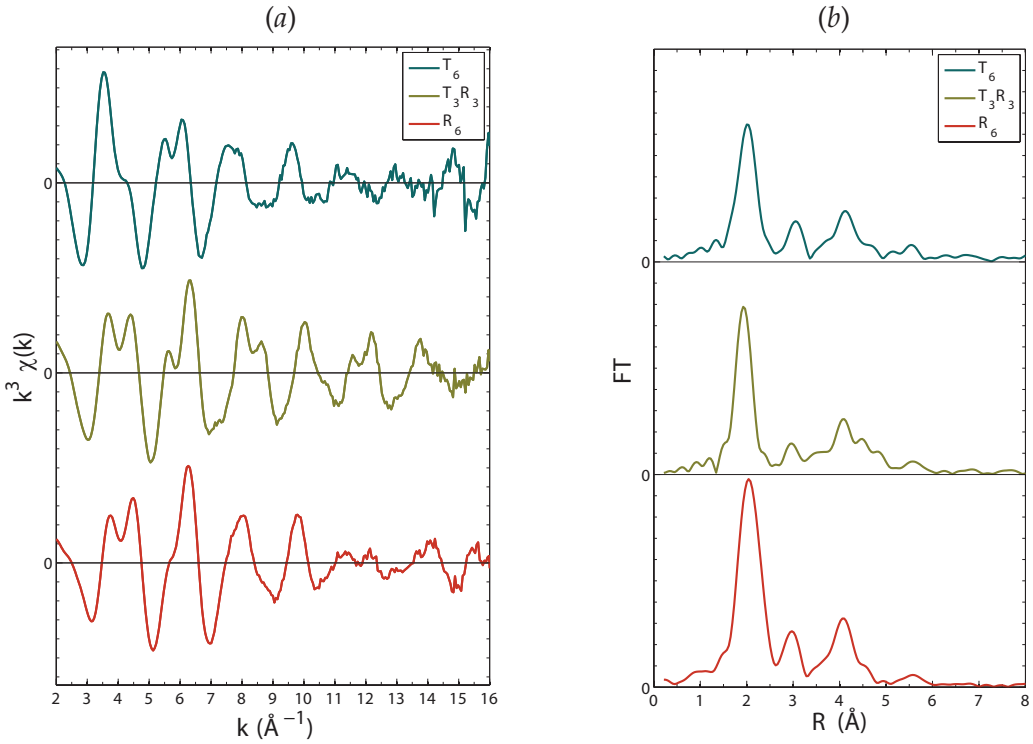


Figure 7.10: (*a*) k^3-weighted EXAFS spectra and (*b*) radial distribution functions calculated as the modulus of the phase corrected Fourier transforms of the three protein samples: T_6 insulin, T_3R_3 insulin and R_6 insulin.

By comparing the k^3-weighted $\chi(k)$ at low k-values (approx. 3–7 $Å^{-1}$) the T_3R_3-conformation have character of both T_6 and R_6 which is in agreement with the dual tetrahedral/octahedral coordination of the zinc ions as seen from the crystal structure and the XANES spectra. An indication for coordination of thiocyanate in the T_3R_3-conformation is seen as a slightly more intense and complex peak around 4.5–5.0 Å in the radial distribution function whereas the chloride which is fulfilling the tetrahedral coordination in the R_6-

conformation is seen as a more intense and slightly more broadened inner coordination shell peak at 2.0–2.2 Å.

The optimized distances and Debye-Waller factors are presented in Table 7.6 and the best fits are shown in Figure 7.11 (dashed lines). Due to the complexity of the double cluster necessary for modelling the T_3R_3-conformation restrained refinement resulted in too many refineable parameters, and was therefore abandoned. Restrained refinement for T_6 and R_6-conformations gave the best fits where residuals down to 11–15 % were reached. Further EXAFS refinement statistics are summarized in Table 7.7.

The distances to the inner coordination sphere as determined by EXAFS were used as restraints for refinement of the T_6- and R_6 insulin XRD-structures. The EXAFS distances were given as the ideal distance and the uncertainty as the standard deviation σ in *PHENIX*. Slightly increased R_{free} were observed for both T_6- and R_6 insulin, when applying the restraints from the EXAFS model.

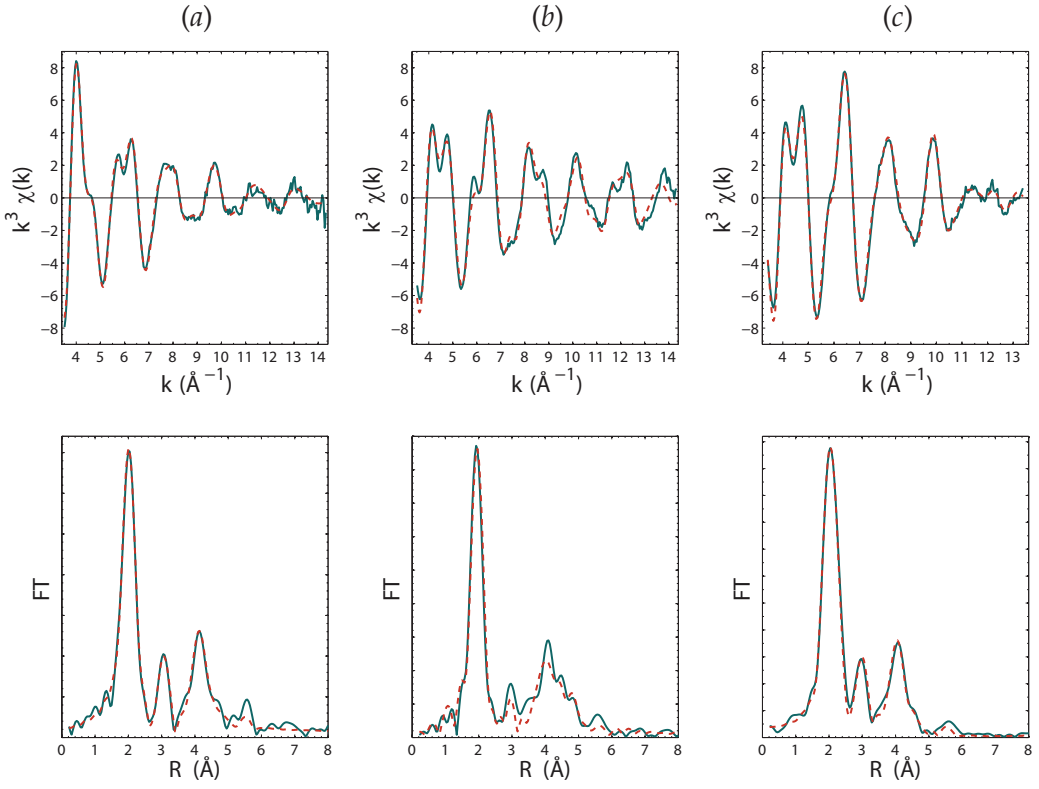


Figure 7.11: k^3-weighted EXAFS (top) and radial distribution functions calculated as the modulus of the phase-corrected Fourier transform (bottom) of the three insulin conformations: (*a*) T_6 (*b*) T_3R_3 and (*c*) R_6. Experimental spectra are shown in blue and simulated in red (dashed), using the parameters from the restrained refinement given in Table 7.6 for T_6 and R_6, and parameters from the constrained refinement for T_3R_3 given in Table 7.6.

Table 7.6: Zn-coordination distances as refined from EXAFS of T_6-, R_6- and T_3R_3 insulin compared to crystallographic values. T_6 and R_6 have been refined with restrained refinement, while T_3R_3 has been refinement using constrained refinement.

T_3-sites						
	T_6 insulin			T_3R_3 insulin		
	XRD	Restrained EXAFS		XRD	Constrained EXAFS	
	$\langle R\rangle^{\S}$ (Å)	R (Å)	$2\sigma^2$ (Å^2)	$\langle R\rangle^{\S}$ (Å)	R (Å)	$2\sigma^2$ (Å^2)
$N^{\varepsilon 2}$ (HisB10)	2.10	2.074(3)	0.012(1)	2.07	2.025(11)†	0.014(1)
$C^{\varepsilon 1}$	3.03	3.07(4)‡	0.020(3)	3.05	2.84	0.018(1)
$C^{\delta 2}$	3.13	3.05(3)‡	0.020(3)	3.05	3.16	0.018(1)
$N^{\delta 1}$	4.17	4.22(2)‡	0.017(3)	4.16	4.03	0.028(4)
C^{γ}	4.26	4.22(3)‡	0.017(3)	4.20	4.21	0.028(4)
C^{β}	5.69	5.55(5)‡	0.017(3)	5.61	5.66	0.030(3)
O^{w1}	2.29	2.135(11)	0.030(1)	2.47	2.289(15)	0.035(5)
O^{w2} (axial)	3.09	2.88(3)	0.021(10)			
R_3-sites						
	R_6 insulin			T_3R_3 insulin		
	XRD	Restrained EXAFS		XRD	Constrained EXAFS	
	$\langle R\rangle^{\S}$ (Å)	R (Å)	$2\sigma^2$ (Å^2)	$\langle R\rangle^{\S}$ (Å)	R (Å)	$2\sigma^2$ (Å^2)
$N^{\varepsilon 2}$ (HisB10)	2.08	2.001(4)	0.007(1)	2.02	1.987(11)†	0.014(1)
$C^{\varepsilon 1}$	3.09	2.98(2)‡	0.010(2)	2.99	2.99	0.018(1)
$C^{\delta 2}$	3.04	3.04(2)‡	0.010(2)	3.03	2.96	0.018(1)
$N^{\delta 1}$	4.18	4.15(1)‡	0.012(2)	4.10	4.09	0.028(4)
C^{γ}	4.19	4.14(2)‡	0.012(2)	4.16	4.11	0.028(4)
C^{β}	5.60	5.53(3)‡	0.012(2)	5.58	5.52	0.030(3)
O (LeuB6)	4.87	4.90(5)	0.020(9)			
Cl (axial)	2.21	2.218(3)	0.006(1)			
N^{SCN} (axial)				1.83	1.802(9)	0.014(1)
C^{SCN}				2.98	2.96	0.018(1)
S^{SCN}				4.72	4.69	0.017(3)

† The rotation angle of the histidine unit around an axis orthogonal to the imidazole plane going through the $N^{\varepsilon 2}$ was included in the refinement.

‡ The ϕ-angle (polar coordinates) was refined in order to allow some movement of the atoms in the restrained imidazole ring.

§ Average distances of all Zn-sites in the structure. The standard deviations among the different sites are below 0.03 Å for T_6 insulin and below 0.1 Å for R_6 insulin.

7.3.5 Quantitative fitting of XANES spectra

Quantitative fitting of XANES spectra was possible for T_6 and R_6 insulin and the calculated XANES spectra are shown in Figure 7.12 for both conformations before and after the fitting. FDM calculation of a XANES spectrum using a 4.5 Å cluster took 1–2 days on the high performance computer available at the Technical University of Denmark.

Table 7.7: EXAFS refinement statistics. T_6 and R_6 have been refined with restrained refinement, while T_3R_3 has been refinement using constrained refinement.

	T_6 insulin	R_6 insulin	T_3R_3 insulin
E_f (eV)	−3.31	−5.47	−2.366
ϵ_v^2	0.4127	0.1920	3.1572
R_{exafs} (%)	14.09	10.17	22.53
R_{dist} (%)	1.14	0.65	
R_{total} (%)	15.23	10.82	
N_{p}	19	19	14
k-range (Å^{-1})	2.8–14.3	2.8–13.3	2.8–14.3
w_{dist}	0.5	0.5	

Using the MT-approach a XANES spectrum could be calculated in minutes on a standard PC, but matches the observed poorly, as seen in Figure 7.12, whereas the spectra calculated by the FDM approach are in better agreement with the experimental spectra.

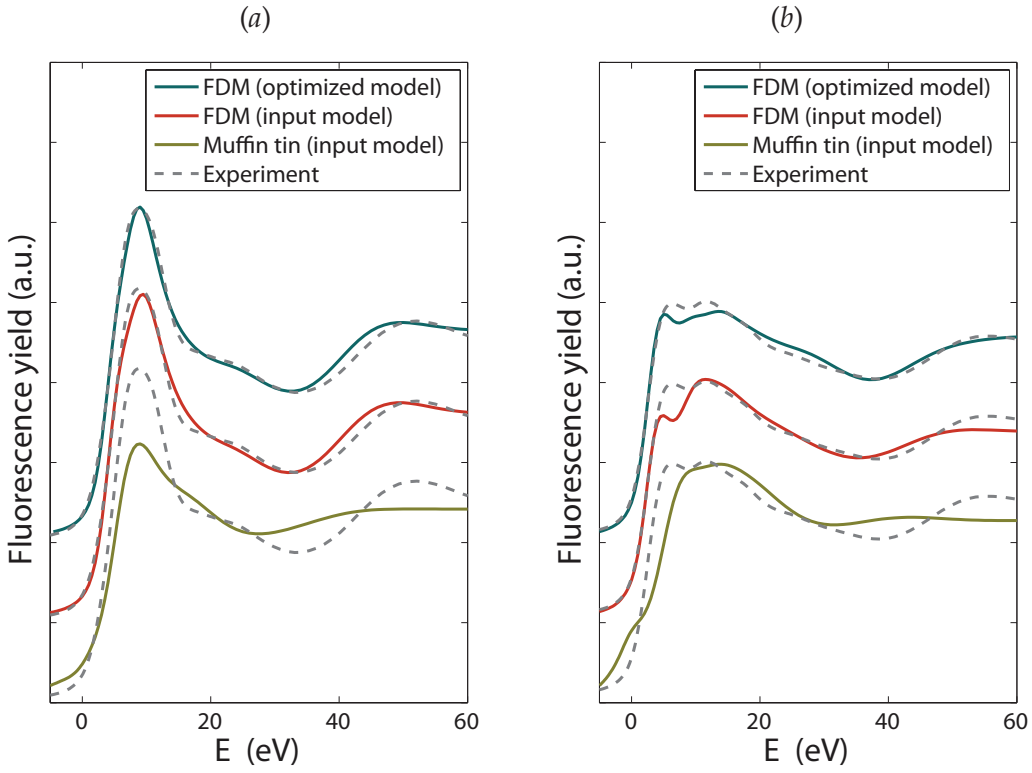


Figure 7.12: XANES calculated on a 4.5 Å cluster around each Zn-atom of (*a*) T_6 insulin, and (*b*) R_6 insulin. Calculation were done on the input model (EXAFS model) using both the MT approximation (green) and FDM approach (red), and on the model optimized by *FitIt* (XANES model) using the FDM approach (blue). The calculated spectra are compared with experimental XANES (dashed). The offset of the energy scale is 9659 eV corresponding to the K-edge position of metallic zinc.

The XANES technique is more sensitive to bond angles among the used techniques. Besides the distances to the inner coordination sphere, two angles were found to have a significant influence on the XANES spectra: $\angle$(Zn–N$^{\varepsilon 2}$–C$^{\varepsilon 1}$) for both conformations and $\angle$(N$^{\varepsilon 2}$–Zn–O^{w1}) for T_6 and $\angle$(N$^{\varepsilon 2}$–Zn–Cl) for R_6. The geometrical parameters, P1–P4, varied during the quantitative fit, are visualized in Figure 7.13 and their values are listed in Table 7.8 together with the R_{xanes} for the fits.

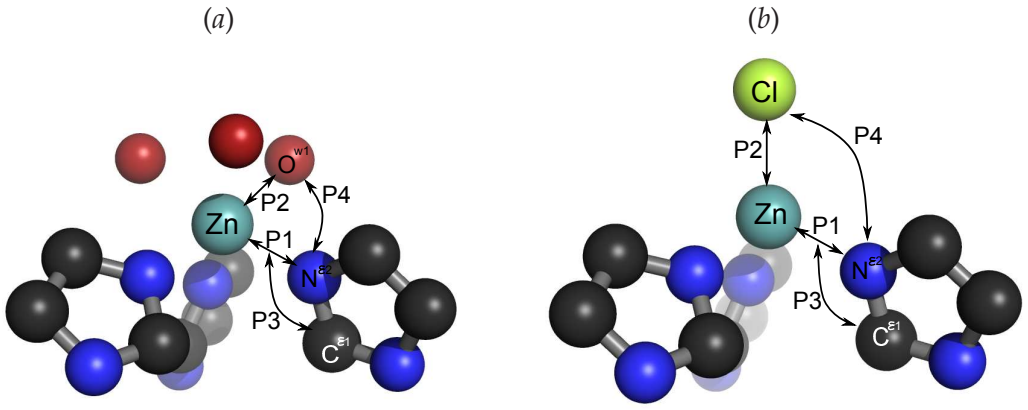


Figure 7.13: The 4.5 Å cluster used for calculation of XANES spectra of (*a*) T_6 and (*b*) R_6 insulin. The parameters which were varied during the quantitative fit P1–P4 are visualized.

Table 7.8: Quantitative XANES fit parameters and R-factors for the initial model (EXAFS model) and the model optimized by *Fit-It* (XANES model).

T_6 insulin	Input model (EXAFS model)	Optimized model (XANES model)
P1: Zn–N$^{\varepsilon 2}$ (HisB10)	2.07 Å	2.08 Å
P2: Zn–O^{w1}	2.14 Å	2.13 Å
P3: $\angle$(Zn–N$^{\varepsilon 2}$–C$^{\varepsilon 1}$)	127°	119°
P4: $\angle$(N$^{\varepsilon 2}$–Zn–O^{w1})	166°	173°
R_{xanes}	3.7 %	2.5 %
R_6 insulin	Input model (EXAFS model)	Optimized model (XANES model)
P1: Zn–N$^{\varepsilon 2}$ (HisB10)	2.00 Å	2.00 Å
P2: Zn–Cl	2.22 Å	2.24 Å
P3: $\angle$(Zn–N$^{\varepsilon 2}$–C$^{\varepsilon 1}$)	130°	141°
P4: $\angle$(N$^{\varepsilon 2}$–Zn–Cl)	109°	107°
R_{xanes}	3.4 %	2.4 %

Regarding the inner coordination sphere of Zn, the XANES analysis of T_6 insulin shows a more regular octahedrally coordinated Zn than determined from the XRD, as the angle N$^{\varepsilon 2}$–Zn–O^{w1} was optimized from 166° to 173°. The tetrahedral Zn-site in R_6 insulin, was

verified without significant optimizations, meaning that the initial R_6-model from the EXAFS optimized XRD structure was better than the T_6-model. However, in-plane rotation of the histidines was optimized from 130° to 141° with significant influence on the shape and the intensity ratio of the main double peak of the spectrum, see Figure 7.12(*b*). The optimized models from XANES were used for another round of EXAFS analysis, which gave fits of similar quality as the initial.

7.3.6 Comparison with other reported Zn-site geometries

The result of the analysis of bond distances involved in the zinc coordination observed in other insulin structures, as well as in small molecules with similar coordination geometries, is summarized Table 7.9. The Zn–N bond distances for the R_3-sites amongst the insulin structures are plotted as a function of the resolution, Figure 7.14.

Table 7.9: Average bond distances and statistical standard deviations for Zn bonds to its inner sphere ligands among 31 insulin structures deposited at PDB and 39 small molecule structures with similar Zn coordination.

PDB-analysis	T_3-sites		R_3-sites	
Number of Zn-sites	25		42	
	Zn–$N^{\varepsilon 2}$ (Å)	Zn–O^{w1} (Å)	Zn–$N^{\varepsilon 2}$ (Å)	Zn–Cl (Å)
All structures	2.06(5)	2.41(17)	2.02(10)	2.18(9)
XRD this work†	2.10(1)	2.29(3)	2.08(5)	2.21(10)
EXAFS this work†	2.07	2.14	2.00	2.22
CSD-analysis	**Octahedral Zn (3N+3X)**		**Tetrahedral Zn (3N+1Cl)**	
Number of structures	26		13	
	Zn–N(sp^2) (Å)	Zn–O (Å)	Zn–N(sp^2) (Å)	Zn–Cl (Å)
	2.16(3)	2.16(4)	2.01(2)	2.26(3)

† Only distances for the T_6 and R_6 insulin are presented.

7.4 Discussion

7.4.1 Single crystal XRD structures

The three conformations of bovine insulin were crystallized in the space group $R3$, and data were collected to 1.40 Å, 1.30 Å and 1.80 Å resolutions for T_6, T_3R_3 and R_6 insulin, respectively. The structures of bovine insulin in the T_3R_3 and R_6 conformations have not been solved before. In analogy to what has been observed for an earlier reported bovine T_6 insulin structure (Smith *et al.*, 2005) the structural consequences of the sequence

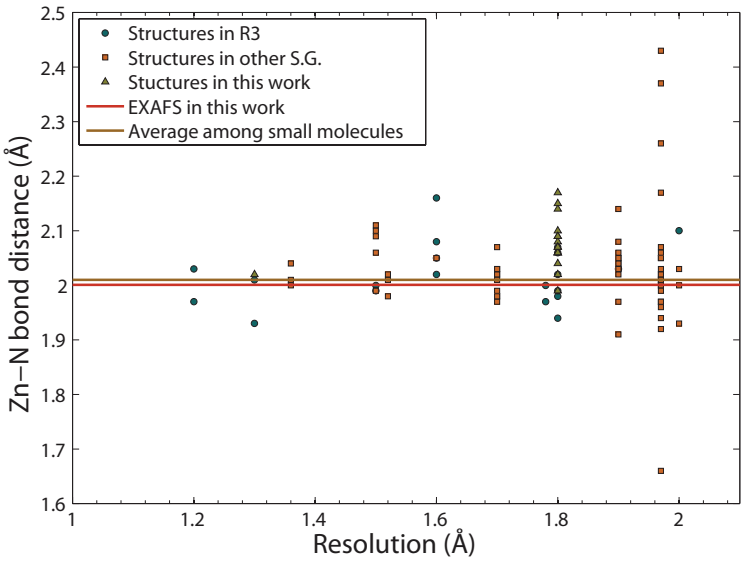


Figure 7.14: Zn–N distances as a function of resolution among 42 R_3-sites deposited in PDB. Insulin structures in $H3$ are distinguished from other space groups. Furthermore the Zn–N distances from the XRD-structures presented in this work, the Zn–N distance in R_6 insulin as determined by EXAFS and the average Zn–N distance among tetrahedral Zn-sites in small molecules are included.

change at A8 and A10 were not significant in any of the three conformations. However there are some differences to previously reported structures. The structures presented here were superposed with other structures deposited in the Protein Data Bank. Independent T_2, TR and R_2 dimers in T_6, T_3R_3 and R_6 insulin, respectively, were superposed using a least-squares procedure using *SUPERPOSE* (Krissinel & Henrick, 2004), where the displacements of C^α-atoms of all residues in common were minimized, and the root mean squares (RMS) were calculated. The T_6-structure shows a high resemblance to the previously published bovine structure (RMS$= 0.282$ Å, PDB-entry 2A3G), (Smith *et al.*, 2005), and the porcine structure (RMS$= 0.239$ Å, PDB entry 4INS), (Baker *et al.*, 1988). A somewhat larger discrepancy to the human (RMS$= 1.013$ Å, PDB-entry 1MSO), (Smith *et al.*, 2003) may be explained a difference of the conformation of the B1.2–B4.2 chain. The T_3R_3 insulin structure is in close resemblance with the porcine structure (RMS$= 0.377$ Å, PDB-entry 2TCI), (Whittingham *et al.*, 1995), which was also crystallized with thiocyanate. The bovine R_6-structure displays a eight-fold cell doubling compared to the human R_6-structure (PDB-entry 1EV3), (Smith *et al.*, 2000). This means that the asymmetric unit contains 16 monomers, whereof the majority (13 out of 16 monomers) of the B-chain N-termini are found to adopt R^g-conformation. Excluding PheB1 and ValB2 from the superposition resulted in RMS displacement of C^αs of 0.413 Å. The R^g conformation has previously been observed for one out of six molecules in a hexamer in monoclinic R_6 insulin structures (Smith *et al.*, 2000). Together with the observed rotation of adjacent hexamers/dimers

along the *c*-direction by 9°, the conformational variation of the B-chain N-termini explain the eight-fold doubling of the unit cell. Using powder diffraction a phase transition was observed for the R_6-conformation when cooled, in analogy to what has been seen in the human T_3R_3-structure by Smith *et al.* (2001) with the large unit cell being the low temperature structure.

Both Zn-sites in T_6 insulin structure are octahedral and coordinate three water molecules at similar, however long, distances to what has been observed in the other T_6-structures (Baker *et al.*, 1988; Smith *et al.*, 2003, 2005). For the T_3R_3, Zn sites, the three water molecules which fulfil octahedral coordination are clearly observed at the T_3-site, in contrast to the porcine (Whittingham *et al.*, 1995) and the human (Smith *et al.*, 2001) structure, where tetrahedral coordination fulfilled by respectively a single water molecule and chloride ion, has been modelled. The R_3-sites in both the T_3R_3 and R_6 conformations are however always tetrahedral, as one lyotropic anion fulfils the coordination geometry.

7.4.2 XAS

Using EXAFS the inner coordination sphere distances are more accurately determined (down to 0.005 Å), whereas distance to outer shells are determined with accuracies around 0.01–0.05 Å. Comparing XAS with XRD the level of accuracy obtained by these two methods is still intensely discussed, but it certainly depends on the actual resolution of the XRD structure, and the inner coordination sphere distances must be significantly better determined by XAS. Not surprisingly, the results demonstrate that the technique is best for mononuclear metalloproteins or metalloproteins containing several metal clusters that all have identical coordination geometry. Samples having different coordinations of the same metal, as in the T_3R_3-conformation, highly increase the complexity of data treatment and reduce the amount of structural information which can be resolved. It was clearly demonstrated that the FDM-methods reproduce the XANES spectra quite well. The coordination geometry of the zinc sites (bond angles in particular) was optimized from XANES. However, very good quality of the EXAFS spectra (high signal-to-noise ratios for data up to $k = 13 - 14$ Å^{-1}) allowed the initial structural model to be quite accurate. Therefore only subtle adjustments of the coordination geometries were possible with quantitative XANES fitting. The optimized bond distances were in agreement with EXAFS, and thereby XANES serves as tool for cross-validation. Furthermore qualitative comparison of XANES spectra is a powerful tool for identification and verification of coordination geometry.

7.4.3 Comparison with other reported Zn-site geometries

The data base search for small-molecule structures with a geometry that resembles the T_6 geometry revealed only a small number of structures with Zn coordinated to three nitro-

gen and three oxygen ligands with facial isomerism. Therefore the three ligands besides nitrogen were unspecified in the search. For the T_3-site Zn–N distances there is a good agreement between the mean XRD and the EXAFS results. The small molecule search shows a somewhat longer distance, maybe due to the difficulty to find matching structures. For the Zn–O distances, however, EXAFS clearly gives a distance much closer to the small-molecule structures, while XRD gives a too large number and a very large standard deviation. A brief look at some difference maps for insulin structures on PDB containing T_3-sites (including the XRD structure presented here) often reveals unstructured densities, modelled with water molecules and at chemically unreasonable bond lengths, in particular in structures with lower resolution.

For the R_3-sites the Zn–N distance found by the EXAFS study is closer to the Zn–N distance found by averaging all the PDB-structures than to the structure determined in this paper. The broad spread amongst the distances found by XRD is seen in Figure 7.14. It is clear that the high-resolution structures show a very good agreement with results from the small-molecule survey and with the EXAFS result.

An explanation for the quite large discrepancy between the XRD structures and the EXAFS results could be the difference in X-ray dose imposed on the samples. The doses for the two experiments were estimated using the photon energy and flux specifications for beam line 911-2 and 811 as well as the sample size and exposure time. The estimation shows a higher dose by more than two orders of magnitude ($2.6 \cdot 10^9$ Gy) for the single crystal used for XRD compared to the XAS sample ($2.5 \cdot 10^7$ Gy), wherefore it must be assumed that the single crystals suffer from a higher degree of radiation damage. It is also reasonable to assume that radiation damage influences the loosely bound water molecules in the T_3-site more severely than the tightly bound chloride ion in the R_3-site.

7.5 Conclusive remarks

The structures of all three conformations (T_6, T_3R_3 and R_6) of bovine insulin were solved by single crystal XRD, whereof two of the structures are new (T_3R_3 and R_6). The developed procedure for verification of crystal forms by XRPD was successfully applied to characterize the insulin samples before the XAS experiments. For R_6 insulin, XRPD was also used to identify a phase transition, in which the unit cell volume is eight-fold doubled when cooled.

The distances between the zinc and its ligands were very accurately determined with EXAFS (within 0.01 Å), in particular those in the first coordination sphere. The EXAFS determinations are in agreement to what has been observed among small molecules with similar coordination geometries, as well as other high resolution insulin structures. These observations may be coupled to a lower radiation dose for XAS experiments. EXAFS may

thus be better suited for studying metal sites in proteins, which are sensitive to radiation damage. Using XANES the coordination geometry was verified by qualitative comparison with reference compounds and realistic XANES spectra were calculated using finite difference methods.

This study on hexameric bovine insulin highly illustrates the complementarity between XRD and XAS for studying metals in proteins.

Chapter 8

XRD and XAS studies of T_6 insulin with Fe, Ni, and Cu

Hexameric insulin is stabilized by other divalent cations than zinc. In this chapter a study of T_6 insulin with iron, nickel and copper is presented, with particular focus on the metal coordination. This work further illustrates the complementarity between X-ray absorption spectroscopy and diffraction. Hexameric insulin can also be stabilized by some monovalent cations. Insulin structures with rubidium and caesium are solved and briefly presented in Appendix B.

8.1 Introduction

Metal ions are known to play a key role in the storage, release and conformations of insulin (Emdin *et al.*, 1980; Sudmeier *et al.*, 1981). The type of metal that binds to insulin influence on the stability of the hexamer, as well as on the conformation and coordination geometry, and it has therefore been suggested that different intermediates in the T_6 to R_6 pathway could be trapped by crystallizing with different metals (Sreekanth *et al.*, 2009).

Metal ions are also known to play a role in amyloid fibrillation, by which peptides convert into insoluble aggregates. The mechanisms behind fibrillation are poorly understood, but studying formation of fibrils is important as fibrils are known to be involved in many diseases also counting type II diabetes. As already mentioned, zinc plays an important role in the stabilization of the insulin hexamer, thereby inhibiting the formation of insulin fibrils. Likewise, copper is known to bind to the fibrils and is believed to be involved

in Alzheimer's disease. Insulin structures containing other divalent transition metal ions are therefore relevant for studying the hexamer stability. These studies can furthermore be complemented by spectroscopy, as the metals have unoccupied $3\,d$-orbitals, which are spectroscopically active.

Hexameric insulin containing copper, nickel and iron has been studied with single crystal diffraction and XAS, with emphasis on conformation and metal coordination. To support the interpretation of the XAS data, *ab initio* Hartree-Fock calculations have been carried out evaluating energies of copper clusters with different geometries.

8.2 Experimental

8.2.1 Growth of insulin single crystals

Crystallization experiments of hexameric insulin with Mn^{2+}, Fe^{2+}, Co^{2+}, Ni^{2+} and Cu^{2+} were carried out. As the affinity to insulin decreases through the series from Zn to Mn, insulin single crystals were only obtained with Ni^{2+} and Cu^{2+}. Insulin crystals containing Fe^{2+} were grown from zinc contaminated media by using insulin solutions where the zinc was not chelated and removed prior to the crystallization experiments.

For the growth of nickel and copper insulin single crystals, excess zinc ions were removed from the insulin before use using the procedure described in Section 7.2.1. The insulin solution was diluted with milli-Q water and pH was adjusted to 2.0 using HCl(aq). The concentration was determined to 7.5 mg ml^{-1} by UV-vis spectrophotometry. Nickel and copper insulin crystals were grown using the vapour diffusion technique. 2 µl of the metal free insulin solution was mixed with 2 µl reservoir solution and equilibrated in a hanging drop against 1 ml reservoir with a composition given in Table 8.1. After 5 days 200–300 µm large single crystals were observed. The crystals resembled those from T_6 Zn insulin in size and shape. Single crystals with dimensions of 200 µm were soaked in cryoprotective solution containing 15 vol.% PEG 400, 25 vol.% glycerol and 60 vol.% mother liquor (Smith *et al.*, 2001, 2003) and stored in liquid nitrogen until diffraction analysis.

For the growth of iron insulin single crystals, lyophilized insulin (Sigma Aldrich, I-5500) was dissolved in 0.02 M HCl and pH was carefully raised with NaOH until all protein was re-dissolved at pH around 7.0. The protein concentration was determined to 5.2 mg ml^{-1}. The crystals were grown using the vapour diffusion technique. 3 µl of the insulin solution was mixed with 2.4 µl reservoir solution and equilibrated in a hanging drop against 1 ml reservoir with a composition given in Table 8.1. After 2 months small crystals with dimensions of 60–80 µm were observed, Figure 8.1(*a*).

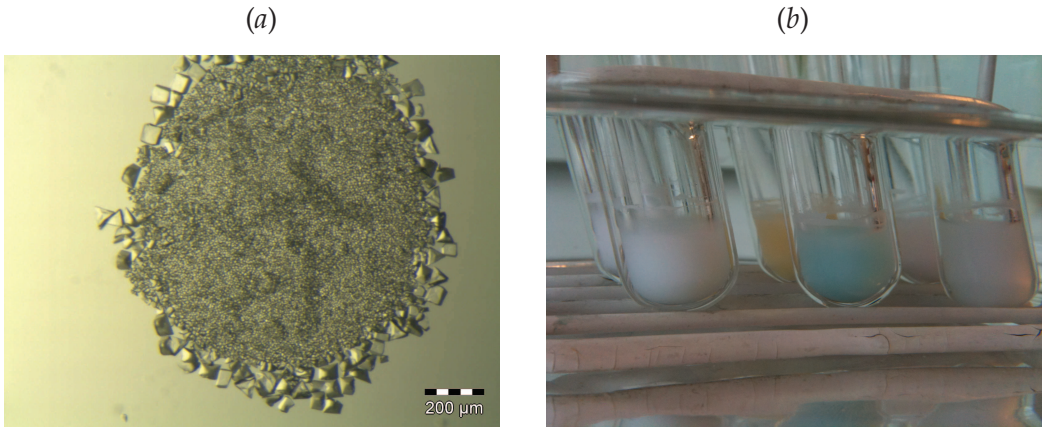


Figure 8.1: (*a*) Single crystals of T_6 insulin crystallized with Fe^{2+}. (*b*) Preparation of micro-crystal samples of Mn^{2+}, Fe^{2+}, Co^{2+}, Ni^{2+} and Cu^{2+} insulin for XAS experiments.

8.2.2 Preparation of insulin micro-crystals

Micro-crystal samples of Mn^{2+}, Fe^{2+}, Co^{2+}, Ni^{2+} and Cu^{2+} insulin were prepared for X-ray absorption spectroscopy measurements using the batch crystallization technique, Figure 8.1(*b*). Solutions containing 5 mg ml^{-1} metal free insulin, 0.01 M HCl, 5 mM M^{2+} acetate, 0.02 M sodium citrate and 15 vol.% acetone were mixed in acid washed tubes. The crystallization solutions were hereafter treated as in the preparation of Zn^{2+} T_6 micro-crystals, described in Section 7.2.2. The samples were analysed with XRPD, which showed that T_6 insulin micro-crystals only were obtained for preparations with Ni^{2+} and Cu^{2+}.

8.2.3 Single crystal diffraction

Single crystal diffraction data were collected at MAX-lab, Lund, Sweden, on beam line 911-2, MAX-II using a MarResearch CCD detector. The scattering properties generally improved after annealing the crystals a few times. Data collection and data processing statistics for all crystals are summarized in Table 8.2. The structures were refined using

Table 8.1: Reservoir compositions for single crystal growth of nickel, copper and iron insulin using the vapour diffusion technique.

	Ni^{2+} insulin	Cu^{2+} insulin	Fe^{2+} insulin[†]
Sodium citrate (M)	0.05	0.05	0.054
M^{2+} acetate (mM)	15	7.5	8
Acetone (vol.%)	15	15	25
pH	7.4	6.8	8.3

[†] Fe^{2+} insulin were grown using insulin solutions where the zinc was not chelated and removed prior to the crystallization experiments.

the peptide chain from the T_6 Zn insulin structure described in Chapter 7 and validated as described in Section 7.2.3. Refinement and validation statistics are listed in Table 8.3.

Table 8.2: Data collection parameters and processing statistics. All data were collected at beam line 911-2, MAX II, MAX-lab, Lund, Sweden.

	Ni^{2+} insulin	Cu^{2+} insulin	Fe^{2+} insulin
Data collection			
Wavelength (Å)	1.040	1.040	1.038
Temperature (K)	100	100	100
No. of frames	107	180	180
Oscillation range (°)	1	1	1
Exposure time per frame (s)	10	10	10
Detector distance (mm)	95	85	120
Resolution collected (Å)	1.45	1.35	1.78
Data processing			
Resolution (Å)	16.78–1.50 (1.54–1.50)†	16.87–1.45 (1.49–1.45)†	24.3–1.82 (1.92–1.82)†
No. of reflections	41817 (1946)†	35631 (2614)†	42583 (6175)†
No. of unique reflections	12710 (752)†	14336 (1069)†	7368 (1086)†
Redundancy	3.29 (2.59)†	2.49 (2.45)†	5.8 (5.7)†
Completeness (%)	97.6 (79.7)†	98.1 (100.0)†	99.9 (100)†
R_{sym} (%)‡	3.9 (13.7)†	5.0 (45.6)†	6.4 (75.4)†
$\langle I/\sigma(I)\rangle$	19.21 (10.45)†	14.00 (2.52)†	14.4 (2.2)†
Space group	$R3$	$R3$	$R3$
No. of molecules per asu	2	2	2
a (Å)	80.85	81.36	81.01
c (Å)	33.36	33.40	33.61
Average mosaicity (°)	0.52	0.29	0.42
Solvent content§ (%)	32.8	33.7	34.0

† Values in parentheses are for the outermost resolution shell.
‡ R_{sym} is defined as $\sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| \,/\, \sum_{hkl}\sum_i I_i(hkl)$, where $\langle I(hkl)\rangle$ is the mean intensity of a set of equivalent reflections.
§ Estimated by the program *Matthews* (Kantarjieff & Rupp, 2003).

T_6 Ni insulin

Two nickel atoms were inserted. The side chains of residues GlnB4.1, ValB12.1, LeuB17.1, CysA11.2 and GlnB4.2 were modelled in two alternating conformations. A total of 80 water molecules were inserted. Restrained refinement was carried out in *PHENIX* and hydrogen atoms were included. The atomic displacement factors for the peptide chain were refined by a combination of TLS refinement and isotropic refinement, using the TLS domains described in Section 7.2.3. Other were refined isotropically. Validation showed that only one residue SerA9.1 fell into the outlier region in the Ramachandran plot using the regions defined by Kleywegt & Jones (1996). This residue is involved in a hydrogen bonding to a HisB5.2 in a neighbouring hexamer.

Table 8.3: Data-refinement and validation statistics.

	Ni^{2+} insulin	Cu^{2+} insulin	Fe^{2+} insulin
Resolution cut (Å)	1.50	1.45	1.90
No. of atoms in the model			
Total non H-atoms	906	922	826
Total H-atoms	792	793	0
Total protein atoms	1616	1615	792
Total ordered water molecules	80	98	32
Total M^{2+}-ions	2	2	2
Type of M^{2+} siteI/siteII	Ni/Ni	Cu/Cu	Zn/Fe
***B*-factors**[§]			
Overall (Å^2)	29	24	38
Main chain (Å^2)	23	19	35
Side chains and water molecules (Å^2)	32	26	41
M^{2+}-ions (Å^2)	14	15	22
RMS deviation from ideal			
Bonds (Å)	0.016	0.014	0.008
Angles (°)	1.484	1.469	1.007
Ramachandran plot[‡]			
In core regions (%)	98.9	98.9	98.9
Outliers (%)	1.1	1.1	1.1
***R*-factors**[†]			
R	0.1819	0.1686	0.1952
R_{free}	0.2243	0.2044	0.2525

[§] The analysis of B-factors was done using *BAVERAGE* included in *CCP4* (Collaborative Computational Project, Number 4, 1994)

[†] R and $R_{free} = \sum ||F_{obs}| - |F_{calc}|| \,/\, \sum |F_{obs}|$, where F_{obs} and F_{calc} are the observed and calculated structure-factor amplitudes, respectively. R_{free} was calculated with a random 5 % subset of all reflections excluded from the refinement.

[‡] The definition of the Ramachandran plot regions according to Kleywegt & Jones (1996).

T_6 Cu insulin

Two copper atoms were inserted. The side chains of residues GlnB4.1, ValB12.1, LeuB17.1, CysA11.2 and ValB12.2 were modelled in two alternating conformations. 98 water molecules were inserted in total, and hydrogen atoms were included. The refinement procedure of coordinates and atomic displacement factors is hereafter analogue to the one for T_6 Ni insulin. Validation showed that only one residue SerA9.1 fell into the outlier region in the Ramachandran plot.

T_6 Fe insulin

Initially, two iron atoms were inserted in the structure. The iron atom placed in site I did not account for all the density, as seen in Figure 8.5(*a*), wherefore this site was remodelled

as a Zn-site, Figure 8.5(*b*). Residue AlaB30.1 and the side chains of LysB29.1 and LysB29.2 were disordered and could not be modelled. Two alternating conformations were modelled in the side chains of residues GluB13.1. In total 32 water molecules were inserted. The atomic displacement factors were refined isotropically. In the Ramachandran plot residue SerA9.1 fell into the outlier region.

8.2.4 X-ray absorption spectroscopy

Fluorescence XAS data were collected for Ni and Cu insulin at the Ni and Cu K-edges respectively, on beam line 811 at the synchrotron at MAX-lab, Lund, Sweden. The data collection and following analysis were carried out using the procedure described in Section 3.4.

8.2.5 Energy calculations on Cu sites

The 3 d^9 electron configuration of Cu^{2+} will predict a tetragonal Jahn-Teller distortion for six-coordinated copper, (Cotton *et al.*, 1995). As this geometry is inconsistent with the C_3 symmetry of the T_3 metal site, in which the metal ions normally are coordinated by six ligands, *ab initio* Hartree-Fock (HF) calculations were performed on a tetrahedral and an octahedral Cu cluster, shown in Figure 8.2. The HF calculations were performed using *SPARTAN'10* (Shao *et al.*, 2006) A hydroxide ion was chosen in the tetrahedral cluster instead of a water molecule to maintain the C_3-symmetry. The most stable conformation of each cluster was found by energy minimizing calculations using the HF method with the basis set 6-31G*. The energy of the individual components ($Cu(imz)_3$, 3 H_2O and OH^-)

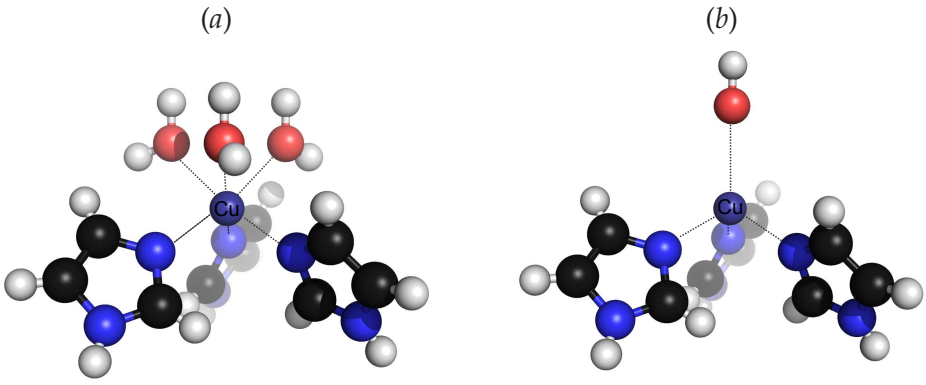


Figure 8.2: Cu clusters for Hartree-Fock calculations (*a*) A copper atom octahedrally coordinated to three imidazole and three water molecules. (*b*) A copper tetrahedrally coordinated to three imidazole molecules and one hydroxide ion. The hydroxide ion, which is isoelectronic with a water molecule, was chosen to maintain the C_3 symmetry. Carbon is shown in black, nitrogen in blue, oxygen in red, hydrogen in white and copper in dark blue.

were hereafter calculated and the energy differences between the complex and the sum of individual components were compared.

8.2.6 Comparison analysis

Searching the PDB reveals hexameric insulin structures crystallized with other metals than zinc. The structure of human insulin with both nickel (PDB-entry 3EXX), (Prugovečki *et al.*, 2009) and copper (PDB-entry 3TT8), have been solved to quite high resolutions 1.35 Å and 1.12 Å, respectively. These structures are compared with the bovine insulin structures presented here.

Nickel and copper coordination in small molecules were analysed by searching the CSD among structures containing nickel or copper coordinating to three monodentate ligands through sp^2-hybridized nitrogen atoms. For the nickel sites, only structures containing Ni coordinated to six atoms (3 nitrogen ligands and 3 oxygen atoms in facial isomerism) were included in the analysis. For copper, structures containing Cu coordinated to respectively four (3 nitrogen ligands and 1 oxygen atom in tetrahedral geometry) and six atoms (3 nitrogen ligands and 3 oxygen atoms) were included in the analysis. The isomerism among all six-coordinated copper structures were meridonial.

8.3 Results

8.3.1 Single-crystal X-ray diffraction structures

An extended conformation of the backbone of residues B1–B8 was observed in all three structures, which thereby adopt the T_6-conformation. The electron densities around residues B1.2–B3.2 were weak in all structures, but it was possible to trace the benzene ring of the PheB1.2, Figure 8.3, which resulted in a conformation of the N-terminal similar to the one observed in the T_6 zinc structure.

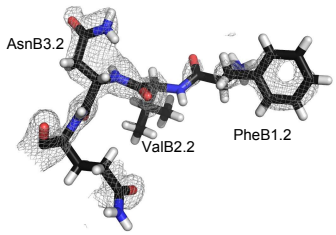


Figure 8.3: N-terminal of the B-chain in monomer 2 of Cu insulin structure. The electron density in the region is poor, but the location of the benzene ring in PheB1.2 is clear. The σ_A-weighted $2F_o - F_c$ maps have been contoured at 1.0 σ. Carbon is shown in black, nitrogen in blue, oxygen in red and hydrogen in white.

The coordination geometries of each metal site are shown in Figure 8.4 for Ni and Cu insulin and in Figure 8.5 for Fe insulin.

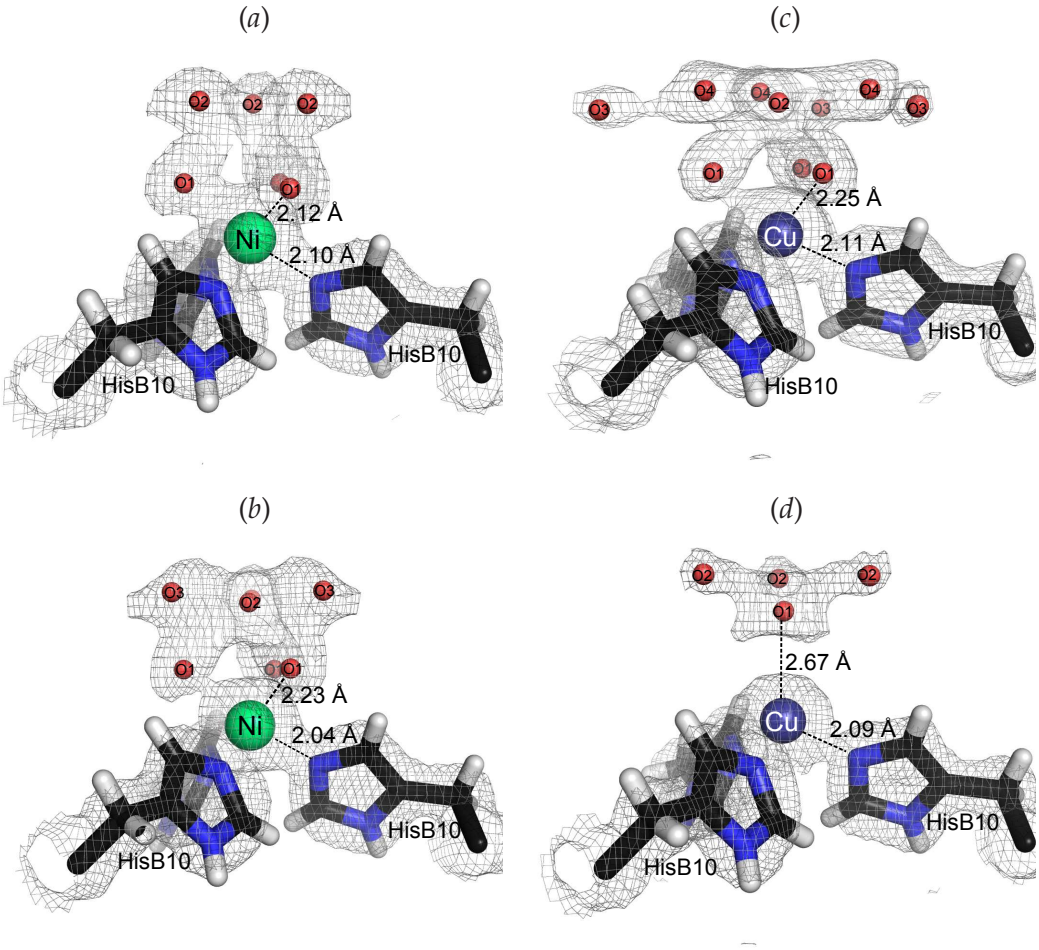


Figure 8.4: (*a*) Nickel site I and (*b*) nickel site II in the 1.50 Å Ni insulin structure. Ni^{2+} is octahedrally coordinated to three HisB10 and three water molecules (red) in both sites. (*c*) Copper site I and (*d*) copper site II in the 1.45 Å Cu insulin structure. Cu^{2+} is octahedrally coordinated in site I and tetrahedrally coordinated in site II. Distances to the first coordination sphere as determined by single crystal X-ray diffraction are shown. The σ_A-weighted $2F_o - F_c$ maps have been contoured at 1.0 σ. Carbon is shown in black, nitrogen in blue, oxygen in red, hydrogen in white, nickel in turquoise and copper in dark blue.

All metal ions bound in the structures are located on the three-fold rotation axis and coordinate to three $N^{\varepsilon 2}$-atoms of the three symmetry-related HisB10 residues. In all structures, electron densities were observed above the metal-ions, which are exposed to the solvent in these open T_3-sites. These densities were modelled with water molecules connected by hydrogen bonding networks, resulting in different coordination geometries for the different insulin derivatives.

In the nickel insulin derivative octahedral coordination geometry was observed at both Ni-sites, fulfilled by three symmetry related water molecules at respectively 2.12 and 2.23 Å from the Ni-atoms in the two sites, Figure 8.4(*a*) and (*b*). Copper was observed to adopt both octahedral (site I) and tetrahedral (site II) coordination geometry in the Cu insulin with Cu–O distances of 2.25 and 2.67 Å, respectively, Figure 8.4(*c*) and (*d*). Tetrahedral coordination of the metal in a T_3-site has been observed in other insulin structures at PDB, predominantly among low resolution structures, e.g. in the bovine T_6 insulin structure at 2.56 Å resolution (PDB-entry 2ZP6).

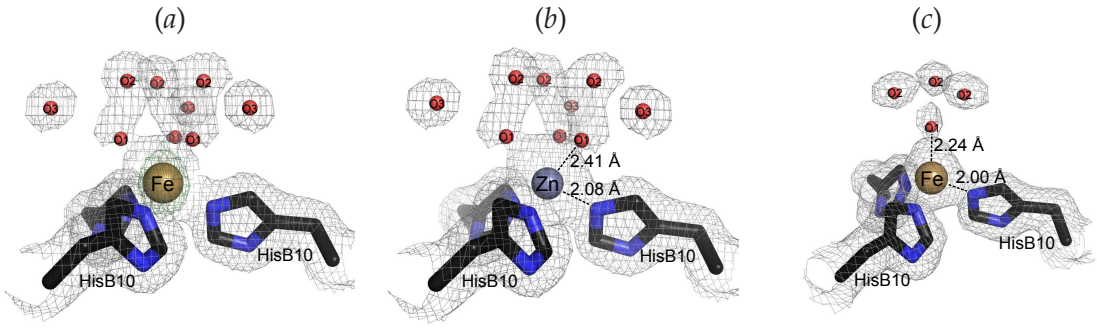


Figure 8.5: Iron and zinc coordination in the 1.90 Å Fe insulin structure. An octahedral coordination was observed in metal site I, which has been modelled with (*a*) Fe^{2+} and (*b*) Zn^{2+}. (*c*) Metal site II contains tetrahedrally coordinated Fe^{2+}. The σ_A-weighted $2F_o - F_c$ maps have been contoured at 1.0 σ and the difference $F_o - F_c$ maps at 3.0 σ. Carbon is shown in black, nitrogen in blue, oxygen in red, hydrogen in white, iron in bronze and zinc in light blue.

Also in the Zn/Fe insulin a dual octahedral/tetrahedral coordination was observed. The zinc ion located in site I is octahedrally coordinated fulfilled by three water molecules 2.41 Å from the zinc ion, Figure 8.5(*b*). In site II tetrahedral coordination is observed with Fe–O distance of 2.24 Å, Figure 8.5(*c*).

8.3.2 X-ray absorption spectroscopy

The extracted k^3-weighted EXAFS spectra and the modulus of the phase-corrected Fourier transforms of Ni and Cu insulin are presented in Figure 8.6. It is seen that the shape of the k^3-weighted $\chi(k)$ for Ni insulin has a high resemblance with that of T_6 Zn insulin indicating an octahedral coordination, whereas for Cu insulin the spectrum have more similarity to that of the R_6 Zn insulin, see Section 7.3.4. The distances and Debye-Waller factors were optimized from a restrained refinement, and are presented in Table 8.4. Both octahedral and tetrahedral models were tested. The best fit of the Ni insulin were obtained from a model with an octahedral geometry ($R = 17.30$ %), whereas a tetrahedral model

gave the best fit for Cu insulin ($R = 19.45$ %), as shown in Figure 8.6. Further EXAFS refinement statistics are summarized in Table 8.4.

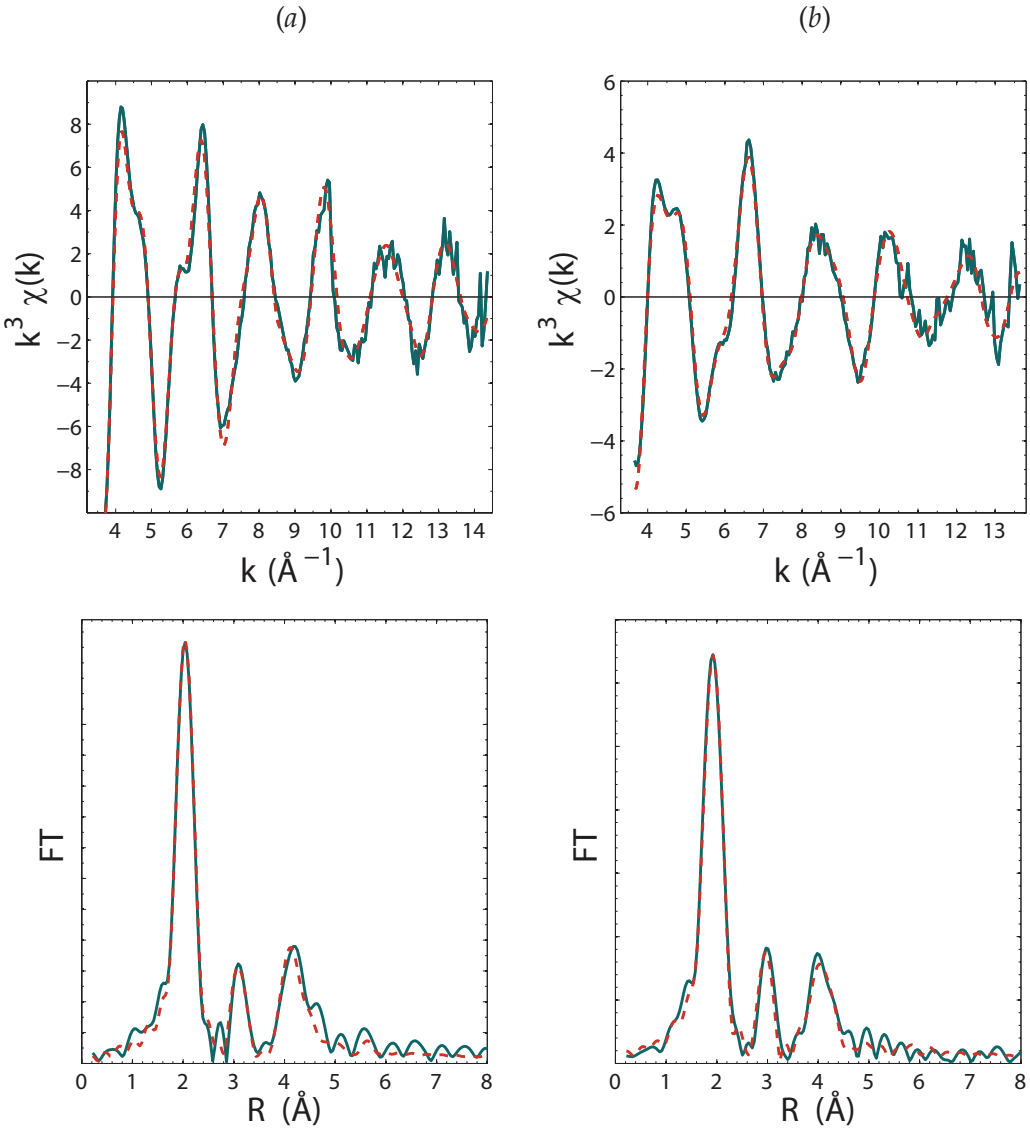


Figure 8.6: k^3-weighted EXAFS (top) and radial distribution functions calculated as the modulus of the phase-corrected Fourier transform (bottom) of (*a*) T_6 Ni insulin and (*b*) T_6 Cu insulin. Experimental spectra are shown in blue and simulated in red (dashed), using the parameters from the restrained refinement given in Table 8.4.

XANES spectra calculated by the FDM method for nickel and copper insulin are shown in Figure 8.7. The nickel spectrum was calculated using the coordinates from the model optimized by EXAFS. As seen in Figure 8.7(*a*) there is a good agreement between experimental and calculated spectra, ($R_{\text{xanes}} = 3.20$ %), and the XANES spectrum is similar to what has been observed for six-coordinated nickel, (Colpas *et al.*, 1991).

For copper insulin, XANES spectra were calculated for both tetrahedral and octahedral coordination with Cu–O distances of 2.18 Å. The weak white line observed in the experimental spectrum suggests a tetrahedral coordination of copper, Figure 8.7(*b*). In analogy to zinc, the XANES for six-coordinated copper gives stronger white lines, (Manceau & Matynia, 2010), which is in agreement with the calculated XANES for the octahedral cluster, see Figure 8.7(*c*). The discrepancy between experimental and calculated spectra for the tetrahedral copper is however larger than for nickel insulin ($R_{\text{xanes}} = 5.65$ %), Figure 8.7(*b*), due to a shoulder occurring at 8981–8984 eV. This shoulder originates from a peak located at 8983 eV, which can be assigned to the $1\,s \rightarrow 4\,p$ electronic transition of copper in oxidation state +I, (Kau *et al.*, 1987). The copper in the sample is thereby partly reduced. Examination of the raw data did not reveal any differences nor indication of evolution among the averaged raw XAS spectra. However, comparison with a fast scan spectrum collected just after mounting in the beam did not show any peak at 8983 eV, meaning that the sample was reduced at the beam line.

Table 8.4: Distances and Debye-Waller factors of Ni and Cu coordination in T_6 insulin as refined from EXAFS. The distances are compared to crystallographic values.

	Ni insulin			Cu insulin		
	XRD	Restrained EXAFS		XRD	Restrained EXAFS	
	$\langle R\rangle^{\S}$ (Å)	R (Å)	$2\sigma^2$ (Å^2)	$\langle R\rangle^{\S}$ (Å)	R (Å)	$2\sigma^2$ (Å^2)
$N^{\varepsilon 2}$ (HisB10)	2.07	2.079(7)	0.004(1)	2.11/2.09	1.984(5)	0.010(1)
$C^{\varepsilon 1}$	3.02	3.06(4)‡	0.011(2)	3.01/3.00	2.93(2)‡	0.011(5)
$C^{\delta 2}$	3.09	3.05(3)‡	0.011(2)	3.16/3.14	3.08(3)‡	0.011(5)
$N^{\delta 1}$	4.15	4.20(3)‡	0.016(4)	4.17/4.16	4.05(4)‡	0.025(6)
C^{γ}	4.21	4.21(3)‡	0.016(4)	4.26/4.24	4.20(3)‡	0.025(6)
C^{β}	5.65	5.59(4)‡	0.016(4)	5.68/5.67	5.64(6)‡	0.030(8)
O^{w1}	2.18	2.09(2)	0.015(3)	2.25/–		
O^{w2} (axial)	3.40	2.82(4)	0.020(9)	3.23/2.67	2.18(3)	0.030(11)
E_f (eV)		−0.50			0.08	
ϵ_v^2		0.5598			0.6951	
R_{exafs} (%)		16.28			18.28	
R_{dist} (%)		1.02			1.17	
R_{total} (%)		17.30			19.45	
N_{p}		21			19	
k-range (Å^{-1})		2.8–14.3			2.8–13.5	
w_{dist}		0.5			0.5	

§ Average distances of both Ni^{2+}-sites in the XRD structure.
† Distances to copper in Cu site I and Cu site II, respectively, in the XRD structure.
‡ The ϕ-angle (polar coordinates) was refined in order to allow some movement of the atoms in the restrained imidazole ring.
§ Average distances of both Ni^{2+}-sites in the XRD structure.

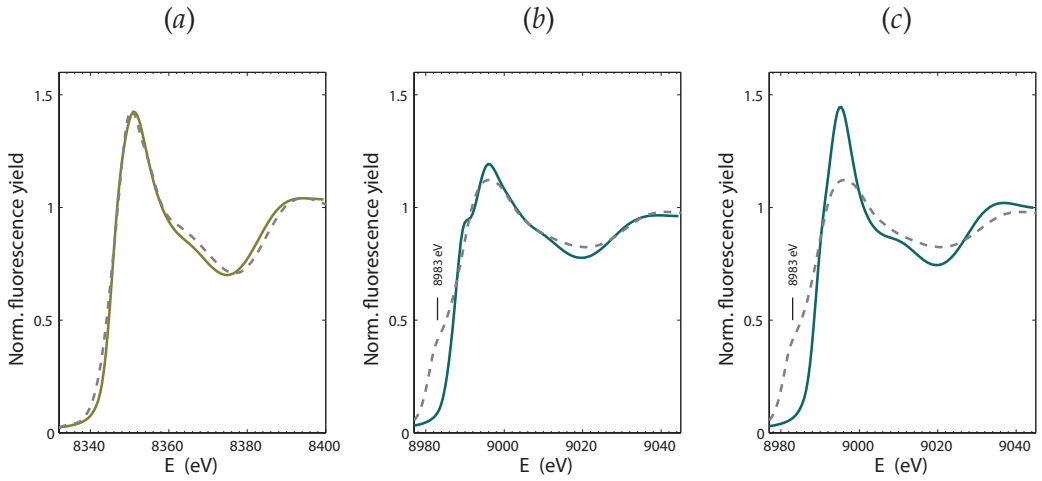


Figure 8.7: XANES calculated on a 4.5 Å cluster around each metal atom in (*a*) octahedrally coordinated nickel, (*b*) tetrahedrally coordinated copper and (*c*) octahedrally coordinated copper. FDM calculations were done on the coordinates from the EXAFS model for octahedral nickel and tetrahedral copper, and on an octahedral copper cluster fulfilled by three oxygen atoms located 2.18 Å from the Cu-atom. The calculated spectra are compared with experimental XANES (dashed). A peak located at 8983 eV indicates the presence of copper in oxidation state +I (Kau *et al.*, 1987).

8.3.3 Energy calculations on Cu sites

The energy minimization of the octahedral complex resulted in unreasonable long Cu–O distances (4.32 Å), whereas a Cu–O distance of 2.10 Å were found in the tetrahedral complex. This indicates that the octahedral geometry is disfavoured. The energy differences between the energy optimized complexes and the sum of individual components showed that the octahedral complex is not significantly stabilized, whereas the tetrahedral complex is stabilized by 6.6 eV.

8.3.4 Comparison with other reported M^{2+}-site geometries

The bond distances involved in the nickel and copper coordination observed in the high resolution structures of human Ni and Cu insulin, as well as in small molecules with similar coordination geometries are summarized in Table 8.5. The distances are compared to those from the XRD structures and EXAFS analysis. The results for the Zn T_6-structure is furthermore included for comparison. The CSD analysis of copper included both tetrahedral and octahedral copper coordination. For the octahedrally coordinated copper a tetragonal Jahn-Teller distortion was observed in all of the 5 structures found. Expanding the search to include polydentate ligands resulted in 163 structures. Most of these structures had meridonial isomerism and were tetragonally distorted. Only 3 structures had non-distorted facial isomerism, which was maintained by chelation from their polydentate ligands.

Table 8.5: Metal coordination geometries and distances to inner sphere ligands in each of the two metal sites for Ni, Cu and Zn insulin. Distances from the XRD structures and the EXAFS analysis are compared to human M^{2+} insulin structures deposited in PDB and small molecules with similar coordination geometries in CSD.

Insulin structures (PDB)	Resolution	Geometry†	M^{2+}–$N^{\varepsilon 2}$ (Å)†	M^{2+}–O^{w1} (Å)†
Nickel sites				
XRD this work	1.50 Å	oct/oct	2.10/2.04	2.12/2.23
Human Ni insulin‡	1.35 Å	oct/oct	2.05/2.06	2.15/2.18
EXAFS this work		oct	2.08	2.09
Copper sites				
XRD this work	1.45 Å	oct/tet	2.11/2.09	2.25/2.67
Human Cu insulin†	1.12 Å	oct/oct	2.04/2.06	2.08/2.24
EXAFS this work		tet	1.98	2.18
HF-calculation this work		tet	2.18	2.10
Zinc sites				
XRD this work	1.40 Å	oct/oct	2.09/2.10	2.31/2.27
Human Zn insulin†	1.00 Å	oct/oct	2.09/2.10	2.22/2.23
EXAFS this work		oct	2.07	2.14

Small molecule structures (CSD)	Number of structures	M^{2+}–N(sp^2) (Å)	M^{2+}–O (Å)
Octahedral Ni (3N+3O)	6§	2.09(3)	2.11(3)
Tetrahedral Cu (3N+1O)	7	2.01(4)	2.36(16)
Octahedral Cu (3N+3O)	5*	2.02(2)	2.29(24)
Octahedral Zn (3N+3X)	26	2.16(3)	2.16(4)

† Metal coordination geometry and M^{2+}–$N^{\varepsilon 2}$ distances in site I and site II, respectively.
‡ Best resolution structures of human insulin with Ni, Cu and Zn deposited in the PDB. PDB entries 3EXX, 3TT8 and 1MSO, respectively.
§ Only Ni-complexes with facial isomerism are included.
* All Cu-complexes showed tetragonal Jahn-Teller distortion and had meridonial isomerism.

8.4 Discussion

Hexameric insulin was successfully crystallized with divalent cations of nickel, copper and iron, and the structures were solved. The affinity of insulin to nickel and copper is large enough to crystallize hexamers with Ni^{2+} or Cu^{2+} in both metal sites, which was seen in the presented structures. The affinity to iron is lower and crystals with Fe^{2+} in both metal sites could not be obtained. Instead, crystals containing one Fe^{2+} and one Zn^{2+} was grown in presence of small amounts of zinc. Moreover, a site selectivity among the Fe^{2+} and Zn^{2+} ions is shown by the crystal structure. Furthermore, T_6 insulin micro-crystals were obtained with nickel and copper. The low affinity metals (cobalt, iron and manganese) are removed by the recrystallization procedure used, by which excess metal was eliminated in these XAS samples.

8.4.1 Conformation

T_6-conformation was observed in all structures, and the structures were compared with the zinc T_6-structure and other structures in PDB by superposing independent T_2 dimers in *SUPERPOSE* (Krissinel & Henrick, 2004) in which the C^{α}-displacements were minimized and the RMS calculated. All structures show high resemblance with the zinc T_6-structure presented in Chapter 7 as well as the bovine T_6 deposited in the PDB, (Smith *et al.*, 2005) with root mean squares below 0.4 Å. In analogy to the bovine zinc T_6-structure, comparison with the structures of human Ni and Cu insulin show larger discrepancies (RMS=1.345 Å and RMS=1.290 Å, for Ni and Cu respectively), due to a different conformation observed for the B1.2–B3.2 chain. The weak determination of the B1.2–B3.2 residues may be a consequence of a partially disordered N-terminal of the B-chain, and may indicate that the insulin partially undergoes conformational change. The dihedral angle A6C^{β}–A6S^{γ}–A11S^{γ}–A11C^{β} of the disulfide bridge CysA6–CysA11 is close to $+100°$ when the monomer adopts T-conformation and close to $-90°$ in R-conformations (Engels *et al.*, 1992). Based on these angles the double conformations of CysA11.2 in the Ni, Cu and Zn structures result in T/R ratio of approx. 70/30 %. For the iron insulin the resolution was not good enough to model alternate conformation for residue CysA11.2. Modelling of an alternate conformation of the entire B-chain N-termini in these structures was attempted, but could not be traced in the electron density. The partial disorder of these residues may be a consequence of the cooling process (Smith *et al.*, 2003), which is in agreement with the enhanced scattering properties of the crystals after annealing. According to Sreekanth *et al.* (2009) different conformations can be trapped by the binding different metals to insulin. This has been studied using a human Arg-insulin, in which an arginine has been attached to the N-terminal of the A-chain. These Arg-insulin structures were compared with other structures of human insulin and showed different conformations of the B-chain N-terminal. The data collection temperature and the crystallization conditions of the compared structures were however not analogous, wherefore the influence from these factors can not be ruled out.

8.4.2 Metal binding

For nickel insulin an octahedral coordination fulfilled by three water molecules is observed in both of the open T_3-sites. The nickel coordination is generally in very good agreement with the human Ni insulin structure (Prugovečki *et al.*, 2009) and other octahedrally coordinated Ni structures deposited in the CSD. The Ni–O^{w} distances as determined by EXAFS are slightly shorter than those from the XRD structure, and thereby reveals an almost undistorted octahedral coordination of nickel, which has been verified by an FDM calculation of the XANES spectrum.

For copper insulin a dual octahedral/tetrahedral coordination of copper was observed in the XRD structure, whereas both copper sites are octahedral in the human Cu insulin structure (PDB-entry 3TT8). In general, regular octahedral coordination of Cu^{2+} is not chemically favoured, since the $3\,d^9$ electron configuration will predict six coordinated Cu^{2+} to be tetragonally distorted by the Jahn-Teller effect. Furthermore such distortions will obstruct the three-fold symmetry. The result from the EXAFS analysis, in which both octahedral and tetrahedral models were tested, supports a tetrahedral coordination, which is also predicted to be more energetic favourable by the HF-calculations. The coordination geometries of Ni and Zn in the other insulin structures have hitherto been verified by XANES calculations. For Cu, XANES spectra calculated for both geometries further confirmed the tetrahedral coordination. A feature in the XANES spectra revealed that reduction of Cu^{2+} to Cu^{+} had taken place. Copper in oxidation state +I does not show any Jahn-Teller effect due to its $3\,d^{10}$ electron configuration, and is usually tetrahedral (Cotton *et al.*, 1995), which is in agreement with the EXAFS results. Thereby copper insulin is more sensitive towards radiation damage compared to nickel and zinc. As concluded in Chapter 7 the radiation doses are about two orders of magnitude lower for the absorption experiments compared to the diffraction experiments, meaning that it is highly possible that copper is reduced in the single crystals used for XRD, as well.

There is a good agreement between the Cu–$N^{\varepsilon 2}$ distances as determined from EXAFS and Cu–N distances in tetrahedrally coordinated copper among small molecule structures in the CSD. Compared to the Ni and Zn T_6-structures, the Cu–$N^{\varepsilon 2}$ are smaller and resemble the Zn–$N^{\varepsilon 2}$ distances observed in the R_6 structure, in which zinc is tetrahedrally coordinated, see Chapter 7. This further confirms the tetrahedral geometry.

8.5 Conclusive remarks

The structures of hexameric bovine T_6 insulin with nickel, copper and iron were solved from XRD data collected to 1.50 Å, 1.45 Å and 1.90 Å, respectively. Trapping different conformations with different metal ions were not observed for the bovine structures presented here, as all structures adopted T_6-conformation similar to that of zinc T_6 insulin.

A dual octahedral/tetrahedral coordination of zinc and iron, respectively, was observed in the XRD structure of iron insulin. In analogy to zinc T_6 insulin, octahedral coordination of nickel was observed in the XRD structure and verified by a calculated XANES spectrum. Distances to nickel were refined from EXAFS showing a very regular octahedral coordination. The XRD structure of copper insulin revealed a dual octahedral/tetrahedral coordination of copper in the two metal sites. According to the EXAFS analysis copper was determined to be tetrahedrally coordinated fulfilled by one water molecule, which is further supported by Hartree-Fock calculations and calculated XANES spectra. The XANES

spectrum further revealed that reduction of Cu^{2+} to Cu^{+} has taken place. The coordination distances as determined by EXAFS are generally similar to what is observed among small molecules with similar coordination geometries.

Chapter 9

Introduction to bone physiology and osteoporosis treatment with Sr

This chapter gives an introduction to bone physiology, osteoporosis and the treatment hereof with strontium containing medicaments, among these strontium ranelate, which is one of several drugs used for treatment of osteoporosis. The mechanism by which strontium is incorporated in bone tissue is not fully understood, but considering the physicochemistry of the bone composites possible locations of strontium have been suggested, and are described in this chapter. X-ray absorption spectroscopy is a well-suited technique for studying the strontium locations and thereby provide information on the incorporation of strontium. The final section presents an overview of X-ray absorption studies performed on bone material. The experimental work is reported in the following two chapters: The crystal structures of polymorphs of strontium ranelate (SrR) are presented in Chapter 10, and the localisation of strontium in bone samples from dogs treated with strontium is studied by XAS and is presented in Chapter 11.

9.1 The structure of bones

Bones have biomechanical and metabolic functions as they both provide structural support to the weight of the body, attachment points for mussels and tendons and make up a natural reservoir for many ions e.g. calcium and phosphate. More detailed information on the function, composition and structure of bone can be found in the reviews by Weiner & Wagner (1998); Seeman & Delmas (2006); Launey *et al.* (2010).

The hierarchically organized structure of bone, as seen in Figure 9.1, provides the material to be stiff in order to carry the weight from the body without breaking, and at the same time flexible to absorb energy by deforming rather that cracking. On the macro-scale bone tissue is classified as either cortical (compact) bone, which is found on the surface of all bones, and trabecular (spongy) bone, which fills the inside of many bones. The compact bone varies in thickness and is present in bone types where an increased strength is important, e.g. the femur bone. The spongy bone is found, where its flexible properties are needed, e.g. in vertebrae and femur neck. Many bones are composed of both types, for instance the skull cap which consists of a layer of trabecular bone between two layers of cortical bone. Such architecture minimizes the total weight of the skeleton, without loosing strength.

The compact bone is composed of osteons which have a concentric lamellar structure and surrounds the Haversian channels, in which blood vessels are protected. The lamellae are composed of sheets of differently orientated fibres, which are made up by linked collagen fibrils. Each fibril is made up of tropocollagen filaments and 50×25 nm plate-shaped crystals of calcium hydroxyapatite arranged in a geometrical pattern (the quarter-stagger model) and make up the basic building block in the bone structure. The tropocollagen consists of three peptide chains, each forming a helix, which coil up in a coiled-coil structure.

9.2 Composition of bones

The mineralized collagen fibrils are the basic building block in bone tissue and primarily consists of three components: Collagen, calcium hydroxyapatite and water, which are described in the following.

9.2.1 Collagen

The collagen in bone tissue is primarily of Bornstein & Traub type I and constitutes approx. 90 % of the proteins present in bone. The sequence of the collagen is made up by amino acid triads often following the patterns Gly–Pro–X or Gly–X–Hyp, in which X is an amino acid other that glycine, proline and hydroxyproline (Hyp). The high fraction of glycine in the sequence facilitates the formation of a left-handed helical structure. Tropocollagen is composed of three helical strands twisted together in a right-handed coiled-coil structure, which further stabilizes the protein. (Miller, 1984) The overall length of each tropocollagen filament is approx. 300 nm and the width is 1.5 nm. In the fibrils, the tropocollagen molecules are lined up head to tail in long rows with a gap of approx. 40 nm. Adjecent rows of tropocollagen are shifted by an offset of 68 nm, known as the quarter-stagger model (Hodge & Petruska, 1963), and the gaps are filled up by small crystals of CaHA, which highly increases the strength and toughness of the material.

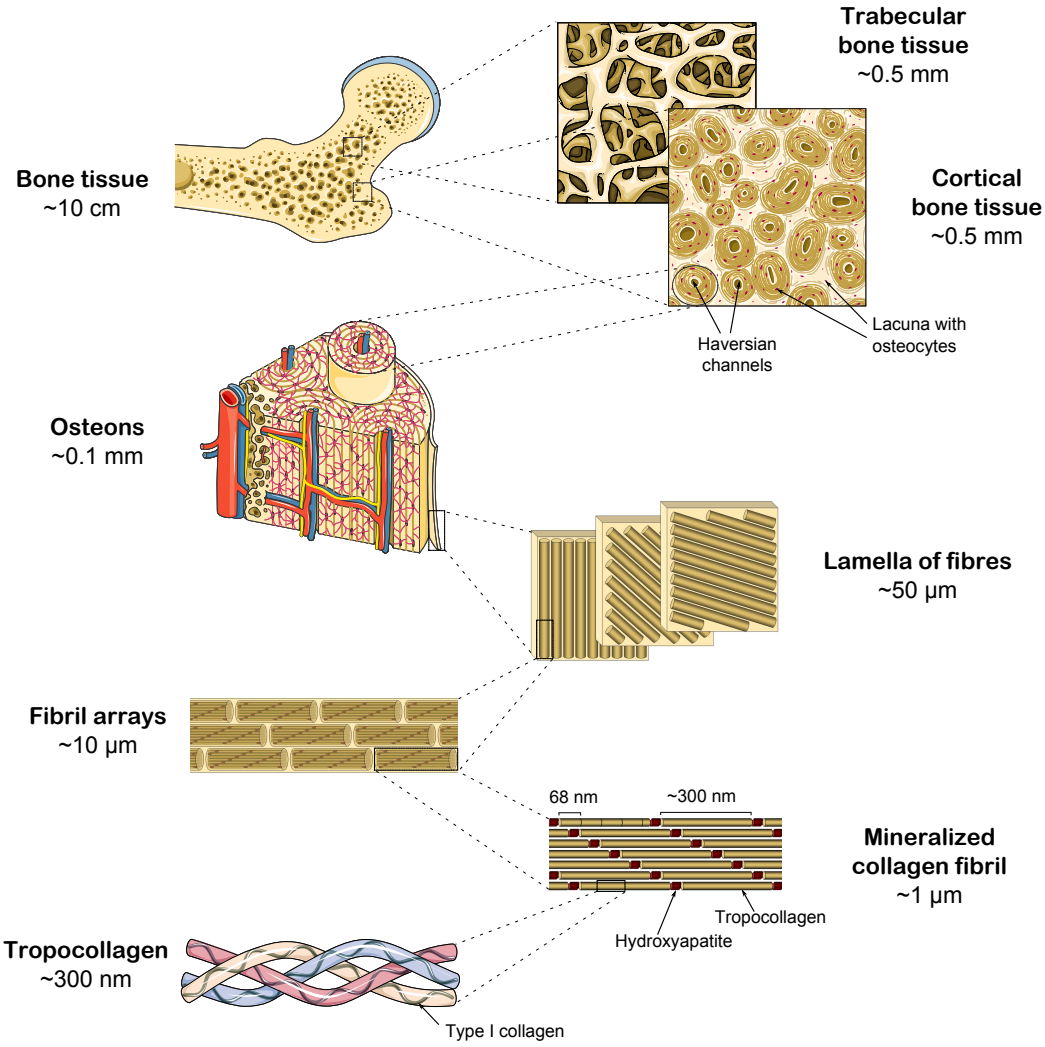


Figure 9.1: The hierarchically organized structure of bone. Bone is classified in two main categories: Cortical (compact) bone and trabecular (spongy) bone. The compact bone is built up of concentric lamellar structures of differently oriented fibres. Each fibre consist of an array of fibrils. Each fibril is composed of mineralized collagen in which tropocollagen and 50×25 nm plate-shaped crystals of calcium hydroxyapatite are systematically arranged. Three collagen molecules coil up to form the tropocollagen. Figure modified from (Launey *et al.*, 2010) using Servier Medical Art (Servier, 2012).

9.2.2 Calcium hydroxyapatite

Calcium hydroxyapatite (CaHA) has the stoichiometric formula $Ca_5(PO_4)_3OH$, but is often written as $Ca_{10}(PO_4)_6OH_2$, as there is two formula units in the crystallographic unit cell. CaHA forms hexagonal crystals belonging to the space group $P6_3/m$, (Kay *et al.*, 1964). The structure is seen in Figure 9.2 and contains two crystallographic different calcium sites, denoted Ca(I) and Ca(II), respectively. Four of the ten calcium ions in the unit cell are located in site I and the remaining six in site II.

The structure can be described as a hexagonal stack of PO_4^{3-} groups. Each phosphor atom coordinates four oxygen atoms in a tetrahedral arrangement, and the phosphate tetrahedra create two different tunnels parallel to the crystallographic c-axis. Ca(I) ions are surrounded by nine oxygen atoms and occupy the sites in the first tunnel, which is coinciding with the c-axis. Ca(II) ions are surrounded by seven oxygen atoms, and are arranged around the second channel, which is occupied by hydroxide ions. More information on biological apatites can be found in the extensive review by Wopenka & Pasteris (2005).

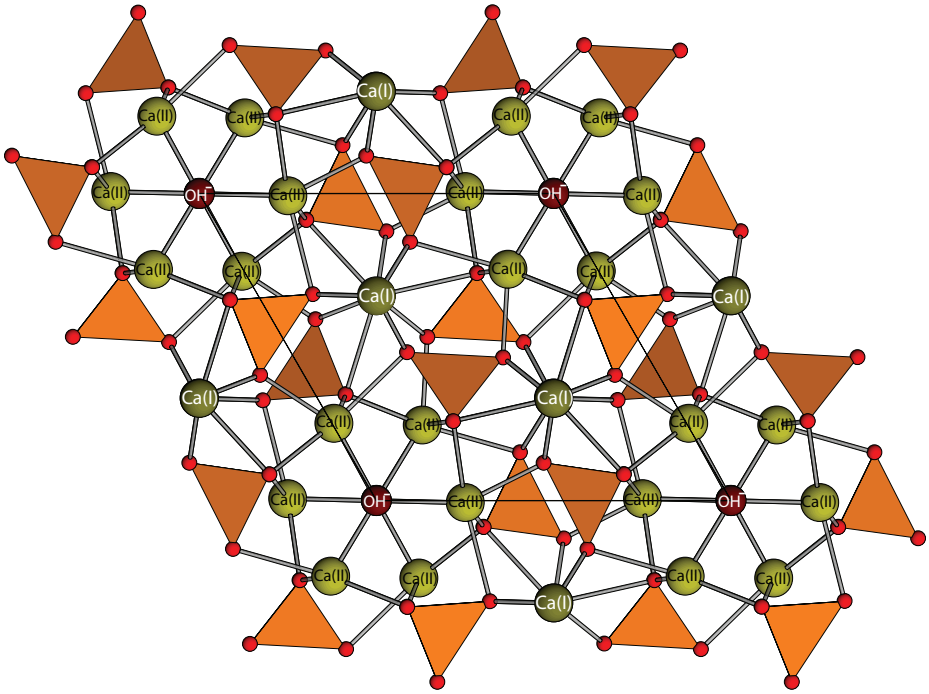


Figure 9.2: The structure of calcium hydroxyapatite viewed along the crystallographic c-axis. The PO_4^{3-} groups are represented by orange tetrahedra. The OH^- ions (dark red) are located on the ternary axis and surrounded by six Ca(II) ions (light green), which accounts for six out of the ten calcium sites in the unit cell. The remaining four calcium Ca(I) (dark green) are located in a channel parallel to the crystallographic c-axis. Figure is drawn from the coordinates by Kay *et al.* (1964).

The calcium ions in the apatitic structure can be substituted by other divalent cations such as Sr^{2+}, Mg^{2+}, Fe^{2+}, Zn^{2+} and Pb^{2+}. For substitution of calcium by strontium the unit cell parameters increase linearly with increasing strontium amounts, and the doping level can be estimated from the cell parameters (O´Donnell *et al.*, 2008). A preferential occupation among the two crystallographic sites upon substitution of Ca^{2+} by other cations depends on the size of the cation, (Tamm & Peld, 2006). Sr^{2+} preferentially occupies site I when present in low doping levels ($<$ 1 atom %), whereas an ideal and equal distribution (Sr(I):Sr(II) = 4:6) is observed at doping levels around 5 %, (Bigi *et al.*, 2007; Terra

et al., 2009). For higher degree of Sr-substitutions the Sr^{2+} preferably occupies the site II (O´Donnell *et al.*, 2008).

Also the anions can be substituted in the apatitic structure. The hydroxide can be replaced by fluoride, which better fits the size of the anion channel position. Fluorapatite occurs in high amount in tooth enamel, where it is advantageous, as it is less soluble than hydroxyapatite in acidic environments. Bone apatite contains approx. 7 %(w/w) carbonate, but the incorporation in apatite is still not fully understood. However, it seems to be generally accepted that CO_3^{2-} substitutes at the phosphate sites in biological apatites. (Wopenka & Pasteris, 2005)

9.2.3 Serum

Water is present between the triple helical molecules and fill up the gaps within and between the collagen fibrils and fibres. Water has an important role on the mechanical properties of bone, which are different for dry bone compared to those of wet bone.

9.3 Bone formation and resorption

Bone tissue is renewed throughout life, where small amounts of bone is removed and replaced by new bone. Most of the skeleton is replaced about every 10 years. The renewing process serves three main purposes: (*i*) reparation of micro-cracks in the skeleton caused by physical stress of the bones, (*ii*) maintenance of the elasticity of the bone, as old bone may become brittle over the years and (*iii*) exchange of calcium and phosphates to and from the blood.

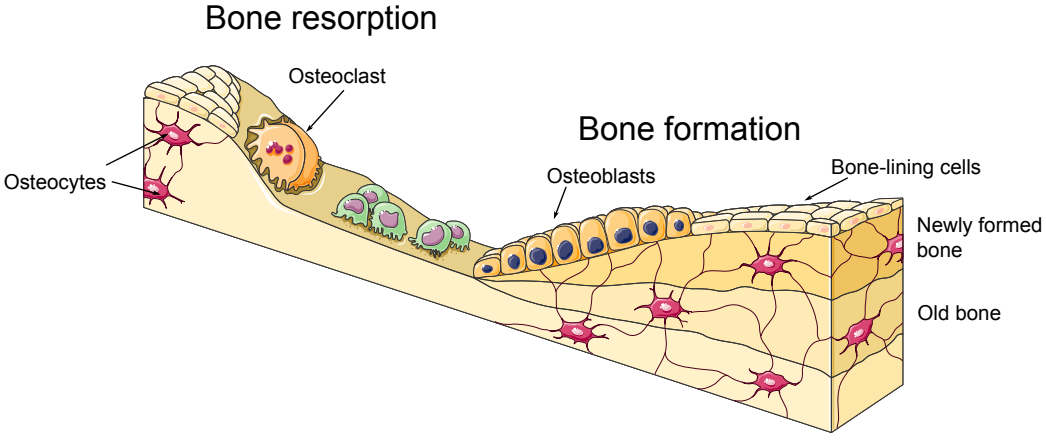


Figure 9.3: Bone remodelling process in which osteoclasts break down old tissue, whereas new bone is formed by osteoblasts. Some osteoblasts are trapped within the newly formed bone and converted into osteocytes. Figure was drawn using Servier Medical Art, (Servier, 2012).

The bone cells are embedded in the organic collagen matrix in the bone tissue, and are involved in the remodelling process, Figure 9.3. In the resorption process osteoclasts break down old tissue e.g. around a micro-crack. The osteoclasts carve tunnels around the damaged areas, which are then filled up by osteoblasts depositing new lamellar bone. After formation of bone some osteoblasts are entombed in the newly formed bone matrix where they are converted into osteocytes, others die or form new bone-lining cells, (Parfitt, 1996). Osteocytes are involved in the bone metabolism and probably sense bone deformation, thereby signalling the need for remodelling, (Han *et al.*, 2004).

9.4 Osteoporosis and treatment with Sr medicaments

Osteoporosis is characterized by an imbalance between the rates of bone resorption and bone formation, leading to a reduced bone mass and increased risk of fractures. The disease is a major worldwide health problem in the industrialized countries, and it has been estimated that more than 10 millions Americans older than 50 years (mostly women) are affected by osteoporosis. Treatment with strontium containing medicaments have shown to increase the bone mass in post-menopausal osteoporosis patients and to reduce resorption of old bone (Marie *et al.*, 2001).

At present, a general understanding of the therapeutic effect of strontium towards bone tissue is not fully achieved, but the combined action on bone metabolism sets Sr^{2+} unique as compared to existing osteoporosis therapies and has lead to an excessive research and development in strontium containing medicament during the last decade, (Schrooten *et al.*, 2003). Strontium ranelate, (SrR), 5-[bis(carboxymethyl) amino]-3-carboxy-methyl-4-cyano-2-thiophenecarboxylate, Figure 9.4, is one pharmaceutical compound for treating osteoporosis marketed as Protelos® by Servier®, (Horvath *et al.*, 2008).

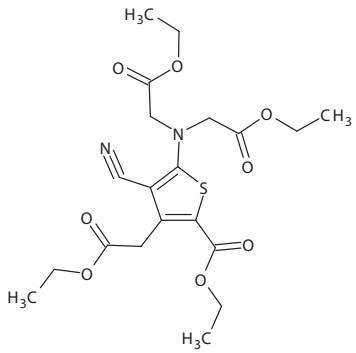


Figure 9.4: The structure of strontium ranelate, (5-[bis(carboxymethyl) amino]-3-carboxy-methyl-4-cyano-2-thiophenecarboxylate).

Clinical trials have shown that SrR reduces the risk of vertebral and non-vertebral fractures in women with osteoporosis (Meunier *et al.*, 2004; Reginster *et al.*, 2005, 2008). In experimental studies, SrR treatment has been shown to prevent loss of trabecular bone strength (Bain *et al.*, 2009) and to improve fracture healing in oestrogen-deficient rats (Ozturan *et al.*, 2011). Ammann *et al.* (2004) reported that 2 years of treatment with SrR increased trabecular and cortical bone mass and bone strength in healthy rodents.

9.5 Incorporation of Sr into bones

Strontium has a great affinity for bone and is readily incorporated into bone tissue. One of the key factors to understand the effect of strontium as a drug against osteoporosis is the mechanism by which strontium is inserted in bone tissue during treatment. Considering the physicochemistry of the main composites of bone, possible pathways of incorporation of ions have been suggested (Rey *et al.*, 2007; Cazalbou *et al.*, 2004), and three hypotheses summarize plausible locations of strontium (Bazin *et al.*, 2011), Figure 9.5. Sr^{2+} cations may be present in the serum surrounding the bones or be adsorbed on the surface of collagen (hypothesis I), for which Sr^{2+} will be surrounded by an inner coordination sphere consisting of oxygen atoms only. The Sr^{2+} cations may then be absorbed in collagen or bind to poorly crystalline Ca phosphates present on the surface of the nano-crystals (hypothesis II) and finally Sr^{2+} may incorporate in the calcium hydroxyapatite CaHA (hypothesis III). Determining the localisation of Sr^{2+} cations may help to confirm or disprove the suggested hypotheses.

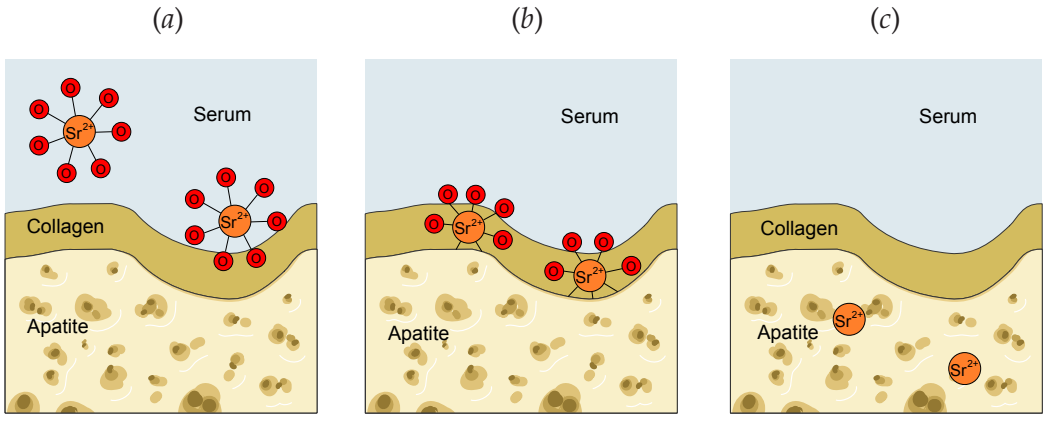


Figure 9.5: Schematic representation of the three hypotheses of the localisation of strontium in bone tissue. (*a*) Hypothesis I: Sr^{2+} is surrounded by an inner coordination sphere consisting of oxygen atoms in the serum, or the surface of collagen. (*b*) Hypothesis II: Sr^{2+} cations are absorbed in collagen or bind to poorly crystalline calcium phosphates present on the surface of the nano-crystals. (*c*) Hypothesis III: Sr^{2+} is incorporated in the apatite crystals.

The amount of strontium incorporated into bones depends on the dose level and the duration of treatment. Gender and bone types also influence on the incorporation since the bone turnover alter in different bone types. For instance the bone remodelling rate is higher in trabecular bone compared to cortical, (Dahl *et al.*, 2001).

9.6 XAS studies on apatites and bone structures

Through the years many different techniques have been used for *in vitro* studies of the distribution and localisation of Sr and other trace elements in calcified tissues. Particularly studies including X-ray fluorescence (XRF) (Behets *et al.*, 2005; Zhang *et al.*, 2005; Zoeger *et al.*, 2008; Bellis *et al.*, 2009; Pemmer *et al.*, 2011) and X-ray diffraction (XRD) (Capuccini *et al.*, 2009) or a combination hereof (Bradley *et al.*, 2007; Oste *et al.*, 2007), have be have been used to characterize the influence of trace metals on the spatial repartition of the apatitic structure and to localise these metals, however mostly on a macro-scale. Also small angle X-ray scattering (SAXS) (Li *et al.*, 2010) and scanning electron microscopy (SEM) (Verberckmoes *et al.*, 2004) count among the methods used for characterization of bone.

X-ray absorption spectroscopy (XAS) is well-suited for studying the local environment of Sr present in small concentrations on the micro- or nano-scale and has been used during the last 30 years for characterization of disorder in calcifications and apatite structures often in combination with other complementary techniques such as diffraction and vibrational spectroscopy (Bazin *et al.*, 2009a).

The first XAS-studies of bone, carried out on the Ca K-edge, revealed a higher disorder of the calcium ion environment present in bones compared to that of a synthetic prepared CaHA sample (Miller *et al.*, 1981; Binsted *et al.*, 1982). To interpret and model EXAFS data from bone samples quantitatively, comparison with well analysed reference compounds (Eanes *et al.*, 1981) were required. An EXAFS analysis on mouse bone (Harries *et al.*, 1988), by comparison with synthetically prepared CaHA with high (Harries *et al.*, 1986) and poor (Harries *et al.*, 1987) crystallinity, revealed that the average calcium location is best described by a partially crystalline hydroxyapatite.

Several years after an EXAFS study on human bone involving different bone types confirmed that calcium is present as disordered hydroxyapatite and that its chemical environment remains the same in all the analysed bone types Peters *et al.* (2000). In these studies the disorder of calcium is included in the EXAFS models as higher displacement factors, which cover both static and thermal (Debye-Waller) disorder. The static disorder among Ca ions is highest in bone tissue containing carbonates, which reduces the crystallite size, whereby a larger fraction of non-apatitic calcium may be present in the gaps between the HA-crystallites. Also the bone resorption process is closely related to the de-

gree of disorder in the CaHA and thereby closely related to the composition of the mineral phase (Peters *et al.*, 2000).

XAS was used to study disordered CaHA nano-crystals (Liou *et al.*, 2004) and calcium phosphates in non-apatitic environments, e.g. dicalcium phosphate, tricalcium phosphate and octacalcium phosphate (Eichert *et al.*, 2005). The subtle features in the XANES spectra allowed discrimination between the different species, which were found to be existent in synthetic poorly crystalline apatites. A recent study demonstrated that XANES could be used to assess and characterize calcium phosphates in human osteoarthritic articular cartilage (Nguyen *et al.*, 2011).

XAS has also been used to study the insertion of other elements in the HA structure, e.g. Sr (Rokita *et al.*, 1993), Cd (Sery *et al.*, 1996), Ga (Korbas *et al.*, 2004), Zn (Bazin *et al.*, 2009b), Ba, Pb (Sugiyama *et al.*, 2001). Both gallium and strontium have been identified to have therapeutic effects on bone, which is seen as inhibitory activity towards bone resorption. XAS has been used to assess these sparsely distributed metals in bone samples and to reveal how they are inserted in the bone structure (Korbas *et al.*, 2004; Bazin *et al.*, 2011).

9.7 Conclusive remarks

As a pharmaceutical agent, strontium has been shown to have positive effect against osteoporosis as it concomitantly reduces bone resorption and stimulates bone formation. Clinical trials have shown that treatment with strontium ranelate reduces the risk of vertebral and non-vertebral fractures in women with osteoporosis. The mechanism by which strontium is incorporated in bone tissue and how it influences on the remodelling processes are not fully understood. In principle, strontium can be localised in the serum, collagen and hydroxyapatite, which constitute the main components of the hierarchically organized structure of bone. X-ray absorption spectroscopy is well-suited for studying the strontium incorporation, and the method has been used during the last 30 years for characterization of calcifications and apatite structures.

Chapter 10

The crystal structures of polymorphs of Sr ranelate

This chapter describes the structures of different polymorphs of strontium ranelate. When crystallized from aqueous media strontium ranelate forms a nonahydrate. The crystal structure of strontium ranelate nonahydrate is presented in Paper IV, (Ståhl *et al.*, 2011) Appendix A.4. Upon dehydration strontium ranelate undergoes a series of phase transitions, resulting in several stable hydrates. These polymorphs have been studied with X-ray powder diffraction and thermogravimetric analysis (TGA) and described in this chapter. A manuscript on the polymorph studies is presently in preparation for publication.

10.1 Introduction

Strontium ranelate (SrR) is one of several drugs used for treatment of osteoporosis, (Schrooten *et al.*, 2003). The structural characterization of the compound has hitherto been sparse, only an XRPD pattern of the nonahydrate is reported as the α-crystalline form in the patent by Horvath *et al.* (2008). The main reason for this is the difficulties associated with growing single crystals large enough for XRD analysis, as crystals of SrR are highly sensitive to humidity and temperature, which is typical for many pharmaceutical compounds. SrR is known to form several hydrates containing a total of nine, eight, seven and four water molecules (Wierzbicki *et al.*, 1990). XRPD and TGA are well-suited methods to study the structural changes, which are observed upon dehydration.

10.2 Experimental

10.2.1 Synthesis of Sr ranelate

The synthesis was performed as an alkaline hydrolysis of ranelic acid tetraethyl ester followed by precipitation by strontium chloride, and was scaled down from the description by Kumar *et al.* (2009). 2.05 g ranelic acid tetraethyl ester was placed in a round bottom flask and suspended in 3 ml tetrahydrofuran (THF). 9.6 ml of an aqueous solution of 10 %(w/w) lithuim hydroxide was added. The LiOH solution was filtered before use. The reaction mixture was stirred for 4.5 hours at room temperature. The mixture was then filtered to remove insoluble material. A solution of 4.9 g strontium chloride hexahydrate in 14.8 ml water was slowly added to the filtrate while stirring vigorously. After a few minutes a white precipitate was formed and the reaction mixture was moderately stirred for 19.5 hours. The precipitate was separated by filtration. The crude product was analysed by XRPD.

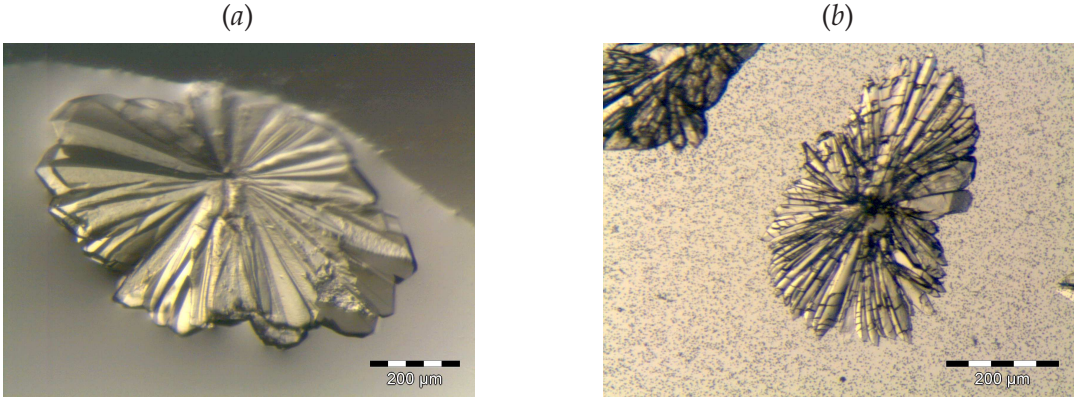


Figure 10.1: Crystals of strontium ranelate nonahydrate (*a*) before and (*b*) after evaporation of the mother liquor.

10.2.2 Crystallization of Sr ranelate polymorphs

Recrystallization was performed by dissolving 0.5 g of the crude strontium ranelate in 100 mL ultra-pure water (18.2 MΩ·cm, milli-Q, Millipore). The solution was heated to 353 K for 10 minutes and saturated with more crude strontium ranelate. The solution was then heated to 363 K for 5 minutes and filtered hot. The filtrate was cooled to room temperature and left for evaporation in a petri dish covered with parafilm with holes in it. After evaporation to almost dryness small crystals of 30–50 µm were observed. Growing larger crystals was performed by successive recrystallizations, where the temperature at which the crystals were redissolved was stepwise decreased. The crystals were washed and recrystallized a second time in 100 ml water at 333 K. A third recrystallization was carried

out at 323 K. After a fourth recrystallization at 313 K and slowly evaporation controlled by covering the petri dish with parafilm with holes in it, clusters of elongated crystals with dimensions up to 500 μm were observed, Figure 10.1(*a*).

When the crystals were removed from the mother liquor they rapidly degraded into smaller fragments with less crystal water, Figure 10.1(*b*). A wetted crystal was therefore quickly mounted on a goniometer and transferred for single crystal X-ray data collection where it was cooled to 120 K, as described in Paper IV included in Appendix A.4.

Samples for XRPD analysis were taken from wet crystals and treated differently in order to obtain different hydrates. Sample I: Crystals were dried at RT at ambient humidity, resulting in disintegration of the crystals which broke into small units of 10–30 μm. Sample II: Crystals were dried over silica in a desiccator for 2 days at RT, resulting in a powder of sub-micrometer crystals. Sample III: Crystals were dried by heating to 400 K.

10.2.3 X-ray powder diffraction

X-ray powder diffraction patterns for the different samples were collected on a Huber G670 diffractometer using Cu$K\alpha_1$ radiation ($\lambda = 1.5406$ Å). Data collection times and temperatures are listed in Table 10.1. The SrR nonahydrate was prepared in a glass capillary (0.5 mm in diameter) and wetted with water to ensure saturation. Samples dried at RT were prepared on flat specimen holders. The sample dried by heating was prepared in a glass capillary and dried at 400 K on the diffractometer using a nitrogen cryostream (Oxford Cryosystems).

An *in situ* XRPD experiment upon dehydration was carried out at beam line 711, MAX-lab, Lund, Sweden, using a wavelength $\lambda = 1.26406$ Å. A powder sample of crystals dried at RT and ambient humidity were mounted in a capillary and heated from 323 K to 473 K in steps of 5 K. The data collection time was 5 min per pattern with intervals of 3 min between the measurements to raise an equilibrate the temperature, resulting in a heating rate of 0.63 K min^{-1}.

10.2.4 Thermogravimetric analysis

Thermogravimetric analysis was performed on a sample of dry crystals on a Netzsch STA 409 PC/PG instrument. The sample was heated from 293 to 773 K with a rate of 1 K min^{-1}.

10.3 Results

An overview over the phase transitions occurring upon dehydration was obtained from the *in situ* XRPD experiment. The diffraction patterns are compared with the TGA results

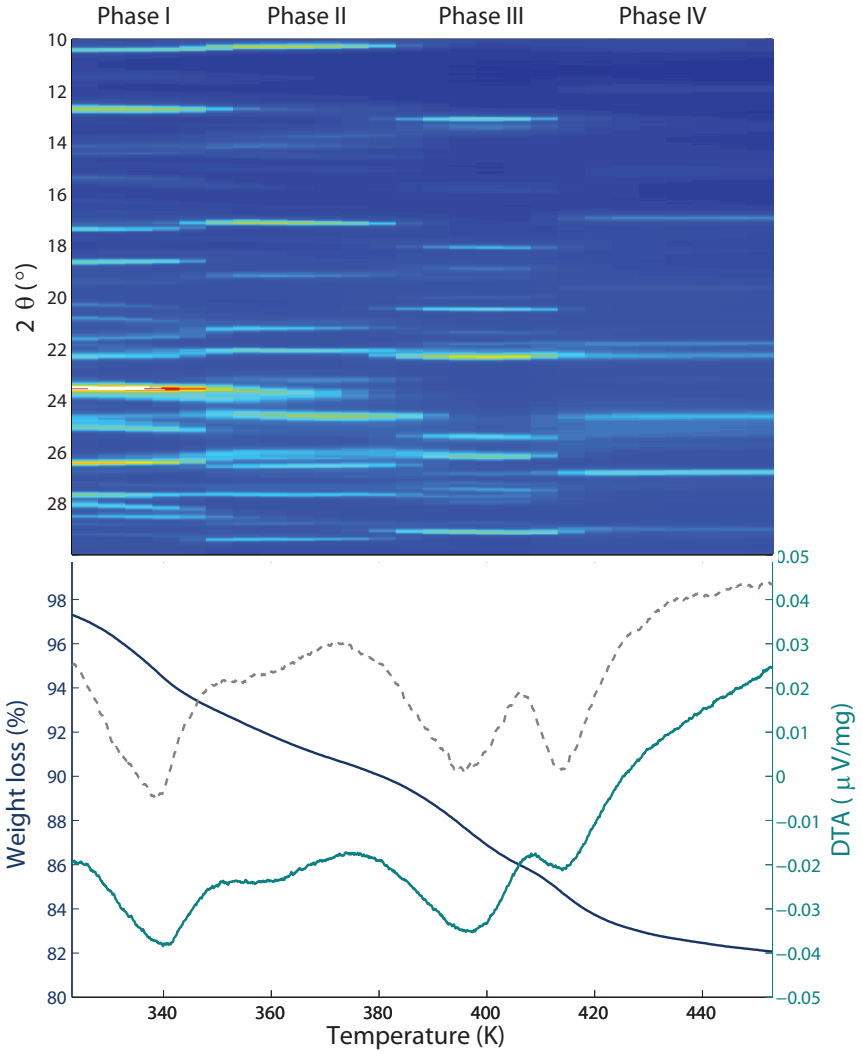


Figure 10.2: XRPD patterns and TGA of strontium ranelate collected upon dehydration by ramping the temperature with a rate of 1 K min^{-1}. The XRPD patterns are shown as a surface plot where white/yellow corresponds to high Bragg peak intensity. The patterns are collected using synchrotron radiation with a wavelength $\lambda = 1.2641$ Å, but are converted to Cu $K\alpha_1$ radiation ($\lambda = 1.5406$ Å). The weight loss in % determined by TGA is plotted in blue (left axis) and the DTA in turquoise (right axis). The first derivative of the weight loss (DTG) is included (dashed line).

and shown in Figure 10.2. Four dry phases were observed, which are referred to as phase I, II, III and IV in the following.

Powder diffraction patterns of the crude product and different hydrates of recrystallized SrR are shown in Figure 10.3. The crude product consists of SrR nonahydrate and impurities of primarily $Sr(OH)_2{\cdot}8H_2O$, Figure 10.3(*a*). It was found crucial to remove these impurities to obtain larger crystals. The wet crystal sample was identified as the nonahydrate, Figure 10.3(*b*). The nonahydrate undergoes phase transitions upon drying. Two

stable phases were observed when drying the crystals at ambient humidity (sample I) and over silica (sample II). These phases were respectively identified as the phase I and phase II, by comparing the XRPD patterns shown in Figure 10.3(*c*) and 10.3(*d*), respectively, with those from the *in situ* experiment. Adding water to sample I and sample II resulted in regeneration of the nonahydrate phase. An intermediate phase occurs when drying the crystals at 400 K, Figure 10.3(*e*). This phase was identified as phase III.

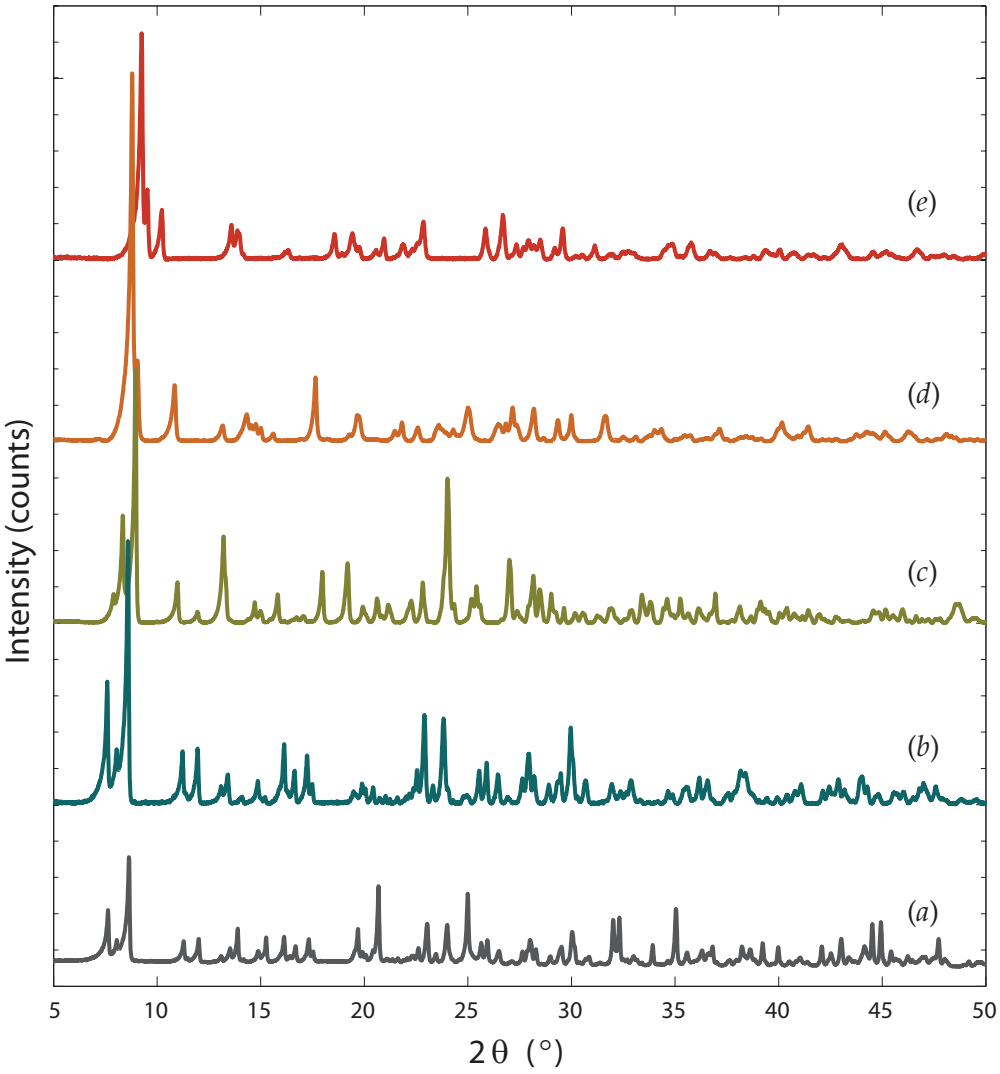


Figure 10.3: XRPD patterns of the crude product and different hydrates of recrystallized strontium ranelate collected on a Huber G670 diffractometer using Cu $K\alpha_1$ radiation ($\lambda = 1.5406$ Å). (*a*) The crude product. (*b*) Wet SrR after recrystallization. (*c*) Recrystallized SrR dried at RT and ambient humidity (sample I). (*d*) Recrystallized SrR dried at RT over silica (sample II). (*e*) Recrystallized SrR dried at 400 K at the diffractometer (sample III).

The unit cell parameters were determined by a full pattern profile fit of the four XRPD patterns shown in Figure 10.3(*b*)–(*e*), and the respective cell volumes were calculated, Table 10.1.

Table 10.1: Data collection specifications for XRPD experiments and unit cell parameters for hydrates of SrR as determined by full pattern profile fits.

	SrR nonahydrate	Sample I	Sample II	Sample III
Temperature (K)	293	293	293	400
Data collection time (min)	120	120	120	72
a (Å)	8.4308(2)	8.2754(2)	8.2085(3)	8.6453(2)
b (Å)	12.4360(3)	12.4669(4)	11.3906(5)	10.5297(4)
c (Å)	12.7717(3)	11.5163(3)	11.3921(6)	10.8499(3)
α (°)	110.239(1)	112.545(1)	117.637(2)	118.305(2)
β (°)	96.998(1)	89.869(2)	94.794(3)	91.538(3)
γ (°)	105.813(1)	102.398(1)	92.535(4)	90.210(3)
V (Å^3)	1173.7(2)	1067.5(2)	936.1(4)	871.3(3)
Estimated no. of H_2O	9	7	5	3

10.4 Discussion

The crystal structure of SrR nonahydrate, presented in Paper IV (Appendix A.4), is a metal organic framework (MOF) with ranelate anions interconnecting layers of polyhedra of eight and nine-coordinated Sr^{2+}-cations, thereby forming channels which contain water. Four of the nine water in the nonahydrate are not directly coordinated to strontium but found in these channels. The nonahydrate dehydrates very easily, which was indicated by the crystal structure, as it only contains 8.79 out of nine molecules of water. Thereby 0.21 water molecules were lost during crystal mounting.

The dehydration was followed by XRPD and TGA during heating and four dry phases were observed. The exact water content present in each phase cannot be determined, as the initial water content is unknown in the experiments. However an estimation can be done by comparing the relative weight losses from the TGA with the XRPD patterns. Each water molecule is estimated to correspond to a weight loss of approx. 2.8 %. Naming the phases after the highest possible number of water, results in an assignment of phase I–IV to heptahydrate, pentahydrate, trihydrate and dihydrate, respectively. The estimated number of water in each phase is included in Table 10.1. The loss of water is also reflected in the decreasing unit cell volume.

During the transitions between nonahydrate, phase I and phase II the unit cell lengths a, b and c are successively decreased as a consequence of the compression of solvent channels. Both phase I and phase II are stable and can be obtained at RT by varying the hu-

midity. The phase transition is fully reversible, as the nonahydrate can be obtained from phase I and phase II by addition of water. Conversely, the transition to phase III is irreversible and heating is required. A larger rearrangement in the structure is indicated by a sudden increase of the unit cell length a. This intermediate phase is undergoes a transformation to phase IV when the equilibrated at 400 K.

10.5 Conclusive remarks

Strontium ranelate was synthesized by an alkaline hydrolysis of ranelic acid tetraethyl ester followed by precipitation by strontium chloride. The growth of larger crystals was optimized by successive recrystallization. The structure of SrR nonahydrate was solved from single crystal XRD. The crystals were found to be very sensitive to humidity and temperature, and deteriorate easily upon dehydration. Besides the nonahydrate four other hydrates were observed from X-ray powder diffraction, and their water contents estimated by thermogravimetric analysis from which the phases were identified as heptahydrate, pentahydrate, trihydrate and dihydrate.

Chapter 11

Studies of incorporation of strontium into bone using XAS

The work presented in this chapter includes a study of the localisation of strontium incorporation into bone tissue from dogs using XAS. Two series of XAS-studies were performed one at room temperature and one at 100 K. The results presented here is for the 100 K study, as the data quality improves when cooling. A manuscript has been written on the results presented in this chapter and will be published soon. The EXAFS analyses and results from the RT reference samples (strontium chloride hexahydrate, strontium disalicylate dihydrate, strontium hydroxyapatite and strontium doped calcium hydroxyapatite) are included in Appendix C. EXAFS results and analyses of RT bone samples series from dog calvaria (52 weeks treatment), dog femur (4 weeks treatment) and rat spine (26 weeks treatment) are presented in Appendix D.

11.1 Introduction

The disorder of calcium in bone material has in previous XAS studies been described by increased Debye-Waller factors, (Harries *et al.*, 1988; Peters *et al.*, 2000). The classical shell-models used for EXAFS interpretation may be insufficient for describing static disorder among calcium or strontium in bone tissue (see Appendix D), wherefore a new approach in which XAS spectra are analysed by a linear combination of spectra of selected composites present in the bone sample, has been used. This may give a more realistic chemical

interpretation of the XAS data as it allows estimation of the relative distribution of strontium in the different bone composites.

11.2 Experimental

11.2.1 Bone samples

Femur (thigh) bones and calvariae (skull-caps) were dissected from beagle dogs treated with Sr malonate for 4 and 52 weeks, respectively (Scantox, Ejby, Denmark). The bone dissection and preparation was performed by Anders C. Raffalt and further described in Raffalt (2011). The Sr contents were determined by ICP-MS and is described in Raffalt *et al.* (2008) and Raffalt (2011). Treatment dose, treatment time and bone Sr concentrations are summarized in Table 11.1.

Table 11.1: Doses and content of Sr in the bone samples analysed.

	Sample	Sr ranelate dose (mg kg^{-1} day^{-1})	Sr concentration (mg g^{-1})
Femur, 4 weeks, low dose	59147-7	300	5.7 ± 0.5
Femur, 4 weeks, high dose	59147-15	1000	10.4 ± 1.6
Calvaria, 52 weeks, low dose	62346-16	100	16.2 ± 0.2
Calvaria, 52 weeks, high dose	62346-25	1000	36.7 ± 2.0

11.2.2 Reference compounds

Three reference compounds representing the Sr coordination in each of the three phases in the incorporation hypothesis (see Section 9.5) was used in this study.

A 0.3 M solution of $SrHPO_4$ was made by dissolving $SrHPO_4$ in 10 % phosphoric acid. Sr-soaked collagen was prepared from a suspension containing 35 mg ml^{-1} Bornstein & Traub type I collagen (Sigma C9879) and 0.1 M $SrCl_2 \cdot 6H_2O$ in isotonic NaCl solution (0.9 %). The collagen was soaked for four weeks under agitation, and recovered by filtration and washed six times with isotonic NaCl solution. Calcium hydroxyapatite with approximately 5 % (w/w) Ca substituted by Sr was prepared by modifying the synthesis of pure calcium hydroxyapatite reported by Kumta *et al.* (2005) to include $SrCl_2$ as a starting material in an amount corresponding to 5 %(w/w) Sr substitution of Ca. The Sr concentration in the apatite was confirmed by ICP-MS analysis and further analysed by X-ray powder diffraction (XRPD) and Rietveld refinement.

11.2.3 X-ray powder diffraction

XRPD data for Sr-doped CaHA were collected on a Huber G670 diffractometer using Cu $K\alpha_1$ radiation ($\lambda = 1.5406$ Å) for 30 minutes at RT. XRPD data for bone samples were collected at MAX-lab beam line 711, Lund, Sweden, on a Huber G670 diffractometer using synchrotron radiation with a wavelength $\lambda = 1.2641$ Å. Data were collected on the bone samples for 20 minutes at 323 K and on a femur sample after heating to 1073 K.

11.2.4 X-ray absorption spectroscopy

For XAS analysis the Sr phosphate solution and the collagen was mounted in a 1 mm thick sample holder for XAS, see Chapter 5. Preparation of Sr-doped CaHA was done by grinding the samples in a mortar and pressing the powder into tablets of approximately 1 mm thickness. A representative fragment of each bone sample was crushed to a powder and pressed to a tablet with a pressure of 10 tons before it was mounted on the sample holder of the cryostat.

Sr K-edge X-ray absorption spectra were recorded on beam line 811 at MAX-lab, Lund, Sweden, using a Si(311) double-crystal monochromator detuned 30 % at 16750 eV, further specifications can be found in Carlson *et al.* (2006). Fluorescence data were collected at 100 K using a liquid nitrogen cryostat (Oxford Instruments, ITC503) and a passive implanted planar silicon detector. Two or three spectra were collected for each sample. The data collection ranges are specified in Table 11.2, and the data collection times were as follows: pre-edge data (150–30 eV before the edge) were collected in steps of 5 eV for 1 s, the edge (from 30 eV before to 30 eV beyond the edge) in steps of 0.3 eV for 1 s and the EXAFS (from 30 eV beyond the edge) in steps of 0.05 Å^{-1} for 1–15 s. No degradation or radiation damage of the samples was observed.

Table 11.2: Number of spectra and data collection ranges for the samples.

	Number of spectra	Data collection range (eV)
Sr phosphate solution	2	15955–17065
Sr soaked collagen	3	15955–16600
Sr doped CaHA	2	15955–17065
Femur, 4 weeks, low dose	3	15955–16900
Femur, 4 weeks, high dose	3	15955–17050
Calvaria, 52 weeks, low dose	3	15955–16900
Calvaria, 52 weeks, high dose	3	15955–17050

Averaging, background subtraction and normalization were carried out using *WinXAS* (Ressler, 1998) as described in Section 3.4. All spectra were energy-calibrated using a simultaneously collected spectrum of strontium chloride hexahydrate for which the main edge inflection point was assigned to 16109.8 eV (Ubgade & Sarode, 1987).

11.3 Results

11.3.1 X-ray powder diffraction

The X-ray powder diffractograms for calvaria, femur bone before and after pyrolysis at 1073 K and 5 % Sr-doped CaHA are seen in Figure 11.1. The powder patterns collected at the synchrotron have been converted to Cu $K\alpha_1$ radiation ($\lambda = 1.5406$ Å). Common for all samples is that they consist of hydroxyapatite, however with different crystallite size, which is seen as different peak broadenings. After heating the femur sample to 1073 K the organic matrix is pyrolysed leaving the hydroxyapatite phase back.

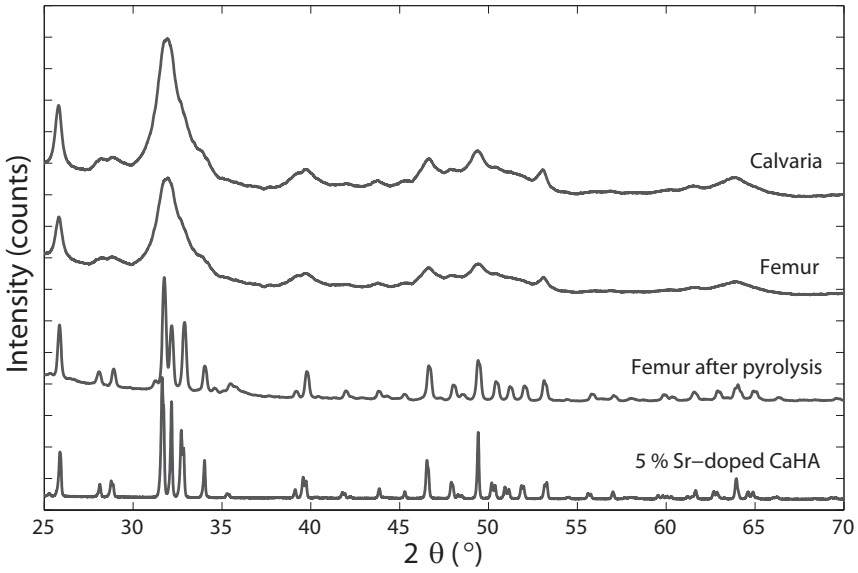


Figure 11.1: XRPD patterns of bone samples and Sr-doped CaHA. From top calvaria, femur bone, femur bone after heating to 1073 K and 5 % Sr-doped CaHA. Patterns collected at $\lambda = 1.2641$ Å are converted to Cu $K\alpha_1$ radiation ($\lambda = 1.5406$ Å).

Rietveld refinement of the 5 % Sr-doped CaHA showed an almost equal occupation of Sr^{2+} (0.048(4) and 0.053(5)) among calcium site I and II, respectively. This further comfirms the 5 % doping level and is in accordance to previous observations (Bigi *et al.*, 2007; Terra *et al.*, 2009).

11.3.2 X-ray absorption spectroscopy

XANES spectra for all bone samples and the three reference compounds are shown in Figure 11.2. In the XANES region, the bone samples are very similar. For SrHA a small feature is clearly observed at 16135 eV, which could be assigned to outer coordination spheres (Bazin *et al.*, 2011) and thereby verify the incorporation of Sr^{2+} cations in apatite.

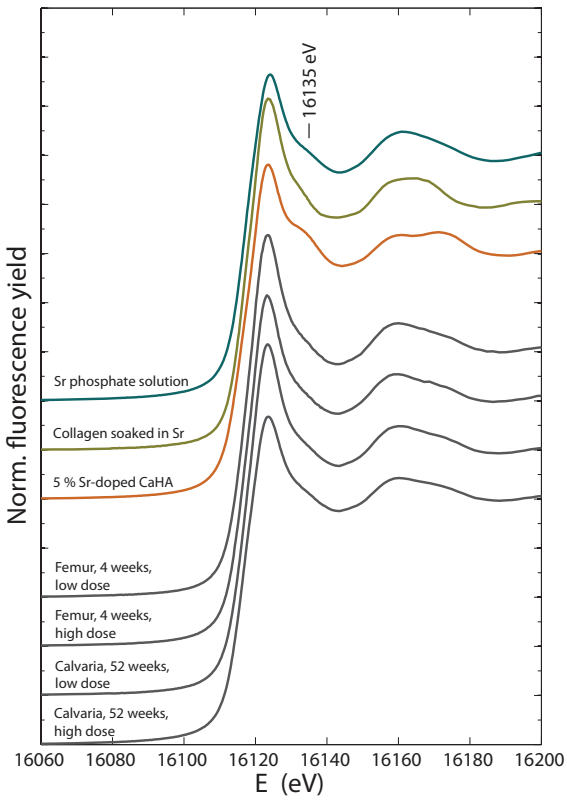


Figure 11.2: XANES spectra collected at the Sr K-edge of Sr-treated bone samples and reference compounds.

The extracted EXAFS and the radial electron density distributions are shown in Figure 11.3. The EXAFS of the three reference compounds are very different. For Sr^{2+} cations in solution, only the inner coordination sphere is observed, whereas for SrHA and Sr-soaked collagen several coordination spheres are observed, implying that the Sr^{2+} cations are coordinated to and inserted in the solid matrix.

It is clearly evident that strontium in the bone samples is partly incorporated in the apatite structure. Likewise, it is also clear that the Sr^{2+} cations are not exclusively present as apatite. As XAS probes the average localisation of the absorbing atom type the EXAFS spectra for the bone samples, $\chi_{\text{Bone}}(k)$, were modelled by a linear combination of EXAFS spectra of the three representative reference compounds (SrHA, Sr soaked collagen and Sr phosphate solution). The modelling was carried out using the linear combination fit procedure included in *ATHENA* (Newville, 2001; Ravel & Newville, 2005), which uses a function of the type

$$\chi_{\text{Bone}}(k) = a_{\text{HA}} \cdot \chi_{\text{HA}}(k) + a_{\text{Col}} \cdot \chi_{\text{Col}}(k) + a_{\text{Sol}} \cdot \chi_{\text{Sol}}(k) \tag{11.1}$$

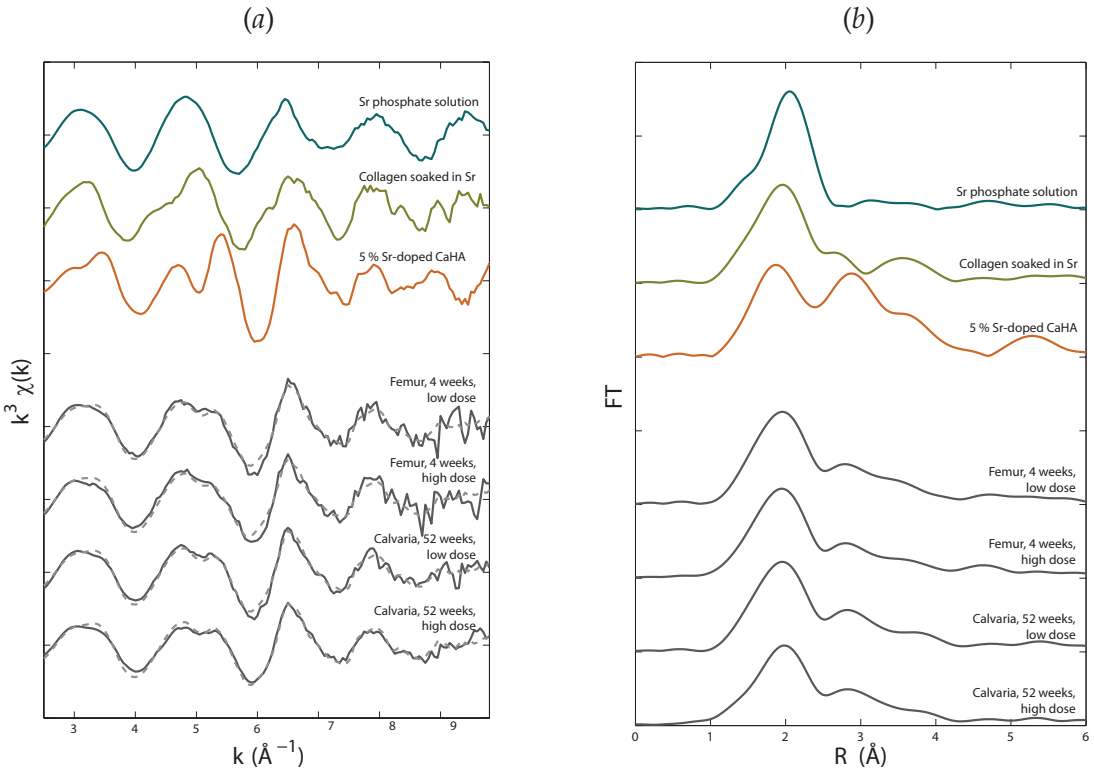


Figure 11.3: (*a*) Extracted and k^3-weighted EXAFS functions, $k^3\chi(k)$. The bone spectra are fitted (dashed curve) with a linear combination of EXAFS spectra for SrHA, Sr-soaked collagen and Sr^{2+} cations in solution. The distribution coefficients are listed in Table 11.3. (*b*) The modulus of the Fourier transform of $k^3\chi(k)$ giving the radial electron density distribution. NB the distances, R, have not been phase corrected.

in which $\chi_i(k)$ and a_i are the EXAFS function and the fraction of reference compound i, respectively. The coefficients a_i were determined by least squares procedure and forced to sum to 1:

$$a_{\mathrm{HA}} + a_{\mathrm{Col}} + a_{\mathrm{Sol}} = 1 \tag{11.2}$$

The fitted EXAFS spectra are shown in Figure 11.3(*a*), and the coefficients describing the distribution of strontium location and the residuals are listed in Table 11.3. The R-factor reported is given as:

$$R_{\mathrm{exafs}} = \frac{\sum \left(\chi_{\mathrm{Exp}}(k) - \chi_{\mathrm{Bone}}(k)\right)^2}{\left(\chi_{\mathrm{Exp}}(k)\right)^2} \cdot 100\% \tag{11.3}$$

11.4 Discussion

The localisation of strontium in physiological calcifications depends on the mechanism of how strontium is inserted into bone tissue. The hierarchical structure of bone tissue and

Table 11.3: Relative distribution of strontium localisation in the bone samples as determined by XAS. The distribution coefficients are given as fractions a_i representing the fraction of strontium incorporated in hydroxyapatite, coordinated to collagen or present in solution. Standard deviations determined from the least squares fit are given in parentheses.

	a_{HA}	a_{Col}	a_{Sol}	R_{exafs}
Femur, 4 weeks, low dose	0.44(2)	0.15(4)	0.41(5)	4.9 %
Femur, 4 weeks, high dose	0.36(2)	0.18(5)	0.46(5)	6.7 %
Calvaria, 52 weeks, low dose	0.44(2)	0.17(4)	0.39(4)	4.0 %
Calvaria, 52 weeks, high dose	0.45(1)	0.21(3)	0.34(3)	3.7 %

its many different composites makes the material highly inhomogeneous and provides many different coordination sites for strontium. In general, the suggested hypotheses agree that the positively charged strontium ions have a high affinity for coordination to oxygen atoms, which are available in high amounts in the bone composites.

In this study presence of Sr^{2+} cations in serum was represented by an aqueous solution of Sr^{2+}. As only one coordination sphere is clearly visible in the XAS spectrum from the Sr phosphate solution, a high static and/or kinetic disorder among the outer coordination shells is well imitated by this sample. Further analysis of this spectrum reveals a strontium coordination number of 8.1 oxygen atoms with a distance of 2.600 Å and a Debye-Waller factor, $\sigma^2 = 0.009$, which match the Sr–O distance of 2.57 Å and coordination number of 8 reported in an EXAFS study on solubilized Sr^{2+} cations performed at 298 K (Seward *et al.*, 1999).

As strontium is inserted into solid material the structural order increases. This is seen for the Sr-soaked collagen, as the outer coordination spheres become visible in the radial electron density distribution and in the hydroxyapatite, in which the structural order is even higher, and the electron density shells become more intense and better resolved.

The hydroxyapatite present in the bone samples are poorly crystalline, as seen from the peak broadening in the XRPD patterns. The XAS fitting results clearly demonstrated that 35–45 % of the total strontium content is incorporated into CaHA. Demonstration of a possible preference towards one of the sites in the bone samples is however beyond the information level of the XAS-data presented here. But the for the 5 % Sr-doped HA used as model compound in this study, an approximately equal distribution of strontium among the two calcium sites was confirmed by XRPD and a following Rietveld refinement.

The EXAFS spectra were fitted by a linear combination of three components. However, a correlation may exist between the collagen component and the solution component. Therefore, it cannot be clearly distinguished by this method, whether Sr^{2+} cations tend to coordinate to the collagen, are present in the serum, or both, but both hypotheses are plausible. Generally, it is evident that strontium occupies highly disordered sites in which only the first coordination shell is visible by XAS and that the distance to this inner coordination

sphere resembles the bond lengths of Sr–O observed in other strontium compounds. Such sites are better described by the aqueous Sr^{2+} cations in solution rather than Sr-soaked collagen, and may account for at least 30–50 % of the total strontium content.

The result from the EXAFS analysis did not show any significant differences of the relative distribution of strontium in the four bone samples. Thus, while the dose levels and treatment time highly influence the bone strontium concentrations, the relative amounts of strontium incorporated into CaHA seem to be independent of these factors.

XAS is a well-suited method to study the localisation of Sr^{2+} cations in bone tissue, as the method is element specific and very sensitive to small concentrations. In general, the concentration of strontium in the analysed bone samples is very high compared with the natural levels of 0.07–0.13 mg g^{-1} found in the corresponding placebo treated animals. However, the high concentrations present in these samples provided XAS spectra with good signal-to-noise ratios.

Obtaining a representative average sample of the entire bone type for XAS experiments depends on the dissections and requires careful sample preparation by which the sample homogenized. Contrarily, working with beam sizes on a macro-scale (approx. 1×1.5 mm) a representative average of the sample is obtained, as a relatively large area of the sample was exposed. The distributions determined by this method may however be connected with some degree of uncertainty, as bone types differs on the macro-scale (cortical and trabecular).

11.5 Conclusive remarks

From this study of bone samples from dogs treated with Sr malonate, around 35–45 % of the strontium present is incorporated into CaHA by substitution of some of the calcium ions, and at least 30 % is located at sites with a high structural disorder similar to Sr^{2+} in solution. Strontium can be absorbed in collagen in which it obtains a higher structural order than when present in solution, but less order than when it is incorporated in CaHA. These observations strongly support that strontium uptake in physiological calcifications follows the hypotheses presented in Section 9.5.

Analysing the EXAFS spectra by fitting a linear combination of spectra of selected composites present in bone, resulted in a new compositional model. This approach allowed the relative distribution of strontium in the different bone composites to be estimated, in contrast to the common shell models. The distribution of strontium among the different composites seems to be independent of treatment period and dose level.

Chapter 12

Conclusions and outlook

This thesis presents the work on three different projects, in which three different X-ray techniques are used, XRPD, XRD and XAS. As structural information is obtained at different resolution, these techniques have different applications and the combination of them provides complementary information, which has been illustrated in this work.

XRPD for fast characterization of protein powders

Among the three techniques, X-ray powder diffraction provides the lowest resolution. This makes the technique suitable for fast characterization of protein powders. A procedure for fast identification and verification of protein powders has been developed and tested on micro-crystals of lysozyme and insulin. Verification of a protein crystal form by XRPD requires calculation of realistic powder diffraction patterns from PDB coordinate data. The key factor to bring the calculated patterns in agreement with the observed patterns is correction for bulk-solvent, but also background, unit cell parameters and geometrical factors have to be taken into account. Experimental powder diffraction patterns were obtained using a standard laboratory X-ray powder diffractometer. Optimization of the sample holder design and collimation of the beam by pinholes were found crucial to increase the signal-to-noise ratios, while minimizing the sample volume. The final sample holder design operate with sample volumes down to 7 μl, and has been successfully applied for both powder diffraction and XAS experiments. The method was successfully applied for verification of micro-crystals of T_6, T_3R_3 and R_6 insulin prior to synchrotron XAS experiments, and used to identify a phase transition in rhombohedral R_6 insulin in which the unit cell

volume is eight-fold doubled when cooled. Furthermore, XRPD was used after the XAS experiments to monitor eventual degradations of the samples from radiation damage.

Combination of XRD and XAS for studying metalloproteins

Combination of XRD and XAS provides a strong tool for studying metalloproteins. The overall three-dimensional structure is obtained at medium/high resolution from XRD, whereas only the local structure around the metal sites are obtained from XAS, however at ultra-high resolution. Hence XAS works as a magnifying glass centred on the metal site. The synergy and complementarity between these two techniques were illustrated in the study of hexameric insulin.

The crystal structures of all three conformations (T_6, T_3R_3 and R_6) of hexameric bovine zinc insulin were solved by single crystal XRD to 1.40 Å, 1.30 Å and 1.80 Å resolution, respectively. T_6 insulin was furthermore crystallized with nickel, copper and iron, and their XRD crystal structures were solved to 1.50 Å, 1.45 Å and 1.90 Å resolution, respectively. The two metal sites in hexameric insulin were occupied by divalent cations of the same metal type in all structures besides the Fe insulin. Crystals with Fe^{2+} in both metal sites could not be obtained, as the affinity to iron is too low. Instead, a structure containing one Fe^{2+} and one Zn^{2+} was obtained from crystals grown in presence of small amounts of zinc. The bovine insulin crystal structures were highly similar to analogue human and porcine insulin structures deposited in the PDB. At resolution better than 1.50 Å the atomic positions were unambiguously clear, and hydrogen atoms were included in the structures. In general, ordered water molecules were easily resolved, with exception of those located in the open metal sites in T_3-trimers.

Octahedral coordination fulfilled by three water molecules was observed in the T_3-sites of zinc and nickel insulin, however with quite long metal–water distances (up to 2.47 Å). For copper insulin, a dual octahedral/tetrahedral coordination of copper was observed in the crystal structure. Likewise, an octahedral zinc and a tetrahedral iron coordination, was observed in the iron insulin structure. Whereas more undefined electron densities were observed in the open T_3-sites and the metal coordination geometry thus could be somehow ambiguously interpreted, the electron densities around the Zn^{2+}-ions in the closed R_3-sites clearly revealed tetrahedral coordinations of zinc, which is caused by the steric hindrance from the α-helical conformation of the three B-chain N-termini.

A more accurate determination of the metal coordination distances, in particular in the T_3-sites, was obtained from EXAFS spectroscopy. XAS spectra with good signal-to noise ratios to 13–14 $Å^{-1}$ in k-space were collected from micro crystalline insulin samples. Using the coordinates from the crystal structures as starting models for the EXAFS refinement the distances between the metal and its first coordination sphere ligands were determined with standard deviations below 0.01 Å. The coordination geometries as determined by

EXAFS were in agreement with those determined by XRD for all three conformations of zinc insulin, as well as nickel T_6 insulin. For copper T_6 insulin, tetrahedral coordination of copper was modelled in both sites. The tetrahedral copper coordination was further supported by *ab initio* Hartree-Fock model calculations and calculated XANES spectra.

The XANES and EXAFS regions contain complementary information, as the XANES region is more sensitive to multiple scattering and thus contains more information about the three dimensional arrangement of atoms. Using finite difference methods, calculated XANES spectra of the models optimized from EXAFS reproduced the experimental XANES spectra very well. The coordination geometries could be verified from the calculated XANES spectra. Furthermore, coordination distances and, in particular, bond angles of T_6 and R_6 zinc insulin were optimized from a quantitative XANES analysis.

Using synchrotron radiation, the energy load per sample area is generally higher for XRD experiments compared to XAS. The concomitant radiation damage particularly influenced the coordination of water molecules in the open T_3-sites, and thus explained the more unstructured electron density observed in the XRD structures in these sites. As the radiation dose was two orders of magnitude lower for the XAS experiments, less damage was induced in these samples. Compared to the medium resolution XRD structures, the EXAFS determinations were in perfect agreement with bond distances observed among small molecules with similar coordination geometries, as well as with other high resolution crystal structures of insulin.

The cluster approach, in which atoms within the same molecule are grouped, was used to refine coordination distances from EXAFS using PDB-coordinates from crystal structures as starting models. This approach enabled a three-dimensional model with inclusion of full multiple scattering calculations to be refined. The constrained and restrained refinement strategies furthermore allowed refinement of distances and displacement factors for many atoms, while keeping the number of refineable parameters to a minimum. This makes the cluster approach superior for metalloprotein applications. EXAFS is best suited for mononuclear metalloproteins or metalloproteins containing several metal clusters that all have identical coordination geometry. Proteins containing different coordinations of the same metal, as in the T_3R_3-conformation, highly increase the complexity of data treatment and reduce the amount of structural information which can be resolved.

Although XAS may be a more gentle method compared to XRD, photoreduction may occur. This was observed in the copper insulin, where the XANES spectrum revealed a reduction of Cu^{2+} to Cu^{+}. Contrarily, Zn^{2+} and Ni^{2+} are harder to reduce and remained stable throughout the experiments.

Combining XRD and XRPD for studying pharmaceuticals

Strontium ranelate is a drug used in the treatment of osteoporosis. Combination of XRD and XRPD were used to study the structure of strontium ranelate and its polymorphs. Strontium ranelate was synthesized and crystallized as a nonahydrate. The structure of strontium ranelate nonahydrate was solved from single crystal XRD. The crystals were found to be very sensitive to humidity and temperature, and deteriorate easily upon dehydration. Four other hydrates were observed from X-ray powder diffraction, and their water contents estimated by thermogravimetric analysis from which the phases were identified as heptahydrate, pentahydrate, trihydrate and dihydrate.

XAS for studying the distribution of strontium in bone tissue

The hierarchical structure of bone tissue and its many different composites makes the material highly inhomogeneous and provides many different coordination sites for strontium. The strontium distribution in bone samples from dogs treated with Sr malonate was studied with XAS. Analysing the EXAFS spectra by fitting a linear combination of spectra of selected composites present in the bone sample, resulted in a new compositional model. This approach allowed the relative distribution of strontium in the different bone composites to be estimated, in contrast to the common shell models.

Around 35–45 % of the strontium present is incorporated into CaHA by substitution of some of the calcium ions, and at least 30 % is located at sites with a high structural disorder similar to Sr^{2+} in solution. Strontium can be absorbed in collagen in which it obtains a higher structural order than when present in solution, but less order than when it is incorporated in CaHA. While the total amount of strontium present in bone depends on treatment period and dose level, the distribution of strontium among the different composites seems to be independent of these factors.

Outlook

X-ray powder diffraction has proven to be a fast technique to characterize protein powders, even with laboratory X-ray diffractometers. XRPD is thus a potent technique for characterization of precipitates occurring in the enzyme production pipelines, where overexpression and subsequently high protein concentrations create new challenges. Identification and verification of crystal forms may provide important information for later purification and thereby optimize the production.

Many structural genomics programmes are currently in progress, and new metalloproteins are being discovered. As metalloproteins constitute a significant fraction of all proteins, the combination of XAS and XRD is highly needed to obtain information about the

structure of metalloproteins and the processes taking place in their metal centres. Newly built synchrotron facilities provide concomitant collection of both XRD and XAS data, and experiments where both techniques are used, as well as *in situ* studies, may be seen more frequently in the near future. Furthermore, XAS may found applications to metalloproteins which are difficult to crystallize. Recent improvements in fitting algorithms against XANES spectra offer a possibility to use information from both the XANES and the EXAFS in the model fitting.

Strontium was found to be incorporation into bone tissue in hydroxyapatite as well as collagen. Substitution of calcium by strontium in the hydroxyapatite structure has already been studied, whereas the influence on collagen remains unknown. Combining XAS, SAXS and fibre diffraction may provide information about how strontium binds to collagen, and could be complemented by mechanical test of how it affects its tensile strength.

Bibliography

Abel, J. J. (1926). Crystalline insulin. *Proc. Nat. Acad. Sci. USA*, **12**, 132–136.

Adams, M. J., Blundell, T. L., Dodson, E. J., Dodson, G. G., Vijayan, M., Baker, E. N., Harding, M. M., Hodgkin, D. C., Rimmer, B. & Sheat, S. (1969). Structure of rhombohedral 2 zinc insulin crystals. *Nature*, **224**, 491–495.

Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L.-W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C. & Zwart, P. H. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Cryst. D*, **66**, 213–221.

Ammann, P., Shen, V., Robin, B., Mauras, Y., Bonjour, J.-P. & Rizzoli, R. (2004). Strontium Ranelate Improves Bone Resistance by Increasing Bone Mass and Improving Architecture in Intact Female Rats. *J. Bone Miner. Res.*, **19**, 2012–2020.

Ankudinov, A.L., Ravel, B., Rehr, J.J. & Conradson, S.D. (1998). Real-space multiple-scattering calculation and interpretation of x-ray-absorption near-edge structure. *Phys. Rev. B*, **58**, 7565–7576.

Ascone, I., Meyer-Klaucke, W. & Murphy, L. (2003). Experimental aspects of biological X-ray absorption spectroscopy. *J. Synchrotron Rad.*, **10**, 16–22.

Ascone, I., Fourme, R., Hasnain, S. S. & Hodgson, K. (2005). Metallogenimics and biological X-ray absorption spectroscopy. *J. Synchrotron Rad.*, **12**, 1–3.

Ascone, I., Girard, E., Gourhant, P., Legrand, P., Roudenko, O., Roussier, L. & Thompson, A. W. (2007). Proxima 1, a New Beamline on the Third Generation SR Source SOLEIL

Combining PX and Single-Crystal BioXAS. *AIP Conference Proceedings: X-ray Absorption Fine Structure - XAFS13*, **882**, 872–874.

Ashley, C. A. & Doniach, S. (1975). Theory of extended x-ray absorption edge fine structure (EXAFS) in crystalline solids. *Phys. Rev. B*, **11**, 1279–1288.

Bain, S. D., Jerome, C., Shen, V., Dupin-Roger, I. & Ammann, P. (2009). Strontium ranelate improves bone strength in ovariectomized rat by positively influencing bone resistance determinants. *Osteoporosis Int.*, **20**, 1417–1428.

Baker, E. N., Blundell, T. L., Cutfield, J. F., Cutfield, S. M., Dodson, E. J., Dodson, G. G., Hodgkin, D. M. C., Hubbard, R. E., Isaacs, N. W., Reynolds, C. D., Sakabe, K., Sakabe, N. & Vijayan, N. M. (1988). The Structure of 2Zn Pig Insulin Crystals at 1.5 Å Resolution. *Phil. Trans. R. Soc. Lond. B*, **319**, 369–456.

Banci, L., Bertini, I. & Mangani, S. (2005). Integration of XAS and NMR techniques for the structure determination of metalloproteins. Examples from the study of copper transport proteins. *J. Synchrotron Rad.*, **12**, 94–97.

Banting, F. G. & Best, C. H. (1922). The internal secretion of the pancreas. *J. Lab. Clin. Med.*, **7**, 251–266.

Basso, S., Fitch, A. N., Fox, G. C., Margiolaki, I. & Wright, J. P. (2005). High-throughput phase-diagram mapping via powder diffraction: a case study of HEWL versus pH. *Acta Cryst. D*, **61**, 1612–1625.

Bazin, D., Chappard, C., Combes, C., Carpentier, X., Rouzière, S., André, G., Matzen, G., Allix, M., Thiaudière, D., Reguer, S., Jungers, P. & Daudon, M. (2009a). Diffraction techniques and vibrational spectroscopy opportunities to characterise bones. *Osteoporosis Int.*, **20**, 1065–1075.

Bazin, D., Carpentier, X., Brocheriou, I., Dorfmuller, P., Aubert, S., Chappard, C., Thiaudière, D., Reguer, S., Waychunas, G., Jungers, P. & Daudon, M. (2009b). Revisiting the localisation of Zn^{2+} cations sorbed on pathological apatite calcifications made through X-ray absorption spectroscopy. *Biochimie*, **91**, 1294–1300.

Bazin, D., Daudon, M., Chappard, C., Rehr, J. J., Thiaudière, D. & Reguer, S. (2011). The status of strontium in biological apatites: an XANES investigation. *J. Synchrotron Rad.*, **18**, 912–918.

Behets, G. J., Verberckmoes, S. C., Oste, L., Bervoets, A. R., Salomé, M., Cox, A. G., Denton, J., De Broe, M. E. & D'Haese, P. C. (2005). Localization of lanthanum in bone of chronic renal failure rats after oral dosing with lanthanum carbonate. *Kidney Int.*, **67**, 1830–1836.

Bellis, D. J., Li, D., Chen, Z., Gibson, W. M. & Parsons, P. J. (2009). Measurement of the microdistribution of strontium and lead in bone via benchtop monochromatic microbeam X-ray fluorescence with a low power source. *J. Anal. At. Spectrom.*, **24**, 622–626.

Benfatto, M., Congiu-Castellano, A., Daniele, A. & Della Longa, S. (2001). MXAN: a new software procedure to perform geometrical fitting of experimental XANES spectra. *J. Synchrotron Rad.*, **8**, 267–269.

Bentley, G., Dodson, E., Dodson, G., Hodgkin, D. & Mercola, D. (1976). Structure of insulin in 4-zinc insulin. *Nature*, **261**, 166–168.

Bentley, G., Dodson, G. & Lewitova, A. (1978). Rhombohedral insulin crystal transformation. *J. Mol. Biol.*, **126**, 871–875.

Bentley, G. A., Brange, J., Derewenda, Z., Dodson, E. J., Dodson, G. G., Markussen, J., Wilkinson, A. J., Wollmer, A. & Xiao, B. (1992). Role of B13 Glu in insulin assembly: The hexamer structure of recombinant mutant (B13 Glu $\rightarrow$ Gln) insulin. *J. Mol. Biol.*, **228**, 1163–1176.

Bigi, A., Boanini, E., Capuccini, C. & Gazzano, M. (2007). Strontium-substituted hydroxyapatite nanocrystals. *Inorg. Chim. Acta*, **360**, 1009–1016.

Binsted, N., Hasnain, S. S. & Hukins, D. W. L. (1982). Developmental changes in bone mineral structure demonstrated by extended X-ray absorption fine structure (EXAFS) spectroscopy. *Biochem. Biophys. Res. Com.*, **107**, 89–92.

Binsted, N., Campbell, J. W., Gurman, S. J. & Stephenson, P. C. (1991). *EXCURV92*. SERC, Daresbury Laboratory, Cheshire, UK.

Binsted, N., Strange, R. W. & Hasnain, S. S. (1992). Constrained and Restrained Refinement in EXAFS Data Analysis with Curved Wave Theory. *Biochemistry*, **31**, 12117–12125.

Blackburn, N. J., Hasnain, S. S., Binsted, N., Diakun, G. P., Garner, C. D. & Knowles, P. F. (1984). An extended-X-ray-absorption-fine-structure study of bovine erythrocyte superoxide dismutase in aqueous solution. *Biochem. J.*, **219**, 985–990.

Bloom, C. R., Heymann, R., Kaarsholm, N. C. & Dunn, M. F. (1997). Binding of 2, 6-and 2, 7-Dihydroxynaphthalene to Wild-Type and E-B13Q Insulins: Dynamic, Equilibrium, and Molecular Modeling Investigations. *Biochemistry*, **36**, 12746–12758.

Blundell, T., Dodson, G., Hodgkin, D. & Mercola, D. (1972). Insulin: the structure in the crystal and its reflection in chemistry and biology. *Adv. Protein Chem.*, **26**, 279–402.

Blundell, T. L., Cutfield, J. F., Cutfield, S. M., Dodson, E. J., Dodson, G. G., Hodgkin, D. C., Mercola, D. A. & Vijayan, M. (1971). Atomic positions in rhombohedral 2-zinc insulin crystals. *Nature*, **231**, 506–511.

Bordas, J., Dodson, G. G., Grewe, H., Koch, M. H. J., Krebs, B & Randall, J. (1983). A comparative assessment of the zinc-protein coordination in 2Zn-insulin as determined by X-ray absorption fine structure (EXAFS) and X-ray crystallography. *Proc. Roy. Soc. Lond. B*, **219**, 21–39.

Brader, M. L., Kaarsholm, N. C., Harnung, S. E. & Dunn, M. F. (1997). Ligand perturbation on a pseudotetrahedral Co (II)(His) 3L site. A magnetic circular dichroism study of the Co (II)-substituted insulin hexamer. *J. Biol. Chem.*, **272**, 1088–1094.

Brader, M.L., Borchardt, D. & Dunn, M.F. (1992). The T to R transition in the copper (II)-substituted insulin hexamer. Anion complexes of the R-state species exhibiting type 1 and type 2 spectral characteristics. *Biochemistry*, **31**, 4691–4696.

Bradley, D. A., Muthuvelu, P., Ellis, R. E., Green, E. M., Attenburrow, D., Barrett, R., Arkill, K., Colridge, D. B. & Winlove, C.P. (2007). Characterisation of mineralisation of bone and cartilage: X-ray diffraction and Ca and Sr K_α X-ray fluorescence microscopy. *Nucl. Instrum. Methods Phys. Res. B*, **263**, 1–6.

Bushnell, G. W., Louie, G. V. & Bryer, G. D. (1990). High-resolution Three-dimensional Structure of Horse Heart Cytochrome *c*. *J. Mol. Biol.*, **214**, 585–595.

Capuccini, C., Torricelli, P., Boanini, E., Gazzano, M., Giardino, R. & Bigi, A. (2009). Interaction of Sr-doped hydroxyapatite nanocrystals with osteoclast and osteoblast-like cells. *J. Biomed. Mater. Res. A*, **89**, 594–600.

Carlson, S., Clausén, M., Gridneva, L., Sommarin, B. & Svensson, C. (2006). XAFS experiments at beamline I811, MAX-lab synchrotron source, Sweden. *J. Synchrotron Rad.*, **13**, 359–364.

Cazalbou, S., Eichert, D., Drouet, C., Combes, C. & Rey, C. (2004). Biological mineralisations based on calcium phosphate. *C. R. Palevol*, **3**, 563–572.

Chen, J., Christiansen, J. Campobasso, N., Bolin, J. T., Tittsworth, R. C., Hales, B. J., Rehr, J. J. & Cramer, S. P. (1993). Refinement of a Model for the Nitrogenase Mo-Fe Cluster Using Single-Crystal Mo and Fe EXAFS. *Angew. Chem. Int. Ed. Engl.*, **32**, 1592–1594.

Cheng, M.-C., Rich, A. M., Armstrong, R. S., Ellis, P. J. & Lay, P. A. (1999). Determination of Iron-Ligand Bond Lengths in Ferric and Ferrous Horse Heart Cytochrome *c* Using Multiple-Scattering Analyses of XAFS Data. *Inorg. Chem.*, **38**, 5703–5708.

Ciszak, E. & Smith, G. D. (1994). Crystallographic evidence for dual coordination around zinc in the T_3R_3 human insulin hexamer. *Biochemistry*, **33**, 1512–1517.

Ciszak, E., Beals, J. M., Frank, B. H., Baker, J. C., Carter, N. D. & Smith, G. D. (1995). Role of C-terminal B-chain residues in insulin assembly: the structure of hexameric Lys^{B28} Pro^{B29}-human insulin. *Structure*, **3**, 615–622.

Coffman, F. D. & Dunn, M. F. (1988). Insulin-metal Ion Interactions: The Binding of Divalent Cations to Insulin Hexamers and Tetramers and the Assembly of Insulin Hexamers. *Biochemistry*, **27**, 6179–6187.

Collaborative Computational Project, Number 4. (1994). The CCP4 Suite: Programs for Protein Crystallography. *Acta Cryst. D*, **50**, 760–763.

Colpas, G. J., Maroney, M. J., Bagyinka, C., Kumar, M., Willis, W. S., Suib, S. L., Baidya, N. & Mascharak, P. K. (1991). X-ray Spectroscopic Studies of Nickel Complexes, with Application to the Structure of Nickel Sites in Hydrogenases. *Inorg. Chem.*, **30**, 920–928.

Corbett, M. C., Tezcan, F. A., Einsle, O., Walton, M. Y., Rees, D. C., Latimer, M. J., Hedman, B. & Hodgson, K. O. (2005). Mo *K*- and *L*-edge X-ray absorption spectroscopic study of the ADP·AlF_4^--stabilized nitrogenase complex: comparison with MoFe protein in solution and single crystal. *J. Synchrotron Rad.*, **12**, 28–34.

Cotton, F. A., Wilkinson, G. & Gaus, P. L. (1995). *Basic Inorganic Chemistry*. John Wiley & Sons.

Crowfoot, D. (1938). The Crystal Structure of Insulin. I. The Investigation of Air-Dried Insulin Crystals. *Proc. Roy. Soc. Lond. A*, **164**, 580–602.

Dahl, S. G., Allain, P., Marie, P. J., Mauras, Y., Boivin, G., Ammann, P., Tsouderos, Y., Delmas, P. D. & Christiansen, C. (2001). Incorporation and Distribution of Strontium in Bone. *Bone*, **28**, 446–453.

D´Angelo, P., Lapi, A., Migliorati, V., Arcovito, A., Benfatto, M., Roscioni, O. M., Meyer-Klaucke, W. & Della Longa, S. (2008). X-ray Absorption Spectroscopy of Hemes and Hemeproteins in Solution: Multiple Scattering Analysis. *Inorg. Chem.*, **47**, 9905–9918.

Davidson, H. W., Rhodes, C. J. & Hutton, J. C. (1988). Intraorganellar calcium and pH control proinsulin cleavage in the pancretic β cell via two distinct site-specific endopeptidases. *Nature*, **333**, 93–96.

DeLano, W. L. (2006). *The PyMOL Molecular Graphics System, v0.99*. http://pymol.org.

Derewenda, U., Derewenda, Z., Dodson, E. J., Dodson, G. G., Reynolds, C. D., Smith, G. D., Sparks, C. & Swenson, D. (1989). Phenol stabilizes more helix in a new symmetrical zinc insulin hexamer. *Nature*, **338**, 594–596.

Derewenda, U., Derewenda, Z., Dodson, E. J., Dodson, G. G., Bing, X. & Markussen, J. (1991). X-ray analysis of the single chain B29-A1 peptide-linked insulin molecule. A completely inactive analogue. *J. Mol. Biol.*, **220**, 425–433.

Dodson, E. J., Dodson, G. G., Hodgkin, D. C. & Reynolds, C. D. (1979). Structural relationships in the two-zinc insulin hexamer. *Biochemistry and Cell Biology*, **57**, 469–479.

Doebbler, J. A. & von Dreele, R. B. (2009). Macromolecular powder diffraction: Structure solution via molecular replacement. *Z. Kristallogr. Suppl.*, **30**, 33–37.

Eanes, E. D., Powers, L. & Costa, J. L. (1981). Extended X-ray Absorption Fine Structure (EXAFS) Studies on Calcium in Crystalline and Amorphous Solids of Biological Interest. *Cell Calcium*, **2**, 251–262.

Edsall, J. T., Felsenfeld, G., Goodman, D. S. & Gurd, F. R. N. (1954). The Association of Imidazole with the Ions of Zinc and Cupric Copper. *J. Am. Chem. Soc.*, **76**, 3054–3061.

Eichert, D., Salomé, M., Banu, M., Susini, J. & Rey, C. (2005). Preliminary characterization of calcium chemical environment in apatitic and non-apatitic calcium phosphates of biological interest by X-ray absorption spectroscopy. *Spectrochim. Acta B*, **60**, 850–858.

Einsle, O., Tezcan, F. A., Andrade, S. L. A., Schmid, B., Yoshida, M., Howard, J. B. & Rees, D. C. (2002). Nitrogenase MoFe-Protein at 1.16 Å Resolution: A Central Ligand in the FeMo-Cofactor. *Science*, **297**, 1696–1700.

Ellis, M. J., Dodd, F. E., Sawers, G., Eady, R. R. & Hasnain, S. S. (2003). Atomic Resolution Structures of Native Copper Nitrite Reductase from *Alcaligenes xylosoxidans* and the Active Site Mutant Asp92Glu. *J. Mol. Biol.*, **328**, 429–438.

Emdin, S. O., Dodson, G. G., Cutfield, J. M. & Cutfield, S. M. (1980). Role of Zinc in Insulin Biosynthesis. *Diabetologia*, **19**, 174–182.

Emsley, P. & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Cryst. D*, **60**, 2126–2132.

Engels, M., Jacoby, E., Krüger, P., Schlitter, P & Wollmer, A. (1992). The T $\rightleftharpoons$ R structural transition of insulin; pathways suggested by targeted energy minimization . *Protein Eng.*, **5**, 669–677.

Engh, R. A. & Huber, R. (1991). Accurate bond and angle parameters for X-ray protein structure refinement. *Acta Cryst. A*, **47**, 392–400.

English, R. B. & Nassimbeni, L. R. (1984). A Redetermination of the Structure of Strontium Chloride Hexahydrate, $SrCI_2.6H_2O$. *Acta Cryst. C*, **40**, 580–581.

Feiters, M. C., Eikelenboom, A. P. A. M., Nolting, H.-F., Krebs, B., van den Ent, F. M. I., Plasterk, R. H. A., Kaptein, R. & Boelens, R. (2003). X-ray absorption spectroscopic studies of zinc in the N-terminal domain of HIV-2 integrase and model compounds. *J. Synchrotron Rad.*, **10**, 86–95.

Fokine, A. & Urzhumtsev, A. (2002). Flat bulk-solvent model: obtaining optimal parameters. *Acta Cryst. D*, **58**, 1387–1392.

Frankær, C. G., Harris, P. & Ståhl, K. (2011). A sample holder for in-house X-ray powder diffraction studies of protein powders. *J. Appl. Cryst.*, **44**, 1288–1290.

Frankær, C.G. (2011). *Laboratory Journal*.

Fricke, H. (1920). The K-Characteristic Absorption Frequencies for the Chemical Elements Magnesium to Chromium. *Phys. Rev.*, **16**, 202–215.

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D. & Bairoch, A. (1995). *The Proteomics Protocols Handbook*. Humana Press. Pages 571–607.

George, S. J., Igarashi, R. Y., Xiao, Y., Hernandez, J. A., Demuez, M., Zhao, D., Yoda, Y., Ludden, P. W., Rubio, L. M. & Cramer, S. P. (2008). Extended X-ray Absorption Fine Structure and Nuclear Resonance Vibrational Spectroscopy Reveal that NifB-co, a FeMo-co Precursor, Comprises a 6Fe Core with an Interstitial Light Atom. *J. Am. Chem. Soc.*, **130**, 5673–5680.

Glatzel, P., Chantler, C.T., Molenbroek, A.M., Newville, M., Rehr, J.J., Sham, T.-K. & Strange, R. (2011). *IUCr Dictionary entries*. http://www.iucr.org/resources/commissions/xafs/xafs-related-definitions-for-the-iucr-dictionary.

Gurman, S. J., Binsted, N. & Ross, I. (1984). A rapid, exact curved-wave theory for EXAFS calculations. *J. Phys. C*, **17**, 143–151.

Gurman, S. J., Binsted, N. & Ross, I. (1986). A rapid, exact, curved-wave theory for EXAFS calculations. II. The multiple-scattering contributions. *J. Phys. C*, **19**, 1845–1861.

Gurman, S.J. (1995). Interprtation of XAFS Data. *J. Synchrotron Rad.*, **2**, 56–63.

Halban, P. A., Mutkoski, R., Dodson, G. & Orci, L. (1987). Resistance of the insulin crystal to lysosomal proteases: implications for pancreatic β-cells crinophagy. *Diabetologia*, **30**, 348–353.

Hammersley, A. P. (1998). *FIT2D V9.129 Reference Manual V3.1*. ESRF Internal Report, ESRF98HA01T.

Hammersley, A. P., Svensson, S. O., Hanfland, M., Fitch, A. N. & Häusermann, D. (1996). Two-dimensional detector software: From real detector to idealised image or two-theta scan. *High Pressure Research*, **14**, 235–248.

Han, Y., Cowin, S. C., Schaffler, M. B. & Weinbaum, S. (2004). Mechanotransduction and strain amplification in osteocyte cell processes. *Proc. Nat. Sci. Acad. USA*, **101**, 16689–16694.

Harding, M. M., Hodgkin, D. C., Kennedy, A. F., O'Connor, A. & Weitzmann, P. D. J. (1966). The crystal structure of insulin II. An investigation of rhombohedral zinc insulin crystals and a report of other crystalline forms. *J. Mol. Biol.*, **16**, 212–226.

Harding, M. M., Nowicki, M. W. & Walkinshaw, M. D. (2010). Metals in protein structures: a review of their principal features. *Cryst. Rev.*, **16**, 242–302.

Harries, J. E., Hukins, D. W. L. & Hasnain, S. S. (1986). Analysis of the EXAFS spectrum of hydroxyapatite. *J. Phys. C*, **19**, 6859–6872.

Harries, J. E., Hukins, D. W. L., Holt, C. & Hasnain, S. S. (1987). Conversion of amorphous calcium phosphate into hydroxyapatite investigated by EXAFS spectroscopy. *J. Cryst. Growth*, **84**, 563–570.

Harries, J. E., Hukins, D. W. L. & Hasnain, S. S. (1988). Calcium Environment in Bone Mineral Determined by EXAFS Spectroscopy. *Calcif. Tissue Int.*, **43**, 250–253.

Hartmann, C. G. (2008). *X-ray Powder Diffraction Studies on Proteins*. M.Phil. thesis, Chemical Department, Technical University of Denmark.

Hartmann, C. G., Nielsen, O. F., Ståhl, K. & Harris, P. (2010). In-house characterization of protein powder. *J. Appl. Cryst.*, **43**, 876–882.

Hartmann, C. G., Harris, P. & Ståhl, K. (2011). In-house characterization of protein powder. *Z. Kristallogr. Proc.*, **1**, 163–168.

Hasnain, S. S. & Hodgson, K. O. (1999). Structure of metal centres in proteins at subatomic resolution. *J. Synchrotron Rad.*, **6**, 852–864.

Hasnain, S. S. & Strange, R. W. (2003). Marriage of XAFS and crystallography for structure-function studies of metalloproteins. *J. Synchrotron Rad.*, **10**, 9–15.

Haumann, M., Müller, C., Liebisch, P. Neisius, T. & Dau, H. (2005). A novel BioXAS technique with sub-millisecond time resolution to track oxidation state and structural changes at biological metal centers. *J. Synchrotron Rad.*, **12**, 35–44.

Heras, B. & Martin, J. L. (2005). Post-crystallization treatments for improving diffraction quality of protein crystals. *Acta Cryst. D*, **61**, 1173–1180.

Hertz, G. (1920). Über die Absorptionsgrenzen in der *L*-Serie. *Z. Phys.*, **3**, 19–25.

Hill, C. P., Dauter, Z., Dodson, E. J., Dodson, G. G. & Dunn, M. F. (1991). X-ray structure of an unusual Ca^{2+} site and the roles of Zn^{2+} and Ca^{2+} in the assembly, stability, and storage of the insulin hexamer. *Biochemistry*, **30**, 917–924.

Hodge, A.J. & Petruska, J. A. (1963). *Aspects of protein structure*. Academic Press, New York. Page 489.

Holton, J. M. (2009). A beginner´s guide to radiation damage. *J. Synchrotron Rad.*, **16**, 133–142.

Hooft, R. W. W., Vriend, G., Sander, C. & Abola, E. E. (1996). Errors in protein structures. *Nature*, **381**, 272.

Horvath, S., Demuynck, I. & Damien, G. (2008). *Alpha crystalline form of strontium ranelate, US Patent No. 7459568*.

Howell, S. L. (1984). The mechanism of insulin secretion. *Diabetologia*, **26**, 319–327.

Huang, X. F. & Arvan, P. (1995). Intracellular Transport of Proinsulin in Pancreatic beta-Cells. *J. Biol. Chem.*, **270**, 20417–20423.

Hwang, J., Krebs, C., Huynh, B. H., Edmondson, D. E., Theil, E. C. & Penner-Hahn, J. E. (2000). A Short Fe−Fe Distance in Peroxodiferric Ferritin: Control of Fe Substrate Versus Cofactor Decay? *Science*, **287**, 122–125.

Jacquamet, L., Traoré, D. A. K., Ferrer, J.-L., Proux, O., Testemale, D., Hazemann, J.-L., Nazarenko, E., El Ghazouani, A., Caux-Thang, C., Duarte, V. & Latour, J.-M. (2009a). Structural characterization of the active form of PerR: insights into the metal-induced activation of PerR and Fur proteins for DNA binding. *Molecular Microbiology*, **73**, 20–31.

Jacquamet, L., Traoré, D. A. K., Ferrer, J.-L., Proux, O., Testemale, D., Hazemann, J.-L., Nazarenko, E., El Ghazouani, A., Caux-Trang, C., Duarte, V. & Latour, J.-M. (2009b).

Structural characterization of the active form of PerR: insights into the metal-induced activation of PerR and Fur proteins for DNA binding. *Mol. Microbiol.*, **73**, 20–31.

Jalilehvand, F. (2000). *Structure of Hydrated Ions and Cyanide Complexes by X-ray Absorption Spectroscopy*. Ph.D. thesis, Royal Institute of Technology, Stockholm.

Jenner, M. J., Wright, J. P., Margiolaki, I. & Fitch, A. N. (2007). Successful protein cryocooling for powder diffraction. *J. Appl. Cryst.*, **40**, 121–124.

Joly, Y. (2001). X-ray absorption near-edge structure calculations beyond the muffin-tin approximation. *Phys. Rev. B*, **63**, 125120.

Joyner, R. W., Martin, K. J. & Meehan, P. (1987). Some applications of statistical tests in analysis of EXAFS and SEXAFS data. *J. Phys. C*, **20**, 4005–4012.

Juers, D. H. & Matthews, B. W. (2004). Cryo-cooling in macromolecular crystallography: advantages, disadvantages and optimization. *Quarterly Reviews of Biophysics*, **37**, 1–15.

Kaarsholm, N. C., Ko, H. C. & Dunn, M. F. (1989). Comparison of solution structural flexibility and zinc binding domains for insulin, proinsulin, and miniproinsulin. *Biochemistry*, **28**, 4427–4435.

Kabsch, W. (1993). Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. Appl. Cryst.*, **26**, 795–800.

Kadima, W. (1999). Role of Metal Ions in the T- To R-Allosteric Transition in the Insulin Hexamer. *Biochemistry*, **38**, 13443–13452.

Kantarjieff, K. A. & Rupp, B. (2003). Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein-nucleic acid complex crystals. *Protein Sci.*, **12**, 1865–1871.

Kau, L.-S., Spira-Solomon, D. J., Penner-Hahn, J. E., Hodgson, K. O. & Solomon, E. I. (1987). X-ray Absorption Edge Determination of the Oxidation State and Coordination Number of Copper: Application to the Type 3 Site in *Rhus vernicifera* Laccase and Its Reaction with Oxygen. *J. Am. Chem. Soc.*, **109**, 6433–6442.

Kay, M. I., A., Young R. & Posner, A. S. (1964). Crystal Structure of Hydroxyapatite. *Nature*, **204**, 1050–1052.

Kimball, G. E. & Shortley, G. H. (1934). The Numerical Solution of Schrodinger's Equation. *Phys. Rev.*, **45**, 815–820.

Kleywegt, G. J. & Jones, T. A. (1996). Phi/psi-chology: Ramachandran revisited. *Structure*, **4**, 1395–1400.

Knudsen, M. V. (2011). *Røntgenkrystallografiske studier af insulin med bundne metalioner*. B.Sc. thesis, Chemical Department, Technical University of Denmark.

Koningsberger, D.C., Prins, R., Bianconi, A., Bunker, B.A., Cramer, S.P., Crozier, E.D., Durham, P.J., Heald, S.M., Ingalls, R., Rehr, J.J., Sayers, D.E., Stern, E. A. & Stöhr, J. (1988). *X-ray Absorption: Principles Applications, Techniques of EXAFS, SEXAFS and XANES*. John Wiley & Sons.

Korbas, M., Rokita, E., Meyer-Klaucke, W. & Ryczek, J. (2004). Bone tissue incorporates *in vitro* gallium with a local structure similar to gallium-doped brushite. *J. Biol. Inorg. Chem.*, **9**, 67–76.

Kossel, W. (1920). Zum Bau der Röntgenspektren. *Z. Phys.*, **1**, 119–134.

Krissinel, E. & Henrick, K. (2004). Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Cryst. D*, **60**, 2256–2268.

Kronig, R. de L. (1931). Zur Theorie der Feinstruktur in den Röntgenabsorptionsspektren. *Z. Phys.*, **70**, 317–323.

Kronig, R. de L. (1932). Zur Theorie der Feinstruktur in den Röntgenabsorptionsspektren II. *Z. Phys.*, **75**, 191–210.

Kumar, B. V. S., Srinivas, K., Kale, S. A. & Pradhan, N. S. C. (2009). *Process for the Preparation of Strontium Ranelate, US Patent Application Publication No. 0214631*.

Kumta, P. N., Sfeir, C., Lee, D.-H., Olton, D. & Choi, D. (2005). Nanostructured calcium phosphates for biomedical applications: novel synthesis and characterization. *Acta Biomater.*, **1**, 65–83.

Laskowski, R. A., MacArthur, M. W., Smith, D. K., Jones, D. T., Hutchinson, E. G., Morris, A. L., Moss, D. S. & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.*, **26**, 283–291.

Launey, M. E., Buehler, M. J. & Ritchie, R. O. (2010). On the Mechanistic Origins of Toughness in Bone. *Annu. Rev. Mat. Res.*, **40**, 25–53.

Lee, P.A. & Pendry, J.B. (1975). Theory of the extended x-ray absorption fine structure. *Phys. Rev. B*, **11**, 2795–2811.

Li, C., Paris, O., Siegel, S., Roschger, P., Paschalis, E. P., Klaushofer, K. & Fratzl, P. (2010). Strontium Is Incorporated Into Mineral Crystals Only in Newly Formed Bone During Strontium Ranelate Treatment. *J. Bone Miner. Res.*, **25**, 968–975.

Liou, S.-C., Chen, S.-Y., Lee, H.-Y. & Bow, J.-S. (2004). Structural characterization of nano-sized calcium deficient apatite powders. *Biomaterials*, **25**, 189–196.

Manceau, A. & Matynia, A. (2010). The nature of Cu bonding to natural organic matter. *Geochim. Cosmochim. Acta*, **74**, 2556–2580.

Margiolaki, I. & Wright, J. P. (2008). Powder crystallography on macromolecules. *Acta Cryst. A*, **64**, 169–180.

Margiolaki, I., Wright, J. P., Fitch, A. N., Fox, G. C., Labrador, A., von Dreele, R. B., Miura, K., Gozzo, F., Schiltz, M., Besnard, C., Camus, F., Pattison, P., Beckers, D. & Degen, T. (2007a). Powder diffraction studies on proteins: An overview of data collection approaches. *Z. Kristallogr. Suppl.*, **26**, 1–13.

Margiolaki, I., Wright, J. P., Wilmanns, M., Fitch, A. N. & Pinotsis, N. (2007b). Second SH3 Domain of Ponsin Solved from Powder Diffraction. *J. Am. Chem. Soc.*, **129**, 11865–11871.

Marie, P. J., Ammann, P., Boivin, G. & Rey, C. (2001). Mechanisms of Action and Therapeutic Potential of Strontium in Bone. *Calcif. Tissue Int.*, **69**, 121–129.

McCoy, A. J, Grosse-Kunstleve, R. W, Adams, P. D, Winn, M. D, Storoni, L. C & Read, R. J. (2007). Phaser crystallographic software. *J. Appl. Cryst.*, **40**, 658–674.

Meunier, P. J., Roux, C., Seeman, E., Ortolani, S., Badurski, J. E., Spector, T. D., Cannata, J., Balogh, A., Lemmel, E. M., Pors-Nielsen, S., Rizzoli, R., Genant, H. K. & Reginster, J. Y. (2004). The effects of strontium ranelate on the risk of vertebral fracture in women with postmenopausal osteoporosis. *New England J. Med.*, **350**, 459–468.

Miller, A. (1984). Collagen: the organic matrix of bone. *Phil. Trans. Roy. Soc. Lond. B*, **304**, 455–477.

Miller, R. M., Hukins, D. W. L., Hasnain, S. S. & Lagarde, P. (1981). Extended X-ray absorption fine structure (EXAFS) studies of the calcium ion environment in bone mineral and related calcium phosphates. *Biochem. Biophys. Res. Com.*, **99**, 102–106.

Moore, P. B. (1998). The three-dimensional structure of the ribosome and its components. *Annu. Rev. Biophys. Struct. Biomol.*, **27**, 35–58.

Murphy, L. M., Strange, R. W. & Hasnain, S. S. (1997). A critical assessment of the evidence from XAFS and crystallography for the breakage of the imidazolate bridge during catalysis in CuZn superoxide dismutase. *Structure*, **5**, 371–379.

Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). Refinement of Macromolecular Structures by the Maximum-Likelyhood Method. *Acta Cryst. D*, **53**, 240–255.

Newville, M. (2001). IFEFFIT: interactive XAFS analysis and FEFF fitting. *J. Synchrotron Rad.*, **8**, 322–324.

Newville, M. (2004). *Fundamentals of XAFS*. http://www.xafs.org/Tutorials.

Nguyen, C., Ea, H. K., Thiaudière, D., Reguer, S., Hannouche, D., Daudon, M., Lioté, F. & Bazin, D. (2011). Calcifications in human osteoarthritic articular cartilage: *ex vivo* assessment of calcium compounds using XANES spectroscopy. *J. Synchrotron Rad.*, **18**, 475–480.

Nicholson, J., Perkins, L. & Körber, F. (2006). The high resolution structure of hexameric T_6 cobalt insulin: A possible pathway for the T to R transition. *Recent Research Developments in Molecular Biology*, **3**, 1–17.

Norrman, M., Ståhl, K., Schluckebier, G. & Al-Karadaghi, S. (2006). Characterization of insulin microcrystals using powder diffraction and multivariate data analysis. *J. Appl. Cryst.*, **39**, 391–400.

O´Donnell, M. D., Fredholm, Y., de Rouffignac, A. & Hill, R. G. (2008). Structural analysis of a series of strontium-substituted apatites. *Acta Biomater.*, **4**, 1455–1464.

Oste, L., Verberckmoes, S. C., Behets, G. J., Dams, G., Bervoets, A. R., Van Hoof, V. O., Bohic, S., Drakopoulos, M., De Broe, M. E. & D'Haese, P. C. (2007). Strontium incorporates at sites critical for bone mineralization in rats with renal failure. *X-Ray Spectrom.*, **36**, 42–49.

Ozturan, K. E., Demir, B., Yucel, I., Cakici, H., Yilmaz, F. & Haberal, A. (2011). Effect of strontium ranelate on fracture healing in the osteoporotic rats. *J. Orthopaedic Res.*, **29**, 138–142.

Padilla, J. E. & Yeates, T. O. (2003). A statistic for local intensity differences: robustness to anisotropy and pseudo-centering and utility for detecting twinning. *Acta Cryst. D*, **59**, 1124–1130.

Parfitt, A. M. (1996). *Osteoporosis*. Academic Press, San Diego. Pages 315–339.

Pemmer, B., Hofstaetter, J. G., Meirer, F., Smolek, S., Wobrauschek, P., Simon, R., Fuchs, R. K., Allen, M. R., Condon, K. W., Reinwald, S., Phipps, R. J., Burr, D. B., Paschalis, E. P., Klaushofer, K., Streli, C. & Roschger, P. (2011). Increased strontium uptake in trabecular bone of ovariectomized calcium-deficient rats treated with strontium ranelate or strontium chloride. *J. Synchrotron Rad.*, **18**, 835–841.

Penner-Hahn, J. E. (2005). Characterization of ´spectroscopically quiet´ metals in biology. *Coord. Chem. Rev.*, **249**, 161–177.

Peters, F., Schwarz, K. & Epple, M. (2000). The structure of bone studied with synchrotron X-ray diffraction, X-ray absorption spectroscopy and thermal analysis. *Thermochim. Acta*, **361**, 131–138.

Phillips, S. E. V. (1980). Structure and refinement of oxymyoglobin at 1.6 Å resolution. *J. Mol. Biol.*, **142**, 531–554.

Prugovečki, B., Dodson, E. J., Dodson, G. G. & Matković-Čalogović, D. (2009). Structure of the T_6 Human Nickel Insulin Derivative at 1.35 Å Resolution. *Croatica Chemica Acta*, **82**, 433–438.

Raffalt, A. C. (2011). *Effects of strontium malonate (NB S101) on compositional, structural and biomechanical properties of calcified tissues in rats and dogs*. Ph.D. thesis, Chemical Department, Technical University of Denmark.

Raffalt, A. C., Andersen, J. E. T. & Christgau, S. (2008). Application of inductively coupled plasma-mass spectrometry (ICP-MS) and quality assurance to study the incorporation of strontium into bone, bone marrow,and teeth of dogs after one month of treatment with strontium malonate. *Anal. Bioanal. Chem.*, **391**, 2199–2207.

Rahuel-Clermont, S., French, C. A., Kaarsholm, N. C. & Dunn, M. F. (1997). Mechanisms of Stabilization of the Insulin Hexamer through Allosteric Ligand Interactions. *Biochemistry*, **36**, 5837–5845.

Ranieri-Raggi, M., Raggi, A., Martini, D., Benvenuti, M. & Mangani, S. (2003). XAS of dilute biological samples. *J. Synchrotron Rad.*, **10**, 69–70.

Ravel, B. & Newville, M. (2005). ATHENA, ARTEMIS, HEPHAESTUS: data analysis for X-ray absorption spectroscopy using IFEFFIT. *J. Synchrotron Rad.*, **12**, 537–541.

Reginster, J. Y., Seeman, E., De Vernejoul, M. C., Adami, S., Compston, J., Phenekos, C., Devogelaer, J. P., Curiel, M. D., Sawicki, A., Goemaere, S., Sorensen, O. H., Felsenberg, D. & Meunier, P. J. (2005). Strontium ranelate reduces the risk of nonvertebral fractures in postmenopausal women with osteoporosis: Treatment of Peripheral Osteoporosis (TROPOS) study. *J. Clin. Endocrinol. Metab.*, **90**, 2816–2822.

Reginster, J. Y., Felsenberg, D., Boonen, S., Diez-Perez, A., Rizzoli, R., Brandi, M. L., Spector, T. D., Brixen, K., Goemaere, S., Cormier, C., Balogh, A., Delmas, P. D. & Meunier, P. J. (2008). Effects of long-term strontium ranelate treatment on the risk of nonvertebral and vertebral fractures in postmenopausal osteoporosis: results of a five-year, randomized, placebo-controlled trial. *Arthritis Rheum.*, **58**, 1687–1695.

Rehr, J. J. & Albers, R. C. (1990). Scattering-matrix formulation of curved-wave multiple-scattering theory: Application to x-ray-absorption fine structure. *Phys. Rev. B*, **41**, 8139–8149.

Rehr, J.J. & Albers, R.C. (2000). Theroetical approaches to x-ray absorption fine structure. *Review of Modern Physics*, **72**, 621–654.

Ressler, T. (1998). WinXAS: a Program for X-ray Absorption Spectroscopy Data Analysis under MS-Windows. *J. Synchrotron Rad.*, **5**, 118–122.

Rey, C., Combes, C., Drouet, C., Sfihi, H. & Barroug, A. (2007). Physico-chemical properties of nanocrystalline apatites: Implications for biominerals and biomaterials. *Mat. Sci. Eng. C*, **27**, 198–205.

Richwin, M., Zaeper, R., Lützenkirchen-Hecht, D. & Frahm, R. (2001). Piezo-QEXAFS: advances in time-resolved X-ray absorption spectroscopy. *J. Synchrotron Rad.*, **8**, 354–356.

Röntgen, W.C. (1896). On a New Kind of Rays. *Nature*, **53**, 274–276.

Rokita, E., Hermes, C., Nolting, H.-F. & Ryczek, J. (1993). Substitution of calcium by strontium within selected calcium phosphates. *J. Cryst. Growth*, **130**, 543–552.

Roy, M., Brader, M. L., Lee, R. W., Kaarsholm, N. C., Hansen, J. F. & Dunn, M. F. (1989). Spectroscopic signatures of the T to R conformational transition in the insulin hexamer. *J. Biol. Chem.*, **264**, 19081–19085.

Ryle, A. P., Sanger, F., Smith, L. F. & Kitai, R. (1955). The disulphide bonds of insulin. *Biochem. J.*, **60**, 541–556.

Rypniewski, W. R., Mangani, S., Bruni, B., Orioli, P. L., Casati, M. & Wilson, K. S. (1995). Crystal Structure of Reduced Bovine Erythrocyte Superoxide Dismutase at 1.9 Å Resolution. *J. Mol. Biol.*, **251**, 282–296.

Sandmark, C. & Brändén, C. (1967). The Crystal Structure of Hexaimidazole Zinc(II) Dichloride Tetrahydrate, $Zn(C_3H_4N_2)_6Cl_2{\cdot}4H_2O$. *Acta Chem. Scand.*, **21**, 993–999.

Sarangi, R., Hocking, R. K., Neidig, M. L., Benfatto, M., Holman, T. R., Solomon, E. I., Hodgson, K. O. & Hedman, B. (2008). Geometric Structure Determination of N694C Lipoxygenase: A Comparative Near-Edge X-Ray Absorption Spectroscopy and Extended X-Ray Absorption Fine Structure Study. *Inorg. Chem.*, **47**, 11543–11550.

Sayers, D.E., Stern, E. A. & Lytle, F.W. (1971). New Technique for Investigating Noncrystalline Structures: Fourier Analysis of the Extended X-Ray Absorption Fine Structure. *Physical Review Letters*, **27**, 1204–1207.

Sayers, D.E., Lytle, F.W., Weissbluth, M. & Pianetta, P. (1975). A short range probe for investigating metalloprotein structures: Fourier analysis of extended x-ray absorption fine structure. *The Journal of Chemical Physics*, **62**, 2514–2515.

Schlichtkrull, J. (1956). Insulin Crystals: I. The minimum Mole-Fraction of Metal in Insulin Crystals Preparred With Zn^{++}, Cd^{++}, Co^{++}, Ni^{++}, Cu^{++}, Mn^{++}, or Fe^{++}. *Acta Chem. Scand.*, **10**, 1455–1458.

Schlichtkrull, J. (1958). *Insulin Crystals: Chemical and Biological Studies on Insulin Crystals and Insulin Zinc Suspensions*. Ejnar Munksgaard Publishers.

Schrooten, I., Behets, G. J. S., Cabrera, W. E., Vercauteren, S. R., Lamberts, L. V., Verberckmoes, S. C., Bervoets A. J., Dams, G., Goodman, W. G., De Broe, M. E. & D'Haese, P. C. (2003). Dose-dependent effects of strontium on bone of chronic renal failure rats. *Kidney Int.*, **63**, 927–935.

Scott, D. A. (1934). Crystalline insulin. *Biochem. J.*, **28**, 1592–1602.

Scott, D. A. & Fisher, A. M. (1935). Crystalline insulin. *Biochem. J.*, **29**, 1048–1054.

Scott, R.A. (1985). Measurement of Metal–Ligand Distances by EXAFS. *Method in Enzymology*, **117**, 414–459.

Seeman, E. & Delmas, P. D. (2006). Bone Quality – The Material and Structural Basis of Bone Strength and Fragility. *New England J. Med.*, **354**, 2250–2261.

Servier. (2012). *Servier Medical Art*. http://www.servier.com.

Sery, A., Manceau, A. & Greaves, G. N. (1996). Chemical state of Cd in apatite phosphate ores as determined by EXAFS spectroscopy. *Am. Mineral.*, **81**, 864–873.

Seward, T. M., Henderson, C. M. B., Charnock, J. M. & Dreisner, T. (1999). An EXAFS study of solvation and ion pairing in aqueous strontium solutions to 300°C. *Geochim. Cosmochim. Acta*, **63**, 2409–2418.

Shao, Y., Molnar, L. F., Jung, Y., Kussmann, J., Ochsenfeld, C., Brown, S. T., Gilbert, A. T. B., Slipchenko, L. V., Levchenko, S. V., O´Neill, D. P. DiStasio, R. A., Lochan, R. C., Wang, T., Beran, G. J. O., Besley, N. A., Herbert, J. M., Lin, C. Y., van Voorhis, T., Chien, S. H., Sodt, A., Steele, R. P., Rassolov, V. A., Maslen, P. E., Korambath, P. P., Adamson, R. D., Austin, B., Baker, J., Byrd, E. F. C., Dachsel, H., Doerksen, R. J., Dreuw, A., Dunietz, B. D.,

Dutoi, A. D., Furlani, T. R., Gwaltney, S. R., Heyden, A., Hirata, S., Hsu, C.-P., Kedziora, G., Khalliulin, R. Z., Klunzinger, P., Lee, A. M., Lee, M. S., Liang, W. Z., Lotan, I., Nair, N., Peters, B., Proynov, E. I., Pieniazek, P. A., Rhee, Y. M., Ritchie, J., Rosta, E., Sherrill, C. D., Simmonett, A. C., Subotnik, J. E., Woodcock III, H. L., Zhang, W., Bell, A. T., Chakraborty, A. K., Chipman, D. M., Keil, F. J., Warshel, A., Hehre, W. J., Schaefer, H. F., Kong, J., Krylov, A. I., Gill, P. M. W. & Head-Gordon, M. (2006). Advances in methods and algorithms in a modern quantum chemistry program package . *Phys. Chem. Chem. Phys.*, **8**, 3172–3191.

Shape Software. (2006). *ATOMS v6.3.1*.

Shulman, R. G., Eisenberger, P., Blumberg, W. E. & Stombaugh, N. A. (1975). Determination of the iron-sulfur distances in rubredoxin by x-ray absorption spectroscopy. *Proc. Nat. Acad. Sci. USA*, **72**, 4003–4007.

Smith, G. D. & Ciszak, E. (1994). The structure of a complex of hexameric insulin and 4'-hydroxyacetanilide. *Proc. Nat. Acad. Sci. USA*, **91**, 8851–8855.

Smith, G. D. & Dodson, G. G. (1992). Structure of a Rhombohedral R_6 Insulin/Phenol Complex. *Proteins: Structure, Function, and Bioinformatics*, **14**, 401–408.

Smith, G. D., Swenson, D. C., Dodson, E. J., Dodson, G. G. & Reynolds, C. D. (1984). Structural stability in the 4-zinc human insulin hexamer. *Proc. Nat. Acad. Sci. USA*, **81**, 7093–7097.

Smith, G. D., Ciszak, E. & Pangborn, W. (1996). A novel complex of a phenolic derivative with insulin: Structural features related to the T→R transition. *Protein Sci.*, **5**, 1502–1511.

Smith, G. D., Ciszak, E., Magrum, L. A., Pangborn, W. A. & Blessing, R. H. (2000). R_6 hexameric insulin complexed with m-cresol or resorcinol. *Acta Cryst. D*, **56**, 1541–1548.

Smith, G. D., Pangborn, W. & Blessing, R. H. (2001). Phase changes in $T_3R_3^f$ human insulin: temperature or pressure induced? *Acta Cryst. D*, **57**, 1091–1100.

Smith, G. D., Pangborn, W. A. & Blessing, R. H. (2003). The structure of T_6 human insulin at 1.0 Å resolution. *Acta Cryst. D*, **59**, 474–482.

Smith, G. D., Pangborn, W. A. & Blessing, R. H. (2005). The structure of T_6 bovine insulin. *Acta Cryst. D*, **61**, 1476–1482.

Smolentsev, G. & Soldatov, A. (2006). Quantitative local structure refinement from XANES: multi-dimensional interpolation approach. *J. Synchrotron Rad.*, **13**, 19–29.

Smolentsev, G. & Soldatov, A. V. (2007). FitIt: New software to extract structural information on the basis of XANES fitting. *Comp. Mat. Sci.*, **39**, 569–574.

Sreekanth, R., Pattabhi, V. & Rajan, S. S. (2009). Metal induced structural changes observed in hexameric insulin. *Int. J. Biol. Macromol.*, **44**, 29–36.

Steiner, D. F., Cunningham, D., Spigelman, L. & Aten, B. (1967). Insulin Biosynthesis: Evidence for a Precursor. *Science*, **157**, 697–700.

Ståhl, K. (2009). Crystal structure of Sr disalicylate dihydrate. *Unpublished results*.

Ståhl, K., Frankær, C. G., Raffalt, A. C., Sørensen, S. R. & Andersen, J. E. T. (2011). Polymeric strontium ranelate nonahydrate. *Acta Cryst. E*, **67**, m471–m472.

Strange, R. W. & Feiters, M. C. (2008). Biological X-ray absorption spectroscopy (BioXAS): a valuable tool for the study of trace elements in the life sciences. *Current Opinion in Structural Biology*, **18**, 609–616.

Strange, R. W., Murphy, L. M., Dodd, F. E., Abraham, Z. H. L., Eady, R. R., Smith, B. E. & Hasnain, S. S. (1999). Structural and Kinetic Evidence for an Ordered Mechanism of Copper Nitrite Reductase. *J. Mol. Biol.*, **287**, 1001–1009.

Strange, R. W., Ellis, M. & Hasnain, S. S. (2005). Atomic resolution crystallography and XAFS. *Coord. Chem. Rev.*, **249**, 197–208.

Strange, R. W., Antonyuk, S. V., Hough, M. A., Doucette, P. A., Valentine, J. S. & Hasnain, S. S. (2006). Variable Metallation of Human Superoxide Dismutase: Atomic Resolution Crystal Structures of Cu-Zn, Zn-Zn and As-isolated Wild-type Enzymes. *J. Mol. Biol.*, **356**, 1152–1162.

Sudarsanan, K. & Young, R. A. (1972). Structure of Strontium Hydroxide Phosphate, $Sr_5(PO_4)_3OH$. *Acta Cryst. B*, **28**, 3668–3670.

Sudmeier, J. L., Bell, S. J., Storm, M. C. & Dunn, M. F. (1981). Cadmium-113 nuclear magnetic resonance studies of bovine insulin: two-zinc insulin hexamer specifically binds calcium. *Science*, **212**, 560–562.

Sugiyama, S., Moriga, T., Hayashi, H. & Moffat, J. B. (2001). Characterization of Ca, Sr, Ba and Pb HAP: X-ray diffraction, photoelectron, EXAFS and MAS NMR spectroscopies. *Bull. Chem. Soc. Jpn.*, **74**, 187–192.

Tamm, T. & Peld, M. (2006). Computational study of cation substitutions in apatites. *J. Solid State Chem.*, **179**, 1581–1587.

Terra, J., Dourado, E. R., Eon, J.-G., Ellis, D. E., Gonzalez, G. & Rossi, A. M. (2009). The structure of strontium-doped hydroxyapatite: an experimental and theoretical study. *Phys. Chem. Chem. Phys.*, **11**, 568–577.

The Mathworks. (2008). *MATLAB R2008a*.

Ubgade, R. & Sarode, P. R. (1987). Study of Strontium Compounds and Minerals by X-Ray Absorption Spectroscopy. *Phys. Stat. Sol. A*, **99**, 295–301.

Umena, Y., Kawakami, K., Shen, J.-R. & Kamiya, N. (2011). Crystal structure of oxygen-evolving photosystem II at a resolution of 1.9 Å. *Nature*, **473**, 55–61.

Verberckmoes, S. C., Behets, G. J., Oste, L., Bervoets, A. R., Lamberts, L. V., Drakopoulos, M., Somogyi, A., Cool, P., Dorriné, W., De Broe, M. E. & D'Haese, P. C. (2004). Effects of Strontium on the Physicochemical Characteristics of Hydroxyapatite. *Calcif. Tissue Int.*, **75**, 405–415.

von Dreele, R. B. (1999). Combined Rietveld and stereochemical restraint refinement of a protein crystal structure. *J. Appl. Cryst.*, **32**, 1084–1089.

von Dreele, R. B. (2001). Binding of *N*-acetylglycosamine to chicken egg lysozyme: a powder diffraction study. *Acta Cryst. D*, **57**, 1836–1842.

von Dreele, R. B. (2005). Binding of *N*-acetylglucosamine oligosaccharides to hen egg-white lysozyme: a powder diffraction study. *Acta Cryst. D*, **61**, 22–32.

von Dreele, R. B. (2009). Characterization of proteins by powder diffraction. *Z. Kristallogr. Suppl.*, **30**, 27–32.

von Dreele, R. B., Stephens, P. W., Smith, G. D. & Blessing, R. H. (2000). The first protein crystal structure determined from high-resolution X-ray powder diffraction data: a variant of T_3R_3 human insulin-zinc complex produced by grinding. *Acta Cryst. D*, **56**, 1549–1551.

von Dreele, R. B., Lee, P. L. & Zhang, Y. (2006). Protein polycrystallography. *Z. Kristallogr. Suppl.*, **23**, 3–8.

Watenpaugh, K. D., Sieker, L. C. & Jensen, L. H. (1980). Crystallographic Refinement of Rubredoxin at 1.2 Å Resolution. *J. Mol. Biol.*, **138**, 615–633.

Weiner, S. & Wagner, H. D. (1998). The Material Bone: Structure-Mechanical Function Relations. *Annu. Rev. Mat. Res.*, **28**, 271–298.

Westenhoff, S., Nazarenko, E., Malmerberg, E., Davidsson, J., Katonaa, G. & Neutze, R. (2010). Time-resolved structural studies of protein reaction dynamics: a smorgasbord of X-ray approaches. *Acta Cryst. A*, **66**, 207–219.

Whittingham, J. L., Chaudhuri, S., Dodson, E. J., Moody, P. C. E. & Dodson, G. G. (1995). X-ray crystallographic studies on hexameric insulins in the presence of helix-stabilizing agents, thiocyanate, methylparaben and phenol. *Biochemistry*, **34**, 15553–15563.

Wierzbicki, M., Bonnet, J., Brisset, M. & Tsouderos, Y. (1990). *Nouveaux sel de metaux bivalents de l'acide N, N-di(carboxyméthyl)amino-2 cyano-3 carboxyméthyl-4 carboxy-5 thiophène leur procede de preparation et les compositions pharmaceutiques les renfermant, EP Patent Specification No. 0415850*.

Wopenka, B & Pasteris, J. D. (2005). A mineralogical perspective on the apatite in bone. *Mat. Sci. Eng. C*, **25**, 131–143.

World Health Organization. (2012). *World Health Organization Diabetes Programme*. http://www.who.int/diabetes/en/.

Wright, J. P., Besnard, C., Margiolaki, I., Basso, S., Camus, F., Fitch, A. N., Fox, G. C., Pattison, P. & Schiltz, M. (2008). Molecular envelopes derived from protein powder diffraction data. *J. Appl. Cryst.*, **41**, 329–339.

Yano, J. & Yachandra, V. K. (2008). Where Water Is Oxidized to Dioxygen: Structure of the Photosynthetic Mn_4Ca Cluster from X-ray Spectroscopy. *Inorg. Chem.*, **47**, 1711–1726.

Yano, J., Kern, J., Sauer, K., Latimer, M. J., Pushkar, Y., Biesiadka, J., Loll, B., Saenger, W., Messinger, J., Zouni, A. & Yachandra, V. K. (2006). Where Water Is Oxidized to Dioxygen: Structure of the Photosynthetic Mn_4Ca Cluster. *Science*, **314**, 821–825.

Zhang, Y., Cheng, F., Li, D., Wang, Y., Zhang, G., Liao, W., Tang, T., Huang, Y. & He, W. (2005). Investigation of Elemental Content Distribution in Femoral Head Slice with Osteoporosis by SRXRF Microprobe. *Biol. Trace Elem. Res.*, **103**, 177–185.

Zoeger, N., Streli, C., Wobrauschek, P., Jokubonis, C., Pepponi, G., Roschger, P., Hofstaetter, J., Berzlanovich, A., Wegrzynek, D., Chinea-Cano, E., Markowicz, A., Simon, R. & Falkenberg, G. (2008). Determination of the elemental distribution in human joint bones by SR micro XRF. *X-Ray Spectrom.*, **37**, 3–11.

Appendices

Appendix A

Publications

A.1 Paper I: In-house characterization of protein powder

Christian Grundahl Hartmann, Ole Faurskov Nielsen, Kenny Ståhl & Pernille Harris
J. Appl. Cryst. (2010). **43,** 876−882

Journal of
Applied Crystallography
ISSN 0021-8898

Received 9 March 2010
Accepted 25 May 2010

In-house characterization of protein powder

Christian Grundahl Hartmann,[a] Ole Faurskov Nielsen,[b] Kenny Ståhl[a] and Pernille Harris[a]*

[a]Department of Chemistry, Technical University of Denmark, Kemitorvet 207, DK-2800 Kongens Lyngby, Denmark, and [b]Department of Chemistry, University of Copenhagen, Universitetsparken 5, DK-2100 Copenhagen, Denmark. Correspondence e-mail: ph@kemi.dtu.dk

X-ray powder diffraction patterns of lysozyme and insulin were recorded on a standard in-house powder diffractometer. The experimental powder diffraction patterns were compared with patterns calculated from Protein Data Bank coordinate data. Good agreement was obtained by including straightforward corrections for background, unit-cell parameters, disordered bulk solvent and geometric factors. In particular the solvent correction was found crucial for a good agreement. A revised Lorentz factor was derived, which gave a minor, but significant, improvement to the fit in the low-angle region. An attempt to include calculated H-atom positions did not improve the overall fit and was abandoned. The method devised was shown to be a quick and convenient tool for distinguishing precipitates and polymorphs of proteins.

1. Introduction

X-ray diffraction on single crystals is the preferable method to solve and characterize protein structures. The results are usually unambiguously clear, when the structure is revealed. One of the bottlenecks in this process is the need to find the right conditions for growing suitable single crystals. The ability to screen early precipitates for crystals and possible polymorphs would save large amounts of time and effort. Furthermore, as structure-based drug design is a growing field and drug candidates are often proteins themselves, there is a growing need for quick characterization of polymorphs and substrate–protein complexes. High-resolution synchrotron X-ray powder diffraction (XRPD) has already proved its ability for structure solution and refinements of small protein structures from powder samples (von Dreele, 1999, 2001, 2005; von Dreele *et al.*, 2000; Margiolaki & Wright, 2008; Wright *et al.*, 2008), but synchrotron XRPD is not available on a daily basis. However, in-house XRPD available in most chemistry laboratories may offer an alternative. A recent study using medium-resolution XRPD demonstrated the ability to quickly and effectively distinguish between different polymorphs of insulin in combination with multivariate data analysis (Norrman *et al.*, 2006). Based on these previous experiences several challenges can be identified:

(*a*) The peak overlap problem. Only the first 15–20° in 2θ with Cu $K\alpha$ radiation will be of interest.

(*b*) The measured intensities fit poorly with those calculated from known structures. Protein structures available from the Protein Data Bank (PDB; Bernstein *et al.*, 1977) do not generally contain H-atom positions, and the contribution from the disordered solvent in the solvent channels, which is the major source for this discrepancy, is not described. Furthermore, the conventional Lorentz correction tends to infinity when approaching $2\theta = 0°$.

(*c*) The measured peak positions sometimes deviate from those derived from the literature. In many cases this is due to a difference in data collection temperature, but small differences in the amount of solvent and the exact composition of the solvent may also slightly affect the unit-cell parameters.

(*d*) The background problem. The relatively poor scattering from the protein combined with the scattering from mother liquor and sample holder result in a poor signal-to-noise ratio.

(*e*) A relatively large amount of sample and/or measuring time is needed.

In this work we present a study of these challenges, and show some solutions based on easily available software, which can be routinely applied to in-house XRPD. Specifically, we discuss the following: background fitting and subtraction, unit-cell refinement, adding H-atom positions and solvent scattering to the structure model, powder pattern calculation, and proper Lorentz correction at low angles.

2. Experimental

Lyophilized lysozyme from chicken egg white, *Gallus gallus* (95%, Sigma Aldrich L-6876), and insulin from bovine pancreas, *Bos taurus* (Sigma Aldrich I-5500), were used in this study. All crystals were grown using the vapour diffusion technique.

2.1. Crystal growth

For the growth of tetragonal lysozyme the protein was dissolved in 0.02 *M* sodium acetate and the protein concentration was determined to be 13.8 mg ml^{-1} using an extinction coefficient at 280 nm, ε_{280}, of 37 470 M^{-1} cm^{-1}. The protein was equilibrated in a sitting drop against a reservoir of 0.1 *M* HEPES [4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid] and 0.8 *M* potassium sodium tartrate ($KNaC_4H_4O_6$) at pH 7.5

(*cf.* Crystal Screen HR-2-110 No. 29, Hampton Research). The drops were formed of 200 µl of protein solution mixed with 200 µl of reservoir solution.

For the growth of trigonal insulin crystals the insulin was dissolved in 0.01 *M* HCl, the pH was adjusted to 6 with NaOH and the protein concentration was determined to be 5.0 mg ml^{-1} using an ε_{280} of 5960 M^{-1} cm^{-1}. Using the sitting-drop method the protein was equilibrated against a reservoir of precipitant solution of 33 vol.% acetone, 0.11 *M* sodium citrate and 16.6 m*M* zinc acetate at pH 6.5 (Smith *et al.*, 1982). The drops were formed of 200 µl of protein solution mixed with 160 µl of reservoir solution.

For the growth of cubic insulin crystals the protein was dissolved in a 0.02 *M* Na_2HPO_4 and 0.01 *M* Na_3EDTA solution (EDTA is ethylenediaminetetraacetic acid) and the protein concentration was determined to be 8.9 mg ml^{-1}. Using the sitting-drop method the protein was equilibrated against a reservoir of precipitant solution comprising a 0.3 *M* Na_2HPO_4/Na_3PO_4 buffer system adjusted to pH 10.0 and 0.01 *M* Na_3EDTA. The drops were formed of 200 µl of protein solution mixed with 200 µl of reservoir solution.

2.2. Single-crystal X-ray diffraction of lysozyme

The structure of lysozyme has been refined from a single crystal from the same crystal batch as the powder sample in order to obtain an optimal basis for the calculated powder diffraction pattern of lysozyme.

A single crystal was soaked in a cryogenic solution of mother liquor with 20 vol.% glycerol and mounted in a loop. Data were collected at MAX-lab in Lund, Sweden, on beamline I911-3, MAX II, using a Marresearch CCD detector. The crystals diffracted beyond 2.05 Å, but owing to the emphasis on solvent and Lorentz correction in this study, high-resolution data were not required and a data set to 2.05 Å was collected. The data were processed in space group $P4_32_12$ and scaled using the programs *XDS* and *XSCALE* (Kabsch, 1993).

The structure was refined with starting coordinates from PDB entry 1lsa (Kurinov & Harrison, 1995), with the ordered water molecules omitted. The program *REFMAC5* (Murshudov *et al.*, 1997) included in the package *CCP4* (Collaborative Computational Project, Number 4, 1994) was used for refinement, and model building was carried out by *WinCoot* (Emsley & Cowtan, 2004). One molecule of HEPES and 116 ordered water molecules were found. Data collection and refinement statistics are summarized in Table 1.

2.3. X-ray powder diffraction

X-ray powder diffraction patterns were recorded for tetragonal crystals of lysozyme and for trigonal and cubic insulin crystals. Batches containing 2–6 mg of protein of the three types of crystals were grown and the crystals, in a slurry with mother liquor, were gently crushed to a powder with a pestle. The diffraction experiments were carried out at an in-house Huber G670 diffractometer using Cu $K\alpha_1$ radiation (λ = 1.54059 Å). The samples were mounted in a specially designed sample holder (details to be published along with new design),

Table 1
Data collection and refinement statistics for the lysozyme single-crystal data set.

Data collection	
Beamline	I911-3, MAX II, MAX-lab
Wavelength (Å)	1.5000
Temperature (K)	100
Resolution (Å)	40.00–2.05 (2.31–2.05)†
No. of reflections	99 201 (29 371)†
No. of unique reflections	7511 (2202)†
Redundancy	13.2 (13.3)†
Completeness (%)	99.9 (100.0)†
R_{sym} (%)‡	4.1 (9.2)†
$\langle I/\sigma(I)\rangle$	43.93 (25.68)†
Processing	
Space group	$P4_32_12$
No. of molecules per asymmetric unit	1
a (Å)	77.05 (5)
c (Å)	37.30 (3)
Solvent content	35.8%
Refinement	
R§	0.1695
R_{free}§	0.1729

† Values in parentheses are for the outermost resolution shell. ‡ $R_{sym} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$ where $\langle I(hkl)\rangle$ is the mean intensity of a set of equivalent reflections. § *R* and $R_{free} = \sum ||F_{obs}| - |F_{calc}|| / \sum |F_{obs}|$, where F_{obs} and F_{calc} are the observed and calculated structure-factor amplitudes, respectively. R_{free} was calculated with a random 5% subset of all reflections excluded from the refinement.

which contains approximately 150 µl of the powder suspension (corresponding to approximately 5 mg of protein). The measurements were performed on rotating samples at room temperature. Lysozyme data were collected at room temperature for 15 × 4 h, and insulin data for 4 × 4 h. None of the powder diffraction series indicated any time decay. The background was treated in the program *PROTPOW* (http://www.xray.kemi.dtu.dk/English/Computer%20Programs/PROTPOW.aspx) by manually picking between ten and 20 background points, fitting to a spline function and subtracting.

3. Corrections to XRPD patterns calculated from PDB coordinates

Powder diffraction patterns were calculated for three different protein crystals: tetragonal lysozyme, and cubic and trigonal insulin. The lysozyme structure was refined from a crystal originating from the same batch as the lysozyme powder (§2.2). For insulin the structures were downloaded from the PDB. The resulting structure factors, before and after corrections, were entered into *PROTPOW* for powder diffraction pattern calculations. The following section describes how application of a solvent correction, including H-atom positions, and using a revised Lorentz factor affects the calculated data. All corrections were carried out on the calculated patterns.

3.1. Optimization of unit-cell parameters

It was necessary to optimize the unit-cell parameters to obtain fully comparable powder patterns. This could be due to

Table 2
Refined cell parameters from experimental patterns, parameters from the solvent and the geometrical corrections, and residual values evaluating the effects of the corrections.

The bulk-solvent parameters are found from refinement of our own data using *PHENIX.refine* (Adams *et al.*, 2002) for lysozyme data and in the *EDS* (Kleywegt *et al.*, 2004) for insulin data.

	Lysozyme			
	Without H	With H	Insulin, trigonal	Insulin, cubic
General parameters				
PDB	This work		2zp6	1bph
Space group	$P4_32_12$		$R3$	$I2_13$
a (Å)	76.699 (3)		82.584 (8)	79.136 (7)
c (Å)	37.112 (2)		33.862 (4)	–
R_p before solvent correction† (%)	92		99	100
Hydrogen and bulk-solvent correction				
k_{sol} (e Å^{-3})	0.39	0.45		
B_{sol} (Å^2)	51	58		
Reference	This work	This work		
R_p after solvent correction† (%)	23.6	26.6		
Bulk-solvent correction				
k_{sol} (e Å^{-3})	0.39		0.34	0.28
B_{sol} (Å^2)	51		90	90
Reference	This work		*EDS*	*EDS*
R_p after solvent correction† (%)	23.6		19.1	23.3
Geometric correction				
η (Å^{-1})	0.031		0.053	0.032
R_p after geometric correction† (%)	23.4		16.5	21.9

† $R_p = \sum |I_{obs} - kI_{calc}| / \sum |I_{obs}|$.

differences in data collection temperatures (lysozyme powder data were collected at room temperature while the single-crystal data set was collected at 100 K) or variations in solvent concentrations. Optimization can be carried out with a full-pattern profile fit in the program *PROTPOW*. Optimal parameters describing the peak width (FWHM) and peak shape of a pseudo-Voigt function are included in the pattern fit. The peak width is modelled by $(v\tan\theta + w)^{1/2}$ and the fraction of Lorentzian and Gaussian character is modelled by a γ parameter. The peak asymmetry effects can be modelled by splitting the peak into a left (l) and a right (r) hand side contribution. The unit-cell and peak-shape parameters were optimized from the experimental patterns. In the present cases the profile parameter set consisted of w_l, w_r, v, γ_l and γ. The resulting unit-cell parameters are given in Table 2.

3.2. Correction of disordered solvent

Protein crystals contain 30–80% of disordered solvent throughout the structure. The contribution to the diffraction pattern from the disordered solvent is most important for low-resolution data (11–40 Å). Several correction models have been developed. The simplest is the exponential scaling model based upon the Babinet (1837) principle, according to which a simple scale factor function is used (Moews & Kretsinger, 1975). This model assumes that the contribution from the disordered solvent is distributed isotropically. However, this is an approximation since protein crystals contain solvent channels along specific crystallographic directions. The flat bulk-solvent correction model is anisotropic and assumes a flat solvent profile in the solvent regions (Phillips, 1980). The model defines the solvent regions by a solvent mask. Other models, which are all based on the solvent mask principle but assume more complex distributions of electron density in the solvent regions, are reviewed by Jiang & Brünger (1994), from which it is concluded that the flat bulk-solvent model is the simplest model that still gives a realistic correction. The model takes two parameters: k_{sol} defining the level of electron density in the solvent region, and B_{sol} defining the steepness of the border between the solvent and the macromolecular regions. The parameters are refined in today's single-crystal X-ray structure refinement software, and based on 1162 structures Fokine & Urzhumtsev (2002) found that they accumulate around $k_{sol} = 0.35$ e Å^{-3} and $B_{sol} = 46$ Å^2.

The flat bulk-solvent correction model is used in the following to correct for the disordered bulk solvent. This is done using *PHENIX.pdbtools* (Adams *et al.*, 2002), which only requires a PDB coordinate file, k_{sol} and B_{sol} as input. The parameters k_{sol} and B_{sol} can be set to the average values in case no other information is found, but refined values or values found from the *Electron Density Server* (*EDS*; Kleywegt *et al.*, 2004) can improve the result. *PHENIX.pdbtools* calculates the structure factors for the model, $\mathbf{F}_{model}$, which include the contributions from both the protein, $\mathbf{F}_{calc}$, and the solvent, $\mathbf{F}_{sol}$. $\mathbf{F}_{model}$ can then be used for powder pattern calculations.

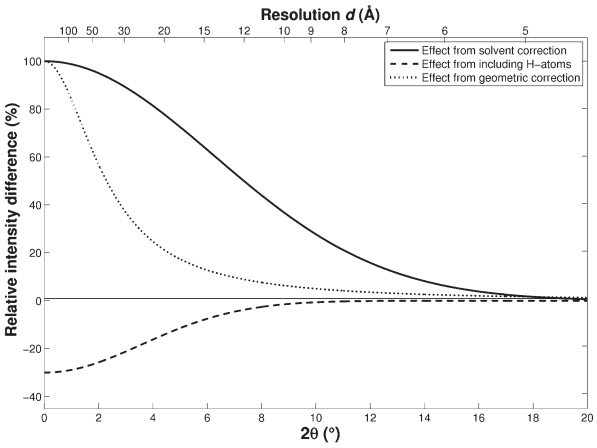


Figure 1
Overview of the trends from the different corrections. The effects are shown as the relative intensity difference $(I_{non\text{-}corr} - I_{corr})/I_{non\text{-}corr}$ plotted as functions of the scattering angle 2θ (using Cu $K\alpha_1$) and resolution $d = \lambda/(2\sin\theta)$. The curves are based on average corrections of lysozyme and insulin data. $I_{non\text{-}corr}$ is the raw intensity from a calculated pattern which has only been Lorentz corrected using the expression in equation (2). I_{corr} for the solvent and hydrogen corrections have been calculated according to the procedures described in §§3.2 and 3.3. The geometric correction curve has been calculated from the expression in equation (3) using $\eta = 0.045$ Å^{-1}.

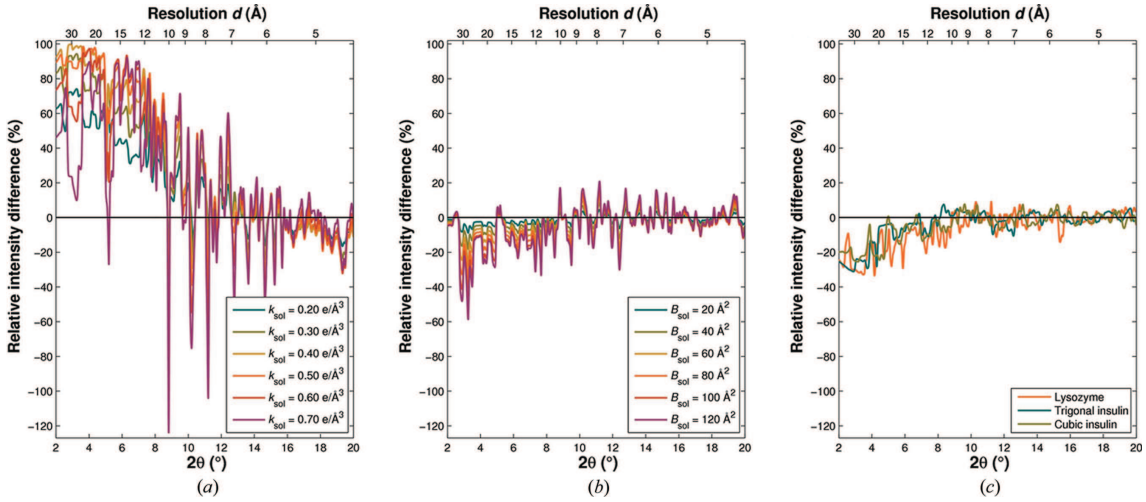


Figure 2
The influence of the solvent correction parameters (*a*) k_{sol} and (*b*) B_{sol} on calculated powder patterns for lysozyme and (*c*) the influence of H atoms on calculated powder patterns for lysozyme and for trigonal and cubic insulin. The effect is plotted as the relative intensity difference $(I_{non\text{-}corr} - I_{corr})/I_{non\text{-}corr}$ against 2θ (using Cu $K\alpha_1$) and resolution $d = \lambda/(2\sin\theta)$. For the variation of k_{sol} (*a*), the parameter B_{sol} has been kept constant to 50 Å^2. For the variation of B_{sol} (*b*), k_{sol} has been kept constant at the average value 0.35 e Å^{-3}.

The effect of the solvent correction was evaluated as the relative intensity difference given as $(I_{non\text{-}corr} - I_{sol})/I_{non\text{-}corr}$, where $I_{non\text{-}corr}$ and I_{sol} are the calculated intensities from non-solvent- and solvent-corrected powder diffraction patterns, respectively. The average effect is shown in Fig. 1 and is based on solvent corrections for lysozyme and for trigonal and cubic insulin. The stepwise effect from each of the solvent parameters k_{sol} and B_{sol} has been investigated in more detail for lysozyme (Figs. 2*a* and 2*b*, respectively). For k_{sol}, the effect has been evaluated relative to a calculated reference powder pattern using k_{sol} = 0.00 e Å^{-3} which corresponds to the case where no solvent correction has been carried out. For B_{sol}, k_{sol} has been kept constant at the average value 0.35 e Å^{-3} and the effect has been evaluated relative to a calculated reference powder diffraction pattern using k_{sol} = 0.35 e Å^{-3} and B_{sol} = 0 Å^2. The bulk-solvent correction is highly anisotropic and both parameters affect the anisotropy. While k_{sol} shows its largest influence at low resolution, B_{sol} seems to scale the correction with a much smaller dependence on resolution.

The PDB coordinates and the bulk-solvent parameters were given as inputs to *PHENIX.pdbtools*, which returned a set of calculated structure factors $\mathbf{F}_{model}$. In this study the bulk-solvent parameters for the two insulin types were taken from *EDS* (Kleywegt *et al.*, 2004), while our own data were used for lysozyme. The parameters used are listed in Table 2. The solvent correction effect is demonstrated in Fig. 3 on cubic insulin and is also seen as a decrease in the high initial R_p values listed in Table 2.

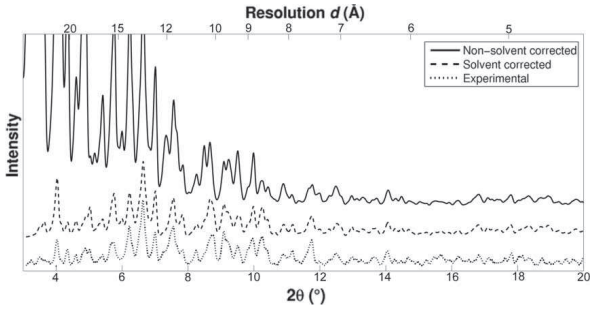


Figure 3
Calculated powder diffraction patterns for cubic insulin before (solid) and after (dashed) solvent correction. The experimental powder pattern (dotted) is included for comparison. The patterns have been scaled in order to obtain comparable intensities for $2\theta > 12°$, meaning that the scale factor k in equation (4) has been determined solely on the reflections where $2\theta > 12°$ [$d = \lambda/(2\sin\theta) < 7.37$ Å]. The calculated powder patterns are corrected by the traditional Lorentz factor [equation (2)]. The powder diffraction patterns have been translated along the y axis.

3.3. Hydrogen correction

Ideal H-atom positions can be calculated using *PHENIX.pdbtools* and then included in the PDB file. The assignment of B factors for the new H atoms is achieved by using the B factor from the adjacent non-H atom. The solvent correction can then be carried out according to the procedure described in §3.2 using the model with the H atoms included. The effect of H-atom inclusion was evaluated as the relative intensity difference given as $(I_{non\text{-}corr} - I_{hyd})/I_{non\text{-}corr}$, where $I_{non\text{-}corr}$ and I_{hyd} are the calculated intensities without and with H-atom positions, respectively.

The average effect is shown in Fig. 1, and the separate effects for lysozyme and for trigonal and cubic insulin are shown in Fig. 2(*c*). As for the solvent correction, the effect is anisotropic and affects mostly the low-resolution data, indicating that there may be a strong correlation between these two corrections.

3.4. Geometric intensity correction for small crystals with large unit cells

The geometric intensity correction, known as the Lorentz factor, L, describes the fraction of a reflection that is in the reflecting condition. In the case of Bragg–Brentano and Debye–Scherrer geometries it is given by

$$L = \frac{1}{\sin(2\theta)} \frac{1}{\sin\theta}. \quad (1)$$

In the case of the Guinier geometry an additional term is required, which is related to the angle between the tangent of the focusing circle and the primary beam, θ_G (typically 45°) (Sas & de Wolff, 1966; Ståhl, 2000):

$$L_G = \frac{1}{\sin(2\theta)} \frac{1}{\sin\theta} \frac{1}{\cos(\theta_G - 2\theta)}. \quad (2)$$

Equations (1) and (2) both assume ideal crystals resulting in infinitesimally small reciprocal lattice points. However, the true size of the reciprocal lattice points depends on the crystallite size, on the internal stress and on defects (Snell *et al.*, 2003), hence the larger and more perfect the crystals, the smaller and more well defined the points. The quality of macromolecular crystals is more erratic than crystals of small molecules, resulting in a considerable smearing of the reciprocal lattice points. This effect has to be included in the Lorentz factor and is introduced by the parameter η, which describes the uncertainty of the length of the scattering vector $|\mathbf{s}|$. A revised Lorentz correction factor L_{rev} for protein powder has been derived using η (the derivation is found in Appendix *A*):

$$L_{rev} = \frac{1}{\sin(2\theta)} \frac{1}{\sin\theta} \frac{1}{\cos(\theta_G - 2\theta)} \frac{\sin^2\theta}{(\sin^2\theta + \lambda^2\eta^2/12)}. \quad (3)$$

With this model the Lorentz factor does not approach infinity when 2θ approaches zero. The corrective effect of the revised Lorentz factor has been calculated as $(I_{non\text{-}corr} - I_{geo})/I_{non\text{-}corr} = 1 - [\sin^2\theta/(\sin^2\theta + \lambda^2\eta^2/12)]$, where $I_{non\text{-}corr}$ and I_{geo} are the calculated intensities from powder diffraction patterns corrected with the traditional Lorentz factor, equation (2), and the revised Lorentz factor using λ = Cu $K\alpha_1$ and η = 0.045 Å^{-1}, equation (3), respectively. Fig. 1 shows that the correction to the Lorentz factor has a much smaller effect than either the solvent or the H-atom corrections.

The parameter η was determined empirically by testing different values of η and evaluating the fit between the calculated and the experimental powder diffraction patterns. Again the largest influence is observed for low-resolution data, although the effect is isotropic. The optimal η values found for the protein powders are presented in Table 2.

4. Discussion

In order to evaluate the improvements of the correction a profile R factor has been calculated before any correction, after solvent correction and after geometrical correction. The values are calculated from the intensities I:

$$R_p = \frac{\sum |I_{obs} - kI_{calc}|}{\sum |I_{obs}|}, \quad (4)$$

where k is a scale factor scaling the calculated intensities, I_{calc}, to the background-subtracted observed intensities, I_{obs}. The R_p values are listed in Table 2.

The effect of the solvent correction for the cubic insulin, which has the highest content of solvent (65.5%) along the (100) and (111) directions, is seen in Fig. 3. The effect on the intensities is, as seen, dramatic and brings the observed and calculated intensities into quite good agreement. However, the solvent parameters are quite sensitive and k_{sol} and B_{sol} may be further refined when experimental structure factors are available.

The inclusion of the H-atom positions in the calculation of a powder diffraction pattern was carried out for lysozyme. Because the effect of including H atoms correlates with the solvent correction, the solvent-correction parameters were refined after refining the lysozyme structure with and without H atoms using *PHENIX.refine* (Adams *et al.*, 2002). Both k_{sol} and B_{sol} increase by 10–15% when H atoms are included (Table 2). As seen from the R factors the hydrogen correction does not improve the fit, rather the contrary, possibly because modelling H atoms in ideal positions is too crude. The hydrogen correction was thus abandoned and it was decided to rely on the bulk-solvent correction only.

The patterns were finally corrected by the Lorentz factor L_{rev}, given in equation (3), and scaled. Comparing the intensities at the smallest 2θ angles, the correction seems to slightly improve the agreement between the calculated and the experimental patterns for all three crystal types. The effect is larger for the insulin data since the first strong reflections occur at $2\theta < 4°$, whereas for lysozyme the first strong measured reflections occur at $2\theta = 4°$. It should be noted that the nature of the Lorentz correction and the parameter η is such that it will correlate with, for instance, the background correction, the general level of the bulk-solvent correction, errors in temperature factors *etc.* Even when all corrections are carefully conducted the parameter η must be cautiously interpreted when obtained from a limited interval in 2θ. The utilization of the parameter η is thus only necessary in the presence of data below 15 Å where the effect of η is demonstrable.

The combined effects of applying background subtraction, optimizing the unit-cell parameters, bulk-solvent correction and the improved Lorentz correction are seen in Fig. 4 for all three samples. The agreement between the experimental and the corrected calculated patterns are generally very good. The profile R factors are in the range 20–25%. Even in the region where 2θ = 12–20° the agreement is good, meaning that for well diffracting crystals it is possible to detect the detailed diffraction pattern down to resolutions of 4 Å on in-house equipment. However, there still remain minor disagreements between the calculated and experimental patterns, which have not been accounted for. The following factors may influence the experimental data:

(i) Preferred crystal orientation, which causes a nonrandom distribution of reflecting planes.

(ii) Insufficient or excess background correction. The overlap problem may result in a background subtraction that also includes some of the signal.

(iii) Despite solvent and hydrogen corrections the structural model may still be deficient.

For the powder patterns the difference in accuracy between lysozyme and insulin is not very large even though the calculated pattern for lysozyme was based on a structure determination from a single crystal from the exact same batch from which the experimental powder pattern was obtained. Fitting our lysozyme powder data with a calculated pattern using the coordinates from PDB entry 2hub (S. Lagziel Simis, Y. Wine, N. Cohen-Hadar, A. Freeman & F. Frolow) gives a similar result (R_p = 27.5%). Matching both $\mathbf{F}_{obs}$ and $\mathbf{F}_{calc}$ with the experimental powder data gives overall similar results. However, generally the low-angle reflections measured in single-crystal experiments suffer from a large uncertainty reflecting the difficulties in integration and data treatment in this region. Furthermore, the low-angle region is often incomplete. Therefore we recommend using $\mathbf{F}_{calc}$ instead of $\mathbf{F}_{obs}$ for powder comparison. This means that PDB coordinate files can be used directly from the Protein Data Bank to calculate realistic powder patterns.

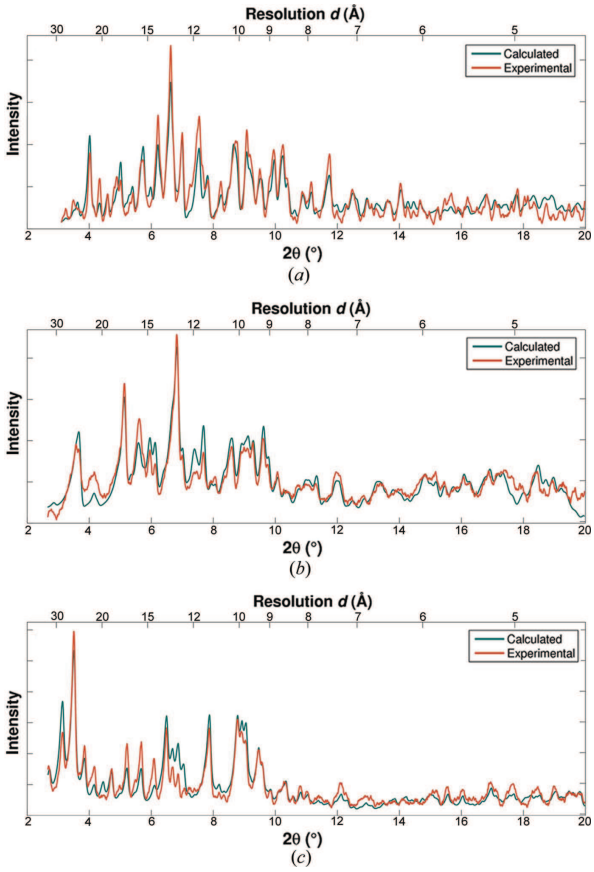


Figure 4
Calculated and experimental powder patterns for (*a*) lysozyme, (*b*) trigonal insulin and (*c*) cubic insulin. The calculated patterns (blue) are corrected for bulk-solvent and geometrical effects using the revised Lorentz factor [equation (3)].

5. Conclusion

Using a standard in-house X-ray powder diffractometer, it has been possible to collect powder patterns up to 4 Å resolution from well diffracting powder samples of lysozyme and insulin. Calculated powder patterns were generated from PDB coordinate data or data presented in this work. By performing corrections for background, unit-cell parameters, disordered bulk-solvent and geometrical factors the agreement between the experimental and the calculated patterns is generally very good, with profile *R* factors in the range 20–25%. The key factor to bring the calculated patterns into agreement with the observed patterns is the bulk-solvent correction. The improved Lorentz factor gives a minor, but significant, improvement to the fit, in particular at the lowest angles. An attempt to include H-atom positions gave no improvement and was abandoned. Moreover, the corrections are all straightforward using easily available standard programs and PDB coordinates.

The main drawback of the method is the large amount of sample needed, about 5 mg presently. It was found that in the vertical mount used the sample tends to settle in the bottom part of the sample compartment, so by an improved sample holder design we aim to reduce the sample amount by an order of magnitude and at the same time to improve the signal-to-noise ratio. Another issue is the data collection time. For this study data were collected for 16–60 h to obtain convincing statistics to 20° in 2θ. However, for polymorph identification 1–2 h is sufficient even in the present setup. The data collection time and the angular resolution of the peaks may furthermore be improved simply by the use of chromium $K\alpha$ radiation instead of Cu $K\alpha$.

APPENDIX *A*
Derivation of a revised Lorentz correction factor

Assuming that the reciprocal lattice points have a certain spatial extension parallel to the scattering vector, **s**, the sphere of reflections will have a thickness 2η. The parameter η thus reflects the distribution of the length of **s**. The fraction f of a reciprocal lattice point P touching the Ewald sphere is given as the area emphasized in grey in Fig. 5, divided by the volume V_{shell}, *i.e.*

$$f = \frac{A_{\text{spherical segment}}}{V_{\text{shell}}} = \frac{A_{\text{large spherical cap}} - A_{\text{small spherical cap}}}{V_{\text{large sphere}} - V_{\text{small sphere}}}, \quad (5)$$

with

$$A_{\text{large spherical cap}} - A_{\text{small spherical cap}} = 2\pi s(2\eta) \quad (6)$$

and

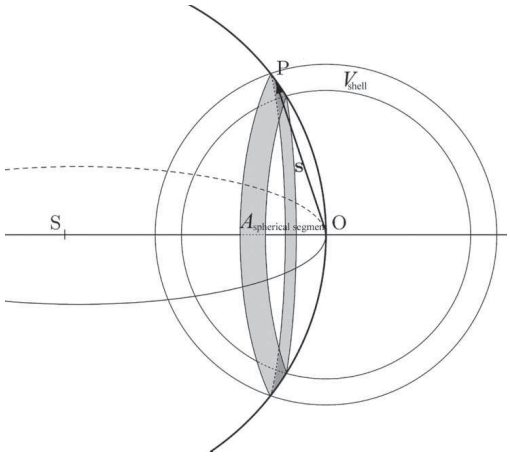


Figure 5
Three-dimensional model of the intersection of the Ewald sphere and the shell of reflections. The intersection is the area of a spherical segment emphasized in grey.

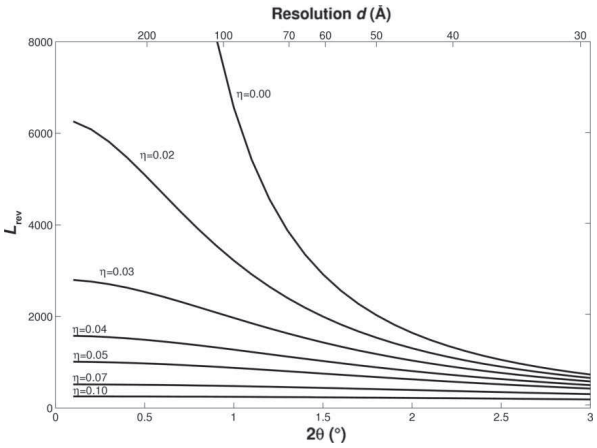


Figure 6
The entire Lorentz factor given in equation (9) as a function of the scattering angle 2θ (using Cu $K\alpha_1$) and resolution $d = \lambda/(2\sin\theta)$, for different sizes of the uncertainty parameter η at small angles.

$$V_{\text{large sphere}} - V_{\text{small sphere}} = (4/3)\pi[(s+\eta)^3 - (s-\eta)^3]. \quad (7)$$

Insertion of these geometrical sizes expressed in terms of s and using the fact that $s = 2\sin\theta/\lambda$, f becomes

$$f = \frac{3}{2}\frac{s}{3s^2+\eta^2} = \frac{\lambda}{4}\frac{\sin\theta}{\sin^2\theta + \lambda^2\eta^2/12}. \quad (8)$$

The revised Lorentz factor L_{rev} then becomes

$$L_{\text{rev}} = \frac{1}{\sin(2\theta)}\frac{\sin\theta}{\sin^2\theta + \lambda^2\eta^2/12}. \quad (9)$$

In contrast to the traditional Lorentz correction the divergence towards infinity for the geometrical term f is prevented by this model. Instead f converges towards zero at low angles as a result of the sine term in the numerator in equation (9). At higher angles f does not deviate much from the simple factor $1/\sin\theta$; this means that the revised expression would not affect normal XRPD patterns, which typically start at $2\theta = 3^\circ$.

Plotting L_{rev} for different values of η it is seen that there is a large effect for small angles compared to the conventional correction ($\eta = 0$ Å^{-1}; Fig. 6). Combining the geometrical term f with the original Lorentz term it is seen that the entire Lorentz factor has a finite value for $2\theta = 0^\circ$.

We are grateful for the beamtime provided at MAX-lab (Lund, Sweden) and for the Danish Natural Science Research Council contribution to DANSYNC. We also acknowledge the support by the European Community – Research Infrastructure Action under the FP6 programme 'Structuring the European Research Area'.

References

Adams, P. D., Grosse-Kunstleve, R. W., Hung, L.-W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W., Read, R. J., Sacchettini, J. C., Sauter, N. K. & Terwilliger, T. C. (2002). *Acta Cryst.* D**58**, 1948–1954.
Babinet, J. (1837). *C. R. Acad. Sci.* **4**, 638–648.
Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F. Jr, Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T. & Tasumi, M. (1977). *J. Mol. Biol.* **112**, 535–542.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Fokine, A. & Urzhumtsev, A. (2002). *Acta Cryst.* D**58**, 1387–1392.
Jiang, J.-S. & Brünger, A. T. (1994). *J. Mol. Biol.* **243**, 100–115.
Kabsch, W. (1993). *J. Appl. Cryst.* **26**, 795–800.
Kleywegt, G. J., Harris, M. R., Zou, J., Taylor, T. C., Wählby, A. & Jones, T. A. (2004). *Acta Cryst.* D**60**, 2240–2249.
Kurinov, I. V. & Harrison, R. W. (1995). *Acta Cryst.* D**51**, 98–109.
Margiolaki, I. & Wright, J. P. (2008). *Acta Cryst.* A**64**, 169–180.
Moews, P. C. & Kretsinger, R. H. (1975). *J. Mol. Biol.* **91**, 201–225.
Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.
Norrman, M., Ståhl, K., Schluckebier, G. & Al-Karadaghi, S. (2006). *J. Appl. Cryst.* **39**, 391–400.
Phillips, S. E. V. (1980). *J. Mol. Biol.* **142**, 531–554.
Sas, W. H. & de Wolff, P. M. (1966). *Acta Cryst.* **21**, 826–827.
Smith, G. D., Duax, W. L., Dodson, E. J., Dodson, G. G., de Graaf, R. A. G. & Reynolds, C. D. (1982). *Acta Cryst.* B**38**, 3028–3032.
Snell, E. H., Bellamy, H. D. & Borgstahl, G. E. O. (2003). *Methods Enzymol.* **368**, 268–288.
Ståhl, K. (2000). *J. Appl. Cryst.* **33**, 394–396.
Von Dreele, R. B. (1999). *J. Appl. Cryst.* **32**, 1084–1089.
Von Dreele, R. B. (2001). *Acta Cryst.* D**57**, 1836–1842.
Von Dreele, R. B. (2005). *Acta Cryst.* D**61**, 22–32.
Von Dreele, R. B., Stephens, P. W., Smith, G. D. & Blessing, R. H. (2000). *Acta Cryst.* D**56**, 1549–1553.
Wright, J. P., Besnard, C., Margiolaki, I., Basso, S., Camus, F., Fitch, A. N., Fox, G. C., Pattison, P. & Schiltz, M. (2008). *J. Appl. Cryst.* **41**, 329–339.

A.2 Paper II: In-house characterization of protein powder

Christian Grundahl Hartmann, Pernille Harris & Kenny Ståhl
Z. Kristallogr. Proc. (2011). **1,** 163−168

Z. Kristallogr. Proc. **1** (2011) 163-168 / **DOI** 10.1524/zkpr.2011.0024

In-house characterization of protein powder

C. G. Hartmann*, P. Harris, K. Ståhl

Department of Chemistry, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark
* Contact author; e-mail: cghar@kemi.dtu.dk

Keywords: in-house powder diffraction, protein, bulk-solvent correction

Abstract. Collecting protein powder diffraction data on standard in-house powder diffractometers requires careful handling of the samples. Specially designed sample holders combined with optimized collimation were found to be the key factors in improving the data quality and reducing the data collection time. For safe identification of the crystal form the experimental patterns have to be compared with patterns calculated from known crystal structures. Very good agreement with Protein Data Bank data was obtained after including corrections for background, unit cell parameters, disordered bulk-solvent, and geometric factors. The data collection and correction procedures were demonstrated by the identification of three different crystal forms of insulin.

1. Introduction

Since the first protein structure was refined from X-ray powder diffraction (XRPD) data ten years ago [1], high-resolution protein powder diffraction at synchrotron sources has been developed and used for refining and even solving small protein structures [2]. Furthermore, it has recently been shown that in-house XRPD can be successfully used for the identification of different crystal forms of proteins [3,4]. It thereby offers a convenient and quick way to characterize crystalline protein precipitate in the laboratory.

Growing suitable protein crystals of sufficient size and quality is one of the major bottlenecks in solving protein structures from single-crystal X-ray diffraction experiments. Screening for the optimal crystallization conditions is trial-and-error based and often gives many hits with precipitate. Early powder diffraction characterization of those precipitates would be valuable when optimizing the crystallization process.

Protein crystals contain 30–70 % of disordered solvent and they must be prevented from drying out and are thus kept in the mother liquor. Due to the large molecules and the high solvent content, protein crystals are soft and in general poor scatterers. Combined with the non-Bragg scattering from the surrounding mother liquor and the sample holder, it results in poor peak-to-noise ratios (P/N). Minimizing the sample amount and improving the P/N requires careful design of the sample holders and collimators.

Furthermore, the large unit cells result in Bragg peaks occurring at very low angles and at higher angles the reflections are suffering from a severe peak overlap limiting the data resolution to about 4 Å (2θ = 2–20° with Cu$K\alpha_1$-radiation in our tests).
Reliable identification of crystal forms requires calculation of realistic protein powder patterns based on for example Protein Data Bank (PDB) coordinate files. Including straightforward corrections for background, unit cell parameters, disordered bulk-solvent and geometric (Lorentz) factors we have obtained very good agreements between measured and calculated patterns [3]. The techniques will be demonstrated with data from three crystal forms of insulin.

2. Collection of protein powder data using in-house equipment

The diffraction experiments were carried out on a standard in-house Huber G670 diffractometer with a Guinier (transmission) geometry setup using Cu$K\alpha_1$-radiation (λ = 1.54059 Å). The data were recorded on an imaging strip in the range 2–100° in 2θ. The diffractometer was equipped with an Oxford Cryo-cooler. Crystals were grown in batches containing 2–6 mg of protein and suitable powder samples were obtained by gently crushing the crystals in the mother liquor resulting in a powder slurry. As cryo-protectant glycerol was added for the sample used in sample holder model 3 (loop). The samples were mounted in specially designed sample holders. Two sample holders for room temperature measurements and a loop for cryo-cooling temperatures (100–120 K) were tested. These are shown in figure 1. The sample holder design must fulfil several requirements. In general the major challenge is to find a way to support a thin film of wet powder. Firstly, the samples must be kept in sealed sample chambers or alternatively kept cryo-cooled in order to prevent the samples from drying out. Secondly, data collections are run in transmission mode meaning that the thickness of the samples must be as low as possible and the sample chamber windows must be made of a material with low absorption and low contribution to the background. Mylar was shown to be an ideal choice. Thirdly, the crystallite packing efficiency must be high and homogeneous and collimation of the X-ray beam must ensure that it hits the homogeneous part of the sample in order to maximize the *P*/*N*.

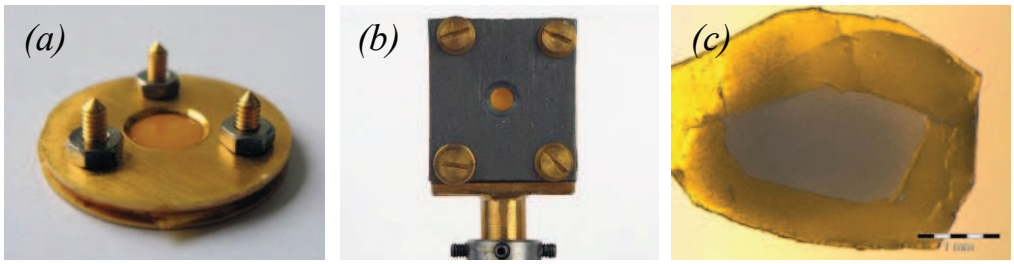


Figure 1. Different sample holders for in-house powder diffraction on proteins. The sample holders support slurries of protein crystallites in mother liquor. (a) Sample holder model 1 contains 120–150 µL of sample. (b) Sample holder model 2 contains 10 µL of sample. (c) Model 3: The loop supports 1–2 µL of cryo-cooled sample.

Sample holder model 1 requires approx. 150 µL of sample in an O-ring between two pieces of mylar foil. The O-ring and the foil are held in place by two metal discs providing a tight fit. Rotation of the sample is possible. Model 2 consists of a polycarbonate plate with a hole with mylar windows which are glued to the plate after the sample has been loaded. The sample size is approximately 10 µL. The polycarbonate holder is mounted on a goniometer head with a lead pinhole placed in front to reduce the background level. It is not possible to apply rotation of the sample. The holder was easily adaptable to synchrotron beamlines for X-ray absorption spectroscopy (XAS) without the lead pinhole. Model 3 is a loop (diameter approx. 2 mm) where small sample amounts are required. A thin film of wetted powder is easily supported on the loop cut from mylar and prevented from drying out by cryo-cooling. The loop was used in combination with the lead pinhole from model 2.
All the sample holders were tested with insulin to compare their performances. Data were collected for 4 h for sample holder model 1 and 2 and for 12 h for model 3 (loop). The raw powder diffraction patterns from trigonal insulin prior to background subtraction are directly compared on figure 2. The peak-to-noise ratios have been calculated for each powder pattern in figure 2 and are given as $P/N = (I_{Bragg}+I_{BG})/I_{BG}$, evaluated at the most intense peak at 2θ = 6.84º.

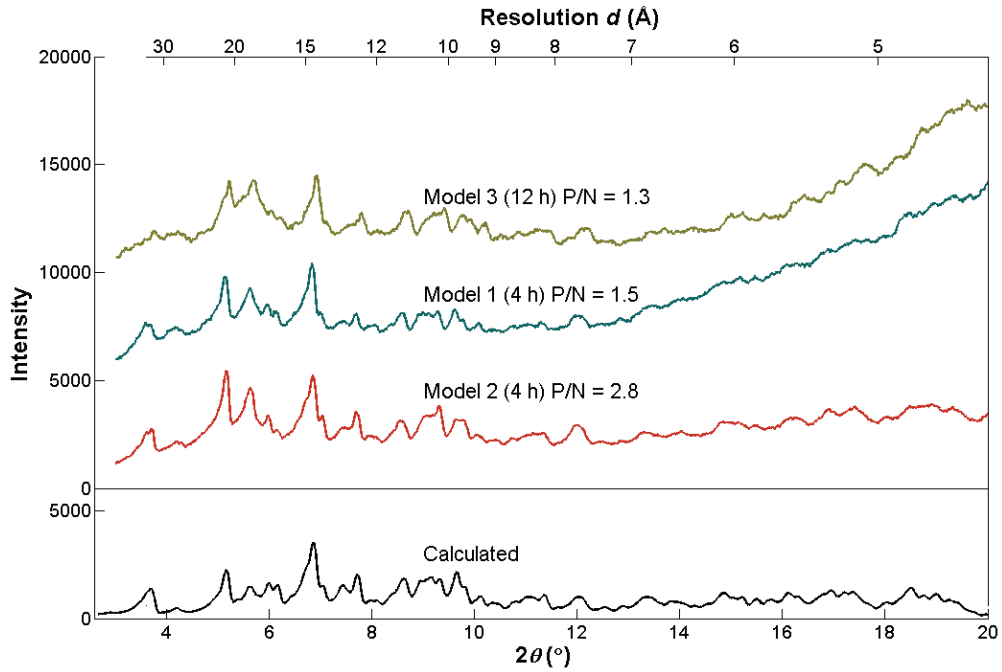


Figure 2. Raw data prior to background subtraction. The patterns have been collected on trigonal insulin for 4 h for sample holder model 1 and 2 and for 12 h for model 3. Comparison with calculated pattern (bottom), where the intensities have been scaled to the pattern from model 2.

The effect of reducing the background using the lead pinhole is clearly demonstrated by comparing model 1 with model 2. Presently 10 µL of sample is required to collect good powder diffraction patterns within hours, but in fact the crystal form can be safely identified

in 15 minutes. The sample volumes can be further reduced to 1–2 µL using the loop (model 3), but at the expense of longer data collection time and poorer *P*/*N*.

3. Data treatment

The background was subtracted in the program *PROTPOW* [3] (available from http://www.xray.kemi.dtu.dk/English/Computer%20Programs/PROTPOW.aspx) by manually picking between 10 and 20 background points, fitting to a spline function and subtracting. The manual procedure is necessary to distinguish background from overlapping signal. Optimization of the unit cell parameters was done by a full-pattern profile fit. Optimized parameters describing the peak width (FWHM) and peak shape of a pseudo-Voigt function are included in the pattern fit by *PROTPOW*.

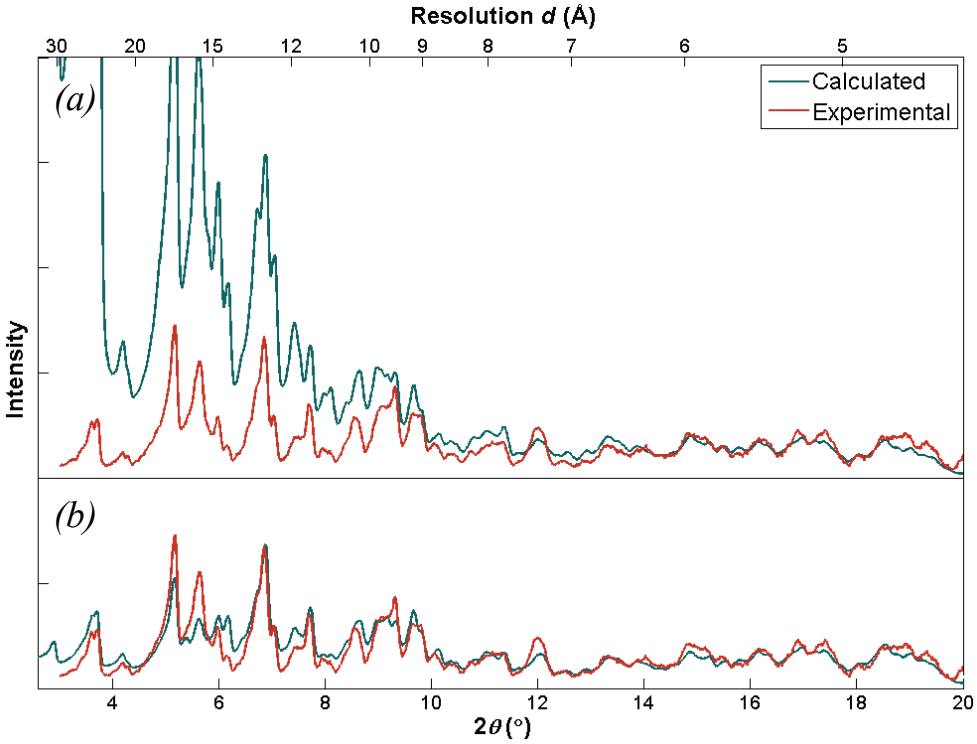


Figure 3. Comparison of calculated and experimental powder diffraction patterns for trigonal insulin. (a) Before correction for disordered bulk-solvent. (b) After correction for disordered bulk-solvent.

Powder patterns were calculated in *PROTPOW* from PDB coordinate files using the optimized unit cell and peak profile parameters. In general the peak positions fit well, but without taking into account the disordered bulk-solvent, the intensities fit poorly with the measured intensities, in particular at low 2θ, as seen in figure 3(*a*). The contribution from the solvent was calculated in *PHENIX* [5] using the flat bulk-solvent model [6] where a constant level of electron density is assumed in the voids between the macromolecule. The bulk-

solvent model takes two parameters k_{sol} defining the level of electron density and B_{sol} defining the steepness of the border between the solvent and the macromolecular regions. The parameters are refined in today's single-crystal X-ray structure refinement software, but in case no other information is available following average values can be used k_{sol} = 0.35 e Å^{-3} and B_{sol} = 46 Å^2 [7]. The contribution from the solvent is added to the calculated structure factors. The improved fit from the powder pattern based on calculated structure factors is seen on figure 3(*b*). Applying an improved geometrical (Lorentz) factor further improves the fit [3].

4. Results and discussion

The applicability of the methods has been demonstrated in a recent study where XRPD was used to verify three different conformations (T_6, $T_3R_3^f$ and R_6) of hexameric Zn-insulin before and after EXAFS experiments. The same samples in model 2 sample holders were used for both experiments. The differences between the crystal forms are obvious in figure 4.

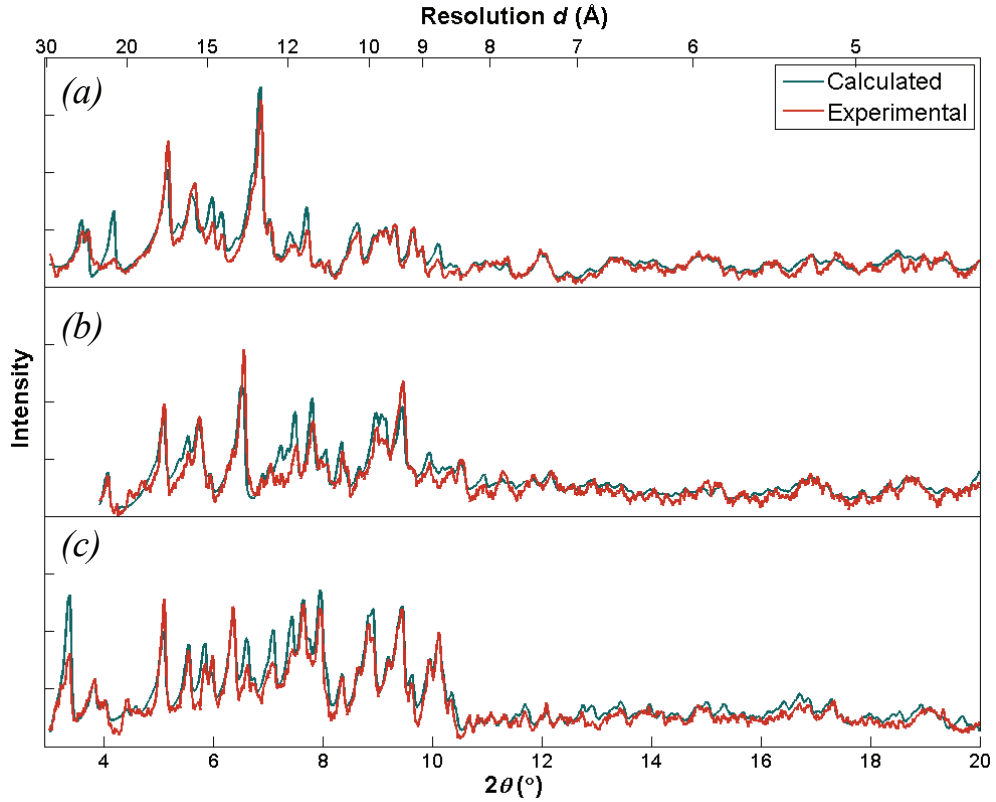


Figure 4. Verification of three different conformations of hexameric insulin before an X-ray absorption spectroscopy study. XRPD patterns for insulin powders (a) T_6 , (b) $T_3R_3^f$ and (c) R_6.

For well known protein structures the XRPD data can be used to identify and verify different protein crystal forms. A reliable identification requires that a comparable powder pattern can be calculated, which requires *a priori* knowledge about the space group and the cell parameters, e.g. from a preliminary indexation. An in-house XRPD pre-characterization could be applied to conclude that a material is crystalline and from the appearance of low angle reflection distinguish between protein and salt crystals from the precipitant – also for unknown structures.

5. Concluding remarks

Collecting XRPD data on well diffracting proteins for identification and verification of known crystal forms can be carried out on in-house equipment. Special sample holders are required in order to prevent the protein crystallites from drying out, and a careful sample preparation as well as an optimized collimation is crucial in order to obtain best possible *P*/*N*. Further, it is essential that unit cell parameters are optimized for the calculated patterns and that the patterns are corrected for solvent scattering. With the sample holder designs presented here 1–10 μL of well-diffracting protein powder are required for identification within 0.5–2 h. The technique may today be applied to screening of well-crystallized protein powders and thus be of interest also to the enzyme production industry where identification could be used as a tool for controlling the enzyme products during manufacturing as well as for quality control of the final product. By reaching even smaller sample sizes and testing "not-so-well" diffracting protein powders the screening of early precipitates for crystals by XRPD is now within reach.

References

1. Von Dreele, R.B., 1999, *J. Appl. Crystallogr.*, **32**, 1084.
2. Margiolaki, I. & Wright, J.P., 2008, *Acta Crystallogr. A*, **64**, 169.
3. Hartmann, C.G., Nielsen, O.F., Ståhl, K. & Harris, P., 2010, *J. Appl. Crystallogr.*, **43**, 876.
4. Norrman, M., Ståhl, K., Schluckebier, G. & Al-Karadaghi, S., 2006, *J. Appl. Crystallogr.*, **39**, 391.
5. Adams, P.D., Grosse-Kunstleve, R.W., Hung, L.-W., Ioerger, T.R., McCoy, A.J., Moriarty, N.W., Read, R.J., Sacchettini, J.C., Sautera, N.K. & Terwilliger, T.C., 2001, *Acta Crystallogr. D*, **58**, 1948.
6. Phillips, S.E.V., 1980, *J. Mol. Biol.*, **142**, 531.
7. Fokine, A. & Urzhumtsev, A., 2002, *Acta Crystallogr. D*, **58**, 1387.

Acknowledgements. We are grateful to the Carlsberg Foundation for funding the cryo-system used for some of these experiments. Special thanks to Ms. M. V. Knudsen for growing R_6 insulin crystals.

A.3 Paper III: A sample holder for in-house X-ray powder diffraction studies of protein powders

Christian Grundahl Frankær, Pernille Harris & Kenny Ståhl
J. Appl. Cryst. (2011). **44,** 1288−1290

Journal of
Applied Crystallography
ISSN 0021-8898

Received 5 July 2011
Accepted 22 September 2011

A sample holder for in-house X-ray powder diffraction studies of protein powders

Christian Grundahl Frankær, Pernille Harris and Kenny Ståhl*

Department of Chemistry, Technical University of Denmark, Kemitorvet 207, DK-2800 Kongens Lyngby, Denmark. Correspondence e-mail: kenny@kemi.dtu.dk

A sample holder for handling samples of protein for in-house X-ray powder diffraction (XRPD) analysis has been made and tested on lysozyme. The use of an integrated pinhole reduced the background, and good signal-to-noise ratios were obtained from only 7 µl of sample, corresponding to approximately 2–3 mg of dry protein. The sample holder is further adaptable to X-ray absorption spectroscopy (XAS) measurements. Both XRPD and XAS at the Zn *K*-edge were tested with hexameric Zn insulin.

1. Introduction

During the past decade, X-ray powder diffraction has proven itself able to solve and refine small protein structures from high-resolution powder data using synchrotron radiation (see *e.g.* Margiolaki & Wright, 2008, and references therein). Powder diffraction patterns with well resolved peaks can also be obtained from highly optimized laboratory equipment (Margiolaki *et al.*, 2007), but at the cost of longer data collection times. However, for a fast characterization and fingerprinting of protein crystal forms, X-ray powder diffraction (XRPD) can be carried out on widely used standard laboratory diffractometers (Hartmann *et al.*, 2010, 2011). This 'rough' technique may be well suited for detection of substrate binding and verification of different protein conformations of already known systems (Norrman *et al.*, 2006), and could thereby become a valuable complementary tool in the laboratory, as well as for product characterization in the pharmaceutical industry and enzyme production.

Metalloprotein structure determinations from single crystals are increasingly complemented by X-ray absorption spectroscopy (XAS). Quick in-house XRPD characterization of crystal forms before and after XAS measurements is indeed possible. In the present note the design of a sample holder for both in-house XRPD and XAS is described and tested on lysozyme and two conformations of hexameric Zn insulin, T_6 and R_6.

2. Sample holder design

The sample holder was developed for the Huber G670 diffractometer equipped with a flat powder specimen holder. The holder has been further improved from the holder used in the study by Hartmann *et al.* (2010) and is also adaptable for XAS experiments, as demonstrated by tests on hexameric Zn insulin. To collect as much signal as possible on the in-house X-ray sources, the general principle behind the design is to support a thin film of protein crystallites with a relatively large surface area. Rotation of the sample around an axis χ orthogonal to the detector is also important to avoid sedimentation on the windows and walls of the sample chamber and so to avoid multiple layers of powder. Schematic drawings and a photograph of the sample holder are shown in Fig. 1.

The sample holder consists of a 1.0 mm-thick polycarbonate plate with a hole (either 4.5 or 3.0 mm in diameter, creating sample chamber volumes of 16 and 7 µl, respectively). Further dimensions are specified in Figs. 1(*a*) and 1(*b*). Metal-free polycarbonate was chosen to avoid a fluorescence signal from the sample holder, which for XAS experiments would conflict with the fluorescence signal from metalloproteins. Small holes have been drilled so that the holder fits the cryostat (Oxford Instruments, ITC503) available at the 8-11 XAS beamline at the synchrotron at MAX-lab, Lund, Sweden. The sample holder takes a small cylindrical volume of protein powder wetted by its mother liquor, and with a sample thickness of 1.0 mm this holder still allows the diffractometer to work in transmission mode. The

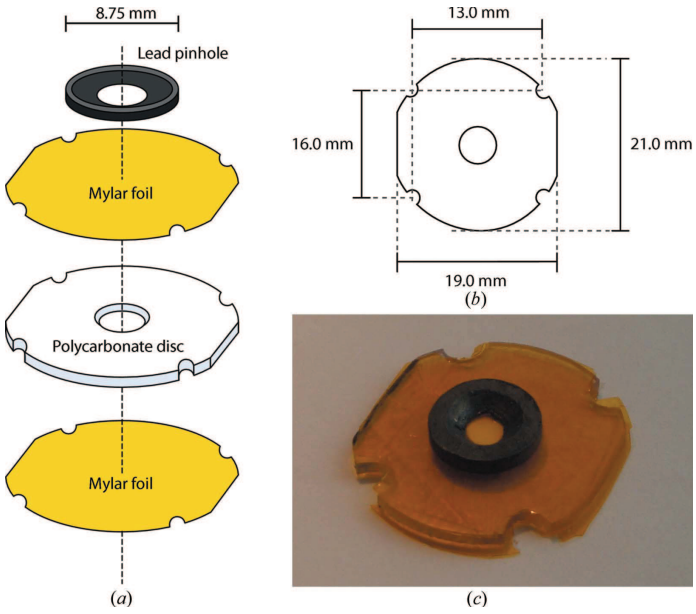


Figure 1
XRPD sample holder for the Huber G670 diffractometer. (*a*) Drawing of the individual parts of the unassembled sample holder. (*b*) Dimensions of the polycarbonate plate. (*c*) Photograph of the sample holder.

sample is sealed by two pieces of mylar foil, which are attached to the polycarbonate plate with either epoxy glue or double-sided adhesive tape. For powder diffraction, a lead plate with a pinhole matching the diameter of the sample chamber is glued on top of the assembled holder and serves for efficient shielding of the background scattering. The inner edge profile of the lead plate has been cut at a 45° angle to prevent shielding of the incoming X-ray beam. The pinhole plate fits exactly into the flat powder specimen holder at the Huber G670 diffractometer, allowing the sample to be centred in the beam. The relatively compact shape of the sample chamber eases mounting of the sample.

3. Experimental

The sample holder was tested on the Huber G670 diffractometer with protein powder of lysozyme from chicken egg white (*Gallus gallus*, 95% Sigma Aldrich L-6876) and Zn insulin from bovine pancreas (*Bos taurus*, Sigma Aldrich I-5500). Tetragonal lysozyme crystals and trigonal T_6 insulin crystals were grown according to the procedures described by Hartmann *et al.* (2010). R_6 insulin crystals were grown with the phenol derivative *m*-cresol according to the procedure of Smith *et al.* (2000). The T_6 and R_6 insulin crystals were filtered using an Ultra-free-MC centrifuge filter (Amicon/Millipore) with a pore size of 0.22 µm, and excess zinc was eliminated by washing the crystals three times with a solution having a composition similar to the mother liquors except for containing zinc. Excess mother liquor was removed and the wetted crystals were gently crushed to a concentrated homogeneous suspension of protein powder.

The sample holders were loaded with the suspensions taking care that no air bubbles were trapped in the sample chambers when gluing the top foil in place. Rotation around χ was applied to the samples and data were collected on two lysozyme samples (16 and 7 µl, corresponding to approximately 5–6 and 2–3 mg of protein) for 4 h at room temperature using Cu $K\alpha_1$ radiation (λ = 1.5406 Å). The raw lysozyme powder diffraction patterns are compared in Fig. 2. For Zn insulin, data were collected for 30 min on 7 µl samples. The raw powder patterns are shown in Fig. 3, from which the crystal forms can be distinguished. For these systems, no radiation damage was observed.

An X-ray absorption spectrum was collected at the Zn *K*-edge (9659 eV) on the mounted insulin sample at the XAS beamline 8-11 at the synchrotron at MAX-lab, using an Si(111) double-crystal monochromator (Carlson *et al.*, 2006). Fluorescence data were collected at 100 K with a passive implanted planar silicon detector as three scans of 60 min, which were subsequently averaged. To check for occurrence of radiation damage the spectra were compared prior to averaging. No radiation damage was observed. The absorption spectra at the Zn *K*-edge are shown in Fig. 4. The extended X-ray

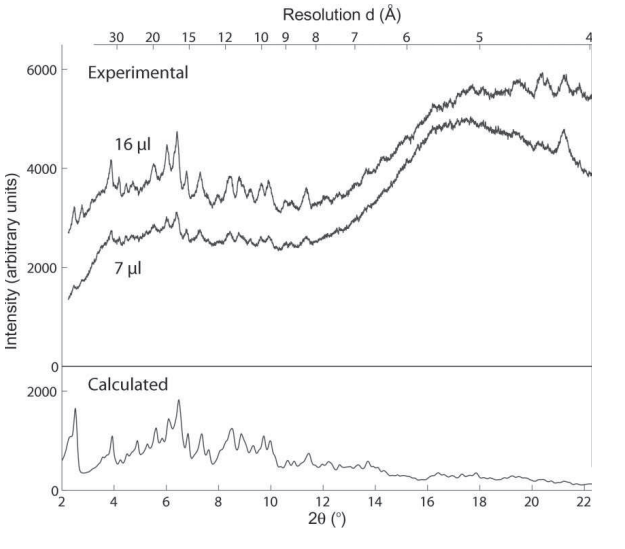


Figure 2
XRPD patterns of tetragonal crystals of hen egg white lysozyme prior to background subtraction. The data collection time was 4 h and Cu $K\alpha_1$ radiation (λ = 1.5406 Å) was used. A pattern calculated using optimized unit-cell and peak-shape parameters is shown for comparison (bottom).

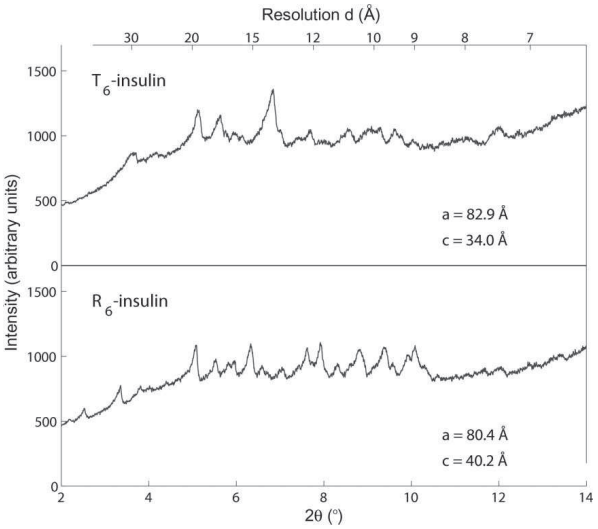


Figure 3
XRPD patterns of trigonal crystals of T_6 and R_6 insulin prior to background subtraction. Data were collected on 7 µl of sample for 30 min using Cu $K\alpha_1$ radiation (λ = 1.5406 Å). Unit-cell parameters were optimized from a full pattern profile fit.

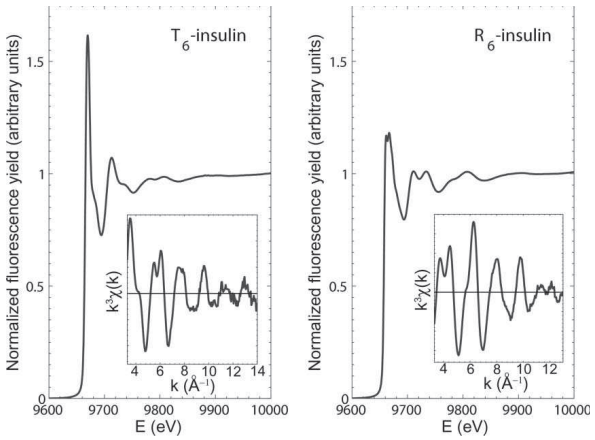


Figure 4
XAS spectra of hexameric Zn insulin collected at the Zn *K*-edge with the combined XRPD and XAS sample holder. Insets: Extracted and k^3-weighted EXAFS spectra. Data with good signal-to-noise ratios were obtained to 13–14 Å^{-1} in *k*-space.

absorption fine structure (EXAFS) spectra were extracted using *WinXAS* (Ressler, 1998).

4. Results

The background intensity levels and the peak-to-noise ratios, *P*/*N*, which are interrelated, were estimated by looking at the raw lysozyme powder patterns presented in Fig. 2. Separating background from diffraction is, however, not straightforward because of the heavy peak overlap. The estimation of the background intensity was therefore done by scaling a calculated pattern to the experimental. Prior to the calculation of a powder pattern, unit-cell and peak-shape parameters were optimized using the procedure described by Hartmann *et al.* (2010). The background level, BG, was evaluated as a sum of intensities I in the 2θ range 2–20° by the following equation:

$$\mathrm{BG} = \frac{\sum I_{\mathrm{Background}}}{\sum \left(I_{\mathrm{Bragg}} + I_{\mathrm{Background}}\right)} 100\% = \frac{\sum \left(I_{\mathrm{Total}} - I_{\mathrm{Bragg}}\right)}{\sum I_{\mathrm{Total}}} 100\%, \quad (1)$$

where I_{Bragg} was evaluated from a calculated pattern that was scaled to the experimental using *PROTPOW* (http://www.xray.kemi.dtu.dk/English/Computer Programs/PROTPOW.aspx). The background levels for lysozyme were 87.4 and 92.3% for the 16 and 7 µl holders, respectively.

The *P*/*N* ratios and the background levels depend on the collimation of the beam. Pinholes integrated into the design of the flat specimen sample holder ensure that the beam irradiates only the homogeneous part of the sample and thereby reduce the background, in particular at low 2θ angles. After an optimization of the beam stop position, it is possible to measure reflections down to 2.0°.

The adaptability of the sample holder was illustrated with insulin, where the same sample holder was used for XRPD and XAS. This ensures that both analyses are carried out on exactly the same sample, which reduces the amount of sample used and the preparation time. Both T_6 and R_6 insulin form trigonal crystals belonging to space group *R*3, but with different Zn coordination and unit-cell parameters. As seen from the powder patterns in Fig. 3, the two crystal forms are clearly distinguishable, and the difference in Zn coordination is clearly evident from the absorption spectra shown in Fig. 4.

5. Conclusions

A sample holder for identification and fast characterization of protein crystal forms by XRPD using a standard laboratory source has been developed. Optimizing the collimation by pinholes, the design was improved to operate with sample volumes down to 7 µl, which is comparable to the performance of capillaries but the samples are easier to prepare. For metal-containing proteins the improved sample holder design was successfully applied to samples used for both powder diffraction and XAS experiments.

Special thanks to Katarina Norén for assistance at the XAS experiments. We acknowledge the workshop at DTU Chemistry for fabrication of the sample holders and Susanne Helmark for illustrations and the photograph.

References

Carlson, S., Clausén, M., Gridneva, L., Sommarin, B. & Svensson, C. (2006). *J. Synchrotron Rad.* **13**, 359–364.
Hartmann, C. G., Harris, P. & Ståhl, K. (2011). *Z. Kristallogr. Suppl.* In the press.
Hartmann, C. G., Nielsen, O. F., Ståhl, K. & Harris, P. (2010). *J. Appl. Cryst.* **43**, 876–882.
Margiolaki, I. & Wright, J. P. (2008). *Acta Cryst.* A**64**, 169–180.
Margiolaki, I., Wright, J. P., Fitch, A. N., Fox, G. C., Labrador, A., Von Dreele, R. B., Miura, K., Gozzo, F., Schiltz, M., Besnard, C., Camus, F., Pattison, P., Beckers, D. & Degen, T. (2007). *Z. Kristallogr. Suppl.* **26**, 1–13.
Norrman, M., Ståhl, K., Schluckebier, G. & Al-Karadaghi, S. (2006). *J. Appl. Cryst.* **39**, 391–400.
Ressler, T. (1998). *J. Synchrotron Rad.* **5**, 118–122.
Smith, G. D., Ciszak, E., Magrum, L. A., Pangborn, W. A. & Blessing, R. H. (2000). *Acta Cryst.* D**56**, 1541–1548.

A.4 Paper IV: Polymeric strontium ranelate nonahydrate

Kenny Ståhl, Christian Grundahl Frankær, Anders Christer Raffalt, Søren Roed Sørensen & Jens Enevold Thaulov Andersen
Acta Cryst. (2011). E**67,** m471−m472

Acta Crystallographica Section E
Structure Reports Online
ISSN 1600-5368

Polymeric strontium ranelate nonahydrate

Kenny Stahl,* Christian G. Frankaer, Anders C. Raffalt, Søren R. Sørensen and Jens E. T. Andersen

Department of Chemistry, Technical University of Denmark, DK-2800 Lyngby, Denmark
Correspondence e-mail: kenny@kemi.dtu.dk

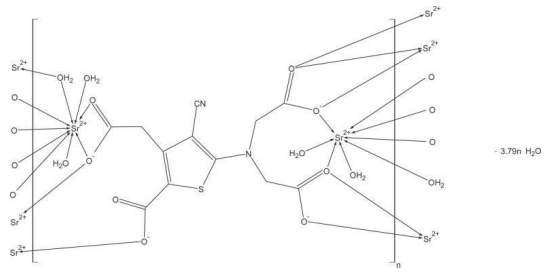
Received 9 March 2011; accepted 17 March 2011

Key indicators: single-crystal X-ray study; T = 120 K; mean σ(C–C) = 0.003 Å; H-atom completeness 99%; disorder in main residue; R factor = 0.033; wR factor = 0.075; data-to-parameter ratio = 17.6.

The title compound, poly[[μ-aqua-tetraaqua{μ-5-[bis-(carboxylatomethyl)amino]-3-carboxylatomethyl-4-cyanothiophene-2-carboxylato}distrontium(II)] tetrahydrate], $[Sr_2(C_{12}H_6N_2O_8S)(H_2O)_5]\cdot 3.79H_2O$, crystallizes with nine- and eight-coordinated Sr^{2+} cations. They are bound to seven of the eight ranelate O atoms and five of the water molecules. The SrO_8 and SrO_9 polyhedra are interconnected by edge-sharing, forming hollow layers parallel to (011). The layers are, in turn, interconnected by ranelate anions, forming a metal–organic framework (MOF) structure with channels along the a axis. The four water molecules not coordinated to strontium are located in these channels and hydrogen bonded to each other and to the ranelates. Part of the water H atoms are disordered. The compound dehydrates very easily and 0.210 (4) water molecules out of nine were lost during crystal mounting causing additional disorder in the water structure.

Related literature

For the effect of strontium on osteroporosis, see Schrooten *et al.* (2003). For a patent describing the synthesis and powder diffraction pattern of the title compound, see Horvath *et al.* (2008). For related strontium carboxylate structures, see, for example: Stahl *et al.* (2006).



Experimental

Crystal data

$[Sr_2(C_{12}H_6N_2O_8S)(H_2O)_5]\cdot 3.79H_2O$
M_r = 671.84
Triclinic, $P\overline{1}$
a = 8.3585 (3) Å
b = 12.3865 (5) Å
c = 12.6474 (5) Å
α = 109.880 (1)°
β = 97.148 (1)°
γ = 105.321 (1)°
V = 1154.00 (8) Å^3
Z = 2
Mo $K\alpha$ radiation
μ = 4.80 mm^{-1}
T = 120 K
0.15 × 0.10 × 0.07 mm

Data collection

Bruker SMART APEX diffractometer
Absorption correction: multi-scan (*SADABS*; Sheldrick, 2002)
T_{min} = 0.574, T_{max} = 0.710
17404 measured reflections
6617 independent reflections
5375 reflections with $I > 2\sigma(I)$
R_{int} = 0.032

Refinement

$R[F^2 > 2\sigma(F^2)]$ = 0.033
$wR(F^2)$ = 0.075
S = 1.02
6617 reflections
375 parameters
21 restraints
H atoms treated by a mixture of independent and constrained refinement
$\Delta\rho_{max}$ = 1.48 e Å^{-3}
$\Delta\rho_{min}$ = −1.23 e Å^{-3}

Table 1
Selected bond lengths (Å).

Sr1—O8	2.4557 (18)	Sr2—O23	2.5921 (18)
Sr1—O3^{i}	2.4782 (19)	Sr2—O2^{i}	2.6222 (17)
Sr1—O5	2.5234 (16)	Sr2—O21	2.6445 (17)
Sr1—O7^{ii}	2.6149 (19)	Sr2—O6	2.6628 (16)
Sr1—O25	2.652 (2)	Sr2—O2^{iv}	2.6848 (16)
Sr1—O22	2.6560 (19)	Sr2—O1^{iv}	2.6944 (17)
Sr1—O27	2.657 (2)	Sr2—O22	2.7108 (18)
Sr1—O8^{ii}	2.7834 (17)	Sr2—O5	2.7228 (16)
Sr2—O6^{iii}	2.5452 (16)		

Symmetry codes: (i) $-x+1, -y+1, -z+1$; (ii) $-x+1, -y+1, -z+2$; (iii) $-x, -y, -z+1$; (iv) $x, y-1, z$.

Table 2
Hydrogen-bond geometry (Å, °).

D—H···A	D—H	H···A	D···A	D—H···A
O21—H21A···O25	0.81 (2)	1.98 (2)	2.781 (3)	169 (3)
O21—H21B···O24^{iii}	0.85 (2)	1.93 (2)	2.765 (3)	169 (3)
O22—H22A···O2^{i}	0.81 (2)	2.16 (3)	2.761 (2)	131 (3)
O22—H22B···O26^{iv}	0.82 (2)	1.94 (2)	2.755 (3)	173 (3)
O23—H23A···O21^{iii}	0.81 (2)	1.96 (2)	2.766 (3)	174 (4)
O23—H23B···O26^{i}	0.80 (2)	2.11 (2)	2.867 (3)	159 (3)
O24—H24A···O1^{v}	0.82 (2)	2.03 (2)	2.760 (3)	148 (3)
O24—H24B···O4	0.84 (2)	1.93 (2)	2.756 (3)	172 (3)
O25—H25A···N1^{ii}	0.82 (2)	2.15 (2)	2.898 (3)	152 (3)
O25—H25B···O27^{ii}	0.80 (2)	1.93 (2)	2.648 (3)	150 (4)

D—H···A	D—H	H···A	D···A	D—H···A
O26—H26A···O28^{vi}	0.85 (2)	1.90 (2)	2.731 (3)	164 (5)
O26—H26C···N1	0.85 (2)	2.37 (4)	3.108 (3)	146 (5)
O27—H27A···O4^{i}	0.83 (2)	1.79 (2)	2.615 (3)	172 (4)
O27—H27B···O29^{vii}	0.82 (2)	1.94 (2)	2.727 (4)	160 (4)
O28—H28A···O24^{v}	0.80 (2)	1.97 (2)	2.756 (3)	167 (4)
O28—H28B···O28^{viii}	0.82 (2)	2.02 (2)	2.835 (4)	174 (8)
O28—H28C···O29	0.82 (2)	2.04 (3)	2.836 (4)	164 (7)
O29—H29A···O7	0.83 (2)	1.78 (2)	2.595 (3)	168 (7)
O29—H29B···O28	0.83 (2)	2.03 (3)	2.836 (4)	164 (7)

Symmetry codes: (i) $-x+1, -y+1, -z+1$; (ii) $-x+1, -y+1, -z+2$; (iii) $-x, -y, -z+1$; (iv) $x, y-1, z$; (v) $-x, -y+1, -z+1$; (vi) $-x+1, -y+2, -z+2$; (vii) $x+1, y, z$; (viii) $-x, -y+2, -z+2$.

Data collection: *SMART* (Bruker, 1999); cell refinement: *SAINT-Plus* (Bruker, 1999); data reduction: *SAINT-Plus*; program(s) used to solve structure: *SHELXS97* (Sheldrick, 2008); program(s) used to refine structure: *SHELXL97* (Sheldrick, 2008); molecular graphics: *ORTEP-3* (Farrugia, 1997) and *ATOMS* (Dowty, 2000); software used to prepare material for publication: *SHELXL97*.

Ms L. Berring and Ms A. Schøneberg are gratefully acknowledged for the data collection and Dr Stephan Christgau for supplying the strontium ranelate.

Supplementary data and figures for this paper are available from the IUCr electronic archives (Reference: SI2343).

References

Bruker (1999). *SMART* and *SAINT-Plus*. Bruker AXS Inc., Madison, Wisconsin, USA.
Dowty, E. (2000). *ATOMS*. Shape Software, Kingsport, Tennessee, USA.
Farrugia, L. J. (1997). *J. Appl. Cryst.* **30**, 565.
Horvath, S., Demuynck, I. & Damien, G. (2008). US Patent No. 7459568.
Schrooten, I., Behets, G. J. S., Cabrera, W. E., Vercauteren, S. R., Lamberts, L. W., Verberckmoes, S. C., Bervoets, A. J., Dams, G., Goodman, W. G., De Broe, M. E. & D Haese, P. C. (2003). *Kidney Intl*, **63**, 927–935.
Sheldrick, G. M. (2002). *SADABS*. University of Göttingen, Germany
Sheldrick, G. M. (2008). *Acta Cryst.* A**64**, 112–122.
Stahl, K., Andersen, J. E. T. & Christgau, S. (2006). *Acta Cryst.* C**62**, m144–m149.

supplementary materials

Acta Cryst. (2011). E**67**, m471-m472 [**doi:10.1107/S1600536811010099**]

Polymeric strontium ranelate nonahydrate

K. Stahl, C. G. Frankaer, A. C. Raffalt, S. R. Sørensen and J. E. T. Andersen

Comment

In recent years it has been found that Sr has a significant influence on the development and growth of bone and the effect of dose on bone structure has been investigated in great detail (Schrooten *et al.* 2003). Strontium ranelate (5-[bis(carboxymethyl) amino]-3-carboxymetyl- 4-cyano-2-thiophenecarboxylate) is one promising pharmaceutical compound for treating osteoporosis marketed as Protelos^{R} by Servier (Horvath *et al.*, 2008). Strontium ranelate is known to form several hydrates with totally nine, eight, seven or four waters (Horvath *et al.*, 2008). The initial dehydration observed here results from an expulsion of O27 or O29 and migration of the remaining water to site O30. As a consequence Sr1 is partially seven-coordinated (c.f. Table 1). The water hydrogen sites connected to O26, O28 and O29 are disordered. In essence, the alternating hydrogen bonding scheme between O28 and O29 is transmitted to a partial O26 - O28 hydrogen bond, and leaves H26B and H29C without hydrogen bond acceptors (c.f. Table 2).

Experimental

Strontium ranelate nona hydrate of 97% purity (Clauson-Kaas A/S) was recrystallized at different temperatures. Recrystallization at temperatures above 353 K appeared to produce the crystals of better quality. Upon cooling to room temperature large crystals of millimeter dimensions were obtained in the saturated solution. However, when the crystals were removed from the solution they rapidly degraded into smaller units of micron dimension. The smaller crystals showed out to contain less crystal water, as compared to the large crystals, presumably seven or five water molecules per formula unit. Thus, wet crystals were quickly transferred to the goniometer for X-ray data collection at 120 K. Several crystal were tried before an acceptable structure refinement was achieved. For all cases of lower quality data the SOF of O30 was about 0.3, confirming its role in the initial dehydration of strontium ranelate and the deterioration of the crystals.

Refinement

The H atoms of the CH_2 groups were placed in calculated positions with C—H = 0.99, and refined as riding atoms. The H atoms of the water molecules were located in difference Fourier maps and refined with restrained O—H distances of 0.82 (2) Å. The H atoms of the partially occupied O30 (SOF=0.210 (4)) could not be located. The H displacement parameters were set to 1.2 (CH_2) or 1.5 (H_2O) times U_{eq} of the corresponding C or O atoms.

Figures

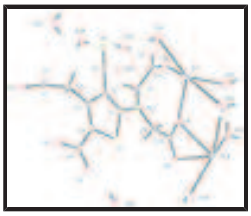


Fig. 1. The asymmetric unit of (I) showing 50% probability displacement ellipsoids and the atomic numbering. Hydrogen atoms are represented by circles of arbitrary size and shows one consistent set of water H atoms.

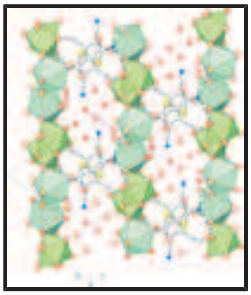


Fig. 2. The crystal packing of (I) viewed down the a-axis. Hydrogen atoms are omitted for clarity.

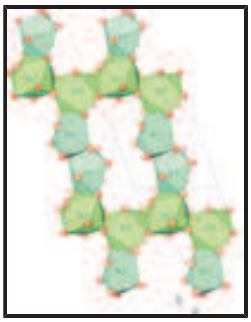


Fig. 3. The polyhedral layer of (I) viewed down the (01$\bar{1}$) direction.

Poly[[μ-aqua-tetraaqua{μ-5-[bis(carboxylatomethyl)amino]-3- carboxylatomethyl-4-cyanothiophene-2-carboxylato}distrontium(II)] tetrahydrate]

Crystal data

$[Sr_2(C_{12}H_6N_2O_8S)(H_2O)_5]\cdot 3.79H_2O$	$Z = 2$
$M_r = 671.84$	$F(000) = 671.8$
Triclinic, $P\bar{1}$	$D_x = 1.933$ Mg m^{-3}
Hall symbol: -P 1	Mo $K\alpha$ radiation, $\lambda = 0.71073$ Å
$a = 8.3585$ (3) Å	Cell parameters from 6752 reflections
$b = 12.3865$ (5) Å	$\theta = 2.6$–$30.6°$
$c = 12.6474$ (5) Å	$\mu = 4.80$ mm^{-1}
$\alpha = 109.880$ (1)°	$T = 120$ K
$\beta = 97.148$ (1)°	Tabular, colorless
$\gamma = 105.321$ (1)°	0.15 × 0.10 × 0.07 mm
$V = 1154.00$ (8) Å^3	

Data collection

Bruker SMART APEX diffractometer	6617 independent reflections
Radiation source: fine-focus sealed tube	5375 reflections with $I > 2\sigma(I)$
graphite	$R_{int} = 0.032$
ω scan, frame data integration	$\theta_{max} = 30.9°$, $\theta_{min} = 2.6°$
Absorption correction: multi-scan (*SADABS*; Sheldrick, 2002)	$h = -12 \rightarrow 12$
$T_{min} = 0.574$, $T_{max} = 0.710$	$k = -17 \rightarrow 17$
17404 measured reflections	$l = -18 \rightarrow 18$

Refinement

Refinement on F^2	Primary atom site location: structure-invariant direct methods
Least-squares matrix: full	Secondary atom site location: difference Fourier map
$R[F^2 > 2\sigma(F^2)] = 0.033$	Hydrogen site location: inferred from neighbouring sites
$wR(F^2) = 0.075$	H atoms treated by a mixture of independent and constrained refinement
$S = 1.02$	$w = 1/[\sigma^2(F_o^2) + (0.0383P)^2]$ where $P = (F_o^2 + 2F_c^2)/3$
6617 reflections	$(\Delta/\sigma)_{max} = 0.002$
375 parameters	$\Delta\rho_{max} = 1.48$ e Å^{-3}
21 restraints	$\Delta\rho_{min} = -1.23$ e Å^{-3}

Special details

Geometry. All e.s.d.'s (except the e.s.d. in the dihedral angle between two l.s. planes) are estimated using the full covariance matrix. The cell e.s.d.'s are taken into account individually in the estimation of e.s.d.'s in distances, angles and torsion angles; correlations between e.s.d.'s in cell parameters are only used when they are defined by crystal symmetry. An approximate (isotropic) treatment of cell e.s.d.'s is used for estimating e.s.d.'s involving l.s. planes.

Refinement. Refinement of F^2 against ALL reflections. The weighted *R*-factor *wR* and goodness of fit *S* are based on F^2, conventional *R*-factors *R* are based on *F*, with *F* set to zero for negative F^2. The threshold expression of $F^2 > 2\sigma(F^2)$ is used only for calculating *R*-factors(gt) *etc*. and is not relevant to the choice of reflections for refinement. *R*-factors based on F^2 are statistically about twice as large as those based on *F*, and *R*- factors based on ALL data will be even larger.

Fractional atomic coordinates and isotropic or equivalent isotropic displacement parameters (Å^2)

	x	*y*	*z*	U_{iso}*/U_{eq}	Occ. (<1)
Sr1	0.52281 (3)	0.38668 (2)	0.836097 (18)	0.01665 (6)	
Sr2	0.26745 (3)	0.029136 (18)	0.557006 (17)	0.01072 (6)	
S1	0.13839 (8)	0.40229 (5)	0.50122 (5)	0.01584 (12)	
C1	0.1878 (3)	0.4919 (2)	0.6477 (2)	0.0160 (5)	
C2	0.3059 (3)	0.6056 (2)	0.6709 (2)	0.0183 (5)	
C3	0.3561 (3)	0.6182 (2)	0.5698 (2)	0.0168 (5)	

C4	0.2738 (3)	0.5158 (2)	0.4721 (2)	0.0169 (5)	
C5	0.3781 (4)	0.7012 (2)	0.7831 (2)	0.0237 (6)	
N1	0.4382 (4)	0.7807 (2)	0.8712 (2)	0.0373 (7)	
C6	0.4862 (3)	0.7305 (2)	0.5736 (2)	0.0187 (5)	
H6A	0.5431	0.7077	0.5091	0.022*	
H6B	0.5745	0.7663	0.6470	0.022*	
C7	0.4063 (3)	0.8258 (2)	0.5642 (2)	0.0140 (4)	
O1	0.3026 (2)	0.85078 (16)	0.62454 (15)	0.0218 (4)	
O2	0.4503 (2)	0.87870 (15)	0.49792 (15)	0.0168 (3)	
C8	0.2835 (3)	0.4871 (2)	0.3498 (2)	0.0208 (5)	
O3	0.3592 (3)	0.57144 (17)	0.32159 (17)	0.0293 (4)	
O4	0.2095 (3)	0.37659 (16)	0.28052 (16)	0.0253 (4)	
N2	0.1127 (3)	0.44571 (17)	0.71912 (16)	0.0138 (4)	
C9	0.0097 (3)	0.3176 (2)	0.6705 (2)	0.0147 (4)	
H9A	−0.0850	0.3036	0.6067	0.018*	
H9B	−0.0415	0.2978	0.7309	0.018*	
C10	0.1092 (3)	0.2303 (2)	0.62385 (19)	0.0125 (4)	
O5	0.2644 (2)	0.25859 (14)	0.67064 (14)	0.0150 (3)	
O6	0.0259 (2)	0.13041 (14)	0.54187 (13)	0.0141 (3)	
C11	0.0937 (3)	0.5241 (2)	0.82968 (19)	0.0167 (5)	
H11A	−0.0207	0.4870	0.8398	0.020*	
H11B	0.0977	0.6033	0.8259	0.020*	
C12	0.2261 (3)	0.5475 (2)	0.9357 (2)	0.0177 (5)	
O7	0.2069 (3)	0.60995 (17)	1.03249 (15)	0.0266 (4)	
O8	0.3467 (2)	0.50536 (16)	0.92535 (15)	0.0213 (4)	
O21	0.1336 (2)	0.04188 (16)	0.73788 (14)	0.0177 (4)	
H21A	0.195 (3)	0.096 (2)	0.7968 (19)	0.027*	
H21B	0.111 (4)	−0.021 (2)	0.753 (3)	0.027*	
O22	0.5486 (2)	0.16837 (18)	0.73157 (15)	0.0218 (4)	
H22A	0.608 (4)	0.162 (3)	0.686 (2)	0.033*	
H22B	0.572 (4)	0.129 (3)	0.768 (3)	0.033*	
O23	0.2011 (2)	−0.02488 (19)	0.33592 (16)	0.0246 (4)	
H23A	0.102 (3)	−0.035 (3)	0.311 (3)	0.037*	
H23B	0.249 (4)	−0.006 (3)	0.291 (2)	0.037*	
O24	−0.0724 (3)	0.17943 (17)	0.24069 (16)	0.0233 (4)	
H24A	−0.138 (4)	0.198 (3)	0.282 (3)	0.035*	
H24B	0.015 (3)	0.240 (2)	0.260 (3)	0.035*	
O25	0.3787 (3)	0.23012 (19)	0.92443 (16)	0.0299 (5)	
H25A	0.430 (4)	0.204 (3)	0.965 (3)	0.045*	
H25B	0.316 (4)	0.259 (3)	0.958 (3)	0.045*	
O26	0.6112 (3)	1.0182 (2)	0.83803 (19)	0.0321 (5)	
H26A	0.686 (5)	1.030 (5)	0.897 (3)	0.048*	0.67
H26B	0.549 (6)	1.037 (5)	0.885 (4)	0.048*	0.67
H26C	0.532 (5)	0.952 (3)	0.823 (5)	0.048*	0.67
O27	0.7186 (3)	0.6180 (2)	0.91336 (19)	0.0194 (6)	0.790 (4)
H27A	0.740 (5)	0.626 (4)	0.854 (2)	0.029*	0.79
H27B	0.811 (4)	0.640 (4)	0.959 (3)	0.029*	0.79
O28	0.1058 (3)	0.92663 (19)	0.99457 (17)	0.0302 (5)	
H28A	0.111 (5)	0.896 (3)	0.9290 (18)	0.045*	

H28B	0.041 (8)	0.966 (6)	1.000 (7)	0.045*	0.50
H28C	0.059 (8)	0.873 (5)	1.015 (6)	0.045*	0.50
O29	0.0120 (4)	0.7429 (3)	1.0829 (2)	0.0287 (7)	0.790 (4)
H29A	0.073 (7)	0.701 (5)	1.058 (6)	0.043*	0.53
H29B	0.045 (9)	0.806 (4)	1.071 (6)	0.043*	0.53
H29C	−0.024 (10)	0.732 (8)	1.016 (3)	0.043*	0.53
O30	−0.0951 (12)	0.6963 (8)	0.9739 (8)	0.023 (2)*	0.210 (4)

Atomic displacement parameters (Å^2)

	U^{11}	U^{22}	U^{33}	U^{12}	U^{13}	U^{23}
Sr1	0.01489 (12)	0.02075 (12)	0.00915 (10)	0.00204 (9)	0.00133 (8)	0.00307 (9)
Sr2	0.01086 (11)	0.01068 (10)	0.00940 (10)	0.00340 (8)	0.00144 (7)	0.00290 (8)
S1	0.0195 (3)	0.0115 (3)	0.0149 (3)	0.0032 (2)	0.0029 (2)	0.0051 (2)
C1	0.0158 (12)	0.0144 (11)	0.0177 (11)	0.0064 (9)	−0.0008 (9)	0.0067 (9)
C2	0.0197 (12)	0.0133 (11)	0.0215 (12)	0.0061 (10)	−0.0004 (10)	0.0078 (10)
C3	0.0133 (11)	0.0122 (11)	0.0255 (12)	0.0049 (9)	0.0006 (9)	0.0087 (10)
C4	0.0164 (12)	0.0138 (11)	0.0241 (12)	0.0060 (9)	0.0057 (10)	0.0105 (10)
C5	0.0287 (15)	0.0139 (12)	0.0256 (13)	0.0010 (11)	−0.0027 (11)	0.0120 (11)
N1	0.0485 (17)	0.0204 (12)	0.0261 (13)	−0.0078 (11)	−0.0058 (12)	0.0083 (10)
C6	0.0129 (11)	0.0152 (12)	0.0301 (13)	0.0046 (9)	0.0032 (10)	0.0122 (10)
C7	0.0120 (11)	0.0104 (10)	0.0167 (11)	0.0015 (8)	0.0008 (9)	0.0042 (9)
O1	0.0269 (10)	0.0231 (9)	0.0262 (9)	0.0140 (8)	0.0147 (8)	0.0152 (8)
O2	0.0174 (9)	0.0160 (8)	0.0201 (8)	0.0055 (7)	0.0065 (7)	0.0101 (7)
C8	0.0220 (13)	0.0200 (13)	0.0260 (13)	0.0100 (11)	0.0119 (11)	0.0111 (11)
O3	0.0386 (12)	0.0209 (10)	0.0317 (11)	0.0058 (9)	0.0167 (9)	0.0141 (9)
O4	0.0348 (11)	0.0179 (9)	0.0252 (10)	0.0072 (8)	0.0152 (8)	0.0092 (8)
N2	0.0185 (10)	0.0084 (9)	0.0109 (9)	0.0030 (8)	0.0002 (7)	0.0016 (7)
C9	0.0142 (11)	0.0121 (11)	0.0150 (11)	0.0029 (9)	0.0014 (9)	0.0038 (9)
C10	0.0157 (11)	0.0109 (10)	0.0106 (10)	0.0022 (9)	0.0027 (8)	0.0056 (8)
O5	0.0125 (8)	0.0144 (8)	0.0140 (8)	0.0033 (6)	−0.0023 (6)	0.0032 (6)
O6	0.0140 (8)	0.0104 (8)	0.0127 (8)	0.0016 (6)	−0.0014 (6)	0.0016 (6)
C11	0.0195 (12)	0.0164 (11)	0.0125 (11)	0.0096 (10)	0.0008 (9)	0.0017 (9)
C12	0.0235 (13)	0.0113 (11)	0.0151 (11)	0.0023 (9)	0.0003 (9)	0.0055 (9)
O7	0.0378 (12)	0.0255 (10)	0.0124 (8)	0.0109 (9)	0.0040 (8)	0.0028 (7)
O8	0.0213 (9)	0.0194 (9)	0.0199 (9)	0.0077 (7)	−0.0036 (7)	0.0056 (7)
O21	0.0200 (9)	0.0184 (9)	0.0120 (8)	0.0043 (7)	0.0025 (7)	0.0046 (7)
O22	0.0186 (9)	0.0305 (10)	0.0139 (9)	0.0109 (8)	0.0024 (7)	0.0043 (8)
O23	0.0172 (9)	0.0421 (12)	0.0195 (9)	0.0106 (9)	0.0070 (8)	0.0164 (9)
O24	0.0270 (11)	0.0201 (9)	0.0198 (9)	0.0028 (8)	0.0098 (8)	0.0066 (8)
O25	0.0413 (13)	0.0248 (11)	0.0151 (9)	0.0009 (9)	0.0046 (9)	0.0060 (8)
O26	0.0299 (12)	0.0322 (12)	0.0283 (11)	0.0011 (10)	0.0029 (9)	0.0131 (10)
O27	0.0243 (13)	0.0161 (11)	0.0129 (11)	0.0003 (9)	0.0054 (9)	0.0042 (9)
O28	0.0422 (14)	0.0295 (12)	0.0163 (9)	0.0133 (10)	0.0022 (9)	0.0060 (9)
O29	0.0313 (15)	0.0330 (15)	0.0249 (14)	0.0161 (12)	0.0070 (11)	0.0107 (12)

Geometric parameters (Å, °)

Sr1—O8	2.4557 (18)	N2—C11	1.462 (3)

Sr1—O3^{i}	2.4782 (19)	C9—C10	1.540 (3)
Sr1—O5	2.5234 (16)	C9—H9A	0.9900
Sr1—O7^{ii}	2.6149 (19)	C9—H9B	0.9900
Sr1—O25	2.652 (2)	C10—O5	1.258 (3)
Sr1—O22	2.6560 (19)	C10—O6	1.258 (3)
Sr1—O27	2.657 (2)	C11—C12	1.517 (3)
Sr1—O8^{ii}	2.7834 (17)	C11—H11A	0.9900
Sr2—O6^{iii}	2.5452 (16)	C11—H11B	0.9900
Sr2—O23	2.5921 (18)	C12—O8	1.253 (3)
Sr2—O2^{i}	2.6222 (17)	C12—O7	1.262 (3)
Sr2—O21	2.6445 (17)	O21—H21A	0.809 (18)
Sr2—O6	2.6628 (16)	O21—H21B	0.845 (17)
Sr2—O2^{iv}	2.6848 (16)	O22—H22A	0.808 (18)
Sr2—O1^{iv}	2.6944 (17)	O22—H22B	0.819 (17)
Sr2—O22	2.7108 (18)	O23—H23A	0.807 (18)
Sr2—O5	2.7228 (16)	O23—H23B	0.798 (18)
S1—C1	1.733 (2)	O24—H24A	0.823 (18)
S1—C4	1.735 (2)	O24—H24B	0.837 (18)
C1—N2	1.358 (3)	O25—H25A	0.819 (18)
C1—C2	1.398 (3)	O25—H25B	0.796 (18)
C2—C5	1.431 (4)	O26—H26A	0.850 (19)
C2—C3	1.439 (4)	O26—H26B	0.849 (19)
C3—C4	1.368 (3)	O26—H26C	0.85 (2)
C3—C6	1.504 (3)	O27—H27A	0.827 (19)
C4—C8	1.482 (4)	O27—H27B	0.823 (19)
C5—N1	1.148 (3)	O28—H28A	0.799 (18)
C6—C7	1.532 (3)	O28—H28B	0.82 (2)
C6—H6A	0.9900	O28—H28C	0.82 (2)
C6—H6B	0.9900	O29—H29A	0.83 (2)
C7—O1	1.256 (3)	O29—H29B	0.83 (2)
C7—O2	1.261 (3)	O29—H29C	0.82 (2)
C8—O3	1.257 (3)	O27—O30^{v}	1.534 (10)
C8—O4	1.275 (3)	O29—O30	1.383 (10)
N2—C9	1.456 (3)		
O8—Sr1—O3^{i}	117.69 (6)	O22—Sr2—O5	67.10 (5)
O8—Sr1—O5	87.39 (5)	C1—S1—C4	92.44 (12)
O3^{i}—Sr1—O5	81.32 (6)	N2—C1—C2	130.8 (2)
O8—Sr1—O7^{ii}	118.97 (6)	N2—C1—S1	119.08 (17)
O3^{i}—Sr1—O7^{ii}	101.51 (7)	C2—C1—S1	110.09 (19)
O5—Sr1—O7^{ii}	146.34 (5)	C1—C2—C5	125.2 (2)
O8—Sr1—O25	85.98 (7)	C1—C2—C3	113.4 (2)
O3^{i}—Sr1—O25	150.13 (6)	C5—C2—C3	121.3 (2)
O5—Sr1—O25	81.94 (6)	C4—C3—C2	111.8 (2)
O7^{ii}—Sr1—O25	79.69 (7)	C4—C3—C6	124.9 (2)
O8—Sr1—O22	147.38 (6)	C2—C3—C6	123.3 (2)

O3^{i}—Sr1—O22	83.58 (6)	C3—C4—C8	131.1 (2)
O5—Sr1—O22	70.82 (5)	C3—C4—S1	112.22 (19)
O7^{ii}—Sr1—O22	76.14 (6)	C8—C4—S1	116.67 (18)
O25—Sr1—O22	67.59 (6)	N1—C5—C2	177.5 (3)
O8—Sr1—O27	74.49 (7)	C3—C6—C7	112.26 (19)
O3^{i}—Sr1—O27	70.14 (7)	C3—C6—H6A	109.2
O5—Sr1—O27	132.72 (6)	C7—C6—H6A	109.2
O7^{ii}—Sr1—O27	77.82 (7)	C3—C6—H6B	109.2
O25—Sr1—O27	137.67 (6)	C7—C6—H6B	109.2
O22—Sr1—O27	138.07 (7)	H6A—C6—H6B	107.9
O8—Sr1—O8^{ii}	70.87 (6)	O1—C7—O2	122.2 (2)
O3^{i}—Sr1—O8^{ii}	130.40 (6)	O1—C7—C6	119.1 (2)
O5—Sr1—O8^{ii}	147.09 (6)	O2—C7—C6	118.7 (2)
O7^{ii}—Sr1—O8^{ii}	48.19 (5)	O3—C8—O4	125.2 (2)
O25—Sr1—O8^{ii}	72.41 (6)	O3—C8—C4	118.9 (2)
O22—Sr1—O8^{ii}	115.58 (5)	O4—C8—C4	115.9 (2)
O27—Sr1—O8^{ii}	65.82 (6)	C1—N2—C9	117.63 (19)
O6^{iii}—Sr2—O23	69.11 (6)	C1—N2—C11	121.82 (19)
O6^{iii}—Sr2—O2^{i}	138.88 (5)	C9—N2—C11	118.5 (2)
O23—Sr2—O2^{i}	71.52 (6)	N2—C9—C10	114.24 (19)
O6^{iii}—Sr2—O21	79.58 (5)	N2—C9—H9A	108.7
O23—Sr2—O21	145.03 (6)	C10—C9—H9A	108.7
O2^{i}—Sr2—O21	141.51 (5)	N2—C9—H9B	108.7
O6^{iii}—Sr2—O6	68.34 (6)	C10—C9—H9B	108.7
O23—Sr2—O6	81.03 (5)	H9A—C9—H9B	107.6
O2^{i}—Sr2—O6	116.53 (5)	O5—C10—O6	123.3 (2)
O21—Sr2—O6	73.03 (5)	O5—C10—C9	119.92 (19)
O6^{iii}—Sr2—O2^{iv}	96.94 (5)	O6—C10—C9	116.7 (2)
O23—Sr2—O2^{iv}	81.54 (6)	N2—C11—C12	115.7 (2)
O2^{i}—Sr2—O2^{iv}	65.68 (6)	N2—C11—H11A	108.3
O21—Sr2—O2^{iv}	118.40 (5)	C12—C11—H11A	108.3
O6—Sr2—O2^{iv}	160.37 (5)	N2—C11—H11B	108.3
O6^{iii}—Sr2—O1^{iv}	79.58 (5)	C12—C11—H11B	108.3
O23—Sr2—O1^{iv}	116.26 (6)	H11A—C11—H11B	107.4
O2^{i}—Sr2—O1^{iv}	108.25 (5)	O8—C12—O7	122.9 (2)
O21—Sr2—O1^{iv}	71.00 (5)	O8—C12—C11	120.5 (2)
O6—Sr2—O1^{iv}	135.19 (5)	O7—C12—C11	116.7 (2)
O2^{iv}—Sr2—O1^{iv}	48.34 (5)	H21A—O21—H21B	106 (3)
O6^{iii}—Sr2—O22	156.57 (6)	H22A—O22—H22B	104 (3)
O23—Sr2—O22	133.65 (6)	H23A—O23—H23B	104 (3)
O2^{i}—Sr2—O22	62.33 (5)	H24A—O24—H24B	108 (3)
O21—Sr2—O22	79.68 (6)	H25A—O25—H25B	109 (4)
O6—Sr2—O22	115.27 (5)	H26A—O26—H26B	87 (5)

O2^{iv}—Sr2—O22	83.49 (5)	H26A—O26—H26C	108 (5)
O1^{iv}—Sr2—O22	83.47 (6)	H26B—O26—H26C	75 (5)
O6^{iii}—Sr2—O5	114.87 (5)	H27A—O27—H27B	107 (4)
O23—Sr2—O5	109.66 (6)	H28A—O28—H28B	111 (6)
O2^{i}—Sr2—O5	89.11 (5)	H28A—O28—H28C	109 (6)
O21—Sr2—O5	69.47 (5)	H28B—O28—H28C	103 (7)
O6—Sr2—O5	48.56 (5)	H29A—O29—H29B	107 (7)
O2^{iv}—Sr2—O5	148.19 (5)	H29A—O29—H29C	87 (6)
O1^{iv}—Sr2—O5	133.93 (5)	H29B—O29—H29C	68 (6)

Symmetry codes: (i) −*x*+1, −*y*+1, −*z*+1; (ii) −*x*+1, −*y*+1, −*z*+2; (iii) −*x*, −*y*, −*z*+1; (iv) *x*, *y*−1, *z*; (v) *x*+1, *y*, *z*.

Hydrogen-bond geometry (Å, °)

D—H···*A*	*D*—H	H···*A*	*D*···*A*	*D*—H···*A*
O21—H21A···O25	0.81 (2)	1.98 (2)	2.781 (3)	169 (3)
O21—H21B···O24^{iii}	0.85 (2)	1.93 (2)	2.765 (3)	169 (3)
O22—H22A···O2^{i}	0.81 (2)	2.16 (3)	2.761 (2)	131 (3)
O22—H22B···O26^{iv}	0.82 (2)	1.94 (2)	2.755 (3)	173 (3)
O23—H23A···O21^{iii}	0.81 (2)	1.96 (2)	2.766 (3)	174 (4)
O23—H23B···O26^{i}	0.80 (2)	2.11 (2)	2.867 (3)	159 (3)
O24—H24A···O1^{vi}	0.82 (2)	2.03 (2)	2.760 (3)	148 (3)
O24—H24B···O4	0.84 (2)	1.93 (2)	2.756 (3)	172 (3)
O25—H25A···N1^{ii}	0.82 (2)	2.15 (2)	2.898 (3)	152 (3)
O25—H25B···O27^{ii}	0.80 (2)	1.93 (2)	2.648 (3)	150 (4)
O26—H26A···O28^{vii}	0.85 (2)	1.90 (2)	2.731 (3)	164 (5)
O26—H26C···N1	0.85 (2)	2.37 (4)	3.108 (3)	146 (5)
O27—H27A···O4^{i}	0.83 (2)	1.79 (2)	2.615 (3)	172 (4)
O27—H27B···O29^{v}	0.82 (2)	1.94 (2)	2.727 (4)	160 (4)
O28—H28A···O24^{vi}	0.80 (2)	1.97 (2)	2.756 (3)	167 (4)
O28—H28B···O28^{viii}	0.82 (2)	2.02 (2)	2.835 (4)	174 (8)
O28—H28C···O29	0.82 (2)	2.04 (3)	2.836 (4)	164 (7)
O29—H29A···O7	0.83 (2)	1.78 (2)	2.595 (3)	168 (7)
O29—H29B···O28	0.83 (2)	2.03 (3)	2.836 (4)	164 (7)

Symmetry codes: (iii) −*x*, −*y*, −*z*+1; (i) −*x*+1, −*y*+1, −*z*+1; (iv) *x*, *y*−1, *z*; (vi) −*x*, −*y*+1, −*z*+1; (ii) −*x*+1, −*y*+1, −*z*+2; (vii) −*x*+1, −*y*+2, −*z*+2; (v) *x*+1, *y*, *z*; (viii) −*x*, −*y*+2, −*z*+2.

Fig. 1

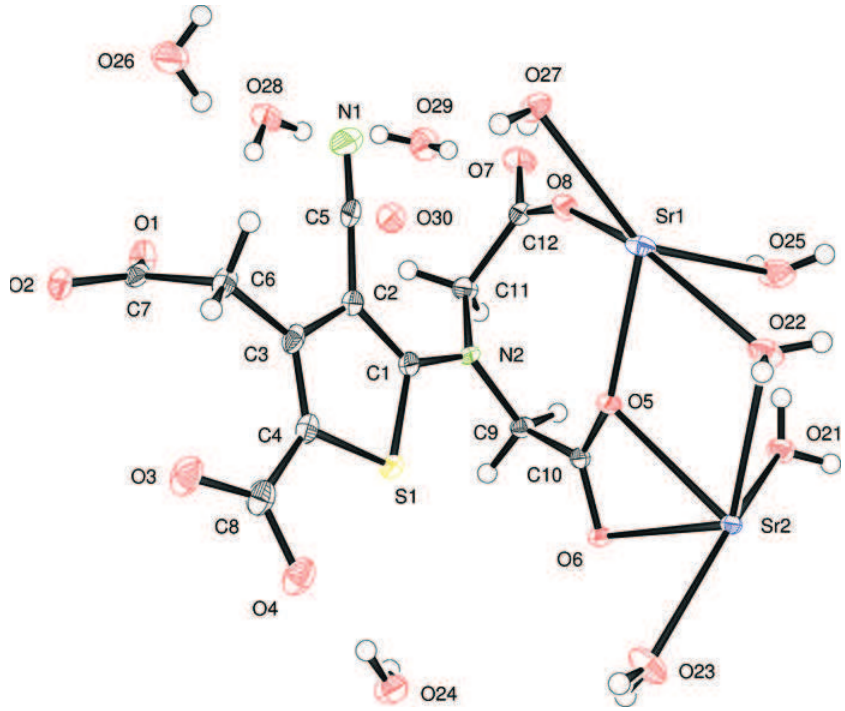

O26
O28
N1
O29
O27
O7
O8
Sr1
O25
O1
C5
O30
C12
O2
C6
C2
C11
C7
O22
C3
C1
N2
O5
C4
C9
O21
O3
C10
S1
Sr2
C8
O6
O4
O24
O23

Fig. 2

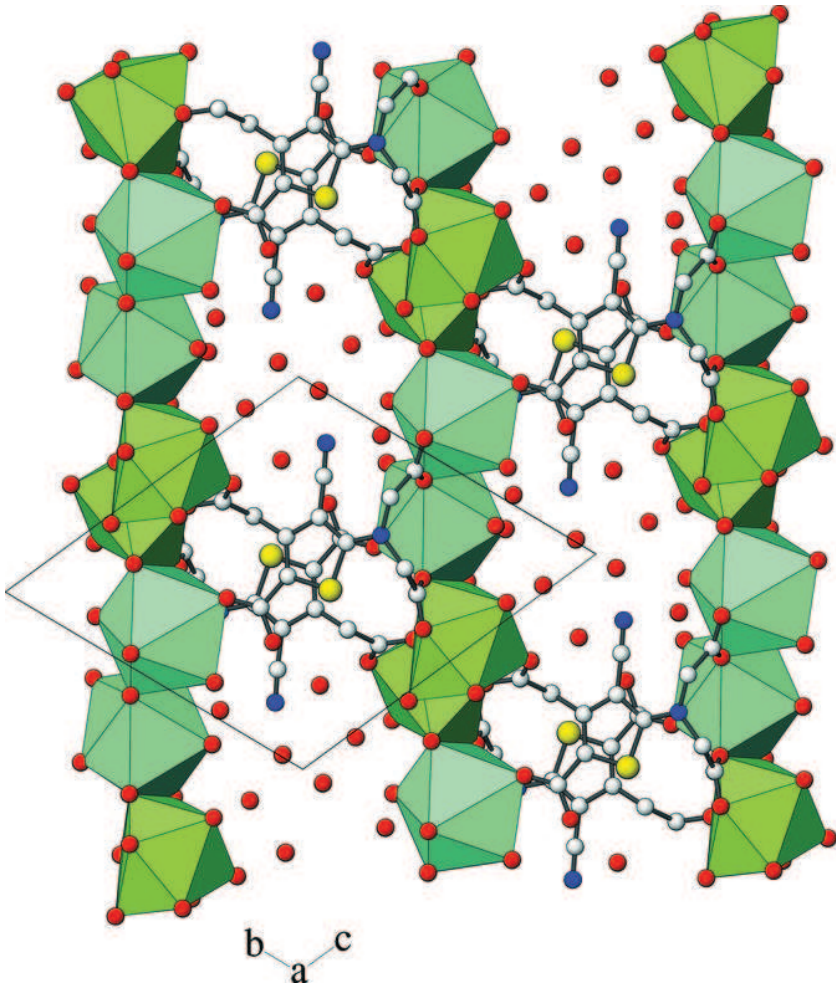
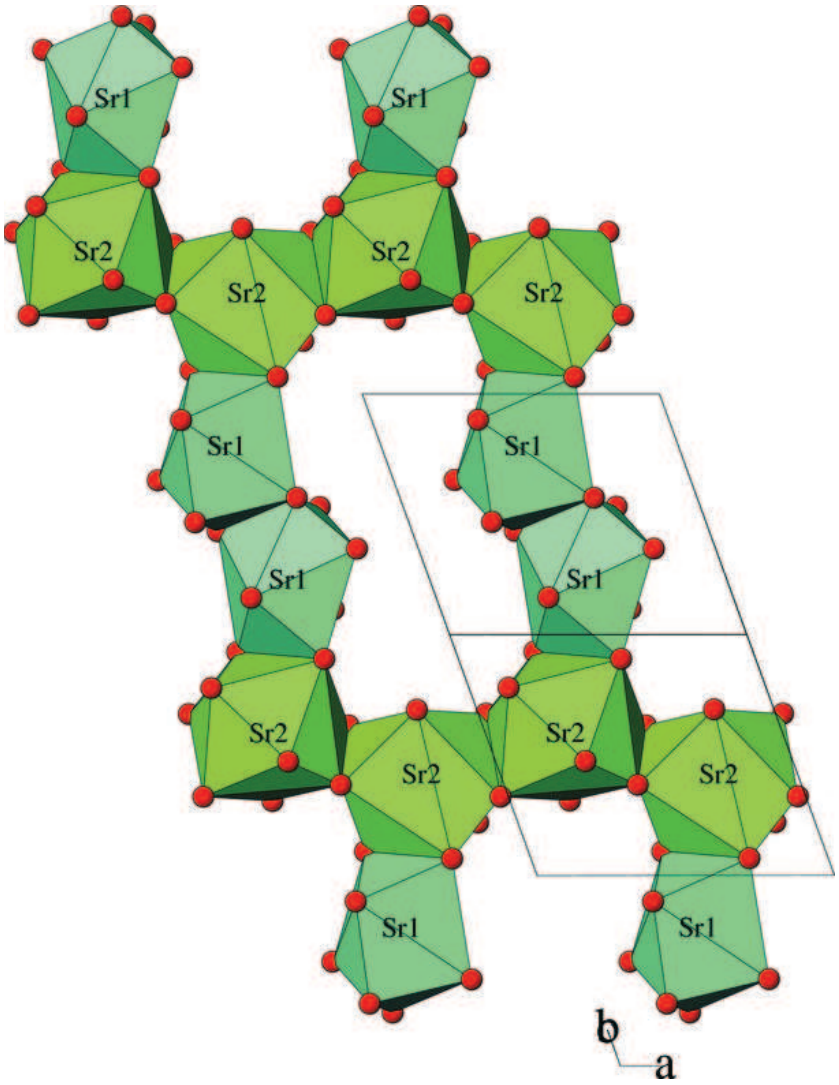


Fig. 3


Sr1
Sr1
Sr2
Sr2
Sr2
Sr2
Sr1
Sr1
Sr1
Sr1
Sr2
Sr2
Sr2
Sr2
Sr1
Sr1
b
a

Appendix B

XRD structures of T_6 insulin with Rb and Cs

Insulin was crystallized with monovalent cations of rubidium and caesium, and the structures were solved. The structures are breifly presented in this appendix.

B.1 Growth of single crystals

The crystals were grown using insulin solutions where the zinc was not chelated and removed prior to the crystallization experiments. Insulin was dissolved in 0.01 M HCl and pH was carefully raised with NaOH until all protein was re-dissolved at pH around 7.0 The protein concentration was determined to 13.0 mg ml^{-1}. The crystals were grown using the vapour diffusion technique. 4 µl of the insulin solution was mixed with 2 µl reservoir solution and 2 µl 0.1 M RbCl or CsCl and equilibrated in a sitting drop against 1 ml reservoir containing of 1 M ammonium sulfate and 0.05 M sodium citrate. After 2 months small crystals with dimensions of approx. 100 µm were observed.

B.2 Single crystal diffraction

Single crystal diffraction data were collected at MAX-lab, Lund, Sweden, on beam line 911-3, MAX-II using a MarResearch CCD detector. The wavelength was 1.00 Å and the data were collected at 100 K. The structures were refined and validated as described in

Chapter 7. Data collection, data processing, refinement and validation statistics for all crystals are summarized in Table B.1.

Table B.1: Data collection parameters, processing-, refinement- and validation statistics. All data were collected at beam line 911-3, MAX II, MAX-lab, Lund, Sweden. Values in parentheses are for the outermost resolution shell.

	Rb^+ insulin	Cs^+ insulin
Data collection		
Detector distance (mm)	263	203
Resolution collected (Å)	2.37	1.92
Data processing		
Resolution (Å)	24.15–2.60 (2.68–2.60)	24.29–2.06 (2.31–2.06)
No. of reflections	10137 (886)	29630 (8618)
No. of unique reflections	2500 (223)	5139 (1505)
Redundancy	4.05 (3.97)	5.76 (5.73)
Completeness (%)	99.9 (100.0)	99.9 (100.0)
R_{sym} (%)	17.5 (79.3)	7.5 (42.8)
$\langle I/\sigma(I)\rangle$	7.81 (1.86)	18.43 (4.45)
Space group	$R3$	$R3$
No. of molecules per asu	2	2
a (Å)	80.93	81.70
c (Å)	33.34	33.40
Average mosaicity (°)	0.52	0.45
Overall B-factors (Å^2)	32	31
No. of atoms in the model		
Total non H-atoms	774	840
Total protein atoms	766	795
Total ordered water molecules	6	43
Total M^+-ions	2	2
Type of M^+ siteI/siteII	Rb/Rb	Cs/Zn
RMS deviation from ideal		
Bonds (Å)	0.01	0.01
Angles (°)	1.23	1.08
Ramachandran plot		
In core regions (%)	98.9	98.9
Outliers (%)	1.1	1.1
R-factors		
R	0.1947	0.1891
R_{free}	0.2747	0.2272

The structures were refined using the peptide chain from the T_6 Zn insulin structure described in Chapter 7. In the rubidium structure both metal sites were modelled with rubidium, whereas in the caesium structure one site was modelled with caesium and the other with zinc. Residues PheB1.1, AlaB30.1, PheB1.2 and AlaB30.2 in the rubidium structure and residues PheB1.1 and AlaB30.1 in the caesium structure were disordered and

could not be modelled. In the caesium structure the side chains of residues GluB21.1 and GluB21.2 were modelled in two alternating conformations. Water molecules were inserted and restrained refinement was carried out in *PHENIX*. The atomic displacement factors for the peptide chain were refined isotropically. Validation of both structures showed that only one residue SerA9.1 fell into the outlier region in the Ramachandran plot using the regions defined by Kleywegt & Jones (1996).

B.3 Results

Both the Rb and the Cs structure adopted same comformation as the Zn T_6 insulin (RMS displacement of $C^{\alpha} < 0.4$), where the conformation of the backbone of residues B1–B8 was extended in both structures. The coordination geometries of each metal site are shown in Figure B.1.

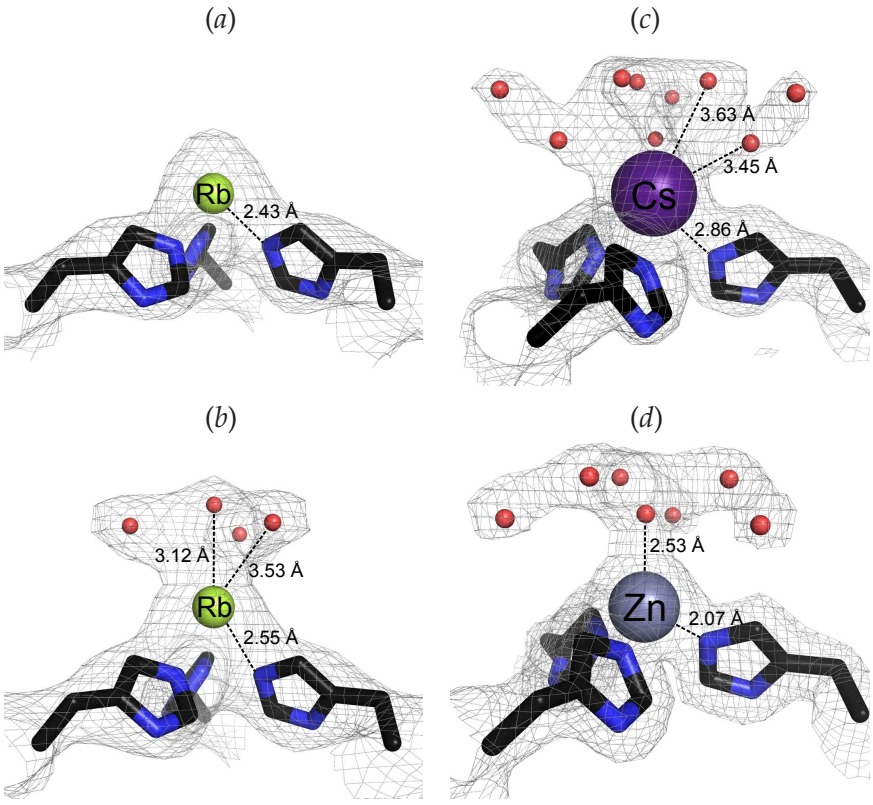


Figure B.1: (*a*) Rubidium site I and (*b*) rubidium site II in the 2.60 Å Rb insulin structure. The Rb^{+} cations are possibly seven-coordinated in both sites, however the electron density is too weak in site I. (*c*) Caesium site I and (*d*) site II, which is best modelled by a zinc atom in the 2.06 Å Cs insulin structure. The Cs^{+} ion is nine coordinated, whereas the Zn^{2+} ion shows tetrahedral coordination geometry. Distances to the first coordination sphere are shown. The σ_A-weighted $2F_o - F_c$ maps have been contoured at 1.0 σ.

Appendix C

EXAFS analysis of Sr reference compounds

This appendix describes the EXAFS analysis performed at room temperature on selected strontium salts, which have been used as reference compounds for the EXAFS studies on strontium containing bone samples.

C.1 Introduction

The positions of Sr-atoms in the pure reference compounds are unambiguously determined from X-ray crystal structures. Analysing well characterized structures with XAS means that factors dependent on equipment and experimental design, e.g. the amplitude reduction factor and the general level of Debye-Waller factors, could be determined.

C.2 Experimental

C.2.1 Samples

The following strontium containing compounds were analysed: Strontium chloride hexahydrate (Merck, analytical grade) strontium disalicylate dihydrate (preparation kindly provided by J. E. T. Andersen, Technical University of Denmark), strontium hydroxyapatite (sample kindly provided by H. Birkedal, Aarhus University) and 5 % strontium

doped calcium hydroxyapatite (preparation kindly provided by J. E. T. Andersen, Technical University of Denmark).

Preparation of the reference samples were done by grinding the samples in a mortar. The samples with high strontium contents were diluted 3 to 4 times with boron nitride in order to obtain a good EXAFS signal. The powder was pressed into tablets and mounted on the XAS sample holder using kapton tape.

C.2.2 X-ray absorption spectroscopy

Sr K-edge X-ray absorption spectra were recorded on beam line 811, MAX-lab, Lund, Sweden in accordance to the procedure described in Chapter 11. Room temperature data were collected in the region 15954–16754 eV except the 5 % Sr-doped CaHA sample where data were collected to 17003 eV. At least three spectra were collected for each sample to ensure reproducibility and that no degradation or radiation damage of the sample had taken place.

The spectra for each sample were averaged, subtracted for background and normalized before the EXAFS, $\chi(k)$ was extracted, as described in Section 3.4. The EXAFS data reduction and modelling was performed in *WinXAS* (Ressler, 1998) using *ATOMS*, which is a part of the *IFEFFIT*-package (Ravel & Newville, 2005), to calculate distances from the strontium absorber to the surrounding atoms and *FEFF8* (Ankudinov *et al.*, 1998; Rehr & Albers, 2000) to calculate the XAFS scattering paths. The model fits were evaluated with the following residual (Ressler, 1998):

$$R = \frac{\sum_{i=1}^{N} \left| y_{\text{exp}}(i) - y_{\text{theo}}(i) \right|}{\sum_{i=1}^{N} \left| y_{\text{exp}}(i) \right|} \cdot 100\% \tag{C.1}$$

C.3 Results and discussion

C.3.1 Strontium chloride hexahydrate

The crystal structure of $SrCl_2 \cdot 6H_2O$ (English & Nassimbeni, 1984), see Figure C.1 was used as a model for the EXAFS refinement. Distances within a radius of 6.3 Å and Debye-Waller factors were determined and listed in Table C.1. The energy scale offset ΔE_0 was refined and finally, by keeping the refined distances constant, the amplitude reduction factor σ_0^2 was refined to 0.71. A residual of 3.8 % was obtained. The calculated k^3-weighted $\chi(k)$ and its Fourier transform are compared with the experimentally obtained and shown in Figure C.2.

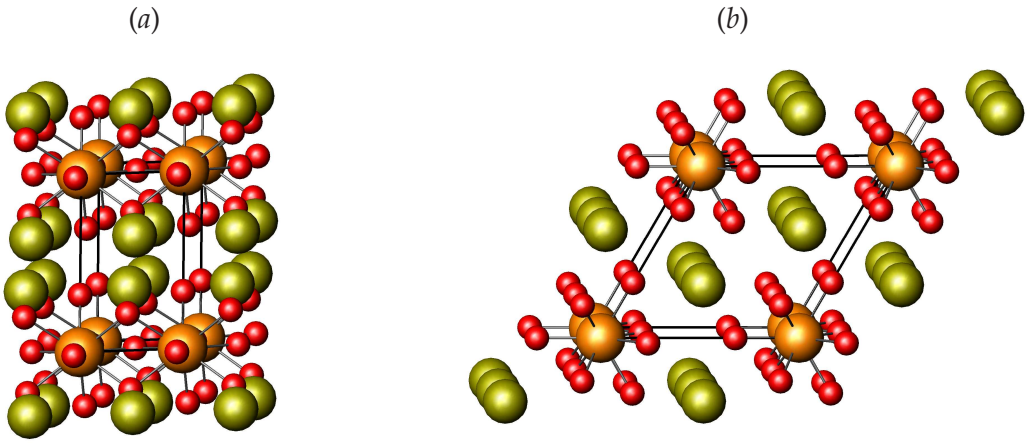


Figure C.1: Strontium chloride hexahydrate crystallizes in space group $P32_1$ and the structure contains Sr^{2+} cations (orange) coordinated to nine oxygen atoms (red). Chlorine atoms are shown in olive. (*a*) The structure viewed along the crystallographic a-axis and (*b*) along the c-axis.

Table C.1: Sr environment in $SrCl_2{\cdot}6H_2O$ within a radius of 6.3 Å, as determined by X-ray crystallography (English & Nassimbeni, 1984) and EXAFS.

Atom type	N_j	R_j(Å) (X-ray)	R_j(Å) (EXAFS)	σ_j^2(Å^2) (EXAFS)
Oxygen	9.0	2.660[a]	2.611(9)	0.016(1)
Strontium	2.0	4.12	4.08(4)	0.017(4)
Chlorine	6.0	4.91	4.95(5)	0.021(7)
Chlorine	6.0	5.18	5.18(3)	0.014(4)
Oxygen	12.0	6.43	6.33(6)	0.019(6)

[a] Weighted average value of 3 oxygen atoms at 2.562 Å and 6 oxygen atoms at 2.709 Å.

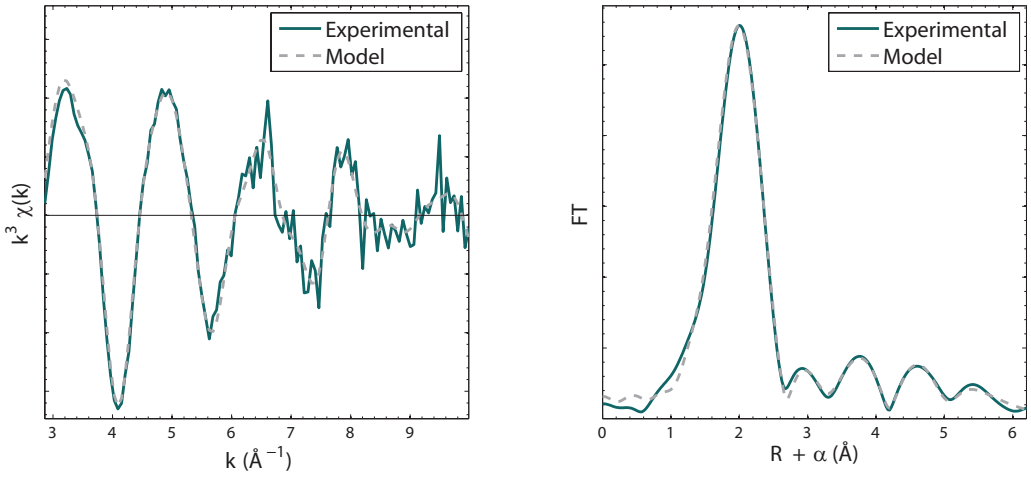


Figure C.2: Left: The background subtracted EXAFS $\chi(k)$ as a function of the photoelectronic wave number k of $SrCl_2{\cdot}6H_2O$. The function is k^3-weighted in order to enhance the details at high k-values. Right: The modulus of the Fourier transform of $\chi(k)$ which gives the radial distribution of electron density around an average Sr. The R-values are not absolute due to a phase shift α.

C.3.2 Strontium disalicylate dihydrate

The crystal structure of strontium disalicylate dihydrate ($SrC_{14}H_{14}O_8 \cdot 2H_2O$) has been solved from XRD (Ståhl, 2009), see Figure C.3 and used as a model for the EXAFS refinement. Distances within a radius of 4.4 Å from the Sr-atom and Debye-Waller factors were determined and listed in Table C.2. The energy scale offset ΔE_0 was refined and finally, by keeping the refined distances constant, the amplitude reduction factor σ_0^2 was refined to 0.86. A residual of 1.7 % was obtained. The calculated k^3-weighted $\chi(k)$ and its Fourier transform are compared with the experimentally obtained and shown in Figure C.4.

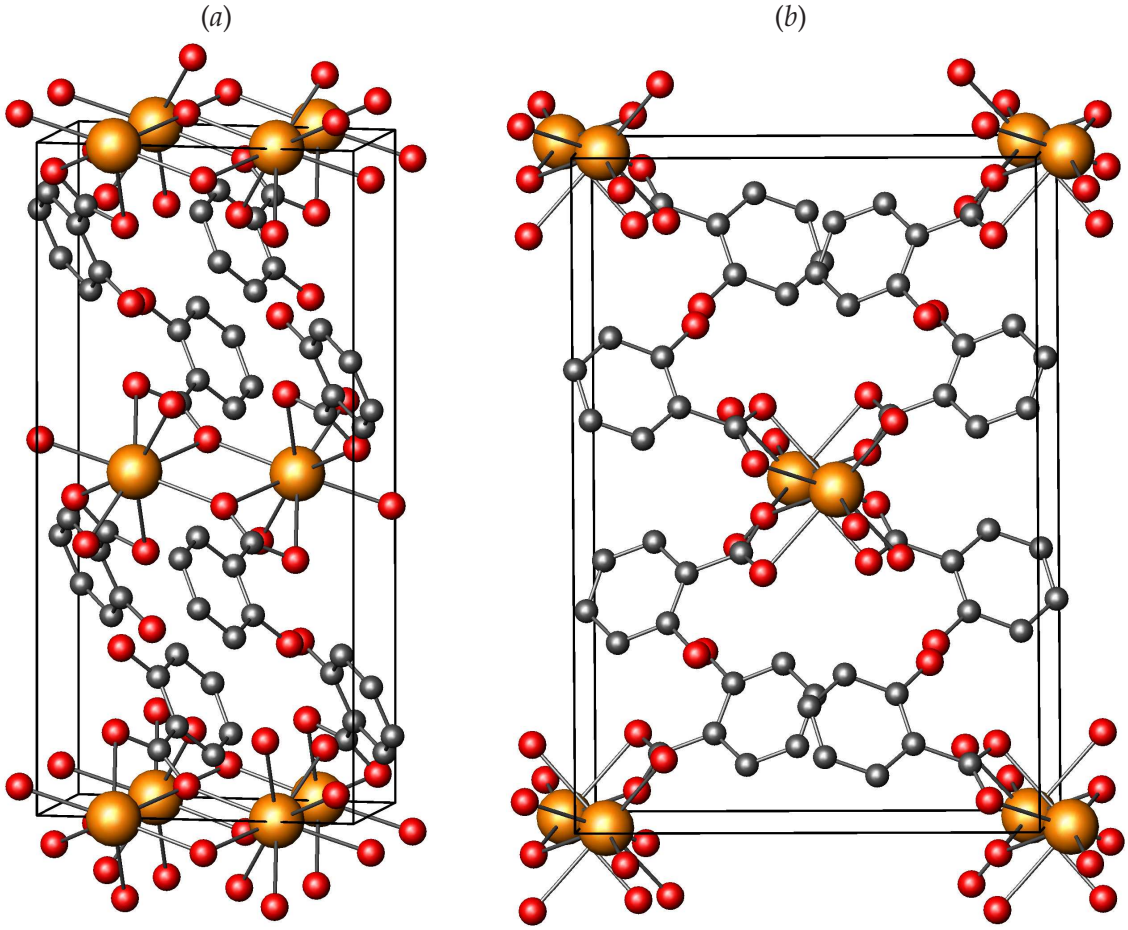


Figure C.3: Strontium disalicylate dihydrate crystallizes in space group $C2/c$ and the structure contains Sr^{2+} cations (orange) coordinated to eight oxygen atoms (red). Carbon atoms are shown in grey. (*a*) The structure viewed along the crystallographic *b*-axis and (*b*) along the *c*-axis.

Table C.2: Sr environment in Sr disalicylate dihydrate within a radius of 4.4 Å, as determined by X-ray crystallography (Ståhl, 2009) and EXAFS.

Atom type	N_j	R_j(Å) (X-ray)	R_j(Å) (EXAFS)	σ_j^2(Å^2) (EXAFS)
Oxygen	6.0	2.57[a]	2.53(1)	0.017(1)
Oxygen	2.0	2.77	2.90(1)	0.005(1)
Carbon	2.0	3.07	3.13(1)	0.005(1)
Carbon	2.0	3.67	3.71(1)	0.002(1)
Carbon	2.0	4.15	4.00(3)	0.002(1)
Strontium	2.0	4.17	4.08(3)	0.010(1)
Carbon/Oxygen	4.0	4.35[b]	4.40(5)	0.017(3)

[a] Weighted average value of 6 oxygen atoms in the range 2.52–2.64 Å.
[b] Weighted average value of 2 carbons at 4.33 Å and 2 oxygen atoms at 4.36 Å. The EXAFS for this shell was modelled by oxygen.

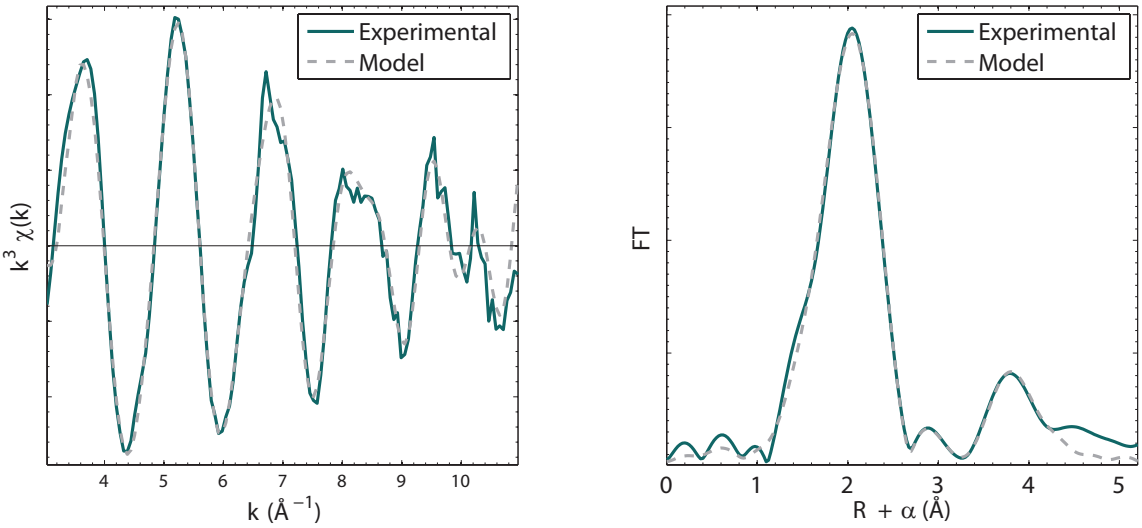


Figure C.4: Left: The background subtracted k^3-weighted $\chi(k)$ as a function of the photoelectronic wave number k of Sr disalicylate dihydrate. Right: The modulus of the Fourier transform of $\chi(k)$ which gives the radial distribution of electron density around an average Sr. The R-values are not absolute due to a phase shift α.

C.3.3 Strontium hydroxyapatite

SrHA contains two different Sr-sites, see Figure C.5. Since it is the average strontium environment which is detected by EXAFS, the average Sr environment have to be calculated as a weighted linear combination of the Sr(I) and Sr(II) environment which make up 40 % and 60 % of the Sr-sites, respectively.

Distances R_j and coordination numbers N_j for all composite shells j were calculated based on the crystal structure (Sudarsanan & Young, 1972). From this analysis a shell model for $R_j < 4$ Å for an average Sr atom in the structure could be suggested, see Table C.3, which as a consequence of the different weighting of Sr-sites, contains non-integral coordination numbers, N_j. Distances to the same type of atom, which are close to each other, were grouped into the same composite shells, in analogy to the EXAFS analysis of CaHA (Harries *et al.*, 1986). The resolution of the different shells was estimated to $\pi/(2\Delta k) = 0.16$ Å, roughly meaning that distances separated with less than 0.2 Å cannot be distinguished and was therefore used as a criterion for grouping close-lying shells.

From the refinement of the shell model the distances and Debye-Waller factors were determined and listed in Table C.3. The energy scale offset ΔE_0 was refined and finally, by keeping the refined distances constant, the amplitude reduction factor σ_0^2 was refined to 0.89. Only single scattering (SS) paths have been included in the model. A residual of 4.0 % was obtained.

Table C.3: Average strontium environment for hydroxyapatite up to $R_j = 4.0$ Å away for the Sr-atom as determined by X-ray crystallography (Sudarsanan & Young, 1972) and EXAFS.

Atom type	N_j	R_j(Å) (X-ray)	R_j(Å) (EXAFS)	σ_j^2(Å^2) (EXAFS)
Oxygen	4.8	2.53	2.561(4)	0.021(1)
Oxygen	3.0	2.78[a]	2.944(4)	0.019(1)
Phosphorus	2.4	3.31	3.53(2)	0.023(2)
Oxygen	1.8	3.47	– [b]	–
Strontium	0.8	3.63	3.679(2)	0.005(1)
Phosphorus	3.0	3.76	3.80(1)	0.010(1)

[a] Weighted average value of 1.2 oxygen atoms at 2.661 Å and 1.8 oxygen atoms at 2.853 Å.

[b] The oxygen shell at 3.47 Å could not be modelled separately, but the low Debye-Waller factor for the 3.63 Å Sr–Sr distance indicate lacking electron density in this region.

The calculated EXAFS and the Fourier transform of the model are compared with the experimentally obtained and shown in Figure C.6. Inconsistencies between the data and the model observed in the low k-region e.g. the shoulder at $k = 3.6$ Å^{-1} may be explained from missing inclusion of path with $R_j > 4$ Å as well as MS-paths.

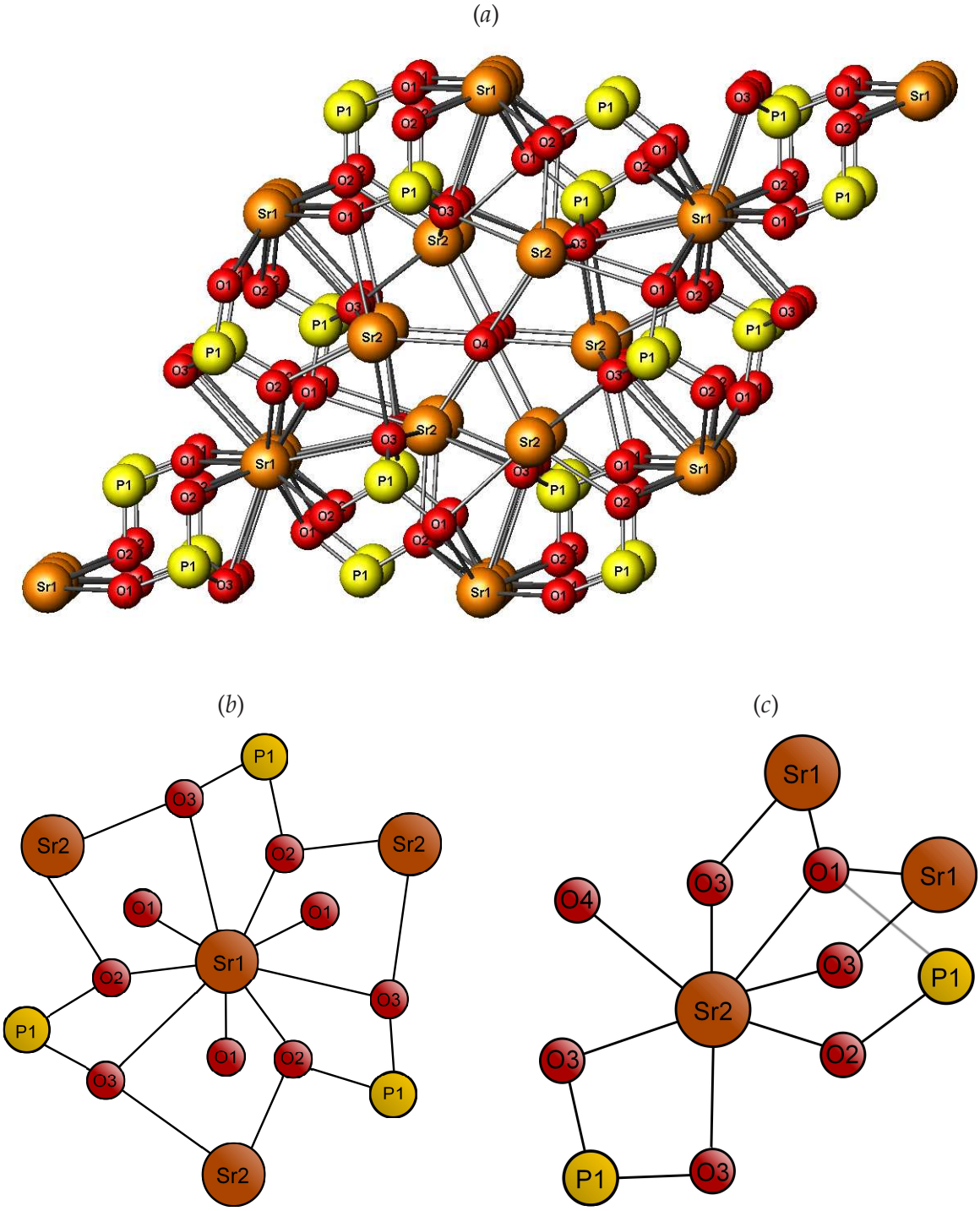


Figure C.5: The structure of Sr hydroxyapatite. (*a*) Overview over the arrangement of atoms, viewed along the crystallographic *c*-axis. (*b*) Schematical representations of the nearest neighbours at the Sr(I) site and (*c*) the Sr(II) site, which occupy 40 % and 60 % of the Sr sites, respectively. The Sr atoms at the Sr(I) and Sr(II) sites are coordinated to nine and seven oxygen atoms, respectively.

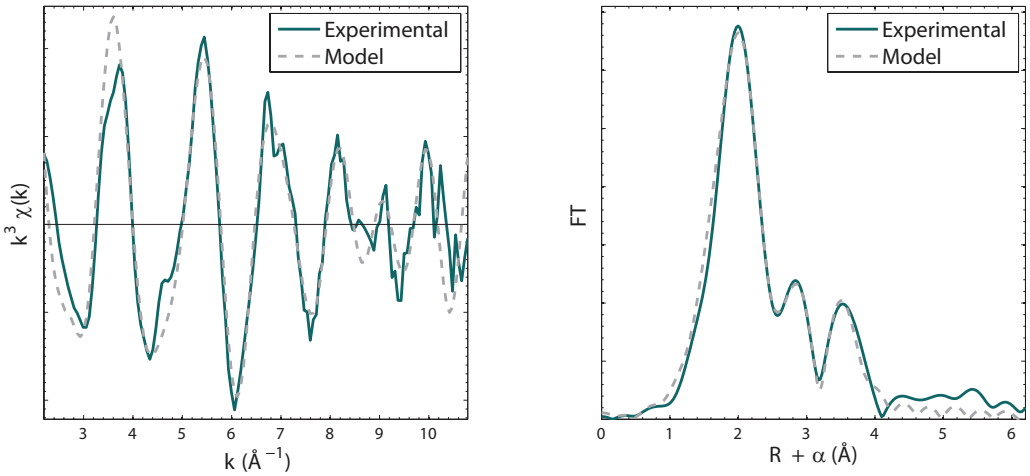


Figure C.6: Left: The background subtracted EXAFS $\chi(k)$ as a function of the photoelectronic wave number k of Sr hydroxyapatite. The function is k^3-weighted in order to enhance the details at high k-values. Right: The modulus of the Fourier transform of $\chi(k)$ which gives the radial distribution of electron density around an average Sr. The R-values are not absolute due to a phase shift α.

C.3.4 5 % Strontium-doped calcium hydroxyapatite

As mentioned in Chapter 9 strontium may have a preference to either of the two crystallographic sites (Bigi *et al.*, 2007; Terra *et al.*, 2009; O´Donnell *et al.*, 2008), depending on the doping level. From Rietveld refinements of X-ray powder diffraction patterns, see Chapter 11 no significant preference for substitution is seen, why a model, assuming Sr to be randomly distributed at both sites respecting the different weighting of Ca sites, has been used. Due to the relatively low doping level (5 %) it is assumed that no Sr atoms have been substituted at adjacent Ca sites. If the Sr-doping is fully randomised the assumption seems meaningful as the most probable Sr–Sr distance will be in the range of 10 Å, which is beyond the limit of what can be detected by EXAFS.

Both CaHA and SrHA form hexagonal crystals belonging to the same crystallographic space group, $P\ 6_3/m$ with a similar structural arrangement. Due to the large radius of a Sr^{2+} ion compared to Ca^{2+} shorter unit cell parameters a and c are observed in CaHA. A linear relationship has been reported between the unit cell parameters and the doping level (Bigi *et al.*, 2007; O´Donnell *et al.*, 2008). With a doping level of 5 % the unit cell parameters could be estimated to $a = 9.43$ Å and $c = 6.88$ Å compared to $a = 9.41$ Å and $c = 6.86$ Å for pure CaHA.

A shell model for $R_j < 6.3$ Å for an average Sr atom in the structure was suggested from the X-ray structure, Table C.4. Distances and coordination numbers for the shell model were calculated based on the CaHA crystal structure (Kay *et al.*, 1964) using the

expanded unit cell parameters. Only the calcium atom at the absorbing atom site were substituted by strontium, thereby calcium occupies all neighbouring M^{2+} sites and the EXAFS contributions from these sites are exclusively modelled by Sr–Ca paths.

Table C.4: Average Sr environment in 5 % Sr-doped CaHA assuming that all neighbouring M^{2+} sites are occupied by calcium. Distances up to $R_j = 6.3$ Å away for the absorbing atom have been calculated from the CaHA structure by Kay *et al.* (1964) and determined by EXAFS.

Atom type	N_j	R_j(Å) (X-ray)	R_j(Å) (EXAFS)	σ_j^2(Å^2) (EXAFS)
Oxygen	7.8	2.50[a]	2.495(2)	0.012(1)
Phosphorus	2.4	3.20	3.233(5)	0.006(2)
Oxygen	1.8	3.42	– [b]	–
Calcium	0.8	3.44	3.46(1)	0.012(1)
Phosphorus	3.0	3.60	3.65(1)	0.008(2)
Oxygen	6.0	3.99	– [b]	–
Calcium	8.4	4.06	4.09(1)	0.014(1)
Oxygen	4.2	4.46	4.54(2)	0.004(1)
Oxygen	6.6	4.81	4.93(2)	0.010(4)
Calcium	2.4	5.84	5.80(2)	0.004(1)
Calcium	6.0	6.30	6.30(2)	0.009(1)

[a] Weighted average value of 4.8 oxygen atoms at 2.40 Å and 1.2 oxygen atoms at 2.51 and 1.8 oxygen atoms at 2.76 Å. Experimental data are best modelled by a single shell at 2.50 Å with $N_j = 7.8$, resulting in a higher Debye-Waller factor.

[b] The oxygen shells at 3.42 Å and 3.99 Å could not be modelled separately, however low Debye-Waller factors in the nearby regions indicate lacking electron density.

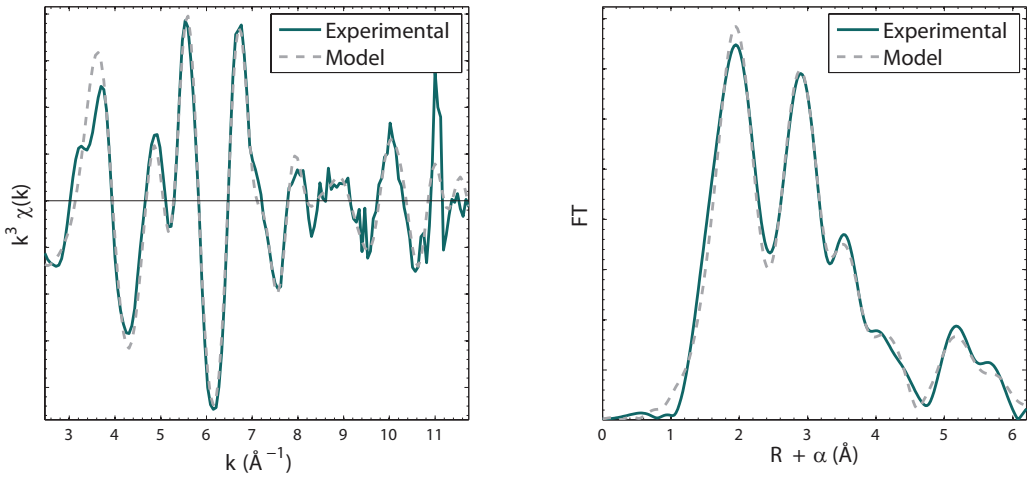


Figure C.7: Left: The background subtracted EXAFS $\chi(k)$ as a function of the photoelectronic wave number k of 5 % Sr-doped CaHA. The function is k^3-weighted in order to enhance the details at high k-values. Right: The modulus of the Fourier transform of $\chi(k)$ which gives the radial distribution of electron density around an average Sr. The R-values are not absolute due to a phase shift α.

The shell model was refined and the distances and Debye Waller factors were determined and presented in Table C.4. σ_0^2 was refined to 0.81 by keeping the distances constant. The calculated EXAFS and the Fourier transform of the model are compared with the experimentally obtained and shown in Figure C.7. Inconsistencies between the data and the model observed in the low k-region may be explained from missing inclusion of MS-paths. Comparing the results with that of pure SrHA a larger k-space (2.5–11.0 Å) is seen because data were collected over a larger range and the shells are thus better resolved. The distances of the shells are different in the two cases although the general arrangement of atoms in the structure are similar, but due to the larger unit cell parameters in pure SrHA the distances are changed. This is found as the explanation for why the oxygen distances between 2.4 and 2.8 Å are grouped into the same shell, whereas the inner O-shell could be resolved in two shells for pure SrHA.

C.4 Conclusions

EXAFS spectra have been collected on four Sr compounds, all with Sr coordination to oxygen. A typical Sr–O bond distances has been determined to 2.5–2.7 Å and the Debye-Waller factors generally fall within a range of $\sigma^2 = 0.005 - 0.025$ Å depending on the degree of static disorder. Amplitude reduction factors, S_0^2 were determined to be around 0.8–0.9, except for $SrCl_2 \cdot 6H_2O$, where it was 0.71.

Appendix D

EXAFS analysis of Sr in bone

This appendix describes a pre-study of strontium location in bone samples with EXAFS performed at room temperature.

D.1 Experimental

The bone samples originated from three different clinical tests: femur from dogs treated with Sr malonate for 4 weeks, calvariae from dogs treated with Sr malonate for 52 weeks (see Chapter 11) and spine from rats treated with Sr malonate for 26 weeks. Each series included three dose levels giving a total number of nine samples. The amounts of strontium have been determined by ICP-MS by Anders C. Raffalt and reported elsewhere, (Raffalt, 2011).

Bone samples from clinical tests with strontium containing medicaments were prepared as tablets. A representative fragment of a bone was crushed to a powder and pressed to a tablet and mounted on the XAS sample holder using kapton tape. Data collection and data treatment was analogue to the procedure described in Chapter C.

D.2 Results and discussion

D.2.1 Qualitative comparison of EXAFS spectra

The background subtracted and k^3-weighted EXAFS spectra and the radial electron density distributions are respectively shown in Figure D.1 for dog femur (*a*) and (*d*), dog cal-

varia (*b*) and (*e*), and for rat spine (*c*) and (*f*). Each test group consists of three samples which have been treated with Sr-medicament in three different dose levels: Highest dose (red), medium dose (orange) and lowest dose (yellow). All spectra are compared with the data from 5 % Sr-doped CaHA (grey spectra in Figure D.1).

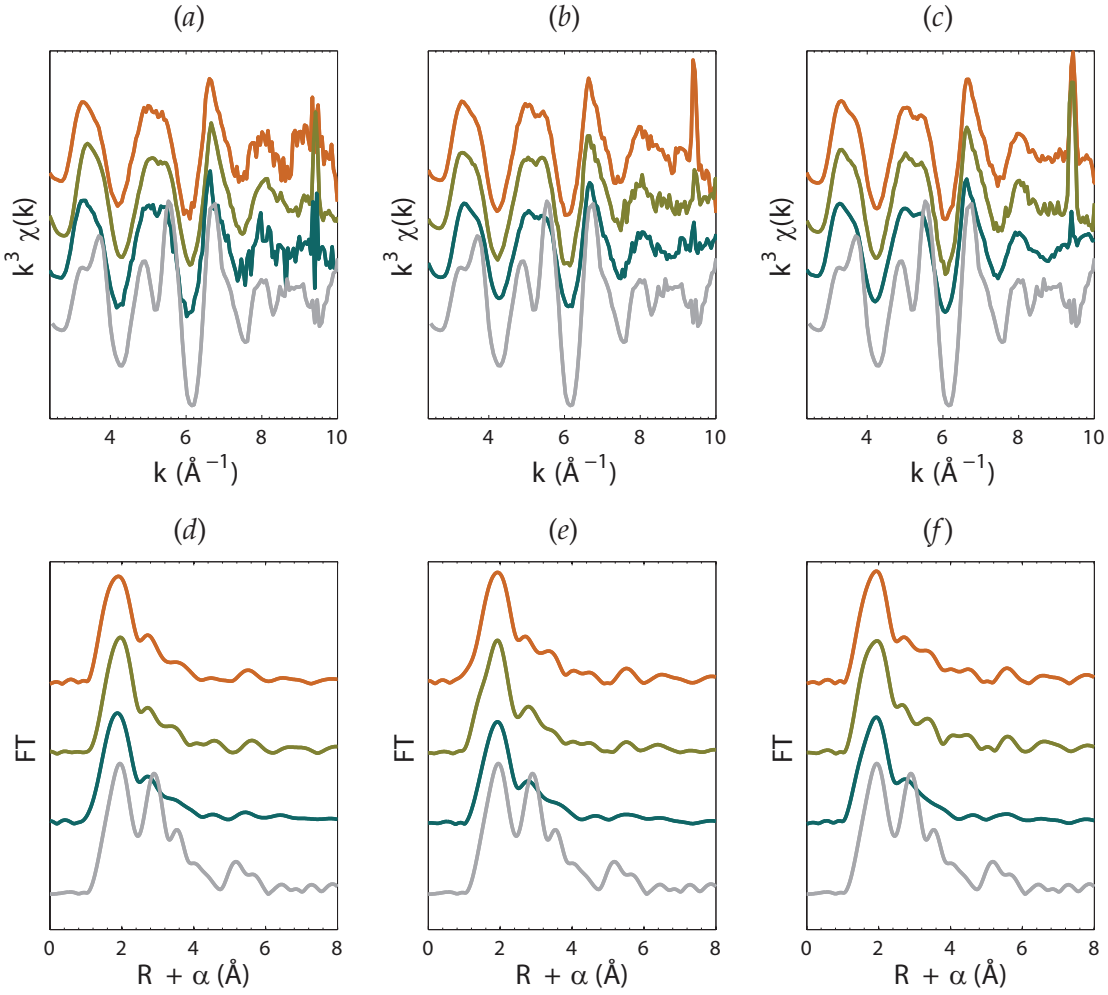


Figure D.1: (*a*)–(*c*) The background subtracted and k^3-weighted EXAFS function, $\chi(k)$, for (*a*) dog femur, (*b*) dog calvaria and (*c*) rat spine. (*d*)–(*f*) The radial electron density distribution given as the modulus of the Fourier transform of $\chi(k)$, for (*d*) dog femur, (*e*) dog calvaria and (*f*) rat spine. The R-values are not absolute due to the phase shift α. Each series consists of three samples with different dose level: Highest dose (orange), medium dose (olive) and lowest dose (blue). The EXAFS function and the radial electron density function for 5 % Sr-doped CaHA are plotted in grey for comparison. Glitches are observed at 9.8 Å^{-1}.

A qualitative comparison of all the spectra reveal a high similarity among the bone samples. The local environment for Sr ions is thereby be expected to be quite similar. Comparing the radial electron distribution functions for the bone samples with the one for the 5 % Sr-doped CaHA, Figure D.1(*d*)–(*f*), it is observed that the coordination shells occur at

similar distances R_j from the absorbing Sr-atom, which gives an evidence for the incorporation of strontium ions in the hydroxyapatite structure by substitution of the calcium. From the relative heights of the peaks it seen that strontium is not exclusively present as hydroxyapatite, but may coordinate to other material with distances around 2.5 Å, e.g. oxygen atoms in collagen and/or water.

A quantitative EXAFS analysis were tried by a linear combination fit in analogy to that described in Chapter 11. However no sample containing a disordered form of strontium (e.g. Sr in solution or Sr absorbed in collagen fibres) was included in this RT-study. Performing a linear combination analysis using the spectra of Sr-doped CaHA collected at RT and the collagen and strontium solution collected at 100 K as reference samples result in bad fits ($R > 10$ %), as the distances to the first coordination shell is different at the two temperatures. The periodical difference of the main oscillation, which originates from the first coordination shell, is directly observed when EXAFS spectra collected at respectively RT and 100 K are compared.

D.2.2 EXAFS modelling using the shell approach

A shell model was fitted to the EXAFS spectrum from dog calvaria (highest dose) using the EXAFS model from Sr-doped CaHA presented in Appendix C as an initial model. The shell model was refined and the distances and Debye-Waller factors were determined and presented in Table D.1. The calculated EXAFS and the Fourier transform of the model are compared with the experimentally obtained and shown in Figure D.2.

Table D.1: Strontium location in dog calvaria (high dose) $R_j < 5.0$ Å as determined by fitting a model to the EXAFS data using the shell approach. The model has been compared with the EXAFS results for 5 % Sr-doped CaHA.

		Dog calvaria		5 % Sr-doped CaHA	
Atom type	N_j	R_j(Å)	σ_j^2(Å^2)	R_j(Å)	σ_j^2(Å^2)
Oxygen	7.8	2.525(2)	0.015(1)	2.495(2)	0.012(1)
Phosphorus	2.4	3.24(2)	0.010(1)	3.233(5)	0.006(2)
Calcium	0.8	– [a]	–	3.46(1)	0.012(1)
Phosphorus	3.0	3.57(4)	0.015(6)	3.65(1)	0.008(2)
Calcium	8.4	4.14(4)	0.019(5)	4.09(1)	0.014(1)
Oxygen	4.2	– [a]	–	4.54(2)	0.004(1)
Oxygen	6.6	4.89(5)	0.011(4)	4.93(2)	0.010(4)

[a] These shells could not be included in the model.

Most of the shells from the Sr-doped CaHA can be adopted in the bone model. However, higher uncertainties as well as higher Debye-Waller factors was observed in the bone model, indicating a higher disorder among strontium. This is in agreement with previously reported EXAFS studies on calcium in which its coordination in bone tissue and in ordered apatites has been compared (Harries *et al.*, 1988). However the chemical informa-

tion from the shell model is limited as it can only be concluded that strontium is disordered and not how it is disordered.

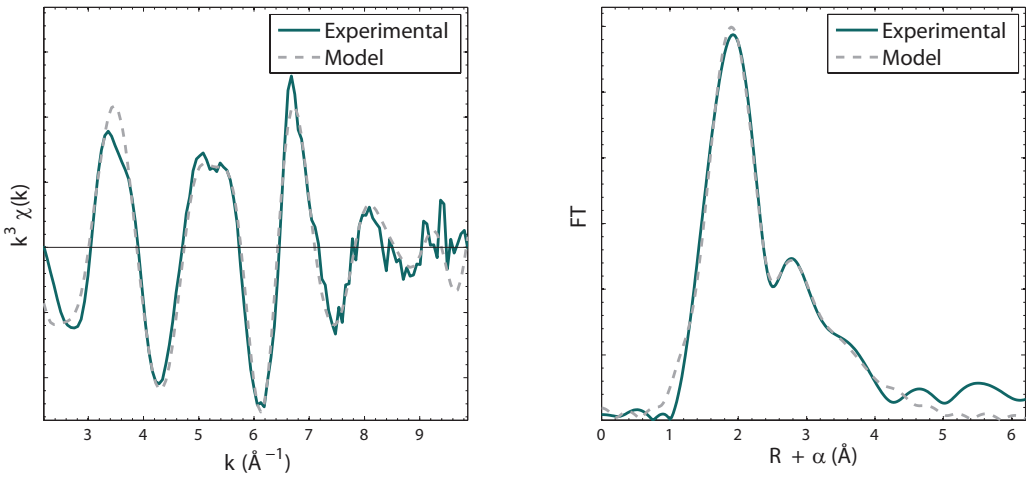


Figure D.2: Left: The background subtracted EXAFS $\chi(k)$ as a function of the photoelectronic wave number k of dog calvaria treated with Sr malonate for 52 weeks. The function is k^3-weighted in order to enhance the details at high k-values. Right: The modulus of the Fourier transform of $\chi(k)$ which gives the radial distribution of electron density around an average Sr. The R-values are not absolute due to a phase shift α.

D.3 Conclusions

XAS data were collected at room temperature to study the strontium location in bone samples were presented. The qualitative comparison of EXAFS spectra shows high similarity in strontium coordination in different bone samples. A comparison with Sr-doped apatite indicates that the strontium ions partly substitute some of the calcium atoms in the apatite structure and partly is present at highly disordered sites, in which it may coordinate to other biological material, probably collagen and water. Using the shell approach to fit a model for a bone sample to the EXAFS resulted in higher Debye-Waller factors, as compared to Sr-doped CaHA. Hence a higher disorder of strontium in the bone tissue can be concluded, but the model does not contain information about how strontium is disordered, which illustrates the weakness of the shell approach.

DTU Chemistry
Department of Chemistry