DEPARTMENT OF
CHEMISTRY

Ultra-high resolution structure determination of transition metal substituted human carbonic anhydrase 2 - inhibitor complexes

JOSÉ PEDRO MALANHO DA SILVA

Master in Biochemistry for Health

JOINT DOCTORATE IN:
CHEMISTRY AT NOVA SCHOOL OF SCIENCE & TECHNOLOGY
STRUCTURAL BIOLOGY, CYCLE XXXIII AT UNIVERSITÀ DEGLI STUDI DI FIRENZE
NOVA University Lisbon
May 2022

DEPARTMENT OF
CHEMISTRY

Ultra-high resolution structure determination of transition metal substituted human carbonic anhydrase 2 – inhibitor complexes

José Pedro Malanho da Silva

Master in Biochemistry for Health

Adviser: Maria dos Anjos Lopez Macedo, Auxiliar Professor, NOVA School of Science and Technology, Portugal
Claudio Luchinat,Full Professor, Università degli Studi Firenze, Italy

Co-advisers: Carlos Frederico de Gusmão Campos Geraldes, Full Professor, University of Coimbra, Portugal

Examination Committee:

Chair: Maria João Lobo de Reis Madeira Crispim Romão, Full Professor, NOVA School of Science and Technology, Portugal

Rapporteurs: Brian James Goodfellow, Assistant Professor, Chemistry Department da University of Aveiro
Mario Piccioli, Associate Professor, Università degli Studi di Firenze

Members: Claudio Luchinat, Full Professor, Università degli Studi di Firenze, Italy
Simone Ciofi Baffoni, Associate Professor, Università degli Studi di Firenze, Italy
Eurico José da Silva Cabrita, Associate Professor with Habilitation, NOVA School of Science and Technology, Portugal
Teresa Sacadura Santos Silva, Assistant Professor at NOVA School of Science and Technology, Portugal

JOINT DOCTORATE IN: CHEMISTRY at NOVA SCHOOL OF SCIENCE AND TECHNOLOGY and STRUCTURAL BIOLOGY, CYCLE XXXIII at UNIVERSITÀ DEGLI STUDI DI FIRENZE
NOVA University Lisbon
May, 2022

Ultra-high resolution structure determination of transition metal substituted human carbonic anhydrase 2 - inhibitor complexes

To my family and friends

Acknowledgments

Four years ago, I could not imagine what I was getting into. In these four years, my life turned around and gained new momentum and catapulted myself to learn, explore and love. To many I must acknowledge my thanks and my gratitude for the amazing journey.

To *Fundação para a Ciência e Tecnologia* for my PhD fellowship (PD/BD/135180/2017) and financial support integrated in the PhD Program in NMR applied to chemistry, materials, and biosciences (PD/00065/2013). To *NOVA School of Science and Technology* for all these years of education, teachers and staff that made my education possible.

To my supervisors, Professor Anjos Macedo, Professor Carlos Geraldes and Professor Claudio Luchinat, thank you for accepting me as your PhD student and giving me the opportunity to explore and learn more about the structural biology world. To professor Anjos Macedo, the supervisor with million times more energy than her student, thank you for the lengthy and interesting on-topic and off-topic discussions on the projects and life advice. Thank you for always pushing me to do and be better.

To Professor Carlos Geraldes, your passion about NMR, your lectures and determination will burn inside me to continue to be a better researcher. I truly have enjoyed your time together online.

To Professor Claudio Luchinat, thank you for the inspiration to continuing pursuing and for the amazing work discussion.

To Doctor Ravera, Professor Fragai and Professor Parigi, thank you for the integration into many interesting projects and work discussions.

Ai miei amici, Linda Cerofolini, Stefano Giuntini, Domenico Rizzo, Giovanni Bellomo, Mercia Sousa and Tommaso Martelli, grazie mille. Mi è piaciuto molto il tempo che abbiamo trascorso insieme in tutte le attività dentro e fuori dal lavoro. Sei voi il motivo per cui ho adorato il mio soggiorno al CERM. Linda, grazie per me insegnare, per aver dedicato tempo a rispondere a tutte le mie domande, sia di lavoro che di curiosità. Stefano, I tuoi milioni di progetti che lavori instancabilmente ogni giorno sono per me un'ispirazione per essere uno scienziato migliore. Domenico, mio hermanito da altri genitori, grazie per tutti i momenti divertenti e per tutto l'aiuto che mi hai dato. Grazie mille.

To *Macromolecular Crystallography Laboratory* group, many thanks for accepting me, teaching me, and the time spent with me during the periods that I was in Portugal. Special thanks to Doctor Marino Santos for all sorts of conversations both work related and life related; as promised you are included in these acknowledgments.

To my PTNMR colleagues, the Gang KONA, thank you very much for the amusing times during lectures and during the periods we were together.

To the unknown old lady walking in the roundabout with two giant shopping bags, although we did not speak, I must thank you for the impactful moment in my life.

Para Doutora Elin Moe e Doutora Célia Romão, tusen takk e muito obrigado. Graças ao vosso encorajamento para eu perseguir a minha paixão eu consegui atingir mais outro patamar da minha vida.

À minha família, obrigado por todo o suporte desde sempre, por todas as dores de cabeça, por todos os momentos bons na vida e pelo amor infinito. Vocês são a razão de eu querer ser sempre melhor.

Para a Aline Fernanda Nakata, futura mulher, digo que quando comecei esta etapa da vida tu não fazias parte dela. Nós conhecemo-nos na Itália quando todas as probabilidades pareciam ínfimas. Ficamos afastados e depois sempre juntos. Começamos uma vida juntos e exploramos Itália e um pouco de Portugal. Muito obrigado por todos os momentos grandes e pequenos, pelo teu apoio e pelo teu amor.

“Life is about balance. Be kind, but don’t let people abuse you. Trust, but don’t be deceived. Be content, but never stop improving yourself.”

(Zig Ziglar).

Abstract

Paramagnetic Nuclear Magnetic Resonance (NMR) is developing to aid the characterization of paramagnetic molecules, whose paramagnetic centers changes the spectroscopic proprieties of said molecules. These paramagnetic centers can be exploited to overcome some troublesome aspects of NMR, such as sensitivity, by increasing the number of experiments. To aid this development, we used human Carbonic Anhydrase II (hCAII), which is a model protein that contains zinc(II) in its active center, which is diamagnetic. hCAII is an enzyme capable of interconverting carbon dioxide to bicarbonate, making it one of the most important proteins in life. Several comprehensive studies have structurally and functionally characterized hCAII making it an excellent model protein. Furthermore, the metal ion in the active center can be substituted by other transition metals ions (see chapter 1), such as cobalt(II) (see chapter 3), nickel(II) (see chapter 4) and copper(II) (see chapter 5), which are paramagnetic that will help answering different problems described in this thesis.

The ion cobalt(II), explored in chapter 3, can induce considerable changes on the NMR observables and is useful to understand the interactions of ligands with proteins. For this we used cobalt(II)-hCAII and used NMR, Electron Paramagnetic Resonance (EPR) and X-ray crystallography to characterize the interaction of thiocyanate under high concentrations with the hCAII. The addition of 500 mM of sodium thiocyanate changes the dynamics of the protein without changing the protein structure.

Solid-state NMR (SSNMR) is another field of NMR where the methodological and practical aspects are currently under development to reach better sensitivity and resolution. We proposed the usage of nickel(II)-hCAII as a paramagnetic molecule to increase the amount of tools in SSNMR, which is explored in chapter 4. The nickel(II) ion is capable of breaking the dipolar bath by changing the frequency of the nuclei close to the paramagnetic center, thus increasing the resolution and sensitivity of the SSNMR experiments. Furthermore, in parallel we discovered that hCAII is capable of binding two nickel(II) ions, one in the active center, as expected and described in literature, and one in the N-terminal site of the protein, a novel discovery.

The description of the paramagnetic effects, such as the Pseudocontact Shifts (PCS), in the NMR observables have been subjected to debate, where different treatments of theoretical equations were clashing. The experimental proof to determine which equation holds true is fully described in chapter 5. For this, we developed copper(II)-hCAII to acquire NMR and EPR data under the same conditions, and determine which equations describe better the PCS. The data interpretation from different techniques (both NMR and EPR) led us to conclude that the original treatment from Kurland and McGarvey equation is the correct one.

Keywords: NMR, Paramagnetism, Protein Characterization

Resumo

A Ressonância Magnética Nuclear Paramagnética (RMN) tornou-se uma importante ferramenta para complementar as técnicas clássicas de RMN em estudos estruturais de metaloproteínas. Os seus centros paramagnéticos intrínsecos ou extrínsecos perturbam suas propriedades espectroscópicas de RMN, incluindo desvios químicos, acoplamentos dipolares residuais (RDCs) e tempos de relaxação nuclear. Os desvios de pseudocontato (PCS), as mudanças de RDC e as contribuições paramagnéticas nas taxas de relaxamento nuclear, induzidas por esses centros, podem ser exploradas para melhorar a qualidade das estruturas de proteínas derivadas de RMN, aumentando o número e o tipo de restrições estruturais que podem ser usadas.

Contribuindo para este desenvolvimento, foi utilizada a Anidrase Carbónica II (hCAII) humana, como uma proteína modelo que contém um ião zinco(II) diamagnético no centro ativo. hCAII catalisa a interconversão de dióxido de carbono em bicarbonato, tornando-se uma das proteínas mais importantes da vida. Vários estudos caracterizaram estruturalmente e funcionalmente a hCAII, tornando-a uma excelente proteína modelo. Além disso, o ião metálico no centro ativo pode ser substituído por outros iões de metais de transição paramagnéticos, conforme apresentado no capítulo 1, como cobalto(II) (ver capítulo 3), níquel(II) (ver capítulo 4) e cobre(II) (ver capítulo 5), cujas contribuições paramagnéticas para os parâmetros de RMN (apresentadas no capítulo 2) podem ajudar a responder aos diferentes problemas descritos nesta tese.

No capítulo 3, cobalto(II)-hCAII foi usada para caracterizar as interações de hCAII com altas concentrações de tiocianato. Ao combinar as alterações induzidas por paramagnéticos nas observáveis de RMN, dados de Ressonância Paramagnética Eletrônica (EPR) e cristalografia de raios-X, concluiu-se que a adição de 500 mM de tiocianato de sódio altera a dinâmica da proteína sem alterar a estrutura da proteína.

A RMN de estado sólido (SSNMR) é outro campo da RMN onde os aspetos metodológicos e práticos estão atualmente em desenvolvimento para alcançar melhor sensibilidade e resolução. No capítulo 4, níquel(II)-hCAII é explorada como uma molécula paramagnética para aumentar a quantidade de ferramentas disponíveis em SSNMR. O ião níquel(II) é capaz de quebrar o banho dipolar alterando a frequência dos núcleos próximos ao centro paramagnético, aumentando assim a resolução e a sensibilidade dos experimentos SSNMR. Além disso, em paralelo, descobriu-se que hCAII é capaz de se ligar a dois iões níquel(II), um no centro ativo, conforme esperado e descrito na literatura, e um novo na sequência N-terminal da proteína.

A descrição teórica dos efeitos paramagnéticos, como os PCS, nas observáveis de RMN tem sido objeto de debate, onde diferentes tratamentos foram conflitantes, fornecendo diferentes equações teóricas. A prova experimental para determinar qual equação é correta está descrita de forma completa no capítulo 5. Para isso, a cobre(II)-hCAII foi preparada e usada para adquirir dados de RMN e EPR nas mesmas condições, e determinou-se qual equação descreve melhor os PCS experimentais. A interpretação dos dados de diferentes técnicas (tanto RMN quanto EPR) levou-nos a concluir que o tratamento clássico da equação de Kurland e McGarvey é o correto e não aquele derivado da mecânica quântica.

Palavras chave: NMR, Paramagnetismo, Caracterização de proteínas

Riassunto

La risonanza magnetica nucleare paramagnetica (NMR) è diventata uno strumento importante per integrare le tecniche classiche NMR negli studi strutturali delle metalloproteine. I centri paramagnetici intrinseci o estrinseci delle metalloproteine perturbano le loro proprietà spettroscopiche NMR, includendo la variazione del valore di chemical shift (PseudoContact Shifts), gli accoppiamenti dipolari residui (RDC) e i tempi di rilassamento nucleare. Pseudocontact Shifts (PCS), RDC e contributi paramagnetici alle velocità di rilassamento nucleare, indotti da tali centri, possono essere sfruttati per migliorare la qualità delle strutture delle proteine ottenute tramite NMR, aumentando il numero e il tipo di vincoli strutturali che possono essere utilizzati.

Per contribuire a questo sviluppo, è stata utilizzata l'anidrasi carbonica umana II (hCAII), come proteina modello, poiché questa contiene uno ione diamagnetico zinco (II) nel suo sito attivo. hCAII catalizza l'interconversione dell'anidride carbonica in bicarbonato, e questo la rende una delle proteine più importanti della vita. Diversi studi multidisciplinari hanno permesso di caratterizzare la struttura e la funzione della hCAII rendendola un'eccellente proteina modello. Inoltre, lo ione metallico nel sito attivo può essere sostituito da ioni di metalli di transizione paramagnetici, come introdotto nel capitolo 1, come il cobalto(II) (vedi capitolo 3), il nichel(II) (vedi capitolo 4) e il rame(II) (vedi capitolo 5), i cui contributi paramagnetici ai parametri NMR (introdotti nel capitolo 2) possono aiutare a rispondere ai diversi problemi descritti in questa tesi.

Nel capitolo 3, la cobalto(II)-hCAII è stata utilizzata per caratterizzare le interazioni della hCAII con alte concentrazioni di tiocianato. Combinando i cambiamenti indotti dallo ione paramagnetico sugli osservabili NMR, i dati di risonanza paramagnetica elettronica (EPR) e la cristallografia a raggi X, si è concluso che l'aggiunta di 500 mM di tiocianato di sodio modifica la dinamica della proteina senza modificare la struttura della proteina.

L'NMR allo stato solido (SSNMR) è un altro campo dell'NMR in cui gli aspetti metodologici e pratici sono attualmente in fase di sviluppo per raggiungere una migliore sensibilità e risoluzione. Nel capitolo 4, la nichel(II)-hCAII viene esplorata come proteina paramagnetica per aumentare la quantità di strumenti disponibili in SSNMR. Lo ione nichel(II) è in grado di modificare la frequenza di risonanza dei nuclei vicini al centro paramagnetico e allo stesso tempo di rompere le interazioni dipolari tra protoni aumentando così la risoluzione e la sensibilità degli esperimenti SSNMR. Parallelamente, inoltre, si è scoperto che hCAII è in grado di legare due ioni nichel(II), uno nel sito attivo, come previsto e descritto in letteratura, ed un secondo metallo nella regione N-terminale della proteina.

La descrizione teorica degli effetti paramagnetici, come il PCS, è ancora oggetto di dibattito, e diverse equazioni teoriche contrastanti sono state proposte per spiegare gli osservabili NMR . La prova sperimentale per determinare quale equazione descrive correttamente i vincoli sperimentali è riportata nel capitolo 5. Per questo obiettivo, la hCAII con il rame(II) nel sito attivo è stata preparata e utilizzata per acquisire dati sperimentali NMR ed EPR nelle stesse condizioni in modo da determinare quale equazione descrive meglio i dati sperimentali PCS. L'interpretazione dei dati con diverse tecniche (sia NMR che EPR) ci ha portato a concludere che il trattamento classico dell'equazione di Kurland e

McGarvey è quello corretto, piuttosto che quello derivante dal trattamento proposto dalla meccanica quantistica.

Parole chiave: NMR, Paramagnetismo, Caratterizzazione di proteine

Table of Contents

List of Figures

List of Tables

List of Abbreviations

bCA	bovine Carbonic Anhydrase
CA	Carbonic Anhydrase
CD	Circular Dichroism
CNS	Central nervous system
CP	Cross Polarization
CSA	Chemical Shift Anisotropy
CSP	Chemical Shift Perturbation
DM-hCA	Double Mutant human Carbonic Anhydrase
EPR	Electron Paramagnetic Resonance
FID	Free Induction Decay
GI	Gastrointestinal
hCA	human Carbonic Anhydrase
HSQC	Heteronuclear Single Quantum Coherence
IPAP	In-Phase Anti-Phase
MAS	Magic Angle Spinning
MR	Molecular Replacement
NMR	Nuclear Magnetic Resonance
PCS	Pseudocontact Shifts
PDA	Point-Dipole Approximation
pnMR	Paramagnetic Nuclear Magnetic Resonance
PRE	Paramagnetic Relaxation Enhancement
PTS	p-toluenesulfonamide
QC	Quantum Chemistry
R_1	Longitudinal Relaxation rate
R_2	Transverse Relaxation Rate
RBC	Red blood cell
RDC	Residual Dipolar Coupling
SAD	Single-wavelength Anomalous Dispersion
SE	Semi Empirical

SSNMR	Solid State Nuclear Magnetic Resonance
TM-hCA	Triple Mutant human Carbonic Anhydrase
WT-hCA	Wild Type human Carbonic Anhydrase
ZFS	Zero Field Splitting

1

Chapter 1 - Introduction

1.1 Carbonic Anhydrases

Carbonic anhydrases (CAs) are zinc(II) metalloenzymes present in almost all living organisms. These enzymes catalyze the interconversion between carbon dioxide and bicarbonate. The zinc(II) ion in the catalytic site acts as a Lewis acid by decreasing the pK_a value of a coordinated water molecule from 14 to 7, which becomes ionized to hydroxyl at physiological pH[1]. The coordinated hydroxide then performs a nucleophilic attack on the substrate CO_2, generating a bicarbonate ion. After the reaction, the active site is regenerated through the addition of a new water molecule and the removal of one of its protons by a buffer-aided transfer to the bulk solvent via proton shuttle residues[1–4]. The mechanism of action of CA can be seen in **Figure 1.1**.

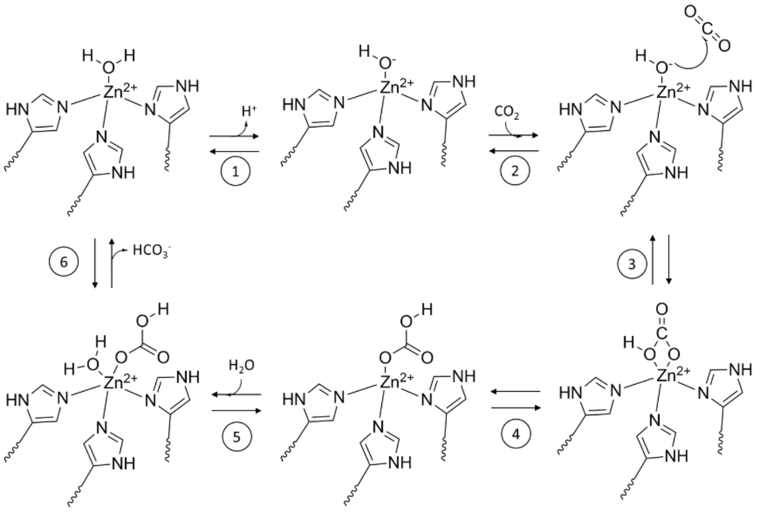


Figure 1.1 Mechanism of action of carbonic anhydrases.[5]

1.1.1 Human Carbonic Anhydrases

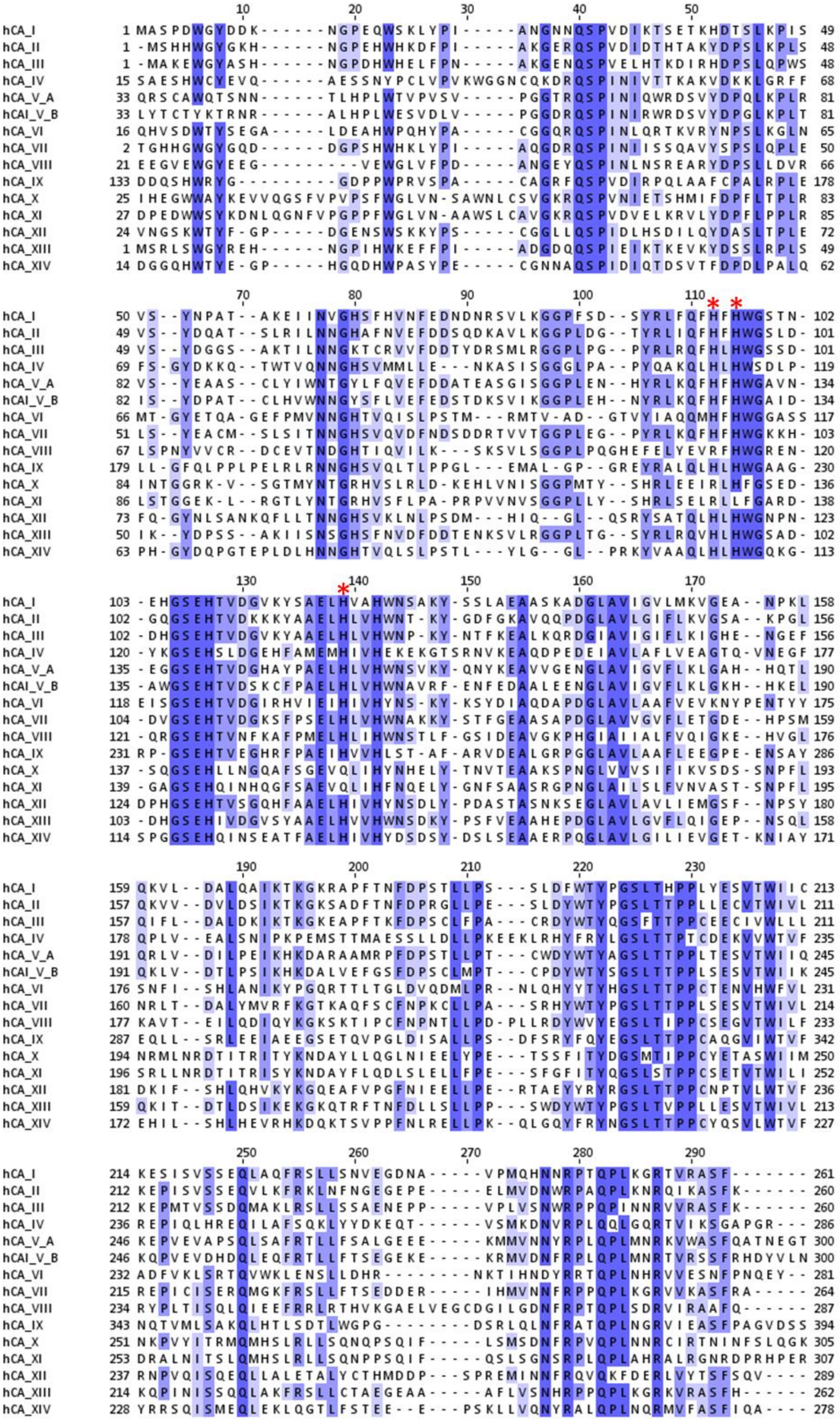
In mammals, the enzyme is present in several tissue-dependent isoforms. There are 15 isoforms of human Carbonic Anhydrases (hCA), out of the 16 mammalian isoforms. The most abundant and the one that has the highest-activity is the second isoform, which is predominantly present in blood[6]. The sequence alignment for the different isoforms of hCA (**Figure 1.2**) shows that their primary sequences are somewhat different, although the isoforms are structurally quite similar (**Figure 1.3**).

Table 1.1 shows the different isoforms of human CA, their type and their localization in the human being. The CAs are present in red blood cells (RBC), gastrointestinal (GI) tract, kidneys, lungs, testes, brain, skeletal muscles, adipocytes, pancreas, capillaries, colon, heart muscles, liver, central nervous system (CNS), GI mucosa, reproductive tract and some are overexpressed in tumors.

Table 1.1: Human carbonic anhydrase isoforms and their classification.[7–9]

Isoform	**Type of isoform**	**Localization**
hCA I	Cytosolic form	RBC, GI tract
hCA II	Cytosolic form	RBC, GI tract, eyes, osteoclasts, kidneys, lungs, testes, brain
hCA III	Cytosolic form	Skeletal muscles, adipocytes
hCA IV	Membrane associated	Kidneys, lungs, pancreas, brain, capillaries, colon, heart muscles
hCA V - A	Mitochondrial form	Liver
hCA V - B	Mitochondrial form	Heart and skeletal muscles, pancreas
hCA VI	Secreted isozyme	Salivary and mammary glands
hCA VII	Cytosolic form	CNS
hCA VIII	“Acatalytic” isoform	CNS
hCA IX	Membrane associated	Tumors, GI mucosa
hCA X	“Acatalytic” isoform	CNS
hCA XI	“Acatalytic” isoform	CNS
hCA XII	Membrane associated	Intestinal, reproductive, epithelia, eyes, tumors
hCA XIII	Cytosolic form	Kidneys, brain, lungs, gut, reproductive tract
hCA XIV	Membrane associated	Kidneys, brain, liver

Inside the RBC, the number of molecules of hCA can reach almost 26 million per RBC[10], where almost 19.5 million molecules are hCAI, 6.3 million are hCAII and the remaining are hCAIII. Assuming that each RBC has a volume of 90 fL, then the total concentration of hCA inside one RBC is around 480 µM. This incredible amount is only surpassed by Hemoglobin, where its subunits can reach up to 200 million copies per RBC[10].



```
                       10        20        30        40        50
hCA_I       1 MASPDWGYDDK------NGPEQWSKLYPI-----ANGNNQSPVDIKTSETKHDTSLKPIS 49
hCA_II      1 -MSHHWGYGKH------NGPEHWHKDFPI-----AKGERQSPVDIDTHTAKYDPSLKPLS 48
hCA_III     1 -MAKEWGYASH------NGPDHWHELFPN-----AKGENQSPVELHTKDIRHDPSLQPWS 48
hCA_IV     15 SAESHWCYEVQ------AESSNYPCLVPVKWGGNCQKDRQSPINIVTTKAKVDKKLGRFF 68
hCA_V_A    33 QRSCAWQTSNN------TLHPLWTVPVSV-----PGGTRQSPINIQWRDSVYDPQLKPLR 81
hCAI_V_B   33 LYTCTYKTRNR------ALHPLWESVDLV-----PGGDRQSPINIRWRDSVYDPGLKPLT 81
hCA_VI     16 QHVSDWTYSEGA-----LDEAHWPQHYPA-----CGGQRQSPINLQRTKVRYNPSLKGLN 65
hCA_VII     2 TGHHGWGYGQD------DGPSHWHKLYPI-----AQGDRQSPINIISSQAVYSPSLQPLE 50
hCA_VIII   21 EEGVEWGYEEG---------VEWGLVFPD-----ANGEYQSPINLNSREARYDPSLLDVR 66
hCA_IX    133 DDQSHWRYG---------GDPPWPRVSPA-----CAGRFQSPVDIRPQLAAFCPALRPLE 178
hCA_X      25 IHEGWWAYKEVVQGSFVPVPSFWGLVN-SAWNLCSVGKRQSPVNIETSHMIFDPFLTPLR 83
hCA_XI     27 DPEDWWSYKDNLQGNFVPGPPFWGLVN-AAWSLCAVGKRQSPVDVELKRVLYDPFLPPLR 85
hCA_XII    24 VNGSKWTYF-GP-----DGENSWSKKYPS-----CGGLLQSPIDLHSDILQYDASLTPLE 72
hCA_XIII    1 MSRLSWGYREH------NGPIHWKEFFPI-----ADGDQQSPIEIKTKEVKYDSSLRPLS 49
hCA_XIV    14 DGGQHWTYE-GP-----HGQDHWPASYPE-----CGNNAQSPIDIQTDSVTFDPDLPALQ 62

                                                                  * *
                       70        80        90        100       110
hCA_I      50 VS--YNPAT--AKEIINVGHSFHVNFEDNDNRSVLKGGPFSD--SYRLFQFHFHWGSTN- 102
hCA_II     49 VS--YDQAT--SLRILNNGHAFNVEFDDSQDKAVLKGGPLDG--TYRLIQFHFHWGSLD- 101
hCA_III    49 VS--YDGGS--AKTILNNGKTCRVVFDDTYDRSMLRGGPLPG--PYRLRQFHLHWGSSD- 101
hCA_IV     69 FS-GYDKKQ--TWTVQNNGHSVMMLLE---NKASISGGGLPA--PYQAKQLHLHWSDLP- 119
hCA_V_A    82 VS--YEAAS--CLYIWNTGYLFQVEFDDATEASGISGGPLEN--HYRLKQFHFHWGAVN- 134
hCAI_V_B   82 IS--YDPAT--CLHVWNNGYSFLVEFEDSTDKSVIKGGPLEH--NYRLKQFHFHWGAID- 134
hCA_VI     66 MT-GYETQA-GEFPMVNNGHTVQISLPSTM---RMTV-AD--GTVYIAQQMHFHWGGASS 117
hCA_VII    51 LS--YEACM--SLSITNNGHSVQVDFNDSDDRTVVTGGPLEG--PYRLKQFHFHWGKKH- 103
hCA_VIII   67 LSPNYVVCR--DCEVTNDGHTIQVILK---SKSVLSGGPLPQGHEFELYEVRFHWGREN- 120
hCA_IX    179 LL-GFQLPPLPELRLRNNGHSVQLTLPPGL---EMAL-GP--GREYRALQLHLHWGAAG- 230
hCA_X      84 INTGGRK-V--SGTMYNTGRHVSLRLD-KEHLVNISGGPMTY--SHRLEEIRLHFGSED- 136
hCA_XI     86 LSTGGEK-L--RGTLYNTGRHVSFLPA-PRPVVNVSGGPLLY--SHRLSELRLLFGARD- 138
hCA_XII    73 FQ-GYNLSANKQFLLTNNGHSVKLNLPSDM---HIQ--GL--QSRYSATQLHLHWGNPN- 123
hCA_XIII   50 IK--YDPSS--AKIISNSGHSFNVDFDDTENKSVLRGGPLTG--SYRLRQVHLHWGSAD- 102
hCA_XIV    63 PH-GYDQPGTEPLDLHNNGHTVQLSLPSTL---YLG--GL--PRKYVAAQLHLHWGQKG- 113

                                *
                       130       140       150       160       170
hCA_I     103 -EHGSEHTVDGVKYSAELHVAHWNSAKY-SSLAEAASKADGLAVIGVLMKVGEA--NPKL 158
hCA_II    102 -GQGSEHTVDKKKYAAELHLVHWNT-KY-GDFGKAVQQPDGLAVLGIFLKVGSA--KPGL 156
hCA_III   102 -DHGSEHTVDGVKYAAELHLVHWNP-KY-NTFKEALKQRDGIAVIGIFLKIGHE--NGEF 156
hCA_IV    120 -YKGSEHSLDGEHFAMEMHIVHEKEKGTSRNVKEAQDPEDEIAVLAFLVEAGTQ-VNEGF 177
hCA_V_A   135 -EGGSEHTVDGHAYPAELHLVHWNSVKY-QNYKEAVVGENGLAVIGVFLKLGAH--HQTL 190
hCAI_V_B  135 -AWGSEHTVDSKCFPAELHLVHWNAVRF-ENFEDAALEENGLAVIGVFLKLGKH--HKEL 190
hCA_VI    118 EISGSEHTVDGIRHVIEIHIVHYNS-KY-KSYDIAQDAPDGLAVLAAFVEVKNYPENTYY 175
hCA_VII   104 -DVGSEHTVDGKSFPSELHLVHWNAKKY-STFGEAASAPDGLAVVGVFLETGDE--HPSM 159
hCA_VIII  121 -QRGSEHTVNFKAFPMELHLIHWNSTLF-GSIDEAVGKPHGIAIIALFVQIGKE--HVGL 176
hCA_IX    231 RP-GSEHTVEGHRFPAEIHVVHLST-AF-ARVDEALGRPGGLAVLAAFLEEGPE-ENSAY 286
hCA_X     137 -SQGSEHLLNGQAFSGEVQLIHYNHELY-TNVTEAAKSPNGLVVVSIFIKVSDS-SNPFL 193
hCA_XI    139 -GAGSEHQINHQGFSAEVQLIHFNQELY-GNFSAASRGPNGLAILSLFVNVAST-SNPFL 195
hCA_XII   124 DPHGSEHTVSGQHFAAELHIVHYNSDLY-PDASTASNKSEGLAVLAVLIEMGSF--NPSY 180
hCA_XIII  103 -DHGSEHIVDGVSYAAELHVVHWNSDKY-PSFVEAAHEPDGLAVLGVFLQIGEP--NSQL 158
hCA_XIV   114 SPGGSEHQINSEATFAELHIVHYDSDSY-DSLSEAAERPQGLAVLGILIEVGET-KNIAY 171

                       190       200       210       220       230
hCA_I     159 QKVL--DALQAIKTKGKRAPFTNFDPSTLLPS---SLDFWTYPGSLTHPPLYESVTWIIC 213
hCA_II    157 QKVV--DVLDSIKTKGKSADFTNFDPRGLLPE---SLDYWTYPGSLTTPPLLECVTWIVL 211
hCA_III   157 QIFL--DALDKIKTKGKEAPFTKFDPSCLFPA---CRDYWTYQGSFTTPPCEECIVWLLL 211
hCA_IV    178 QPLV--EALSNIPKPEMSTTMAESSLLDLLPKEEKLRHYFRYLGSLTTPTCDEKVVWTVF 235
hCA_V_A   191 QRLV--DILPEIKHKDARAAMRPFDPSTLLPT---CWDYWTYAGSLTTPPLTESVTWIIQ 245
hCAI_V_B  191 QKLV--DTLPSIKHKDALVEFGSFDPSCLMPT---CPDYWTYSGSLTTPPLSESVTWIIK 245
hCA_VI    176 SNFI--SHLANIKYPGQRTTLTGLDVQDMLPR--NLQHYYTYHGSLTTPPCTENVHWFVL 231
hCA_VII   160 NRLT--DALYMVRFKGTKAQFSCFNPKCLLPA---SRHYWTYPGSLTTPPLSESVTWIVL 214
hCA_VIII  177 KAVT--EILQDIQYKGKSKTIPCFNPNTLLPD-PLLRDYWVYEGSLTIPPCSEGVTWILF 233
hCA_IX    287 EQLL--SRLEEIAEEGSETQVPGLDISALLPS--DFSRYFQYEGSLTTPPCAQGVIWTVF 342
hCA_X     194 NRMLNRDTITRITYKNDAYLLQGLNIEELYPE---TSSFITYDGSMTIPPCYETASWIIM 250
hCA_XI    196 SRLLNRDTITRISYKNDAYFLQDLSLELLFPE---SFGFITYQGSLSTPPCSETVTWILI 252
hCA_XII   181 DKIF--SHLQHVKYKGQEAFVPGFNIEELLPE--RTAEYYRYRGSLTTPPCNPTVLWTVF 236
hCA_XIII  159 QKIT--DTLDSIKEKGKQTRFTNFDLLSLLPP---SWDYWTYPGSLTVPPLLESVTWIVL 213
hCA_XIV   172 EHIL--SHLHEVRHKDQKTSVPPFNLRELLPK--QLGQYFRYNGSLTTPPCYQSVLWTVF 227

                       250       260       270       280       290
hCA_I     214 KESISVSSEQLAQFRSLLSNVEGDNA-----VPMQHNNRPTQPLKGRTVRASF------- 261
hCA_II    212 KEPISVSSEQVLKFRKLNFNGEGEPE-----ELMVDNWRPAQPLKNRQIKASFK------ 260
hCA_III   212 KEPMTVSSDQMAKLRSLLSSAENEPP-----VPLVSNWRPPQPINNRVVRASFK------ 260
hCA_IV    236 REPIQLHREQILAFSQKLYYDKEQT------VSMKDNVRPLQQLGQRTVIKSGAPGR--- 286
hCA_V_A   246 KEPVEVAPSQLSAFRTLLFSALGEEE-----KMMVNNYRPLQPLMNRKVWASFQATNEGT 300
hCAI_V_B  246 KQPVEVDHDQLEQFRTLLFTSEGEKE-----KRMVDNFRPLQPLMNRTVRSSFRHDYVLN 300
hCA_VI    232 ADFVKLSRTQVWKLENSLLDHR--------NKTIHNDYRRTQPLNHRVVESNFPNQEY-- 281
hCA_VII   215 REPICISERQMGKFRSLLFTSEDDER-----IHMVNNFRPPQPLKGRVVKASFRA----- 264
hCA_VIII  234 RYPLTISQLQIEEFRRLRTHVKGAELVEGCDGILGDNFRPTQPLSDRVIRAAFQ------ 287
hCA_IX    343 NQTVMLSAKQLHTLSDTLWGPG--------DSRLQLNFRATQPLNGRVIEASFPAGVDSS 394
hCA_X     251 NKPVYITRMQMHSLRLLSQNQPSQIF-----LSMSDNFRPVQPLNNRCIRTNINFSLQGK 305
hCA_XI    253 DRALNITSLQMHSLRLLSQNPPSQIF-----QSLSGNSRPLQPLAHRALRGNRDPRHPER 307
hCA_XII   237 RNPVQISQEQLLALETALYCTHMDDP---SPREMINNFRQVQKFDERLVYTSFSQV---- 289
hCA_XIII  214 KQPINISSQQLAKFRSLLCTAEGEAA-----AFLVSNHRPPQPLKGRKVRASFH------ 262
hCA_XIV   228 YRRSQISMEQLEKLQGTLFSTEE--E---PSKLLVQNYRALQPLNQRMVFASFIQA---- 278
```

Figure 1.2: Sequence alignment of the different human Carbonic Anhydrase isoforms. The red asterisks represent the histidines coordinating the metal ion.

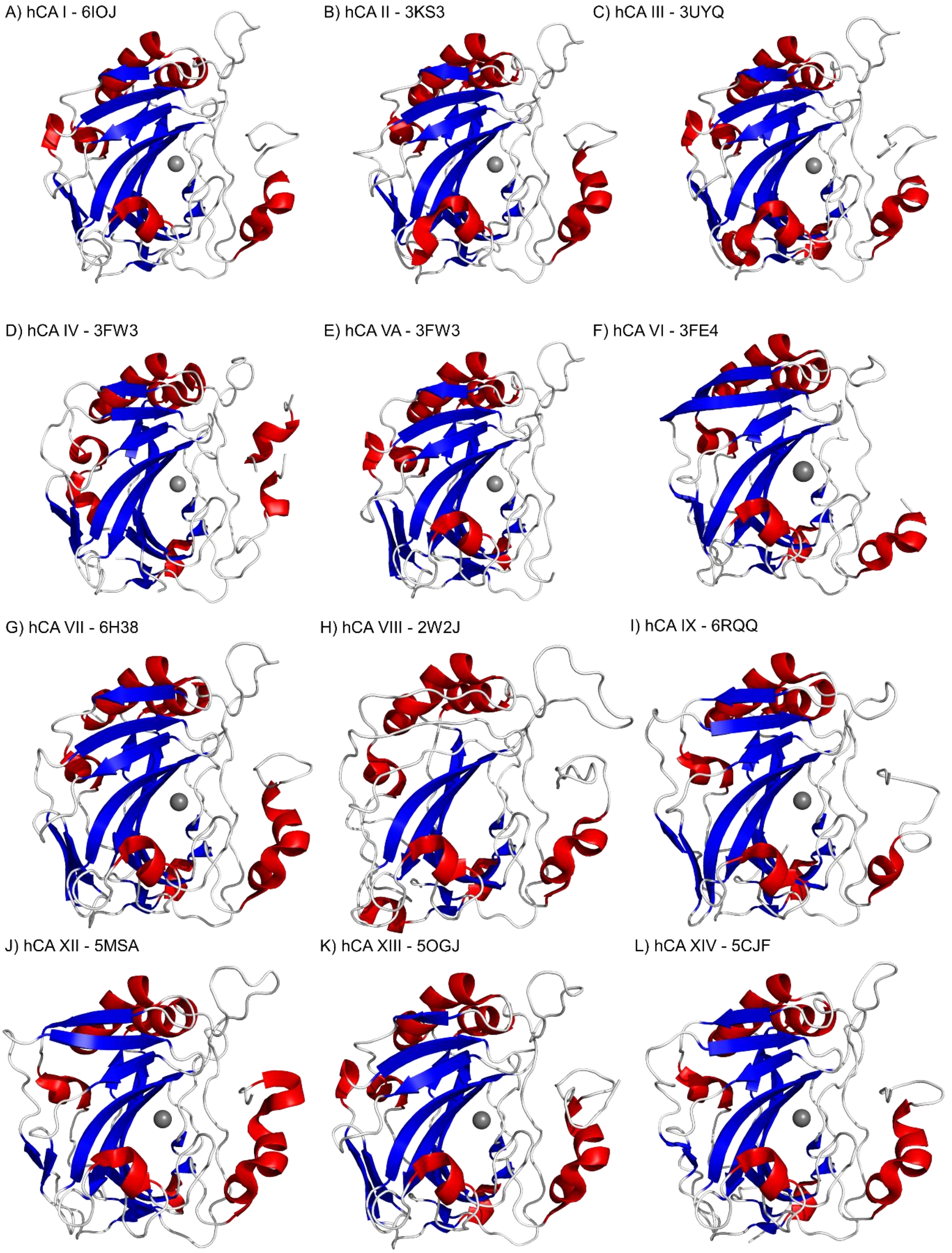


Figure 1.3: Structure of the different human carbonic anhydrase isoforms, except isoforms VB, X and XI and their respective PDB codes. A) hCA I (6I0J)[11]; B) hCA II (3KS3)[12]; C) hCA III (3UYQ)[13]; D) hCA IV (3FW3)[14]; E) hCA VA (1KEQ)[15]; F) hCA VI (3FE4)[16]; G) hCA VII (6H38)[17]; H) hCA VIII (2W2J)[18]; I) hCA IX (6RQQ)[19]; J) hCA XII (5MSA)[20]; K) hCA XIII (5OGJ)[21] and L) hCA XIV (5CJF)[22]. The grey sphere represents the zinc(II) metal ion present in the active center.

The sequence alignment shown in **Figure 1.2** illustrates the sequence similarity of the 15 human isoforms, and most of the enzymes have the active site conserved, except for isoforms VIII, X and XI, which do not have three histidines capable of binding the zinc(II) ion, making them acatalytic.

All the isoforms are monomeric, except for hCA VII, hCAII IX and hCAII XII, which form dimers. The isoforms hCA IV, hCA IX, hCA XII and hCA XIV are associated with membranes. As seen in **Figure 1.3**, most hCA isoforms share the same structural characteristics, such as seven right-handed α-helices and a twisted β-sheet composed by 10 β-strands, where two are parallel and 8 are anti-parallel[7].

The primary function of hCA is to catalyze the interconversion of carbon dioxide to bicarbonate. Furthermore, hCA also has esterase activity towards esters of carboxylic, sulfonic and phosphoric acid derivatives[9], and can oxidize aldehydes to alcohols.

1.1.2 Physiopathology and therapeutics

hCAs are involved in a variety of physio-pathological processes such as respiration, transport of CO_2 or bicarbonate ion, CO_2 homeostasis, electrolyte secretion, calcification, tumorigenicity, and several others[6,23]. These enzymes are responsible for producing HCO_3^-, which is involved in many cellular processes. Therefore, their mutations and malfunctions can lead to disease. hCAs are indeed involved in many diseases and their inhibition can lead to a better management of the symptoms. hCAs are involved in cancer and their metastasis, glaucoma and associated eye disorders, rheumatoid arthritis, osteoporosis, obesity, epilepsy, sleep apnea, neuropathic pain, cerebral ischemia and cardiac dysfunction[9].

To obtain energy, cancer cells burn glucose through glycolysis at an abnormal rate. This higher rate of glycolysis generates a large amount of lactic acid, which is excreted by the cells, and results in an acidic microenvironment around the cancer cells. Generally lactic acid is cleared by the vascular system. However, the vascular system is inefficient at removing the lactic acid due to the higher rate of glycolysis. This acidic microenvironment threatens the survival of cells because it changes the intracellular pH, which is under normal physiological conditions close to pH 7.0[24]. A deviation from this intracellular pH threatens the cell's survivability because cell metabolism is tightly dependent on pH. To ensure survival, cancer cells developed strategies to regulate their intracellular pH. Some strategies include the expression of transporters that are responsible to export H^+ and import Na^+-HCO_3^- or cotransport Cl^-/HCO_3^-. Another strategy involves the overexpression of CAs, which generally are the membrane associated hCAIX and hCAXII. The hCAIX is found in several malignant tumors such as: cervix, brain head and neck, colon breast and bladder cancer, whilst the hCAXII is found in cervical, breast and renal cell carcinomas[9]. These hCA's isoforms are highly controlled by oxygen levels in multiple epithelial tumor types[25], which are lower due to the constant division of tumor cells[9]. To date there is a potent inhibitor of hCAIX and hCAXII called SLC-0111 that is in clinical phase II trials for treatment of aggressive tumors and their metastasis. SLC-0111 is a ureido substituted benzosulfonamide (**Figure 1.4**)[9].

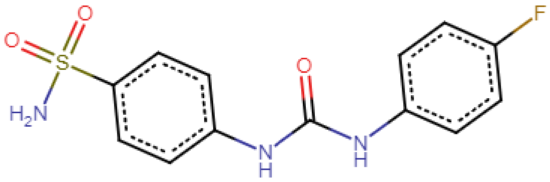


Figure 1.4: Chemical structure representation of SLC-0111 hCAIX and hCAXII inhibitor.

The hCAI, hCAII and hCAIV isoforms are widely distributed in the human eye and play an important role in the aqueous humor secretion. These isoforms control the intra-ocular pressure; therefore, a deregulation of these enzymes can lead to problems. In the case of glaucoma, the isoform hCAXII is overexpressed and increases intra-ocular pressure. The treatment of glaucoma is achieved by inhibiting hCAXII using dorzolamide and brizolamide (**Figure 1.5**) that are administered directly to the eye, which leads to a decrease of intra-ocular pressure and limits the secondary effects compared when administered via other routes[9].

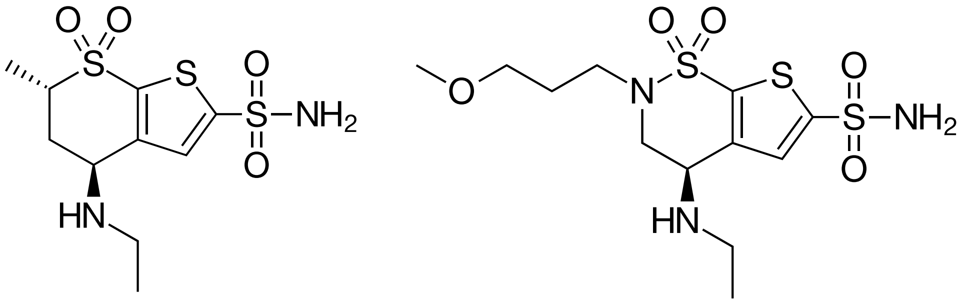


Figure 1.5: Chemical structure representation of Dorzolamide (left) and Brinzolamide (right).

The human bone matrix has 90-95 % collagen fibers that are rigidified by deposition of inorganic mineral salts, such as calcium phosphate. Although calcium phosphate is the main mineral present in the bone, other ions such as sodium, potassium, magnesium and carbonate ions are also present in the bone[26]. Studies involving patients with rheumatoid arthritis determined that hCAI, hCAIII and hCAIV are overexpressed, although only antibodies were detected. *In vitro* studies have shown that hCAI plays an important role in $CaCO_3$ formation, which precipitates during the process of bone formation. Furthermore, since rheumatoid arthritis patients have an overexpression of hCAI, this isoform may be responsible for the improper mineralization of the joints, thus leading to their inflammation. The isoenzyme hCAII is abundantly expressed in osteoclasts and the enzyme's activity provides the hydrogen ions required for the mobilization of calcium ions by an ATP-dependent proton pump. Furthermore, isoenzymes hCAIV and hCAXIV are also expressed in osteoclasts. An imbalance of these human isoforms can lead to osteoporosis and rheumatoid arthritis; therefore, these isoenzymes are potential drug targets[9].

Mitochondrial isoenzymes, hCAVA and hCAVB, are involved in several metabolic processes such as ureagenesis, gluconeogenesis and lipogenesis. The relation between these isoforms and these metabolic pathways is carbon dioxide and the carbonate ion. The carbonate ion is involved in ureagenesis, gluconeogenesis and lipogenesis, among many other metabolic pathways. hCA became a target for obesity after the administration of topiramate, an anticonvulsive drug, to obese epileptic patients, leading to the weight loss of these patients. Inadvertently inhibiting the mitochondrial isoforms of hCA led to a decrease of lipogenesis and to the discovery of an obesity treatment. Furthermore, there are other drug molecules, zonisamide and trifluoromethansulfonamide (**Figure 1.6**) that decrease lipogenesis in adipocytes[9].

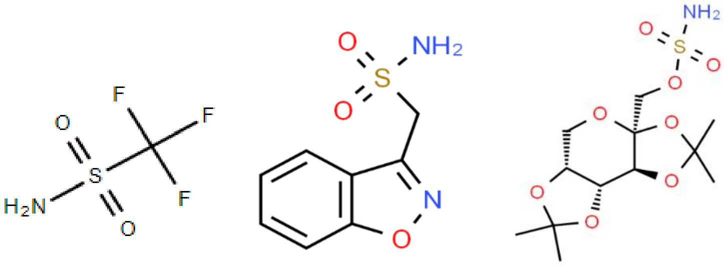


Figure 1.6: Chemical structure representation of trifluoromethansulfonamide(left), zonisamide (center) and topiramate (right).

As mentioned above, targeting of hCA isoforms leads to the treatment of epilepsy, by using topiramate. The isoforms hCAII, hCAIV, hCAVA, hCAVB, hCAVII, hCAXII and hCAXIV are present in the human brain, specifically in oligodendrocytes, choroid plexus, astrocytes and myelinated tracts. In the brain tissue, the carbonate ion is involved in voltage-gated channels, which actively participate in seizures and are also related with the excitatory GABAergic transmission. Methazolamide, acetazolamide (**Figure 1.7**), and zonisamide (**Figure 1.6 center**) can be used as inhibitors to treat epilepsy. Furthermore, the inhibition of hCA present in the CNS and in the peripheral nervous system can help with the neuropathic pain[9].

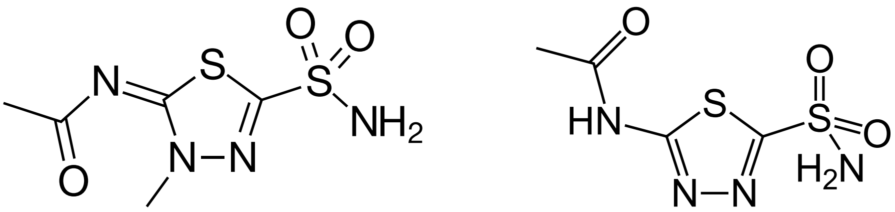


Figure 1.7: Chemical structure representation of methazolamide (left) and acetazolamide (right).

The hCA isoenzymes are also related with brain ischemia, and the targeting of hCA can lead to lower brain edema, lower neuronal death and a lower mortality rate. However, the mechanisms of action involved are not fully understood.

The acatalytic isoforms (hCA VIII, X and XI) are present in the central nervous system and their function is not understood. Furthermore, isoform VIII is related to a rare form a hereditary cerebellar ataxia, mental retardation and disequilibrium syndrome 3 which causes cerebellar atrophy, impaired intellectual development and quadrupedal gait[27,28].

1.1.3 Transition metal substitution

In this thesis, the focus is on the second isoform of human Carbonic Anhydrase (hCAII). The zinc(II) in hCAII is coordinated by four ligands, His94 and His96 through their Nε2 atoms, His 119 through its Nδ1, and a hydroxide ion or a water molecule depending on pH[2]. It has been shown that if zinc(II) is substituted with cobalt(II), the enzyme maintains a good fraction of its enzymatic activity[29,30]; while, manganese(II), iron(II), copper(II), and nickel(II) hCAII derivatives are inactive[30]. The coordination geometries of the metal-substituted enzymes are different from that of zinc, except for cobalt(II), which remains essentially tetra-coordinated in its active form[29,31–34]. For manganese(II) and copper(II), the metal center is penta-coordinated[29], and for nickel(II), it is either five[35] or six-coordinated[29,35,36]. These features are traditionally explained in terms of the different preferences of the different metals for different coordination geometries. The ions nickel(II) and copper(II) also bind to the hCAII in a second region, the N-terminal. In this secondary site, the copper(II) ion is coordinated by two Nε2 atoms from His 4 and His 64. In the case of nickel (II), this ion is coordinated by two backbone nitrogens from His 3 and His 4, one Nδ1 from His 4 and one water molecule.

Up to date, there are over 900 structures of hCAII deposited in the Protein Data Bank. In the vast amount of these deposited data, different metal ligands are present in the structure and/or in the active center. There are a few structures that have the non-native transition metal in its active site, as can be seen in **Figure 1.8**.

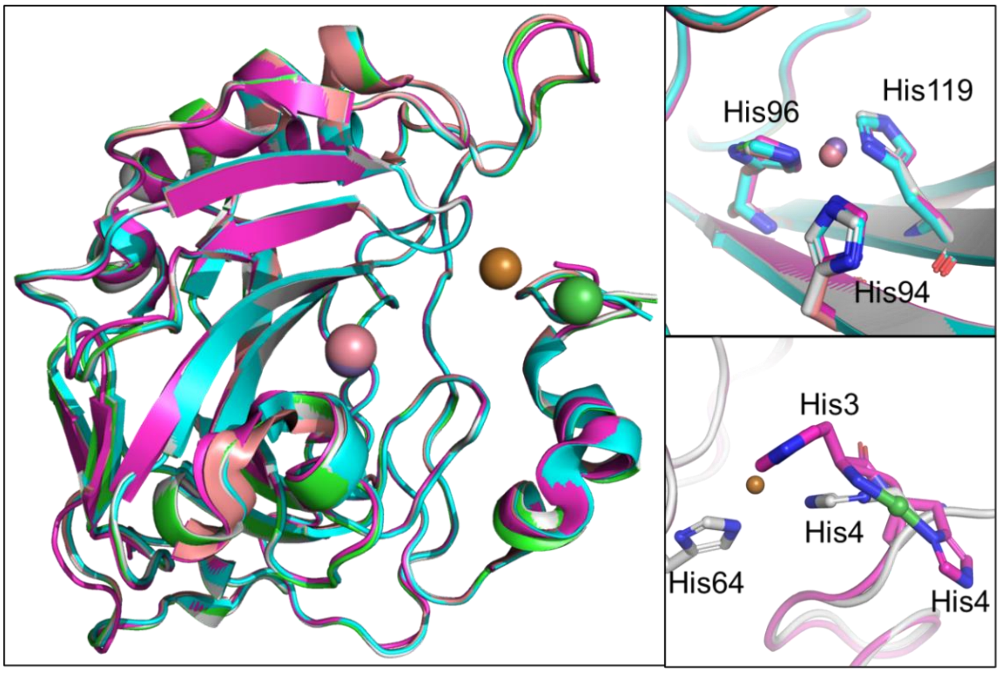


Figure 1.8: Structures of human Carbonic Anhydrase II and their metal derivatives (manganese (green)(1RZD)[29], cobalt (cyan)(3KON)[31], nickel (purple)(6H6S)[35], copper (salmon)(1RZC)[29] and zinc (white)(3KS3)[12]). The purple sphere represents the manganese metal ion, pink sphere represents the cobalt metal ion, brown sphere represents copper metal ion and green sphere represents the nickel metal ion. The manganese, cobalt, nickel, copper and zinc metal ions overlap in the active center and coordinated by His 94, His 96 and His 119. Nickel(II) and copper(II) metal ions can bind at a secondary binding center composed of either His 3 or His 4 and His 3 and His 64, respectively.

1.1.4 The hidden activities of carbonic anhydrase

The metal ion present at the active center of hCAII can be substituted by other transition metal ions, as previously mentioned. Generally, the metal substitution either slows down the hydration of carbon dioxide (i.e. cobalt (II)) or abolishes the activity of hCAII (other transition metal ions). However, hCAII presents some interesting properties when substituted with other transition metals.

Manganese(II)-hCAII acts as a peroxidase and catalyzes a selective epoxidation of olefins in the presence of an amino-alcohol buffer. The mechanism of action is different from the heme peroxides, which perform the same type of catalysis, and do not produce aldehydes as a side product. This type of reaction comes with the cost of damaging the enzyme and making the turnover of the enzyme low[37,38].

The hCAII isoform can bind up to two copper(II) ions, one in the active center and another in the N-terminal region[29,39]. However, the physiological relevance of the binding of these two transition metal ions remained elusive. Copper(II)-hCAII binds copper (II) in the active center, as well to the secondary binding site (His 3, His 4 and His 64). Studies show that copper(II)-hCAII has two different enzymatic activities, it can oxidize 2-aminophenol using hydrogen peroxide, and it can reduce nitrite[40]. For both chemical reactions, both copper (II) centers are involved.

One interesting case is the rhodium(I)-hCAII. This substituted enzyme is capable of hydrogenating olefins[41]. To date, there is no structure of rhodium(I)-hCAII and no mechanism for how rhodium catalyzes the hydrogenation reaction of olefins. However, there is an *in-silico* study showing that rhodium(I)-hCAII can act as a reductase and directly hydrogenate carbon dioxide to formate[42].

1.2 Thesis objectives

The hCAII isoform is an enzyme that has been thoroughly studied for many decades and has proven to be an excellent model protein. This enzyme is involved in several pathologies; therefore, it is a prototypical drug target. The metal ion present in hCAII can be substituted by several other transition metal ions. Some of these transition metal ions have unpaired electrons, making them paramagnetic. The second mammalian CA isoform is one of the most explored enzymes using paramagnetic metal ions[3,43–46]. One of the most explored paramagnetic ions in hCAII is cobalt(II). It is known that different ligands can cause extreme changes in the cobalt electronic structure, which are observed by electronic spectroscopy[47,48] as well as Electron Paramagnetic Resonance (EPR)[44].

In the current days, Nuclear Magnetic Resonance (NMR) has evolved to the point of acquiring almost “error-free” data due to the increased resolution it has achieved. Since NMR and EPR observables can be related through quantum chemistry equations[49], these techniques can be used as a tool to study the changes in the coordination sphere of the paramagnetic metal ion.

Taking this into consideration, in this thesis, the hCAII isoform will be explored as a model to study the paramagnetic effects arising from the paramagnetic transition metal ion that will be incorporated into the protein’s active site.

The thesis is divided according to the metal ion present in the active center of the protein. The transition metal ions explored were cobalt(II), nickel(II) and copper(II). To explore the paramagnetic effects, the main techniques used were: Nuclear Magnetic Resonance (NMR) and EPR, sometimes complemented by X-ray crystallography. The incorporation of the different metal ions in CAII had different objectives, as summarized below.

The cobalt(II) chapter, Chapter 3 – Structural studies of cobalt(II)-DM-hCAII and its thiocyanate adduct, explores the effect of sodium thiocyanate concentration on the protein. In this chapter there are several goals:

- Acquire paramagnetic restrains (PCS and RDC) at different concentrations of sodium thiocyanate
- Study the effect that sodium thiocyanate exerts on the protein at different ligand concentrations
- Crystallize and determine the structure of cobalt(II)-hCAII using X-ray crystallography bound to sodium thiocyanate at high concentration
- Integrate paramagnetic restraints and structure of cobalt(II)-hCAII into a quantum chemistry program to predict paramagnetic shifts

The nickel(II) chapter, Chapter 4. – Investigation of nickel(II)-WT-hCAII, explores the paramagnetic effect of the nickel(II) ion in solid-state NMR. This chapter can be divided into several goals:

- Crystallize and determine the structure of nickel(II)-hCAII using X-ray crystallography
- Use solid-state NMR to acquire structural information of crystalline nickel(II)-hCAII such as PCS
- Optimize solid-state NMR strategies through the usage of a nickel(II) probe

The copper(II) chapter, Chapter 5. – The pseudo-contact shifts of Copper(II)-hCAII, aims to solve a discrepancy between the frameworks that attempt to describe the origin of the pseudocontact shifts. To determine which framework is valid it is required to:

- Acquire PCS and EPR data of copper(II) substituted hCAII bound to oxalate
- Determine which framework predicts better the PCS and EPR parameters.

2

Chapter 2 - Structural Biology Methodologies

2.1 Introduction

The field of Structural Biology studies the structure, function and interactions of biomolecules such as proteins, nucleic acids, lipids and carbohydrates. This field tries to understand how the atomic structure of these biomolecules translates to function. To achieve this goal, the field of Structural Biology uses an array of techniques that provide complementary information. The main techniques used for this purpose are: Nuclear Magnetic Resonance (NMR) Spectroscopy, both in solution and in the solid-state, X-ray Crystallography, Small-Angle X-ray Scattering and Cryogenic Electron Microscopy. Individually, these techniques provide a vast amount of data that give an insightful view of the biomolecular structure, however NMR is the technique on that studies dynamics at an atomic level. Electron Paramagnetic Resonance (EPR) is often used as a complementary spectroscopy that provides information of the radical environment. However, only when integrated with other techniques, the full picture of the role of these biomolecules can be obtained.

The determination of a biomolecule structure requires the measurement of many structural restraints. These molecular restraints can be bond lengths, bond angles, dihedral angles, torsion angles and non-bonded atomic distances. In the case of NMR, the presence of a paramagnetic center causes hyperfine shifts of the nuclear spin signals, paramagnetic residual dipolar couplings, paramagnetic relaxation enhancements and cross-correlation effects. All these effects have distance and angular dependencies, therefore allowing structural information to be extracted.

In this thesis, several Structural Biology techniques were employed to understand the origin of the observed paramagnetic effects and how they can be exploited to extract structural information. To achieve these goals, we used human Carbonic Anhydrase II (hCAII). a metal containing enzyme. This enzyme is a model protein with a native zinc(II) metal ion that can be replaced by other paramagnetic metal ions. In this thesis, we explore the paramagnetic effects of cobalt (II), nickel (II) and copper (II) when substituting the native zinc(II) in hCAII. To study the paramagnetic effects arising from these metal ions, we used mainly NMR and EPR, but these techniques were in some cases complemented by X-ray Crystallography and Circular Dichroism (CD). In this chapter, a brief introduction to the basic principles of paramagnetic NMR and EPR is presented, focusing on the cases of cobalt(II), nickel(II) and copper(II) containing systems. This is complemented by a very brief introduction to the use of X-ray Crystallography in Structural Biology.

2.2 Paramagnetic Nuclear Magnetic Resonance

2.2.1 Introduction

Nuclear Magnetic Resonance (NMR) spectroscopy is an essential technique for structural characterization in Chemistry, Biology and Materials Sciences. NMR spectroscopy studies the interaction between the different magnetic moments of the nuclei of a molecule, having a spin quantum spin **I** ≥ ½, with an external magnetic field (B_0), that splits their energies E of I in a 2I+1 manifold in equal intervals (defined by M_I quantum numbers from -I, (-**I**+1), …, to +I) given by

$$\Delta E = g_N \mu_N B_0 \qquad \text{(Eq. 2.1)}$$

where g_N and μ_N are the g factor and nuclear magneton of the nucleus N with spin I. These quantities are related with another nuclear specific constant, the nuclear magnetogyric ratio (γ_I), through the relation

$$\gamma_I = 2\pi g_N \mu_N / h \qquad \text{(Eq. 2.2)}$$

where h is Planck's constant. The different nuclear magnetic moments present in a molecule can be visualized as magnetic dipoles that also interact with each other through the generated dipolar magnetic fields. These nuclear spins then absorb energy from applied radio frequency (RF) pulses through allowed spin transitions between contiguous energy states ($\Delta M_I = 1$) when the resonance condition applies ($\Delta E = h\nu$). After the excitation pulse, the macroscopic spin magnetization M relaxes back to the equilibrium state M_o, where the corresponding transverse spin component, M_{xy}, is detected through the time-dependent voltage signals (free induction decay, FID) that it induces in a tuned RF coil in quadrature. Fourier transformation of the digitalized FIDs leads to the frequency spectrum of the molecule, which reflects all the unique chemical environments of those nuclei.

When a material is placed in a static magnetic field (H_0), there is an induction of the electronic magnetic moments within the material, resulting in a macroscopic magnetization M_χ, which is proportional to the strength of the magnetic field with a proportionality constant know as magnetic susceptibility of the sample (χ)[50]. While in a vacuum the magnetic induction field (magnetic flux density) is

$$B_0 = \mu_0 H_0 \qquad \text{(Eq. 2.3)}$$

where μ_0 is the magnetic permeability of vacuum, for a material sample

$$B_0 = \mu_0 (1+\chi) H_0 = \mu H_0 \qquad \text{(Eq. 2.4)}$$

where μ is the magnetic permeability of the material.[51]

Some objects have paired electrons (diamagnetic), while others have unpaired electrons (paramagnetic). In diamagnetic materials, the electrons are in closed shells and generate a local magnetic field that repels the external static magnetic field ($\chi^{dia} < 0$), thus contributing to a different chemical shielding. A different effect occurs in paramagnetic materials, where the unpaired electrons tend to partly align with the external magnetic field and contribute to the magnetic susceptibility ($\chi^{para} > 0$). The total molecular magnetic susceptibility (χ^{mol}) can be described by the tensorial sum between the paramagnetic and diamagnetic susceptibility[36,50].

$$\chi^{mol} = \chi^{para} + \chi^{dia} \qquad \text{(Eq. 2.5)}$$

If a molecule is paramagnetic, the paramagnetic contribution is much larger and overcomes the diamagnetic contribution to the magnetic susceptibility.

The magnetic susceptibility can be mathematically represented by a tensor that describes the intensity of the interaction of the magnetic dipole moments with the applied magnetic field along each of its directions. The χ tensor of a paramagnetic molecule has its origin at the paramagnetic metal center (M), and the reference frame is defined as the main directions of the anisotropy tensor. To describe the χ tensor, directions and angles are required. Considering the magnetic field to be along a general direction **κ** described by the spherical angles α and β where the χ tensor is diagonal but coinciding with the z axis of this reference frame (**Figure 2.1**), the main components of the tensor are χ_{xx}, χ_{yy} and χ_{zz} and its angles are α, β, θ and φ. α is the angle between the reference frame's z-axis and the external magnetic field **κ** direction; β is the angle between the reference frame's x-axis and the projection of the external magnetic field on the xy plane; θ and φ are the angles between the paramagnetic metal (M)-nucleus (N) vector **r** and the and z axis of the tensor, and the projection of **r** onto the xy plane of the tensor and the tensor x-axis, respectively.

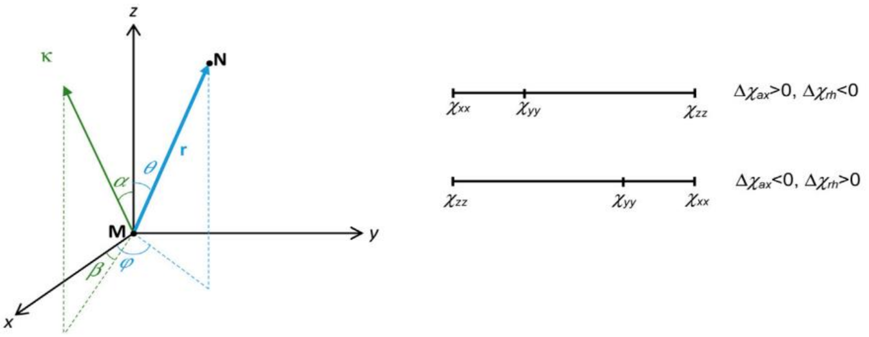


Figure 2.1. Representation of an anisotropic magnetic susceptibility tensor of a molecule, where the reference frame (x,y,z) is defined by the main directions of the anisotropy tensor. Its origin in the metal (M), θ and φ are the angles between the metal (M)-nucleus (N) vector **r** and the z-axis, and from the x-axis to the projection of **r** on the xy plane, respectively; α and β are the angles between the external field direction **κ** and the z-axis, and from the x-axis to the projection of **κ** on the xy plane, respectively.[50]

A χ tensor can have both isotropic (χ_{iso}) and anisotropic ($\Delta\chi$) contributions. An χ_{iso} tensor corresponds to a magnetic moment which is equal in every direction, where χ_{xx}, = χ_{yy} = χ_{zz}, while the $\Delta\chi$ tensor has axial ($\Delta\chi_{ax}$) and rhombic ($\Delta\chi_{rh}$) components:

$$\Delta\chi_{ax} = \chi_{zz} - \tfrac{1}{2}\left(\chi_{xx} + \chi_{yy}\right) \quad ; \quad \Delta\chi_{rh} = \chi_{xx} - \chi_{yy} \qquad \text{(Eq. 2.6 and Eq. 2.7)}$$

The paramagnetic susceptibility has a great impact on the overall magnetic susceptibility of a molecule due to the large magnetic moment of the electron, which is 658 times larger than the magnetic moment of the proton.

The unpaired electron(s) in a paramagnetic molecule can occupy a set of degenerate orbitals. If these orbitals are related by rotation and are not half-filled, the electron motion generates orbital angular momentum. Upon molecular rotation, the electron-nucleus vector changes orientation with respect to the magnetic field. This change in orientation modifies the induced magnetic field intensity. If the induced magnetic field is orientation dependent, then the rotation of the molecule will not average the dipolar coupling between the electron and nucleus. Therefore, the magnetic susceptibility becomes anisotropic.

In paramagnetic molecules, attention must be taken to the coupling between the nuclei and the unpaired electron(s) because they further affect the chemical shifts and relaxation times of the nuclei in the molecule. The chemical shift resulting from the coupling between a nucleus and the unpaired electron(s) is called hyperfine shift.

The hyperfine shift results of two interactions, the contact or Fermi contact (scalar) shift and the pseudocontact (dipolar) shift. The unpaired electrons of the paramagnetic center can be delocalized into the ligands' molecular orbitals and can reach several bonds away from that paramagnetic center. This delocalization of the electrons causes the ligands to have different spin densities at different points of the molecule. Furthermore, this spin density is affected by the molecular geometry and overlap of the molecular orbitals, leading to different chemical shifts. Generally, the spin density distribution is difficult to evaluate, making it hard to know its contribution to the dipolar coupling, and therefore on the pseudocontact shift, thus making it necessary to use quantum calculations to determine the electronic structure of the molecule. However, this problem can, in favorable cases, be simplified thought the Point-Dipole Approximation (PDA), which states that for distances over 3-4 Å from the paramagnetic center, the magnetic moment of the unpaired electron(s) can be assumed to be centered at the paramagnetic center[50].

As mentioned before, the paramagnetic center causes further relaxation to the nuclei in the molecular system. This happens because the time-dependent fluctuations of local magnetic fields at the nucleus originating from the paramagnetic center's magnetic moment further enhances the nuclear relaxation, which is also called paramagnetic relaxation enhancement.

Relaxation plays an important role in paramagnetic molecules. The sources of relaxation can be electron spin relaxation of the metal, molecular rotation of the complex containing metal and chemical exchange of the metal between bound and unbound states. For all cases, the relaxation originates from the fluctuations of local magnetic fields at the nucleus, with correlation times describing these motions represented by $T_{1,2e}$, τ_R and τ_M. The first mechanism is easy to understand, as $T_{1,2e}$ describe the fluctuation of the electron dipolar field. When the source of relaxation is caused by the molecular rotation, meaning it is faster than electron relaxation ($\tau_R < T_{1,2e}$), the nucleus interacts with the electron in random spatial positions, which causes fluctuating electron-nucleus dipolar coupling leading to resonance broadening. Chemical exchange between different sites also causes fluctuating magnetic fields at the nucleus through contact and dipolar mechanisms. Another mechanism, also modulated by molecular rotation, is Curie-spin relaxation (also known as magnetic susceptibility relaxation). This mechanism results from the interaction of the nuclear spins with the static time- averaged magnetic moment $<\mu>$ of the molecule over the different electron spin states, which have small differences in population according to the Boltzmann distribution. The Curie-spin relaxation mainly contributes to the transverse relaxation of the molecule, which increases with increasing strength of the external magnetic field. This contribution leads to the line broadening of the resonances at high fields, especially for large proteins (long τ_R), which explains why sometimes it is better to measure NMR spectra of paramagnetic biomolecules at lower magnetic field strengths.[50]

In nature there are many metalloproteins or proteins capable of extrinsically binding metal ions. These metal ions can be paramagnetic or can be substituted by a paramagnetic analog. Even if the protein systems cannot bind metal ions, it can be chemically modified with a metal binding tag[45,52,53].

The presence of a paramagnetic center in a protein can be extremely useful, especially due to the use of extractable long-range paramagnetic restraints. These restraints can be pseudocontact shifts (PCS)[5,6], self-orientation residual dipolar couplings (RDCs)[54] arising from a paramagnetic center, and paramagnetic relaxation enhancements (PRE)[55]. All these observables are dependent on the distances and angles between the observed nucleus, the χ^{mol} tensor and the paramagnetic center. In this thesis the focus was to acquire pseudocontact shifts and residual dipolar couplings of a paramagnetic protein, hCAII with the zinc(II)-containing diamagnetic active center replaced by the paramagnetic transition metal ions cobalt(II), nickel(II) and copper(II) and use them in structural studies.

2.2.2 Pseudocontact shifts

The dipole-dipole interaction between a nucleus with magnetic moment μ_1 and an unpaired electron with magnetic moment μ_2 in a molecule with an anisotropic magnetic susceptibility tensor $\boldsymbol{\chi}$ changes intensity with their relative orientation in respect to the external magnetic field, and thus its average over all orientations caused by molecular rotation is not averaged to zero. The remaining orientational average causes a chemical shift of the nucleus, known as the pseudocontact (or dipolar) shift, which, together with the scalar (or contact) shift, constitutes the nuclear hyperfine shift. This work focusses on the use of the PCS of transition metal ions. PCS were firstly described by Bloembergen and Dickinson[56], McConnell and Robertson[57] and Kurland and McGarvey[58], and have been extensively used for many decades to study molecules containing paramagnetic transition metal ions[53,59–65] or lanthanides ions[52,53,66–71]. Considering the direction of the external magnetic field (**κ**, **Figure 2.1**) to be coincident with the z axis of the $\boldsymbol{\chi}$ tensor, the PCS is described by Kurland and McGarvey equation:

$$\Delta\delta^{pcs} = \frac{1}{12\pi}\frac{1}{r^3}\left[\Delta\chi_{ax}(3cos^2\theta - 1) + \frac{3}{2}\Delta\chi_{rh}\, sin^2\theta\cos\,(2\varphi)\right] \qquad \text{(Eq. 2.8)}$$

This equation has several components: $\Delta\delta^{pcs}$ is the calculated value of PCS of a nucleus in a molecule, r is the distance between the observed nuclear spin and the unpaired electron at the paramagnetic center; $\Delta\chi_{ax}$ and $\Delta\chi_{rh}$ are the axial and rhombic components, respectively, of the magnetic susceptibility anisotropy tensor (Eqs. 2.6 and 2.7), which in turn also have an angular dependence relative to the magnetic field, defined by θ and φ, which are the polar angles of the metal-nucleus (**r**) vector, where θ describes the angle between **r** and the z-axis of the tensor, and φ is the angle between the projection of **r** on the xy plane and the x-axis[45,50]. As described by the equation, the PCS are information rich in terms of structural restraints, as it describes the positions between the observed nuclei and the paramagnetic center. From the equation above, one can also extract distance information for the nuclei to the metal center, as the PCS decreases with *r* according to $1/r^3$.

For an isotropic χ tensor, where $\chi_{xx} = \chi_{yy} = \chi_{zz}$,, the terms $\Delta\chi_{ax}$ and $\Delta\chi_{rh}$ are equal to zero, therefore canceling the second term of the equation, meaning that no PCS are observed.

The experimental determination of $\Delta\delta^{pcs}$ requires NMR data acquisition in two samples, a paramagnetic and the corresponding diamagnetic one, which is used as reference. Furthermore, the structure between these two samples must remain the same. Then, it is necessary to subtract the chemical shifts of the nuclei in the diamagnetic sample (δ_{dia}) from those of the same nuclei in the paramagnetic sample (δ_{para}), as seen in Eq. 2.9:

$$\Delta\delta^{pcs} = \delta_{para} - \delta_{dia} \qquad \text{(Eq. 2.9)}$$

The PCS values (Eq. 2.8) are independent of the type of observed nucleus and depend only on the position of the mentioned nucleus relative to the paramagnetic center. The acquisition of the PCS for a protein system is easily done using a heteronuclear single quantum coherence spectrum (HSQC), such as the ^{1}H-^{15}N-HSQC spectra, for both diamagnetic and paramagnetic samples (**Figure 2.2**). For this purpose, it is mandatory to label the protein with the ^{15}N isotope, which nowadays is inexpensive. Since the ^{1}H and ^{15}N nuclei of a protein amide group are close in space, it is expected that both nuclei present similar PCS values. The norm is to use only the PCS arising from the ^{1}H nuclei in the structural determination of the protein due to their high gyromagnetic ratio and small spectral window. Furthermore, the chemical shifts of the ^{15}N nuclei are more sensitive than those of the ^{1}H nuclei to small variations in the torsion angles arising from the backbone of the protein, the hydrogen bonding and the electrical fields. With this increased sensitivity, incorrect PCS measurements of the ^{15}N nuclei can occur due to small local variations of these parameters between the diamagnetic and paramagnetic forms of the protein[72].

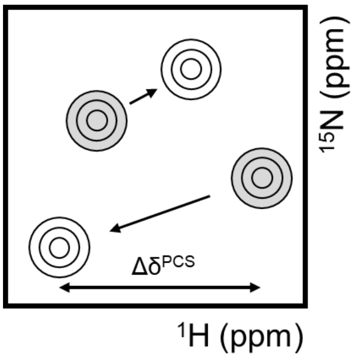


Figure 2.2: Representation of the pseudocontact shift in the NMR ^{1}H-^{15}N-HSQC spectrum. In grey are represented the diamagnetic chemical shifts and in white the paramagnetic chemical shifts

2.2.3 Residual Dipolar Couplings

The dipolar coupling between two nuclei describes the through-space interaction between the magnetic fields associated with their nuclear magnetic moments. The dipolar coupling is measured in Hertz, it can be expressed as a vector connecting the two nuclei and contains information regarding distance and orientation of the said interacting nuclei, which is useful to explore. In a system with many nuclei, such as a protein with thousands of protons, the huge number of those vectors corresponding to all the possible combinations between the different nuclei, together with other interactions, will drastically broaden the solid-state spectrum, overcrowded with overlapping resonances. In solution, the dipolar couplings from the different positions of the nuclear pairs relative to the external magnetic field are averaged to zero by the fast, isotropic molecular rotations of the molecules. This simplifies the NMR spectrum substantially, leaving only sharp lines.

Since the dipolar coupling contains geometrical information between the coupled nuclei, it can be a powerful source of structural restraints. In solution, it is possible to partially reintroduce the dipolar couplings using different strategies to make the distribution of the molecular reorientations to become anisotropic. One of the strategies to align the molecule with the static external magnetic field (B_0) is to have it interact with bicelles or filamentous phages that strongly align with B_0. Another strategy is to introduce a paramagnetic center with anisotropic magnetic susceptibility. In a static external magnetic field (B_0), the paramagnetic center will generate an induced magnetic field that will align with B_0, thus forcing the whole molecule to also partially align to B_0. This generated alignment is residual, of the order of 10^{-3} compared to permanent alignment. However, it generates some NMR effects related to the anisotropy of the nuclear interactions. The most widely used effect is the so-called residual dipolar coupling (RDC), which is shown as a perturbation of the *J*-splitting in aligned samples relative to the value measured in unaligned samples (**Figure 2.3**). Partial alignment of a molecule causes RDCs because when some orientations are preferred relative to others, the dipole-dipole interaction between the nuclear magnetic moments does not average to zero.

The mathematical representation of the RDC, due to the partial alignment induced by an anisotropic paramagnetic center in a molecule, is represented by the following equation:

$$\Delta J_{ij}^{RDC}(Hz) = \frac{-B_0^2}{15kT} \times \frac{\gamma_i \gamma_j h}{16\pi^3 r_{ij}^3} \times \left[\Delta\chi_{ax}(3cos^2\alpha - 1) + \frac{3}{2}\Delta\chi_{rh} sin^2\alpha\, cos2\beta\right] \qquad \text{(Eq. 2.10)}$$

This equation has several components: ΔJ^{RDC} is the calculated value of RDC between the paramagnetic sample and the diamagnetic sample that is generated between the dipolar coupling of *i* an *j* nuclei; *r* is the distance between the observed coupled nuclear spins; $\Delta\chi_{\mathrm{ax}}$ and $\Delta\chi_{\mathrm{rh}}$ are the axial and rhombic components, respectively, of the magnetic susceptibility anisotropy tensor, which in turn also have an angular dependence relative to the magnetic field; α and β are the polar angles, where α describes the angle between the internuclear vector and the z-axis of the tensor, and β is the angle between the projection of the internuclear vector on the xy plane and the x-axis[45,50]; B_0 is the strength of the external magnetic field; γ_i and γ_j are the gyromagnetic ratios of the observed nuclei and *T* is the temperature of the system in Kelvin.

Similarly to the PCS, in an isotropic situation, where $\chi_{xx} = \chi_{yy} = \chi_{zz}$,, the terms $\Delta\chi_{ax}$ and $\Delta\chi_{rh}$ are equal to zero, the second term of the equation is canceled, and no RDCs are observed.

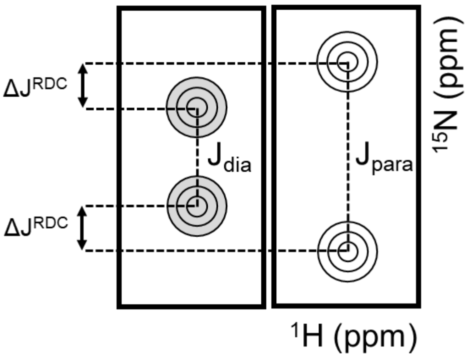


Figure 2.3: Representation of the residual dipolar couplings in the NMR ^{1}H-^{15}N-HSQC IPAP spectrum. In grey are represented the diamagnetic decoupled resonances and in white the paramagnetic decoupled resonances.

The experimental determination of ΔJ^{RDC} requires two samples, one paramagnetic and the corresponding diamagnetic, which is the reference. It is necessary to subtract the *J*-coupling between *i* and *j* nuclei in the diamagnetic sample from the *J*-coupling of the same nuclei in the paramagnetic sample, as seen in Eq. 2.11:

$$\Delta J_{ij}^{RDC} = J_{ij}^{para} - J_{ij}^{dia} \qquad \text{Eq. 2.11}$$

The acquisition of RDC data for a protein system requires acquiring an In-Phase Anti-Phase (IPAP) spectrum, where the coupling between different nuclei is observed. For proteins, the ^{1}H-^{15}N-HSQC-IPAP spectra are used to calculate the RDCs. These spectra are acquired without decoupling in the ^{15}N dimension and the multiplets have positive and negative contributions. From these positive and negative contributions, one can measure the ^1J coupling constant between the ^{1}H and ^{15}N nuclei present in each amide bond.

Equation 2.10 shows that the RDCs are directly proportional the square of the strength of the magnetic field (B_0^2), implying that the molecule will have a higher degree of alignment at higher magnetic fields, which will lead to higher splitting between the up and down components and thus an easier measurement of the RDC. However, if the alignment is generated by a paramagnetic center, then at a higher magnetic field the molecule will experience stronger relaxation effects due to Curie spin relaxation, which may substantially broaden the doublet resonances.

2.2.4 Collecting paramagnetic NMR restraints in biomolecules

Generally, to obtain the paramagnetic restraints, one must prepare two different samples, one paramagnetic and one diamagnetic. The diamagnetic reference can be prepared by removing the paramagnetic center from the molecule, substituting the paramagnetic metal for a diamagnetic one, or by changing the oxidation state of the molecule. For example, low spin iron(III), when reduced to low spin iron(II), becomes diamagnetic, and the same happens for copper(II) when reduced to copper(I). As mentioned above, the paramagnetic metal ion can be substituted by a diamagnetic one. For paramagnetic proteins, the paramagnetic center can be oxidized to become diamagnetic, if possible, removed, or substituted by a diamagnetic analog. If the protein has a metal center that is not paramagnetic, it can be substituted by a paramagnetic one. A third case can also happen where a protein does not have a metal center. In this last case, an artificial metal-containing functional group can be added to the protein by chemically reacting the protein with a metal ion complex[52,53,66–70].

The next step to obtain the paramagnetic restraints is to acquire the paramagnetic and diamagnetic spectra and compare them. For proteins, the best experiments are the ^{1}H-^{15}N-HSQC and its variations. For large biomolecules, TROSY, CRIPT, CRINEPT and methyl-TROSY can be used. Other correlations, such as CC and NC, can be used to acquire the paramagnetic restrains, however their interpretation is less straightforward and might require several structure refinements or pseudocontact shift predictions. Regardless of the correlation, since all measurements are based on comparison, a full assignment for both forms, diamagnetic and paramagnetic, is required.[50]

The hyperfine shift is calculated through the difference in the chemical shifts between the paramagnetic sample and the diamagnetic sample. For nuclei far removed from the paramagnetic center, the expected PCS is modest, and the assignment of these resonances is straightforward, as one can transfer the assignments between spectra. Once a subset of PCS data is acquired and if a structural model is available, an initial magnetic susceptibility anisotropy tensor can be calculated. From this initial tensor, other PCS can be predicted and be used to assign the remaining resonances. As mentioned above, for proteins the best experiment is the ^{1}H-^{15}N-HSQC due to the abundance of amide bonds. When comparing the resonances from the paramagnetic spectrum and the diamagnetic spectrum, the resonances, for the same residue, shift in a diagonal line. This diagonal shift happens because both ^{1}H and ^{15}N nuclei are roughly at the same distance from the paramagnetic center, and therefore the expected shift should be similar.

2.2.5 Use of PCSs and RDCs for χ-tensor calculation

To calculate the anisotropy and aligned tensors derived from PCS and RDC, it is required to compare the experimental and the back-calculated PCS and/or RDC using a structural model. There are many available programs capable of determining the anisotropy tensors, and the one used in this work was FANTEN[73].

To obtain the traceless χ tensor, it is necessary to calculate five independent components (χ_{xx}, χ_{zz}, χ_{xz}, χ_{yz}, χ_{xy}; with $\chi_{yy} = -\chi_{zz} - \chi_{xx}$). The program FANTEN will perform a linear fit of the mentioned components to obtain the best agreement between the back-calculated and experimental PCS/RDC data. FANTEN also requires the protein structure, as well as the metal coordinates if they are not included in the protein structure coordinates. After the calculation of the anisotropy tensor, through the best fit minimization, it is necessary to analyze the correlation plots of the experimental and back-calculated PCS/RDC. This process is iterative and requires the removal of outliers and/or the inclusion of new data to refine the anisotropy tensor values. At each iteration part, there is the calculation of the quantification of the agreement between the experimental and back-calculated data, which is called the Q-factor.

$$Q_{PCS} = \sqrt{\frac{\sum_i (PCS_i^{exp} - PCS_i^{calc})^2}{\sum_i (PCS_i^{exp})^2}} \quad \text{(Eq. 2.12)} \qquad Q_{RDC} = \sqrt{\frac{\sum_i \left(RDC_i^{exp} - RDC_i^{calc}\right)^2}{\sum_i \left(RDC_i^{exp}\right)^2}} \quad \text{(Eq. 2.13)}$$

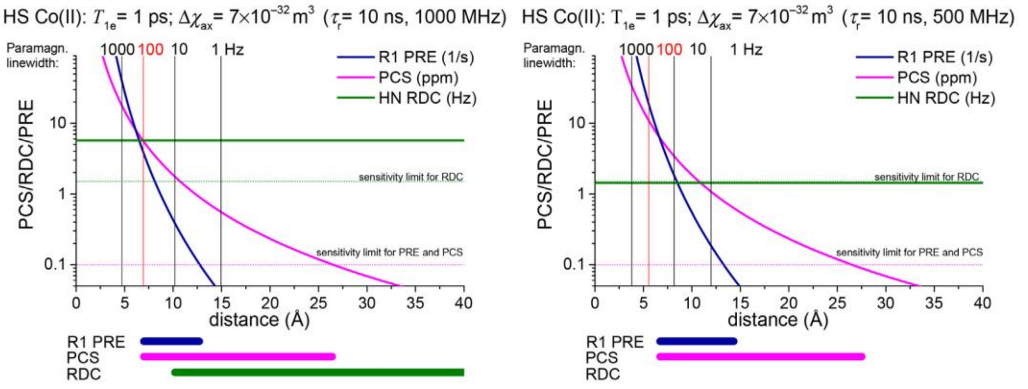
The Q-factor is a measure of the agreement between the observed and the back-calculated data. The Q-factor ranges between 0 and 1 and the closer to 0 the better the agreement between the observed and the back-calculated data.

2.2.6 PCS, PRE and RDC effects in cobalt(II), nickel(II) and copper(II) metalloproteins

The observation and interpretation of solution NMR spectra of paramagnetic metalloproteins depends on the knowledge of several basic parameters, namely the electron relaxation times and mechanisms of the metal center, their effects on the relaxation times (PRE) of the nuclei to be observed, followed by measurement of the hyperfine shifts, in particular PCS, and RDC. Only then, structural and dynamic information on those systems can be obtained. These properties depend critically on the electronic structure and coordination properties of the metal ion site in the protein. Here, these basic properties for copper(II), nickel(II) and copper(II) metalloproteins, directly relevant for this thesis, are briefly summarized.[50]

Cobalt(II), with a d^7 electronic configuration, with low or high spin. In the low spin case, $T_{1,2e}$ are long (0.1-1 ns), allowing EPR spectra to be obtained at room temperature but proton NMR spectra are broad due to the large energy differences between excited states and ground states. This large energy gap decreases the efficiency of the electron relaxation and the relaxation caused by the molecular rotation (if it is also in the same order of magnitude) starts to dominate and broaden the observed resonances. However, high spin cobalt(II), with an electronic spin S = 3/2 is much more interesting for NMR. Cobalt (II), when bound to proteins, has a low symmetry environment that removes the orbital degeneracy and causes the excited states to be close in energy to the ground states. This proximity in energy makes the electron relaxation to be very efficient, and depending on the geometry of the cobalt complex, $T_{1,2e}$ range from 1 to 10 ps. This short electron relaxation time originates large shifts on the observed nuclei, even at large distances from the metal center/NMR probe.

Ligand field: *O_h*:



Ligand field: *T_d*:

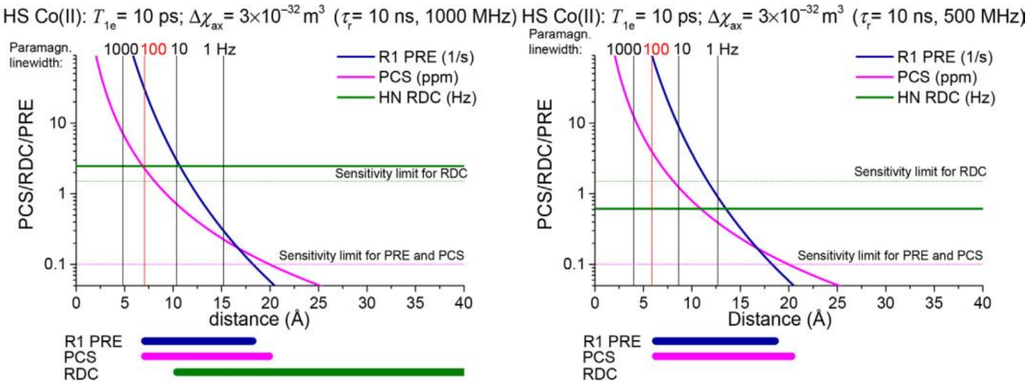
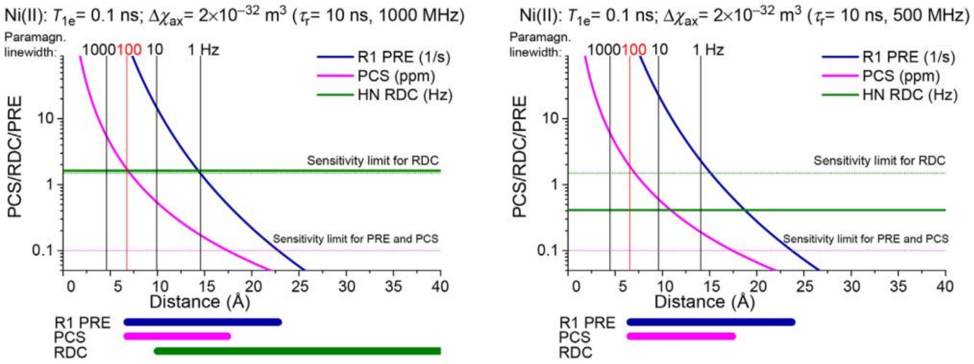


Figure 2.4: Calculated values of ^{1}H PRE, ^{1}H PCS, linewidths and ^{1}H-^{15}N RDC as a function of distance for a cobalt(II) molecule with a τ_R of 10 ns at two different fields, 1000 MHz (left) and 500 MHz (right), and octahedral (O_h) (above) and tetrahedral (T_d) coordination (below). Values of PREs are represented in blue, PCS in pink and RDC in green. The plots also show vertical lines, which are representative of line broadening expected at certain distances from the paramagnetic center. The bars bellow the graph represent the range where the PRE, PCS and RDC are generally measurable (adapted from[50]).

Figure 2.4 shows the expected values of the PRE, PCS, linewidths and RDCs for a protein with a τ_R = 10 ns, containing cobalt(II) with different coordination geometries, which conditions its electron spin relaxation time (T_{1e}). In an octahedral site (O_h), T_{1e} is typically 1 ps, with a typical $\Delta\chi_{ax}$ of 7 x 10^{-32} m^3, while in a tetrahedral site (T_d), T_{1e} is about ten times higher, 10 ps, with a typical $\Delta\chi_{ax}$ of 3 x 10^{-32} m^3. Depending on the geometry, the PCS are measurable between 6 and 27 Å and 6 and 25 Å distances from the metal ion, for octahedral and tetrahedral environments, respectively, and PREs up to 13 and 18 Å distances. The detection limits for ^{1}H NMR peaks is set from that corresponding to a paramagnetic linewidth of 100 Hz up to the sensitivity limit of 0.1 ppm for PCS and 0.1 s^{-1} for PRE.[50] Using a high magnetic field, it is possible to acquire RDC data for both environments, the octahedral environment producing higher RDC values. These are the expected values for the cobalt(II)-hCAII, where cobalt present in the enzyme is in the high spin state.

Nickel(II), with a d^8 electronic configuration, has two unpaired electrons (S =1) in 6- or 4-cordination environments. In tetrahedral coordination, the presence of low-lying excited states leads to very short T_{1e} values, of the order of 1 ps, leading to sharp proton NMR spectra. However, octahedral coordinated nickel(II) has high energy excited states high, with significantly slower electron spin relaxation, with T_{1e} values of about 0.1 ns and much broader proton NMR spectra. Typical values of the magnetic susceptibility anisotropy are in both cases around $\Delta\chi_{ax} = 2 \times 10^{-32}$ m^3. **Figure 2.5** shows the calculated maximum ^{1}H PCS, ^{1}H PRE, linewidths and amide proton paramagnetic RDCs in octahedral and tetrahedral symmetries.

Ligand field: O_h:



Ligand field: T_d:

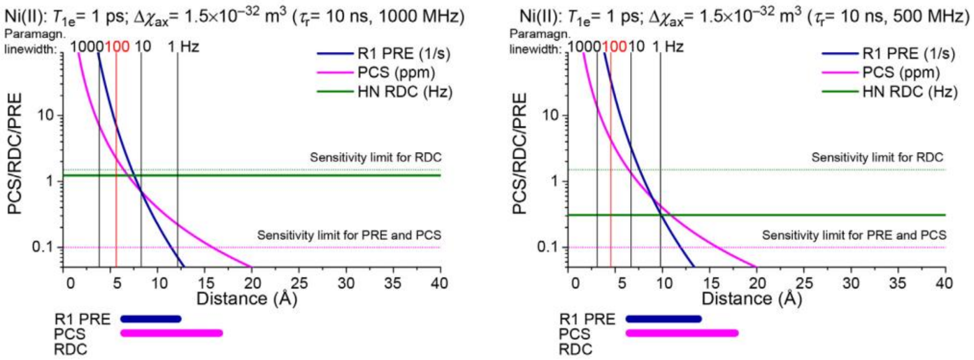


Figure 2.5: Calculated values of ^{1}H PRE, ^{1}H PCS, linewidths and ^{1}H-^{15}N RDC as a function of distance for a nickel(II) molecule with a τ_R of 10 ns at two different fields, 1000 MHz (left) and 500 MHz (right) magnetic field, and octahedral (O_h) (above) and tetrahedral (T_d) coordination (below). Values of PREs are represented in blue, PCS in pink and RDC in green. The plots also show vertical lines, which are representative of line broadening expected at certain distances from the paramagnetic center. The bars bellow the graph represent the range where the PRE, PCS and RDC are generally measurable (adapted from[50]).

The PCS are measurable between 7 and 17 Å and 6 and 16 Å distances from the nickel(II), for octahedral and tetrahedral geometries, respectively, and PREs up to 23 and 12 Å distances. Using a high magnetic field, it is possible to acquire RDC data for the octahedral environment, but for the tetrahedral environment RDC observation is at or below the sensitivity limit at 1000 and 500 MHz, respectively. Nickel(II)-substituted azurin is an example of a protein with a sharp ^{1}H NMR spectrum[74].

Finally, copper(II), with a d^9 electronic configuration (S=1/2), occurs in copper-proteins in mononuclear form in proteins in two different environments: type 2 and type 1 copper(II). It cannot be observed in pure octahedral and tetrahedral symmetries, due to Jahn–Teller distortions, which remove the orbital degeneracy of its ground state. The separation of the electronic energy levels depends on the coordination number and stereochemistry, as well as on the nature of the ligands. In type 2 Cu centers, in the absence of bound sulfur atoms, the most probable geometries, of D_{4h} or C_{4v} type, lead to a large separation of the ground state from the first excited state, leading to long electron spin relaxation times (T_{1e}) of copper(II), of the order of 1 ns, and relatively sharp EPR signals, which are often observable at room temperature. An example is copper-zinc Superoxide dismutase. In type 1 copper centers, copper(II) is bound to two nitrogen atoms of two histidines and at least one cysteinyl sulfur atom, leading to a tetrahedrally distorted- trigonal ligand coordination sphere. Then, the low energy of the excited energy levels leads to a decrease of the T_{1e} values by one order of magnitude relative to type 2 copper, to about 0.5 ns. Their EPR spectra are still relatively sharp at low temperatures, showing smaller hyperfine splittings in the parallel region of the spectrum ($A_{\|}$) relative to type 2 Cu. An example is the blue copper protein azurin.

The magnetic susceptibility anisotropy for copper proteins is small, with estimated values for $\Delta\chi_{ax}$ of about 6 x 10^{-33} m^3.

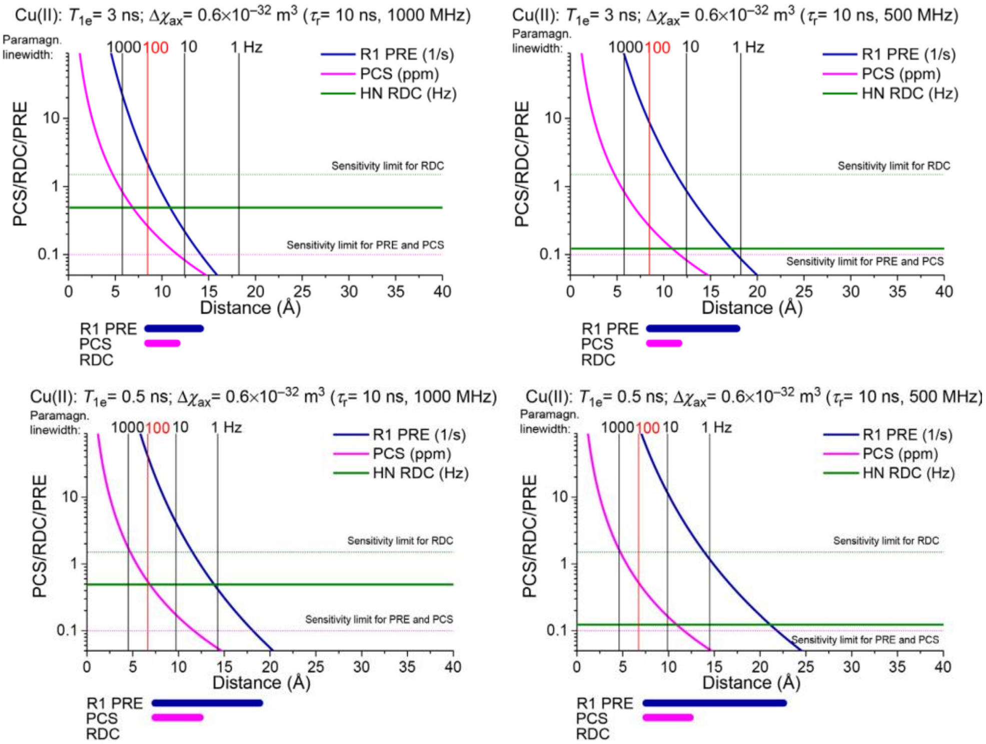


Figure 2.6: Calculated values of ^{1}H PRE, ^{1}H PCS, linewidths and ^{1}H-^{15}N RDC as a function of distance for a type 2 copper(II) center (top) and a type 1 copper(II) center (bottom) in a protein with a τ_R of 10 ns at two different fields, 1000 MHz (left) and 500 MHz (right) magnetic field. Values of PREs are represented in blue, PCS in pink and RDC in green. The plots also show vertical lines, which are representative of line broadening expected at certain distances from the paramagnetic center. The bars bellow the graph represent the range where the PRE, PCS and RDC are generally measurable (adapted from[50]).

Figure 2.6 shows the calculated maximum ^{1}H PCS, ^{1}H PRE, linewidths and amide proton paramagnetic RDCs as a function of the distance from the metal ion, at 500 or 1000 MHz, for a type 2 (top) and type 1 (bottom) copper protein with τ_R of 10 ns. The PCS are in both cases quite small, due to the very small $\Delta\chi_{ax}$ values and the ^{1}H NMR spectra are usually quite broad, especially for type-2 copper proteins, which is the case of copper(II)-hCAII.[75] The PCS are measurable between 7-8 and 15 Å distances in both cases, and PREs up to 16-20 and 20-25 Å distances, for type 2 and type 1 copper, depending on the magnetic field. RDC values cannot be obtained in both cases.

2.2.7 Solid-state NMR of paramagnetic proteins

As previously described, solution NMR is a powerful technique to study proteins. In parallel to the development of solution NMR, solid-state NMR (SSNMR) has been progressively improved and achieved the same sensitivity and resolution as solution NMR. The main difference between the solution NMR and SSNMR results from the stochastic molecular reorientation in liquids. The absence of molecular reorientation in solid samples has great consequences in the spectra obtained, due to the anisotropic interactions of the magnetic moment of each nucleus (and electric quadrupole for $I>1/2$) with the magnetic field and other nuclei in the sample, such as dipole-dipole interactions, chemical shift anisotropy (CSA) and quadrupolar interactions, which are not averaged out in solids. Besides increased relaxation, all these phenomena broaden the resonances beyond detection and interpretation.

In SSNMR, these hurdles have been surpassed by technological and methodological developments. The introduction of complex pulse sequences and fast sample spinning abolish the dipolar coupling. By spinning the sample fast enough, the dipolar contribution will decrease and be averaged out when the spin rate is fast relative to the magnitude of the dipolar coupling; furthermore, it will average out the chemical shift anisotropy thus obtaining isotropic resonances. For this, the sample must undergo spin at a specific angle, the magic angle $\theta \approx 54.74$, also known as Magic Angle Spinning (MAS), which nullifies many anisotropic interactions that have ($3*\cos^2\theta - 1$) spatial dependencies.[50] Cross polarization (CP) will transfer magnetization from usually abundant spins, like protons, to diluted spins (e.g. ^{13}C), increasing the sensitivity, complemented by high power decoupling. To completely abolish the quadrupolar interaction, double rotation techniques are used in SSNMR of quadrupolar nuclei.[76] However, many standard techniques for SSNMR of diamagnetic samples, have a very poor performance in paramagnetic systems, in particular those requiring long RF irradiation periods, such as cross-polarization (CP) and heteronuclear decoupling, which are inefficient because the RF field power does not dominate the PCS and CSA interactions. The most successful new techniques applied to paramagnetic SSNMR use high-power RF pulses, such as the spin-echo and transferred-echo double-resonance (TEDOR) sequences.[77]

For SSNMR using MAS and these efficient techniques, the same principles of solution NMR for PCS measurement can be applied. The PCS calculation for the solid state follows the same procedures, although sometimes deuteration, fast MAS or both are required. After the MAS averaging of the dipolar shielding, the NMR resonance will be centered at a new isotopically averaged position. In solution NMR the paramagnetic contribution will only arise from the paramagnetic center within the molecule. However, for the calculation of the PCS in the solid state one needs to take into consideration that the molecules are compacted and close in space. Therefore, the contribution to the PCS will arise from the paramagnetic center of the molecule (intramolecular), as solution NMR, as well as from the paramagnetic centers from the neighboring molecules (intermolecular) if the paramagnetic probe is close to the surface of the molecule. To determine the contributions from the intermolecular PCS the paramagnetic sample must be diluted with unlabeled diamagnetic molecule. However, the signal-noise to ratio will be worse, therefore a rational approach is required.[36]

2.3 Electron Paramagnetic Resonance

Electron Paramagnetic resonance (EPR) spectroscopy is a method to study molecules or materials that have at least one unpaired electron. The basic concepts from NMR also apply to EPR, however EPR studies the electron spins instead of the nuclear spins. Many of the atomic elements have unpaired electrons, thus they are suitable to be studied by EPR. This spectroscopy is extremely sensitive to the variations of the environment around the unpaired electron. The basic theory of EPR spectroscopy has been described in many textbooks.[78,79]

From the practical point of view of EPR spectroscopy, all paramagnetic transition metal complexes can be divided into two classes: low-spin (LS) ($S = ½$, e.g. copper(II) complexes) and high-spin (HS), with $S > 1/2$. HS complexes can be separated into half-integer spin numbers ($S = 3/2, 5/2$, etc., known as Kramers, e.g. HS cobalt(II) complexes) and integer-spin numbers ($S = 1, 2$, etc., known as non-Kramers, e.g. nickel(II) complexes). For both Kramers and non-Kramers systems the unpaired electrons' magnetic moment, at zero field, lifts the degeneracy of the spin states and separates them. This separation originates half-integer spin states doublets for the Kramers systems and integer spin doublets and one singlet $|0\rangle$ for the non-Kramers systems. The EPR properties of these three categories are distinctly different and require different experimental approaches.[80]

The EPR spectrum of a d block transition-metal ion in a ligand field can be discussed using the spin Hamiltonian $\boldsymbol{H}$ describing its electronic spin system and its interactions (the nuclear Zeeman interaction has a negligible energy):

$$\boldsymbol{H} = \boldsymbol{H}_{LF} + \boldsymbol{H}_{SO} + \boldsymbol{H}_{Z} + \boldsymbol{H}_{HF} + \boldsymbol{H}_{SS} \qquad \text{(Eq. 2.14)}$$

Here, $\boldsymbol{H}_{LF}$, representing the interaction of the ion with the ligand field, is the largest contribution to $\boldsymbol{H}$, giving rise to a splitting in the energy levels of the order of 10^4 cm^{-1}. The first term of the smaller contributions, $\boldsymbol{H}_{SO}$, represents the spin-orbit coupling interaction. This represents the coupling between the electronic magnetic moment, $\boldsymbol{\mu}_S = -\mu_B g_e \boldsymbol{S}$, which is proportional to the effective spin $\boldsymbol{S}$ through the free electron g factor g_e, with the orbital contribution to the magnetic moment, $\boldsymbol{\mu}_L = -\mu_B \boldsymbol{L}$, where $\boldsymbol{L}$ is the operator representing the total orbital angular momentum, that is orientation-dependent. In the limit of LS coupling, the total angular momentum operator $\boldsymbol{J}$ can be defined as the sum of $\boldsymbol{L}$ and $\boldsymbol{S}$, $\boldsymbol{J} = \boldsymbol{L} + \boldsymbol{S}$, and total magnetic moment operator $\boldsymbol{m}$ as $\boldsymbol{m} = -\mu_B (\boldsymbol{L} + g_e \boldsymbol{S})$. Therefore, the vectors representing the total angular momentum and total magnetic moment are not parallel. The orbital angular momentum is quantized and represented by the quantum number L which takes integer values. Each level L includes a manifold of $2L + 1$ states that are labeled by the orbital magnetic quantum number M_L which takes integer values from $-L$ to $+L$.

For LS coupling and assuming that Hund's rules are obeyed,

$$\boldsymbol{H}_{SO} = \lambda\, \boldsymbol{L}\cdot\boldsymbol{S} \qquad \text{with} \qquad \lambda = \zeta/2S \qquad \text{(Eq. 2.15)}$$

where the many-electron SO coupling parameter λ is given in terms of the single-electron SO coupling parameter ζ, where for the free ion the + sign applies to a less than half filled shell and the - sign for a shell that is more than half filled. Values of ζ have values between 10^2 and 10^3 cm^{-1} for 3d metal ions, making the SO interaction the largest after the ligand-field splitting.

The electronic Zeeman interaction, $\boldsymbol{H_Z}$**,** which is the sum of the orbital Zeeman $\mu_B \boldsymbol{B_0} \cdot \boldsymbol{L}$, and spin Zeeman $\mu_B g_e \boldsymbol{B_0} \cdot \boldsymbol{S}$, terms:

$$\boldsymbol{H_Z} = \mu_B \boldsymbol{B_0} (\boldsymbol{L} + g_e \boldsymbol{S}) \qquad \text{(Eq. 2.16)}$$

The Zeeman splitting varies with the applied magnetic field B_0, where $\mu_B B_0$ takes a maximum value of 13 cm^{-1} at the maximum field of 28.18 T that is currently commercially available ($\mu_B g_e B_0 = 26$ cm^{-1}).

The anisotropic electronic Zeeman interaction, $\boldsymbol{H_Z}$**,** can be also represented as:

$$\boldsymbol{H_Z} = \mu_B \boldsymbol{B_0} \cdot \boldsymbol{g} \cdot \boldsymbol{S} \qquad \text{(Eq. 2.17)}$$

where $\boldsymbol{g}$ (g-factor) is a dimensionless anisotropic tensor that describes the interaction of the coupled electron magnetic and angular momentum with the external magnetic field. To date, the g-factor for the free electron (g_e) is the most accurately measured constant. In molecules, the observed deviation of the values of the $\boldsymbol{g}$ tensor components from g_e is caused by the spin-orbit coupling between the ground and excited states, which reflect the molecular orbitals or chemical environment. One of the most important information that can be extracted from EPR spectra are the $\boldsymbol{g}$ tensor components.

These molecular environments are generally anisotropic, meaning that the EPR spectra and respective parameters depend on the orientation between the applied magnetic field and the molecular framework represented by the main axis of $\boldsymbol{g}$ tensor, in a similar way to what has been described above for the χ tensor. Therefore, the g-factor can be defined by three main components, g_x, g_y and g_z. Depending on the molecular symmetry, three possibilities may occur. When $g_x = g_y = g_z$, the environment is isotropic. In axial symmetry, $g_x = g_y \neq g_z$, and it has two components, $g_\parallel$ and $g_\perp$, where $g_\parallel = g_z$ and $g_\perp = g_x = g_y$. Finally, when $g_x \neq g_y \neq g_z$, the environment is rhombic and defined by these three components. The g-values will reflect the environment of the system and the type of EPR spectra obtained. The spectral features can be visualized schematically in **Figure 2.7**.

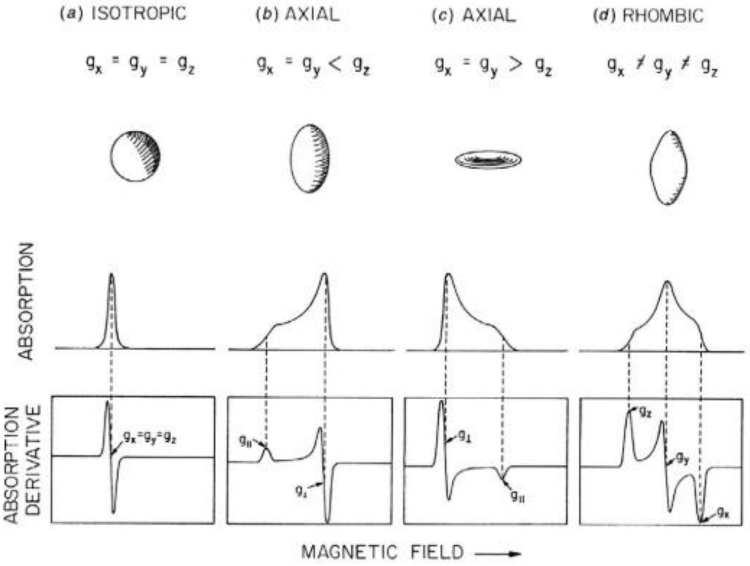


Figure 2.7: Schematic representation of the **g** tensor and the consequential EPR spectra. The upper solid bodies show the shapes associated with isotropic (a), axial (b, c) and rhombic (d) magnetic moments. Underneath are shown the absorption curves (top) and the corresponding EPR derivative curves (bottom).

The term HHF is the hyperfine interaction Hamiltonian:

$$\boldsymbol{H_{HF}} = \boldsymbol{S} \cdot \boldsymbol{A} \cdot \boldsymbol{I} \qquad \text{(Eq. 2.18)}$$

which describes the interaction between the electronic magnetic moment, $\boldsymbol{\mu_S} = - \mu_B\, g_e \boldsymbol{S}$ and the nuclear magnetic moment, $\boldsymbol{\mu_N} = - \mu_B\, g_N \boldsymbol{I}$. The hyperfine tensor ***A*** is also anisotropic. In EPR, the electron senses magnetic moments of nuclei in the molecule with non-zero spin, either through bonds (Fermi contact interactions) or through space (dipolar interactions), included in ***A***. When the unpaired electron senses the nuclear spin of its own element, it is called hyperfine coupling, present as splittings in the EPR spectrum. When the unpaired electron senses the nuclear spin of neighboring nuclei, then it is called super-hyperfine coupling, which is responsible for further splitting the hyperfine resonances in the EPR spectrum. The calculation of these hyperfine coupling constants is useful to determine the environment around the paramagnetic center.

Finally, the $\boldsymbol{H_{SS}}$ term is the zero-field splitting (ZFS) interaction, represented by:

$$\boldsymbol{H_{SS}} = \boldsymbol{S} \cdot \boldsymbol{D} \cdot \boldsymbol{S} \qquad \text{(Eq. 2.19)}$$

which describes the interaction between two or more electronic magnetic moments present in the molecule hyperfine through the ZFS tensor ***D***, which is also anisotropic. The main difference between the LS and HS complexes, as defined for EPR, lies in the phenomenon of ZFS, which appears only for $S \geq 1$, i.e., in high-spin states. The mixing of excited electronic states with the ground state causes the different M_S spin sublevels to split in the absence of a magnetic field in conditions of low symmetry. ZFS is mediated by spin–orbit coupling and its magnitude varies strongly between different metal ions and different possible geometries. Therefore, the ZFS magnitudes vary accordingly, from very small values for highly symmetric complexes of Mn(II), where it is of the order of 10^{-2} cm^{-1}, to octahedral complexes of cobalt(II), where it reaches values of the order of 10^{2} cm^{-1}. The ZFS of the magnitude lying at the lower end of the above range is usually easy to measure by conventional EPR, usually for non-Kramers (integer spin) species with very small ZFS. However, when the ZFS is near or exceeds the conventional EPR quantum energies (~0.3 cm^{-1} for X-band and ~1.2 cm^{-1} for Q-band), its determination becomes very difficult, and often impossible to determine using conventional methods, such as for Kramers (half-integer spin) species. Then the use of EPR at high frequencies and fields (HFEPR) becomes necessary.[81]

As any other spectroscopy, the absorption of electromagnetic radiation is required, and for EPR it occurs in the microwave region. The classic EPR experiment requires irradiating the paramagnetic sample with a fixed frequency of microwaves while sweeping the magnetic field. During the field sweep, the absorption of energy will happen when the energy separation between spin states which are allowed by selection rules matches the irradiating microwave frequency. It is impractical to sweep all the frequencies because it would require sweeping GHz of frequencies. The EPR experiments can be acquired at different magnetic fields, which are called bands. In the biochemistry field, the X-band is commonly used and has a central frequency at 9.5 GHz. However, in some cases, Q-band and HF EPR need to be used, as described above.

2.4 X-ray Crystallography

For many decades, X-ray crystallography has been a formidable tool to extensively characterize the structure of molecules. It has been used to understand the chemical basis for the molecular mechanisms of protein function. To successfully determine the crystal structure of a molecule, using X-ray crystallography, it is mandatory to prepare crystals of that molecule. Furthermore, these crystals need to be single and of a certain size, however small or twinned crystals may be suitable for structure determination if enough effort is employed. In this thesis the crystallized molecules were proteins.

A crystal is an ordered array of molecules that are held together by noncovalent interactions.[82] For protein crystals this means, the protein molecules need to rearrange themselves in a repeated pattern to produce a crystal. Obtaining a protein crystal is a non-linear process that involves many variables and, although many decades of research are present, the process of crystallization is not fully understood[33]. There are many processes to crystallize a protein, e.g., vapor diffusion and dialysis, and there are many physicochemical factors that influence the process of crystallization, e.g. temperature, pH, type of buffer and precipitant.[83]

Once a crystal is obtained, the crystal is irradiated with X-rays. These incident X-rays are diffracted by the electron cloud of the molecule, thus producing the diffraction pattern, which contains all structural information regarding the electron density of each atom of the molecule. To extract this information, it is necessary to know the intensity of the reflections and the angle where the X-ray was diffracted. The intensity of each reflection is easy to determine by simply measuring how dark the reflection is; however, determining the phase of the diffracted X-ray is the problem, the phase problem. There are several strategies that are used to solve the phase problem, and in this thesis two strategies were used, the Molecular Replacement (MR) and Single-wavelength Anomalous Dispersion (SAD).

The MR strategy requires a homologous molecule whose phases are already calculated. These known phases are used as estimates to determine the phases from the target molecule and determine the structure. This strategy is best applied to molecules with a high degree of homology or even for the same molecule bound to different ligands.

The SAD strategy requires the presence of a heavy atom. This heavy atom will generate an anomalous contribution to the diffraction pattern, with respect to the lighter atoms, and create a unique X-ray phase. This contribution originates from the absorption and emission of X-rays by the heavier atom. The absorption of X-rays will depend on the heavy atom element, and for the absorption to happen it is necessary to tune the X-ray source, which is easily done in a beamline. The anomalous contribution is used to calculate the substructure of the molecule, in this case it will determine the position of the anomalous scatterer in the crystalline structure.

3

Chapter 3 – Structural studies of cobalt(II)-DM-hCAII and its thiocyanate adduct

3.1 Introduction

The development of biomolecular NMR has included a parallel evolution of NMR of paramagnetic biomolecules. Many proteins have structural or functional metal ions as cofactors (metalloproteins), which are diamagnetic (e.g. zinc(II)) or paramagnetic (e.g. cobalt(II)). Proteins containing diamagnetic metal ions can be substituted by paramagnetic metal ions, whose presence is useful to extract further structural information from the system in study.

The enzyme Carbonic Anhydrase (CA), from mammals, has been extensively studied, both from the structural and functional points of view. Carbonic Anhydrase has a zinc(II) ion in its active center[12], which can be replaced by a variety of metal ions, such as paramagnetic cobalt(II), nickel(II) and copper(II). Paramagnetic Carbonic Anhydrase derivatives have been studied throughout several decades, and they are still being used to solve questions regarding the origin of paramagnetic NMR (pNMR) shifts[84]. In particular, the useful electronic and magnetic properties of high-spin cobalt(II) make it a powerful UV-vis, pNMR and EPR spectroscopic probe in biological systems. Besides this, the cobalt(II)-CA enzyme displays catalytic properties, including their pH dependence, similar to those of the native enzyme.[48]

As discussed in chapter 2, cobalt (II), with a d^7 electronic configuration, has commonly a high-spin configuration when bound to proteins, with an electronic spin $S = 3/2$ (see **Figure 3.1**). The low symmetry environment of cobalt(II) in proteins (e.g. in a distorted square pyramidal geometry) removes its orbital degeneracy, making the excited states close in energy to the ground state. This proximity in energy, causes the Orbach electron spin relaxation mechanism to be very efficient and, depending on the geometry of the cobalt complex, and thus on the closeness of the low-lying excited states to the ground state, the electron spin relaxation time (T_{1e}) ranges from 1 to 10 ps. Thus, cobalt(II) is a highly valuable probe for protein NMR, originating large shifts on the observed nuclei, due to the large magnetic susceptibility anisotropies ($\Delta\chi$),[46] which induce large PCS and RDC values. Furthermore, the cobalt(II)'s short electron spin relaxation time in distorted protein environments does not severely broaden the paramagnetic resonances. Paramagnetic NMR has been applied in solution to study cobalt(II)-substituted proteins, such as Cu_2Co_2SOD[85,86] and the catalytic domain of cobalt(II)-MMP-12[49,87], as well as in the solid state (SSNMR) for the latter.[88,89] 2D NMR studies in solution of cobalt(II)-CAs started in the 1980's, with the ClO_4^- and NO_3^- adducts of cobalt(II)-bCAII[90] and the cobalt(II)-hCAII-NCS^- adduct[91], followed by recent studies of the cobalt(II)-hCAII adducts of oxalate and furosemide at different pH values[46], and by SSNMR of its furosemide and sulpiride adducts.[23]

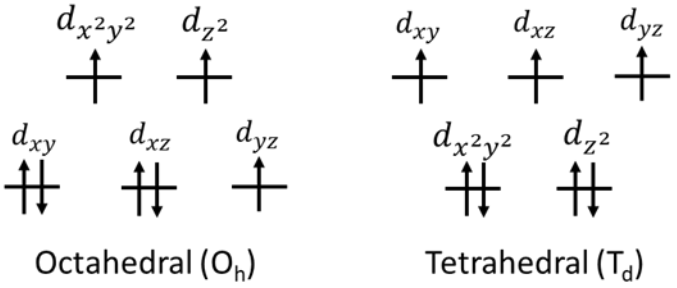


Figure 3.1: Electronic configuration of high-spin cobalt(II) in idealized octahedral and tetrahedral geometries.

Although some studies are decades old, some questions remain unanswered. The UV-vis d-d electronic absorption spectra of cobalt(II)-bCAII and cobalt(II)-hCAI and cobalt(II) are dramatically pH dependent in the 5.8-9.5 pH range. For cobalt(II)-bCAII and cobalt(II)-hCAI there is an increase in the d-d electronic absorption spectra band at 640 nm which was reasonably fitted assuming a double deprotonation step[92]. For cobalt(II)-hCAII a similar behavior is observed and the fitting assumes a single deprotonation step[48], however, the data could also be fitted with a double deprotonation step. The first deprotonation corresponds to a ionizing water molecule present in the coordination sphere of CA and has a pK_a value of 5.9 and 6.9 for cobalt(II)-bCAII and cobalt(II)-hCAI, respectively[92], and 7.35 for cobalt(II)-hCAII.[48] The second ionizable group is the His64 for bCAII and His64 or His200 for hCAI, where His64 is responsible to transfer the proton to the bulk solvent and His200, in hCAI, is close to the active cavity. This second ionizable group has a pK_a 8.7 and 8.0 for cobalt(II)-bCAII and cobalt(II)-hCAI, respectively.[92] The low pK_a values results from the partially hydrophobic nature of the active site cavity. Since there are two ionizable groups, it means there are four dissociation microconstants, as seen in **Figure 3.2**, however at the macroscopic level only two apparent constants are measurable.

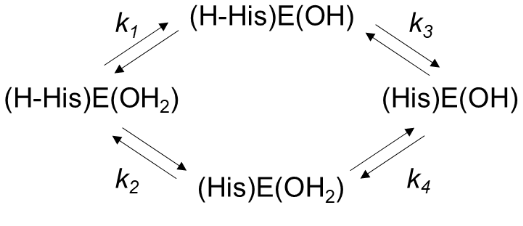


Figure 3.2: Scheme of ionization microconstants of the enzyme (E) CAs.

The overall process in the coordination sphere of CA can be described as an equilibrium between a five-coordinate species and a pseudo-tetrahedral species containing H_2O as the fourth ligand in a pH dependent equilibrium with the species with H_2O ionized to OH^-:

$$[\mathrm{Co(II)N_3(H_2O)_2}] \overset{-\mathrm{H_2O}}{\Longleftrightarrow} [\mathrm{Co(II)N_3(H_2O)}] \overset{-\mathrm{H^+}}{\Longleftrightarrow} [\mathrm{Co(II)N_3(OH^-)}] \qquad \text{(Eq. 3.1)}$$

Studies of the ligand field d-d transitions of cobalt(II)-CA and of its inhibitor derivatives made it possible to group the spectra of all the derivatives into three different classes: (i) spectra characterized by intense d-d transitions in the visible region (ε_{max} > 300 M^{-1} cm^{-1}), which are assigned to pseudo-tetrahedral species; (ii) spectra with low-intensity bands in the visible region (ε_{max} < 200 M^{-1} cm^{-1}), which are indicative of five-coordinate complexes; (iii) spectra with features intermediate between those of the previous two classes, which are consistent with equilibria between four and five coordinate species in solution.[48] The pH dependence of the UV-vis spectra cobalt(II)-bCAII and cobalt(II)-hCAII show that at low pH, the cobalt(II)-hCAII has predominantly five-coordinated species and cobalt(II)-bCAII has predominantly pseudo-tetrahedral coordination species, which is consistent with their relative pK_a values. [48]

The addition of inhibitors (In) to cobalt(II)-CA can lead to pseudo-tetrahedral species, five-coordinate species, or equilibria between the two:

$$[\mathrm{Co(II)N_3(H_2O)(In)}] \overset{-\mathrm{H_2}O}{\Longleftrightarrow} [\mathrm{Co(II)N_3(In)}] \qquad \text{(Eq. 3.2)}$$

The intensity of the absorption spectrum indicates whether the equilibrium is shifted toward four- or five-coordination or lies in between. Therefore, based on the observed ranges of ε_{max} values, inhibitors like SCN^-, I^-, formate, acetate, oxalate, nitrate and 1,2,3-triazole (ε_{max} < 200 M^{-1} cm^{-1}) give rise to five-coordinate species; sulfonamides, NCO^-, CN^-, aniline, imidazole (high pH), and 1,2,4-triazole (ε_{max} > 300 M^{-1} cm^{-1}) give rise to pseudo-tetrahedral species; finally HCO_3^-, N_3^-, F^-, Cl^-, Br^-, phosphate, benzoate and imidazole (low pH) (ε_{max} = 200 -300 M^{-1} cm^{-1}) lead to four-to-five coordinate equilibria[48].

The shape of the EPR spectrum has also been related to the coordination number of the species present[44]. The study of cobalt(II)-bCAII bound to several inhibitors led to a classification of type A, B and C ligands according to the observed EPR spectral features. Type A spectra are characterized by two transitions with g values around 6.1-6.8 and 2.3-2.9, and in some cases a third weaker transition at g = 1.6-1.8. The signal with g ~ 6 is generally broad and in the case of the I^- derivative, a hyperfine structure is observed. Some examples of type A ligands are acetate, nitrate, thiocyanate, azide and iodide. Type B spectra, show a sharp feature in the low-field region of the spectrum, with g values in the ranges 5.8-6.2, 2.2-2.8, and 1.5-1.8. The ligands able to give a type B spectrum are acetazolamide, sulfonamides, cyanate and cyanide. Class C spectra are characterized by a single quasi-symmetrical line centered at $g_\perp$ = 3.84, as $g_\parallel$ is too broad to be detected. Spectra of A and B types are obtained from solutions containing close to stoichiometric amounts of inhibitor, while in the presence of a large excess (> 100:1 molar ratio) of inhibitor a type C spectrum is sometimes obtained. The type A ligands correspond to those that lead to a 5-coordinated species, the type B to a 4-coordinated one and the type C arises from the excess of ligand, such as iodide, thiocyanate and imidazole, that may be able to bind in a secondary site inside the active site pocket, thus further modifying the spectral features.[44]

Finally, water proton T_1^{-1} relaxation rate measurements at various magnetic fields (NMRD = nuclear magnetic resonance dispersion) established that pseudo-tetrahedral species (type B) display T_{1e} values in the 10^{-11} s range, whereas five-coordinate species (type A) have shorter T_{1e} values around 10^{-12} s, leading to sharper NMR spectra[93].

To date, cobalt(II)-hCAII and cobalt (II)-bCAII adducts of type A and B ligands have been explored by NMR[46,90,91], specifically oxalate, perchlorate and nitrate, which are type A ligands, and sulfonamides that are type B ligands. The type A and B ligands have very different effects on the susceptibility tensor anisotropy ($\Delta\chi$) of the cobalt(II) site in the proteins. Type A ligands produce $\Delta\chi$

values that are at least twice the absolute value of those induced by type B ligands, with the correspondingly larger PCS values. However, there is so far no study of the NMR effect of type C ligands. One example of a type A and type C ligand is thiocyanate. When thiocyanate is added at the same concentration as the cobalt(II)-CA protein, it generates a type A EPR spectrum and when thiocyanate is in excess, the EPR spectral features correspond to a type C spectrum.[44]

To investigate the origin of these differences in the observed EPR spectra, and to understand if the type C ligands produce different $\Delta\chi$ values from the others, a double mutant of hCAII (DM-hCAII) capable of binding only one metal ion is used. In this study, we aim at determining the magnetic susceptibility tensor of cobalt(II)-DM-hCAII in the presence of increasing amounts of thiocyanate and investigate these differences using NMR, EPR, CD and X-ray-crystallography. Furthermore, this project also attempts to determine if DM-hCAII is suitable to be used as a model protein to study paramagnetic effects. The previously mentioned study[44] uses bCAII, while in this project the explored protein is the human isoform because its expression and purification systems are already implemented in our laboratory.

3.2 Materials and methods

3.2.1 Double mutant construct

The wild-type form of hCAII was mutated in two positions, H3N and H4N, to avoid the binding of transition metals in the secondary site. The sequence of double mutant human Carbonic Anhydrase 2 (DM-hCAII) is:

MS**NN**WGYGKHNGPEHWHKDFPIAKGERQSPVDIDTHTAKYDPSLKPLSVSYDQATSLRILN
NGHAFNVEFDDSQDKAVLKGGPLDGTYRLIQFHFHWGSLDGQGSEHTVDKKKYAAELHLVHWNTKY
GDFGKAVQQPDGLAVLGIFLKVGSAKPGLQKVVDVLDSIKTKGKSADFTNFDPRGLLPESLDYWTYP
GSLTTPPLLECVTWIVLKEPISVSSEQVLKFRKLNFNGEGEPEELMVDNWRPAQPLKNRQIKASFK.

The human carbonic anhydrase II (hCAII) double mutant (H3N, H4N) was cloned following a standard PCR protocol. For the H3N and H4N mutations the used primers were: forward: 5’- GA AGG AGA TAT ACC ATG GCC AAT AAC TGG GGG TAC GGC AAA CAC -3’ and reverse: 5’ - GTG TTT GCC GTA CCC CCA GTT ATT GGC CAT GGT ATA TCT CCT TC - 3’.

3.2.2 Expression and purification of DM-hCAII

The expression and purification of DM-hCAII was based on the published protocol.[23] However, the protein expression was unrestricted, and the labeling efficiency was close to 33 %. In addition, during the cellular growth, the cell strain (*Escherichia coli* BL21-CodonPlus (DE3)-RIPL) presented growth problems. Therefore, the protocol was optimized to express DM-hCAII.

The expression vector, pCAM coding the DM-hCAII, was inserted in competent *Escherichia coli* BL21 (DE3) using a standard heat shock protocol. The culture was in LB-Agar supplemented with 1 % glucose and ampicillin. The plates were incubated at 37 °C overnight. The following day a successful transformant colony was selected and resuspended in 500 µL of sterile water. 50 µL of culture were plated into 2 plates of LB-Agar supplemented with 1 % glucose and Ampicillin.

The plates were incubated overnight at 37 °C. In the subsequent day, the cells were scraped from one plate to 1 L of LB media supplemented with 1% glucose and ampicillin. The $OD_{600\ nm}$ was measured until 0.6 and then the cultures of LB were centrifuged at 4000 rpm for 15 minutes. The pellet was resuspended in M9 media, supplemented with $MgSO_4$, $CaCl_2$, 0.5 mM $ZnSO_4$, 1.2 g/L of ^{15}N-enriched $(NH_4)_2SO_4$, 3g/L glucose and ampicillin. The culture was incubated at 37 °C 160 rpm for 30 minutes. Afterwards, 1 mM ITPG was added and the temperature decreased to 25 °C. The culture was incubated overnight. The protein expression was halted by centrifuging the cultures at 7500 rpm for 20 minutes. The harvested cell pellet was stored at -20 °C.

The cell pellet was resuspended in a final volume of 70 mL of buffer, 20 mM Tris-SO_4, 0.5 mM $ZnSO_4$ pH 8. The cells were sonicated for 30 seconds with a resting period on ice for 3 minutes and this was repeated for a total of 10 times. The lysate was ultracentrifuged at 40000 rpm for 40 minutes and the supernatant was filtered with a 0.45 micra filter.

The lysate was firstly purified in a 5 mL Histrap FF, using an AKTA Prime System, previously equilibrated with 20 mM Tris-SO_4, 0.5 mM $ZnSO_4$ pH 8. The elution of the protein is done with a gradient of 20 column volumes from 0 to 50 % buffer containing 20 mM Tris-SO_4, 500 mM imidazole, 0.5 mM $ZnSO_4$ pH 8. Eluted fractions were analyzed by SDS-PAGE, and the fractions containing hCAII were pooled. Pooled fractions were concentrated using a 10,000 MWCO Amicon centrifugation device to 10 mL.

As a second purification step a 320 mL Superdex 75 pg 26/60 size exclusion column was used in an AKTA Prime system, with a running buffer of 50 mM Sodium Phosphate pH 7, at 2.5 mL/min. Purified protein was pooled and stored at 4 °C.

3.2.3 Demetallation and metalation protocol of DM-hCAII

For demetallation of DM-hCAII, a solution containing 200 mM sodium phosphate, 50 mM pyridine-2,6-dicarboxylic acid (PDA), pH 7, was added to the protein solution and incubated overnight at 4 °C. The apo-hCAII was buffer exchanged to 10 mM HEPES, pH 6.8, using 10,000 MWCO Amicon centrifugation devices. Protein samples were in a water buffered solution [10 mM HEPES, pH 8] with 10 % D2O, with a protein concentration ranging from 350 µM.

The paramagnetic cobalt(II)-DM-hCAII samples were prepared by titrating the apo-DM-hCAII (350 µM in 10 mM HEPES, pH 6.8) sample with cobalt (II) sulfate solution. The procedure was followed by 1D ^1H and 2D ^1H-^{15}N HSQC spectra by monitoring the disappearance of the resonances of the apo protein and the appearance of the paramagnetic resonances of the metalated protein at different chemical shift values. The titration was stopped when the resonances, which correspond to the residues close to the paramagnetic center, fully disappeared and when the new resonances, at a different chemical shift, did not increase in intensity any longer. The same procedure was followed for the diamagnetic reference, where apo-DM-hCAII (350 µM in 10 mM HEPES, pH 6.8) was titrated with zinc(II) sulfate until the resonances matched to the assignment of zinc(II)-WT-hCAII. Buffer for both protein solutions was exchanged to 10 mM HEPES pH 6.3 and a solution of sodium thiocyanate was added to the protein sample until it reached a final concentration of 500 mM sodium thiocyanate.

3.2.4 NMR measurements

Solution 2D ^{1}H-^{15}N HSQC NMR spectra of ^{15}N isotopically enriched zinc(II) and cobalt(II) substituted DM-hCAII, for the evaluation of PCS, were recorded on a Ascend™ 500 MHz BRUKER NMR Spectrometer, equipped with a PRODIGY CryoProbe and on a Ascend™ 1.2 GHz Bruker NMR Spectrometer, equipped with a TCI CryoProbe. Protein samples were prepared in a water buffered solution (10 mM HEPES, pH 6.3) and were titrated with sodium thiocyanate up to 475 mM.

All the spectra were processed with Bruker TopSpin software and analyzed with the program CARA (the program can be downloaded for free from www.nmr.ch).[94]

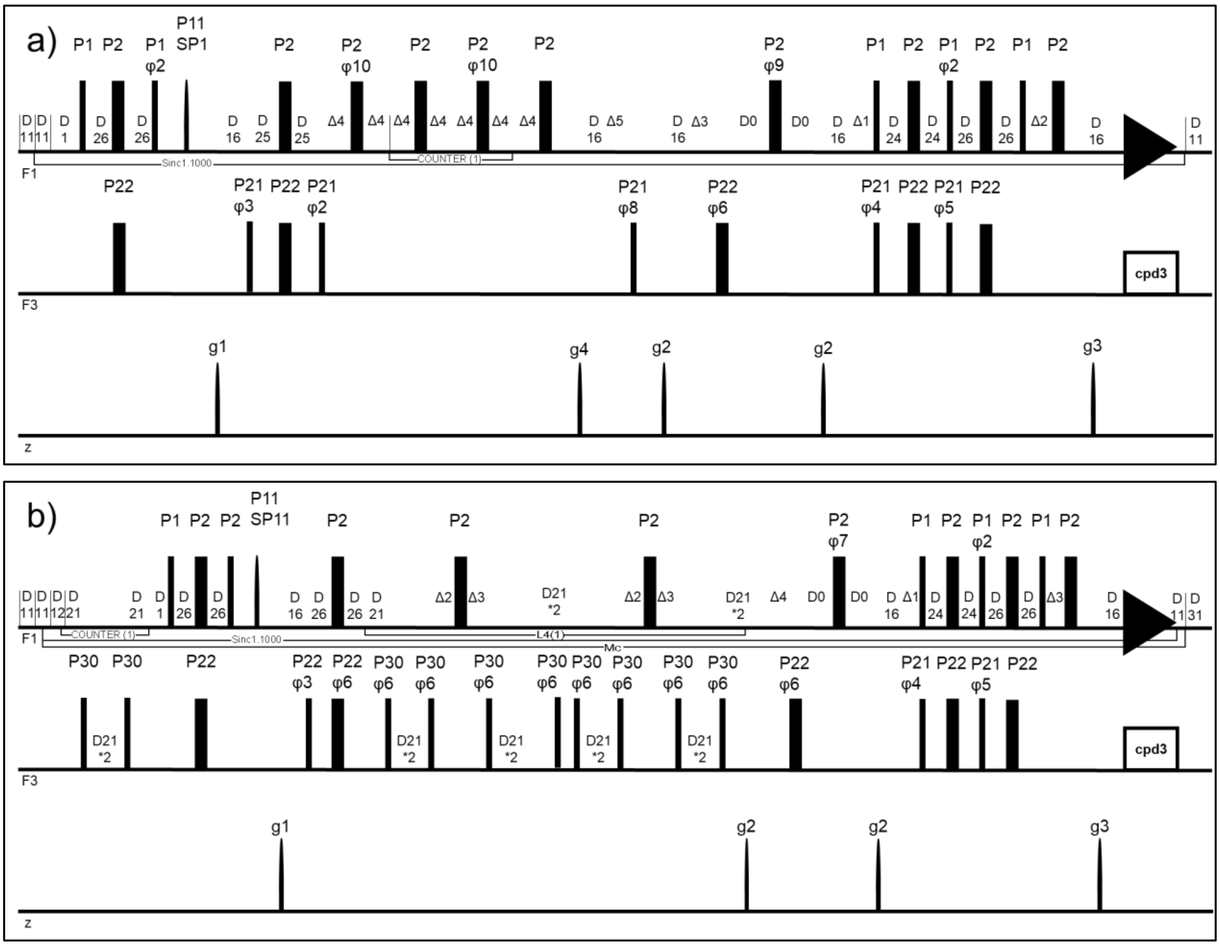
The assignments of the diamagnetic zinc(II) and paramagnetic cobalt(II) spectra were previously published.[46] The assignment of both metal forms in the presence of sodium thiocyanate was done by following the observed shifts. The PCSs were calculated from the difference of the chemical shifts between the paramagnetic cobalt (II) and diamagnetic zinc(II) spectra of the protein.

The ^{1}H-^{15}N RDC were measured at 298 K at 1.2 GHz Bruker NMR Spectrometer, equipped with a TCI CryoProbe, for both metallic forms reacted with 700 µM sodium thiocyanate.

The $\Delta\chi$ tensors were calculated using FANTEN[73] by fitting the PCS (and sometimes also RDC) data against the highest resolution (0.9 Å) structure of zinc(II)-WT-hCAII (PDB 3KS3)[12]. The quality of the fitting was evaluated using the Q-factor. The Q-factor is a quantification of the agreement between the experimental and back-calculated data, it ranges between 0 and 1 and the closer to 0 the better the agreement between the experimental and back-calculated data is. The Q-factor equations are defined in Chapter 2.2.5.

3.2.4.1 Relaxation Experiments

The experiments for the determination of the ^{15}N longitudinal and transverse relaxation rates (R_1 and R_2) and NOEs of the backbone ^{15}N nuclear spins were recorded at 298 K and at 700 MHz on ^{15}N enriched zinc(II)-DM-hCAII. The respective pulse sequences used, based on the ^{15}N-^{1}H HSQC pulse sequence, are represented in **Figure 3.3**[95,96]. The sequence for R_1 determination (**Figure 3.3 a**) used the inversion-recovery technique with a refocused INEPT sequence for the preparation period. The inversion-recovery relaxation times (T) for ^{15}N R_1 measurements ranged from 2 ms to 3000 ms, with a recycle delay of 3.5 s. Phase alternation of ϕ_1 alternates the sign of the ^{15}N z-magnetization, suppressing the contribution of steady-state magnetization to relaxation. Decoupling of the ^{1}H spins during the relaxation period T suppresses ^{1}H-^{15}N dipolar cross relaxation and ^{1}H-^{15}N dipole-dipole (DD)/^{15}N chemical shift anisotropy (CSA) relaxation interference. Decoupling used a composite pulse or a train of ^{1}H 180° pulses spaced at short 5-ms intervals. The delays used in the pulse sequence are as follow: d11 = 30 ms; d12 = 20 µs; d21= 450 µs; d24 = 2.65 ms; d25 = 2.65 ms; Δ1 = 1.4256 ms; Δ2 = 1.208 ms; Δ3 = 440.2 µs and Δ4 = 1.45 µs. The sequence for R_2 determination (**Figure 3.3b**) also used a refocused INEPT sequence for the preparation period, followed by a Carr-Purcell-Meiboom-Gill (CPMG) sequence with delays given by $T = 2n\tau_{cp}$, where 2n and τ_{cp} are the number of spin echo periods and the length of each period, respectively. The delay T ranged between 8.48 ms and 135.68 m. The delays used in the pulse sequence are as follow: d0 = 3 μs; d1 = 3.5 s; d11 = 30 ms; d12 = 20 μs; d21= 450 μs; d24 = 2.65 ms; d25 = 2.65 ms; d26 = 2.65 ms; d28 = 2.65 ms; d31 = 84.8 ms; Δ1 = 1.4256 ms; Δ2 = 1.208 ms; Δ3 = 440.2 μs and Δ4 = 1.45 μs. The steady-state NOE sequence (**Figure 3.3c**) obtains two spectra, one without and another with saturation of the ^{1}H spins during the initial part of the pulse sequence. Saturation used a composite decoupling pulse or a train of ^{1}H 180º pulses spaced at short 5-ms intervals. The delays used in the pulse sequence are as follows: d0 = 3 μs; d1 = 7 s; d11 = 30 ms; d12 = 20 μs; d16 = 200 μs; d24 = 2.65 ms; d26 = 2.65 ms; Δ1 = 1.44 ms; Δ2 = 1.21 ms; Δ3 = 1.46 ms. The hsqcnoef3gpsi spectrum was processed and split into NONOE and NOE using the integrated AU program (split 2) of TopSpin 4.1.1.



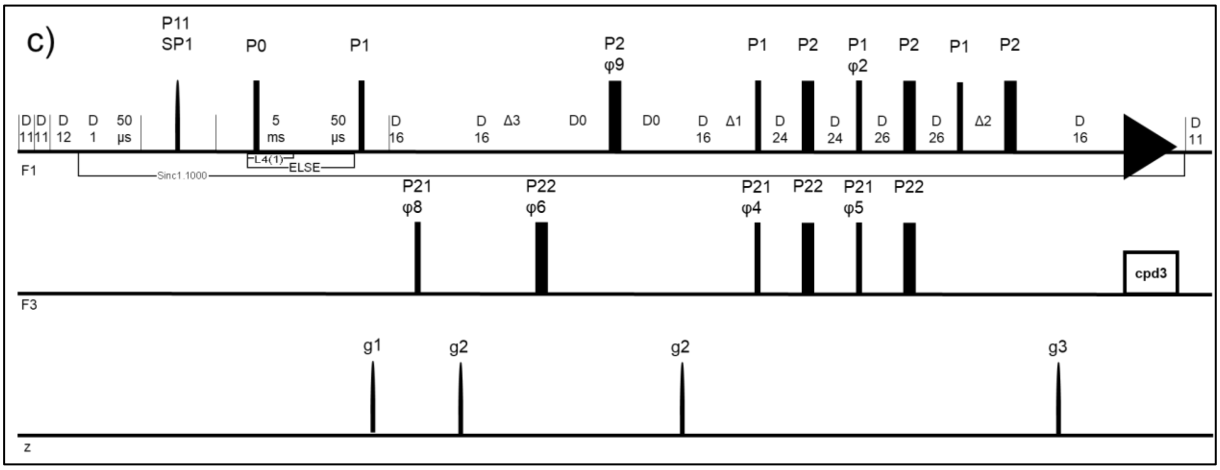


Figure 3.3: Pulse sequences for ^{15}N (a) T_1 (hsqct1etf3gpsi.2) (b) T_2 (hsqct2etf3gpsi.large) and (c) NOE (hsqcnoef3gpsi) spin relaxation measurements of ^{15}N backbone amide spin systems. R_1 and R_2 intensity decay curves are recorded by varying the ^{15}N relaxation period T in a series of two-dimensional experiments. The NOEs are measured by acquiring each spectrum with and without ^1H magnetization saturation. Narrow and wide bars represent 90° and 180° pulses, respectively. All pulses have *x*-phases unless indicated otherwise. Decoupling during acquisition used the cpdprg3 sequence. The phases used in the pulse sequences are: a) φ2 = 1; φ3 = 0 0 0 0 2 2 2 2; φ4= 0 0 2 2; φ5 = 3 3 1 1; φ6 = 0; φ8 = 1 3; φ9= 0 0 2 2; φ10 = 2; φ31 = 0 2 2 0 2 0 0 2; b) φ2 = 1; φ3 = 0 0 0 0 2 2 2 2; φ4= 0 0 2 2; φ5 = 3 3 1 1; φ6 = 0; φ7 = 3; φ8 = 1 3; φ9= 0 0 2 2; φ10 = 2; φ31 = 0 2 2 0 2 0 0 2; c) φ2 = 1; φ4= 0 0 2 2; φ5 = 3 3 1 1; φ6 = 0; φ8 = 1 3; φ9= 0 0 2 2; φ10 = 2; φ31 = 0 2 2 0 2 0 0 2.

The T_1 and T_2 were calculated by fitting the resonance intensities of the different residues using equation 3.3, where T is either T_1 or T_2.

$$f(t) = I * e^{(-\frac{t}{T})} \qquad \text{(Eq. 3.3)}$$

The R_1, R_2 and NOE of each residue were obtained using the Dynamics Center software from Bruker. From these, assuming an isotropic tumbling model for the protein[97], the global rotational correlation time (τ_c^{iso}) experienced by all NH group was calculated from those parameters using the equations 3.4-3.6:

$$\tau_C^{iso} = \left(\frac{1}{2\omega_N^2}\right) * \left[\frac{(6*R_2')}{R_1'} - 7\right]^{\frac{1}{2}} \qquad \text{(Eq. 3.4)}$$

$$R_1' = R_1 * \left[1 - 1.249 * \left|\frac{\gamma_N}{\gamma_H}\right| * (1 - NOE)\right] \qquad \text{(Eq. 3.5)}$$

$$R_2' = R_2 - 1.079 * \left|\frac{\gamma_N}{\gamma_H}\right| * R_1 * (1 - NOE) \qquad \text{(Eq. 3.6)}$$

3.2.4.2 Chemical shift perturbation analysis

The chemical shifts of zinc(II)-DM-hCAII's residues obtained from the 2D ^1H-^{15}N-HSQC spectra were used to calculate their chemical shift perturbation (CSP) index after the addition of sodium thiocyanate at different proportions. The CSP indexes (Δδ) were calculated by using the combined proton and ^{15}N chemical shifts, according to Equation 3.7[98–100]. The standard deviations were calculated by including all the combined chemical shifts (Equation 3.8), and then excluding the residues with a value three times higher than the standard deviation. Afterwards, all residues with a $\Delta\delta_{comb}$ higher than the standard deviation are considered as interacting residues.

$$\Delta\delta_{comb} = \sqrt{(\Delta\delta_H)^2 - 0.1 * (\Delta\delta_N)^2} \qquad \text{(Eq. 3.7)}$$

$$\sigma = \sqrt{\frac{\sum \delta_{comb}}{n-1}} \qquad \text{(Eq. 3.8)}$$

3.2.5 EPR measurements and spectral fitting parameters

Electron Paramagnetic Resonance (EPR) spectroscopy characterization of cobalt (II)-DM-hCAII in the ligand free form, and then with 1:1 and 1:1000 protein/SCN$^-$ ratios, was performed using a Bruker EMX spectrometer equipped with an Oxford Instruments ESR-900, with a microwave frequency of 9.40 GHz, in a continuous flow helium cryostat, at 4 Kelvin, and a high sensitivity perpendicular mode rectangular cavity. Protein samples were prepared to final concentrations of 50 μM. EPR spectra were fitted using the EasySpin[101] program to obtain the g-value and the hyperfine coupling constants parameters.

As explained in Section 2.3 when two unpaired electrons interact there is a phenomenon called ZFS that explains the dipolar spin-spin interaction and is described by Eq. 2.19. This equation can be further expanded to:

$$Hss = S \bullet \boldsymbol{D} \bullet S = D * \left(S_z^2 - \frac{1}{3}S(S+1)\right) + E * \left(S_x^2 - S_y^2\right) \qquad \text{(Eq 3.9)}$$

Where *S* is the effective spin, ***D*** is a traceless tensor and that is anisotropic, D and E are the axial and rhombic parameters, respectively, that describe the system.

$$D = \frac{3}{2}D_z \text{ and } E = \frac{1}{2}\left(D_x - D_y\right) \qquad \text{(Eq. 3.10 and 3.11)}$$

EasySpin[101] program is a software that uses the MATLAB suite's functions to solve numerical problems using several simulation algorithms. EasySpin includes algorithms that solve the Spin-Hamiltonian, thus determining the **g-values** and ***A*** values of the system. Furthermore, this package is able to plot the simulation for visual inspection, thus allowing an easier evaluation of the simulation. For cobalt(II)-DM-hCAII, the used function was "pepper" which is optimized for solid-state continuous wave EPR spectra simulation. The g-values for the difference samples of cobalt(II)-DM-hCAII were determined using an effective g-value (g_{eff}) of $^1/_2$. This approximation holds true because the spin-orbit coupling largely splits the Kramers' doublets and the energy of the X-band (0.3 cm^{-1}) is not sufficient to excite and therefore observe the transitions between the $^1/_2$ and $^3/_2$ states. This approximation was used for bCAII[44], regardless of the cobalt(II) geometries.

The rhombic (E) and axial (D) parameters were back-calculated using the Pilbrow equation[102] (Equation 3.12) using a developed in-house program that minimizes the parameters to reach the measured molecular g-values.

$$g_i' = g_i\left\{1 \pm \frac{2(\alpha_i D+3\gamma_i E)}{D'} \pm \left(\frac{g_i\mu_B B}{D'}\right)^2 \left(\frac{\alpha_i D+3\gamma_i E}{D'}\right)\left[1-\left(\frac{\alpha_i D+3\gamma_i E}{D'}\right)\right]\right\} \qquad \text{(Eq. 3.12)}$$

where **g_i'** is the effective molecular g-values (which for cobalt(II) the effective g-values vary according to **Figure 3.4**), **g_i** is the g-values of the system, without perturbation; **μ_B** is the Bohr's magneton; ***B*** the magnetic field; **α** and **γ** are parameters that describe the x, y, and z components of the system; and ***D'*** is a correlation between *E* and *D* parameters, that is described by Eq. 3.13. For the fittings **α** and **γ** had the values [-1/2, -1/2, 1] and [1/2, -1/2, 0], respectively.

$$D' = \sqrt{D^2 + 3E^2} \qquad \text{(Eq. 3.13)}$$

To calculate the ZFS values one can use the Equation 3.14

$$ZFS = 2\sqrt{D^2 + 3E^2} \qquad \text{(Eq.3.14)}$$

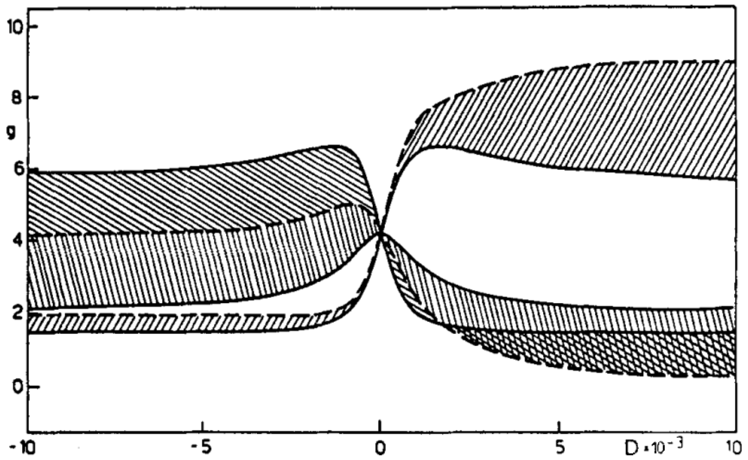


Figure 3.4: Range of effective molecular g-values[103] for distorted octahedral cobalt(II) complexes as a function of D for E/D = 0 (- - -) and E/D = 1/3 (—), \\\\g_x, **|||** g_y, ///g_z.

3.2.6 Circular dichroism experiments

Circular dichroism spectra were recorded on a Chirascan™ qCD spectrometer. Each CD spectrum represents the average of five scans obtained by collecting data at 1 nm bandwidth. Near-UV spectra were collected in the following regions: 200 - 250 nm using a cuvette with a 0.2 mm optical pathlength and 250 - 320 nm using a 2 mm optical pathlength cuvette. Protein samples of zinc(II)-WT-hCAII were diluted to 0.5 mg/mL and to 1 mg/mL for the 200 to 250 nm and the 250 to 320 nm measurements, respectively. The protein buffer consisted of 5 mM HEPES pH 6.3. The sodium thiocyanate ligand concentrations used were the following: 0 µM, 0.066 mM, 33.34 mM and 500 mM. Blank samples containing the ligand at different concentrations were used to subtract the protein's CD spectrum.

3.2.7 X-ray crystallography

The initial solutions of the metalated forms, cobalt(II)- and zinc(II)-DM-hCAII, had a concentration close to 15 mg/mL, in 10 mM HEPES, pH 6.8 buffer. Crystallizations were performed using the sitting drop vapor diffusion method by mixing equal amounts of sample volume and of a solution containing 100 mM HEPES pH 7.5, 2.9 M ammonium sulfate and 1 mM 4-hydroxymercuric benzoic acid sodium salt. cobalt(II)-DM-hCAII was crystallized using seeds from zinc(II)-DM-hCAII. Crystals for both metallic forms were soaked overnight with a solution containing 100 mM HEPES pH 6.3, 2.9 M ammonium sulfate, 1 mM 4-hydroxymercuric benzoic acid sodium salt and 500 mM sodium thiocyanate.

X-ray diffraction data obtained for the crystal of zinc(II)-DM-hCAII was collected at “xtal Crystallography Group” at UCIBIO, FCT-NOVA, using a D8 Venture diffractometer equipped with a copper anode, with a maximum resolution of 1.75 Å. X-ray diffraction data from the crystal of cobalt(II)-DM-hCAII was collected at ALBA synchrotron using the XALOC beamline, with a maximum resolution of 1.46 Å.

Several software packages were used for the structure determination. The datasets were processed using PROTEUM3 software[104], the electron density maps were obtained using CCP4.The refinement of the obtained structures was done using the PHENIX suite[105], water molecules were added between refinement cycles, and the models were subjected to adjustments using COOT[106].

3.3 Results and Discussion

3.3.1 EPR studies

As explained in the introduction, one example of type A and type C ligands is thiocyanate, as shown by the EPR spectra of the cobalt(II)-bCAII protein published by Bertini et.al.[44] in the presence of this ligand (**Figure 3.5**). When thiocyanate was added at the same concentration as the protein, it generated a class A spectrum, with $g_1 = 6.7$, $g_2 = 2.9$ and $g_3 = 1.8$. However, when thiocyanate was in large excess (1 M), the EPR spectral features corresponded to a class C spectrum, with $g_\perp = 4.0$.[44] Type A ligands like thiocyanate were interpreted to form a 5-coordinate adducts around the cobalt(II) ion, most probably distorted square pyramidal. At higher concentrations, thiocyanate becomes a type C ligand which could result from the binding of a second ion within the active cavity, thus further modifying the EPR spectrum of the cobalt enzyme. The higher thiocyanate concentration may form a more regular coordination polyhedrum, as reflected in the axial type epr spectrum.[44]

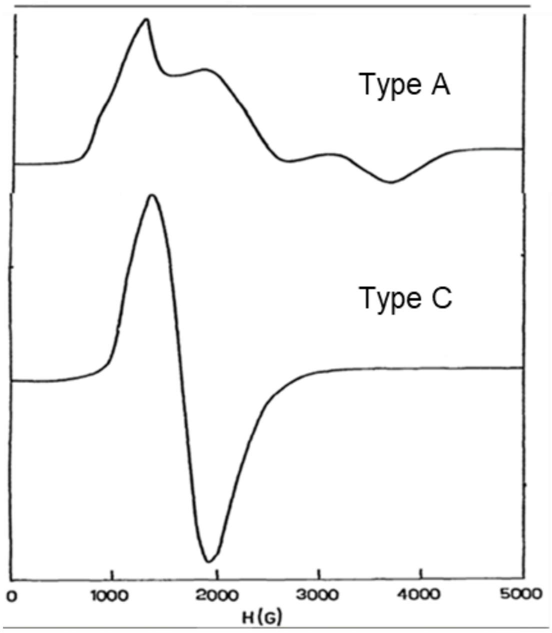


Figure 3.5: EPR spectra of glassy aqueous solution of cobalt(II)-bCAII bound to different thiocyanate ratios: A) 1:1 protein : ligand ratio; B) 1:1000 protein : ligand ratio. Spectra collected at 9.4 GHz at 4.2 Kelvin.[44]

In our study, to understand if the human isoform behaves in the same way as the bovine isoform (bCAII) in the presence of increasing amounts of thiocyanate, the cobalt(II)-DM-hCAII protein was used. **Figure 3.6**. shows the EPR spectra cobalt(II)-DM-hCAII in the absence and presence of sodium thiocyanate in 1:1 and 1:1000 protein: ligand ratios and **Figure 3.7** shows the respective simulations.

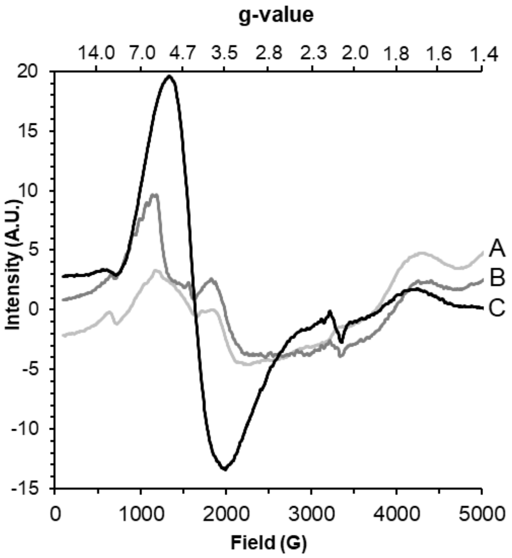


Figure 3.6: EPR spectra of glassy aqueous solution of cobalt(II)-DM-hCAII in the absence and presence of increasing concentrations of sodium thiocyanate. A) Free-cobalt(II)-DM-hCAII B) cobalt(II)-DM-hCAII-thiocyanate (1:1); C) cobalt(II)-DM-hCAII-thiocyanate (1:1000). Spectra collected at 9.4 GHz at 4.0 Kelvin.

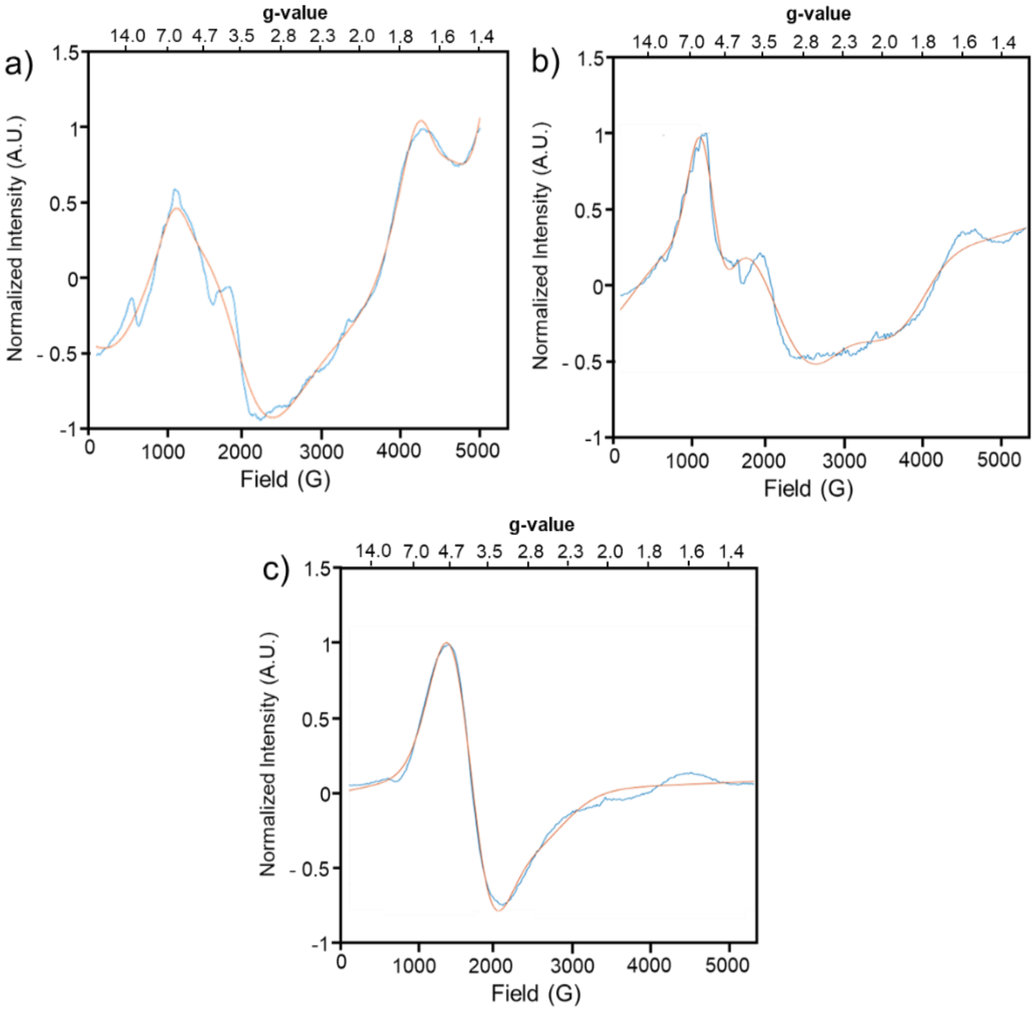


Figure 3.7: Glassy aqueous solution EPR spectra at 4.0 Kelvin of EPR spectra of cobalt(II)-DM-hCAII in the absence and presence of increasing concentrations of sodium thiocyanate (blue) and respective fittings (orange) obtained by EasySpin. A) Free-cobalt(II)-DM-hCAII B) cobalt(II)-DM-hCAII-thiocyanate (1:1); C) cobalt(II)-DM-hCAII-thiocyanate (1:1000).

Table 3.1: EPR values and calculated parameters for the different cobalt(II)-DM-hCAII samples.

Cobalt(II)-DM-hCAII-Free						
	g-values	A (MHz)	Weight	D (cm^{-1})	E (cm^{-1})	ZFS (cm^{-1})
$g_{\|\|(1)}$	4,997	113,94	0,52	34,23	0	68.5
$g_{\perp(1)}$	1,972	295,65				
$g_{\|\|(2)}$	4,757	5,35	0,32	83,36	0	167.0
$g_{\perp(2)}$	1,096	123,21				
Cobalt(II)-DM-hCAII-SCN (1:1.3)						
	g-values	A (MHz)	Weight	D (cm^{-1})	E (cm^{-1})	ZFS (cm^{-1})
g_1	1,62	0,01	0,49	31,3	-8,28	34.4
g_2	2,772	154,11				
g_3	6,81	467,44				
$g_{\|\|}$	4,452	26,61	0,06	79,69	0	159.0
$g_{\perp}$	0,977	448,89				
Cobalt(II)-DM-hCAII-SCN (1:1357)						
	g-values	A (MHz)	Weight	D (cm^{-1})	E (cm^{-1})	ZFS (cm^{-1})
$g_{\|\|(1)}$	4,005	24,25	0,4	----	----	----
$g_{\|\|(2)}$	3,52	1,38	0,37	----	----	----
$g_{\|\|(3)}$	2,8	1,73	0,27	----	----	----

The EPR spectra of cobalt(II)-DM-hCAII changes with increasing concentration of sodium thiocyanate. The EPR spectra of cobalt(II)-DM-hCAII are similar to those of cobalt(II)-bCAII[44]. All the samples presented a copper contamination, which is easily recognizable at 3200 G, as it was determined that the cavity was the source of that copper contamination.

For the fitting of the EPR spectra, a “pseudospin” of ½ was used, instead of 3/2, because the energy separation between the Kramer’s doublets far exceeds the energy of the X-band[44,102,107].

The EPR spectrum of the free form of cobalt(II)-DM-hCAII (**Figure 3.6 A**) has several resonances where the most intense one around 1000 Gauss and a less intense resonance, that is split around 1800 G. This split resonance could be a hyperfine interaction, which for cobalt nucleus (I = 7/2). Furthermore, there is an intense resonance at 4000 G. The fitting of the free form (**Figure 3.7 A**) revealed that there are two axial cobalt(II) species with different g-values (see **Table 3.1**). The presence of two species in the sample could mean that there is conformational exchange in the active site, e.g. tetracoordinated (one water molecule) and pentacoordinate (two water molecules).

The spectrum for cobalt(II)-DM-hCAII in the presence of one equivalent of sodium thiocyanate (**Figure 3.6 B**) has some similarities to that published for cobalt(II)-bCAII[44] (**Figure 3.5 A**). While the EPR spectrum of cobalt(II)-bCAII presents a transition with a shoulder, in cobalt(II)-DM-hCAII the shoulder seems resolved and the higher transition shows eight hyperfine lines, due to the hyperfine interaction with the cobalt(II) nucleus. After the most intense resonance, there is a second less intense one and above 2000 G there should be two other resonances, however in the cobalt(II)-DM-hCAII spectrum these features are severely broadened. The strong signals at 4000 – 5000 G present for 1:1 cobalt(II)-DM-hCAII:SCN$^-$ are weaker than in the absence of SCN$^-$ and different from the high field feature of cobalt(II)-BCAII. The fitting of the 1:1.3 cobalt(II)-DM-hCAII-thiocyanate revealed that the system also may have two different cobalt(II) systems, one rhombic (***D*** = 31,3 cm^{-1} and ***E*** = -8,28 cm^{-1}) and one axial (***D*** = 79,69 cm^{-1}) (see **Table 3.1**) suggesting that, the active center may have different geometries. The first geometry corresponding to the rhombic system may belong to a penta-coordinated system, which is similar spectrum **A** in **Figure 3.5**. The presence of a second species could arise from an incomplete reaction of the enzyme with thiocyanate since the ligand binding affinity could be different for the cobalt(II) ion.

The ZFS values of the free form of cobalt(II)-DM-hCAII and the stoichiometric sample of cobalt(II)-DM-hCAII-thiocyanate, (167.0 cm^{-1} and 159.0 cm^{-1}, respectively) are similar, reinforcing the hypothesis that the geometry of those forms is similar and that there was not enough thiocyanate to fully react with the enzyme. The other ZFSs, 68.5 cm^{-1} and 34.4 cm^{-1}, respectively, are also close in value, meaning that the cobalt(II) ion may have a similar coordination geometry, where the presence of different ligands might explain the different ZFS values.

The EPR spectrum of cobalt(II)-DM-hCAII in the presence of a thousand-fold concentration excess of sodium thiocyanate presents a quasi-symmetrical line centered around 1600 G, very similar to that present in cobalt(II)-bCAII[14], as well as the two features at 4000-5000 G much less intense than for the stoichiometric spectrum. The spectrum C was fitted with three isotropic g values, meaning that there may be many conformations of the active center in the presence of a thousand-fold thiocyanate.

For the fitting it might be required to introduce more variables, e.g. another cobalt(II) spin system, or optimize the weights of the parameters. Furthermore, the reacquisition of data with a higher protein concentration and cryo-protectants can improve the quality of the spectra, meaning a could be obtained.

3.3.2 Paramagnetic NMR studies

The EPR spectral features of cobalt(II)-DM-hCAII in the presence of sodium thiocyanate in 1:1 ratio and 1:1000 ratio reflect its behavior as a type A and type C ligand, respectively, in accordance with what has been observed for cobalt(II)-bCAII[44]. Next, we proceeded to the calculation of the NMR PCS values for cobalt(II)-DM-hCAII in the absence and presence of increasing concentrations of sodium thiocyanate to determine if this ligand provides different $\Delta\chi$ tensor values when it behaves as a class A and class C ligand.

To extract the PCS values, the paramagnetic ^{15}N-cobalt(II)-DM-hCAII protein and the corresponding diamagnetic ^{15}N-zinc(II)-DM-hCAII analogue were prepared. Then, the addition of the ligand to both proteins was monitored by ^{1}H-^{15}N HSQC NMR until above a thousand-fold ligand:protein concentration ratio, in order to calculate their ^{1}H and ^{15}N PCS values The corresponding $\Delta\chi$ values were calculated at several points of the titration, to determine the changes of the magnetic susceptibility tensor during the addition of sodium thiocyanate.

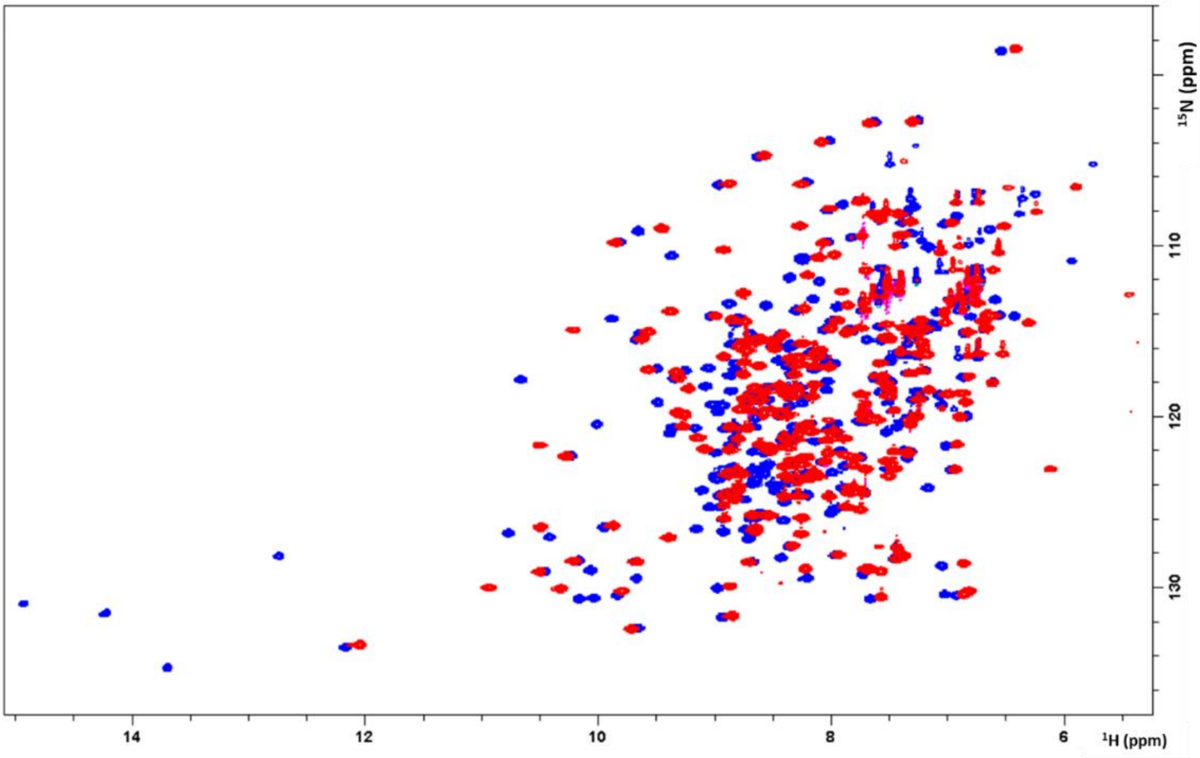


Figure 3.8: ^{1}H-^{15}N HSQC spectra of free-cobalt(II)-DM-hCAII (red) and free-zinc(II)-DM-hCAII (blue) in 10 mM HEPES, pH 6.3. Spectra were recorded at 500 MHz with a protein concentration is 350 µM.

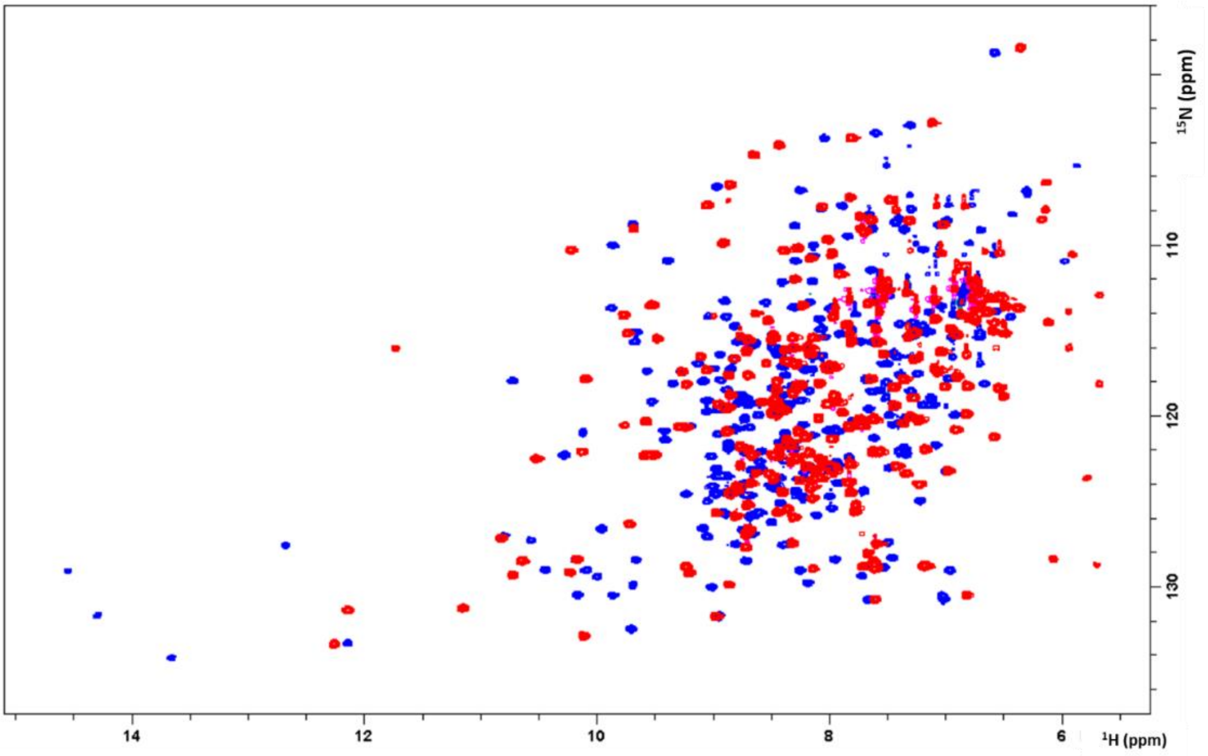


Figure 3.9: ^{1}H-^{15}N HSQC spectra of cobalt(II)-DM-hCAII (red) and zinc(II)-DM-hCAII (blue) in the presence of 475 mM sodium thiocyanate (1:1357), in 10 mM HEPES, pH 6.3. Spectra were recorded at 500 MHz with a protein concentration is 350 µM.

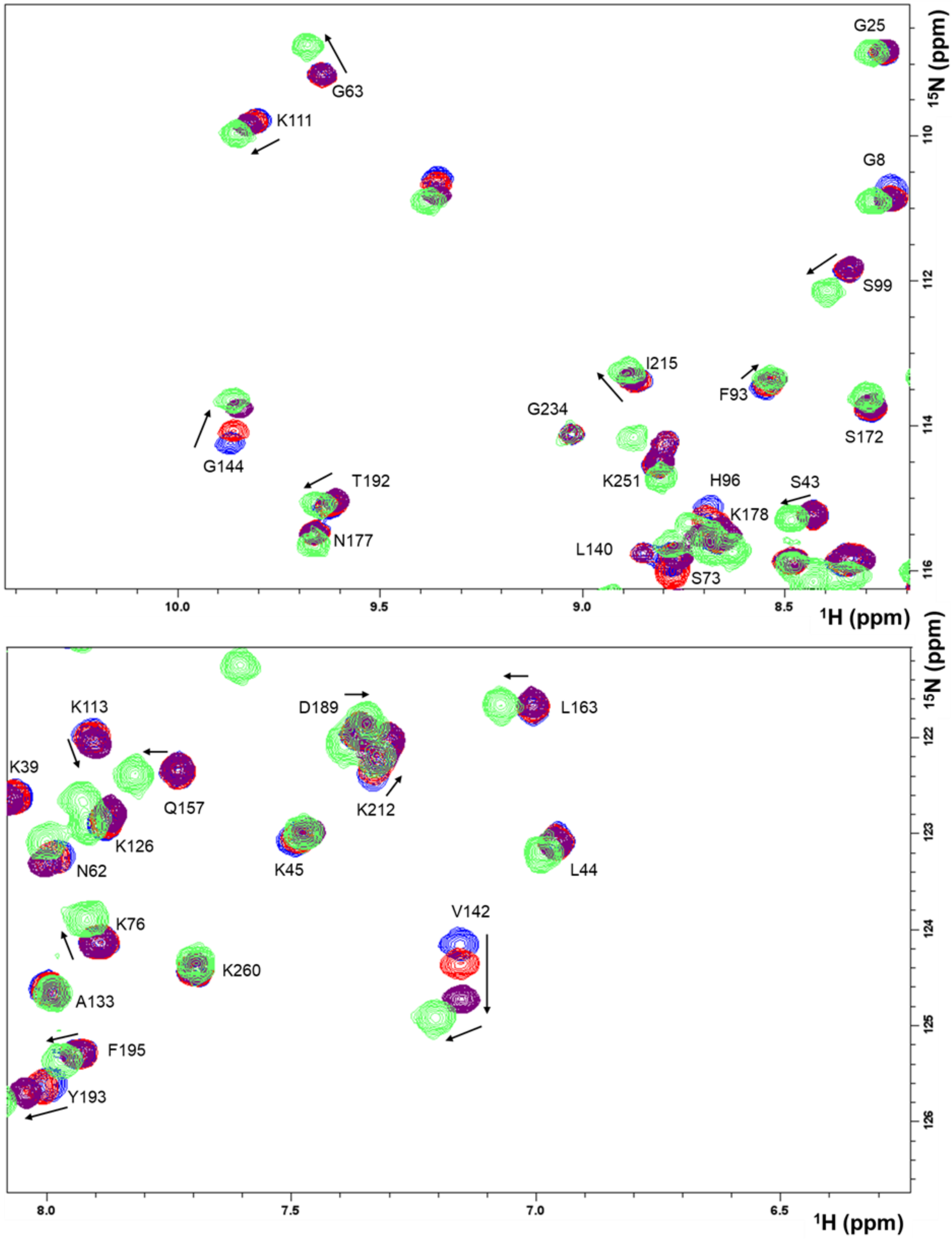


Figure 3.10: ^{1}H-^{15}N HSQC spectra of zinc(II)-DM-hCAII titrated with sodium thiocyanate, in 10 mM HEPES, pH 6.3. (Blue - free enzyme; Red - protein:ligand ratio 1:1.3; purple - protein:ligand ratio 1:13; green – protein:ligand ratio 1:1357). Spectra were recorded at 500 MHz with a protein concentration is 350 µM.

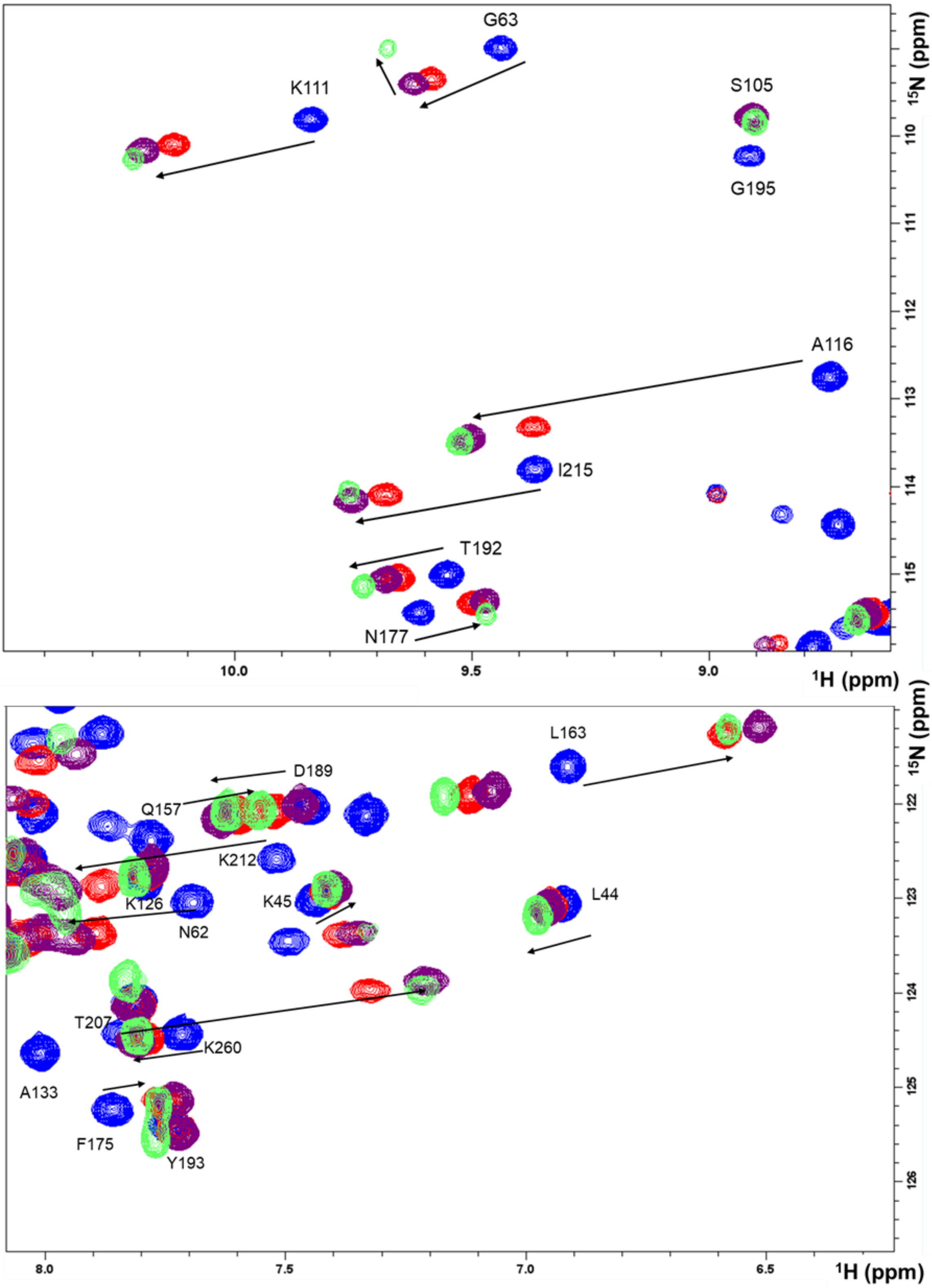


Figure 3.11: ^{1}H-^{15}N HSQC spectra of cobalt(II)-DM-hCAII titrated with sodium thiocyanate, in 10 mM HEPES, pH 6.3. (Blue - free enzyme; Red - protein:ligand ratio 1:1.3; purple - protein:ligand ratio 1:13; green – protein:ligand ratio 1:1357). Spectra were recorded at 500 MHz with a protein concentration is 350 µM.

^{1}H-^{15}N-HSQC spectra represent an NMR fingerprint of the protein because each structure has a unique chemical environment for the N-H bonds. The spectra for the free cobalt(II)-DM-hCAII and zinc(II)-DM-hCAII proteins (**Figure 3.8**) are similar to those obtained in previous studies[4], therefore it can be assumed that the effect of the mutations at the N-terminal (H3N and H4N) do not affect the protein structure. This is helpful in the assignment of the protein spectra because it remains mostly unchanged. **Figure 3.9** shows the spectra of both zinc(II) and cobalt(II)-DM-hCAII at the highest used concentration of sodium thiocyanate. The acquired spectra are different from the free forms in **Figure 3.8**, meaning the sodium thiocyanate binds to the both metal forms of DM-hCAII. The assignment of the diamagnetic zinc(II)-DM-hCAII enzyme with increasing concentrations of sodium thiocyanate was direct due to the small shifts caused by the presence of the ligand, as seen in **Figure 3.10.** The assignment of the paramagnetic enzyme proved to be more challenging due to the large PCS induced after the addition of the ligand (**Figure 3.11**). The assignment of the paramagnetic enzyme required an interactive approach for the different concentrations of sodium thiocyanate. This approach requires the calculation of an initial magnetic susceptibility tensor with the program FANTEN[73], which informs us about the predicted shifts of the nuclei that were not used in the calculation of the tensor. For this initial estimation only residues that were easy to reassign were used, e.g. some glycines that are isolated in the spectrum. Knowing this, it is possible to determine where the resonances have shifted, either to a lower field or higher field. For the ^{15}N-cobalt(II)-DM-hCAII, **Figure 3.11** shows that some resonances shift by over 0.5 ppm, where the major shifts occur after the addition of a stoichiometric amount of ligand. Increasing the ligand concentration up to a thousand-fold also shifts the resonances, although less dramatically. In both proteins, the addition of sodium thiocyanate up to a thousand-fold continued shifting the resonances (**Figure 3.10** and **Figure 3.11**), reflecting that this ligand further interacted with the protein residues at those high concentrations.

The PCS and RDCs for the cobalt(II)-DM-hCAII were calculated using the ^{1}H-^{15}N-HSQC and ^{1}H-^{15}N-HSQC-IPAP spectra of the diamagnetic and paramagnetic proteins. The agreement for correlation plots between the experimental and calculated data (Appendix A: Cobalt Chapter) was good in all cases, as illustrated in **Figure 3.12**, where the Q values are reported, resulting in the best fit values for the axial and rhombic anisotropy components ($\Delta\chi_{ax}$ and $\Delta\chi_{rh}$) and respective isosurfaces and tensor orientations (expressed as angles of the main axes directions with the metal–ligand directions, illustrated in **Figure 3.12**) reported in **Table 3.2** and **Table 3.3**.

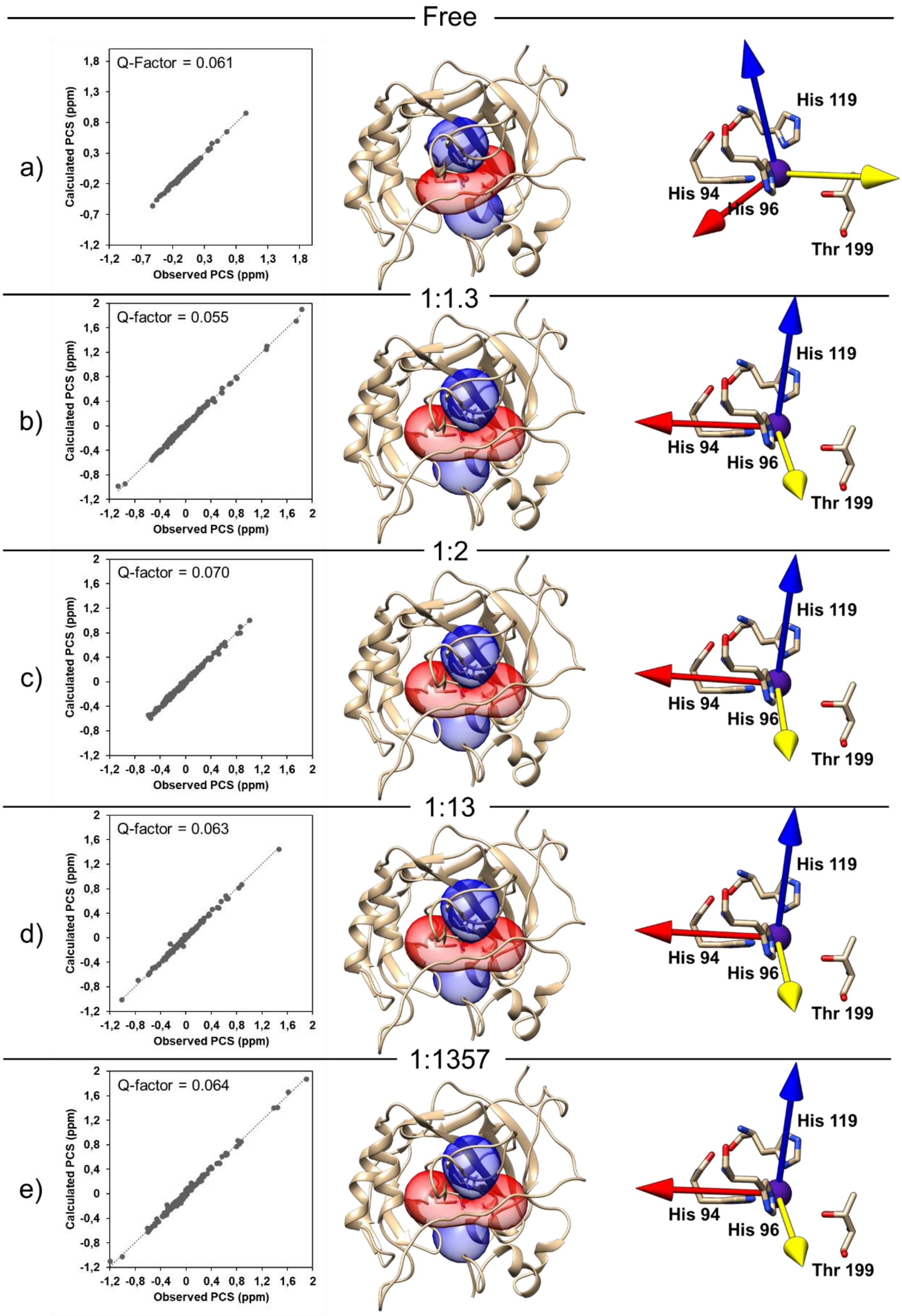


Figure 3.12: Correlation plots between experimental and calculated PCS and respective representation of the PCS surfaces 1 (blue) and -1 (red) and the magnetic susceptibility anisotropy main axis directions in the absence and presence of different concentrations of sodium thiocyanate at pH 6.3. a) Free cobalt(II)-DM-hCAII; b) cobalt(II)-DM-hCAII-thiocyanate 1:1.3; c) cobalt(II)-DM-hCAII-thiocyanate 1:2; d) cobalt(II)-DM-hCAII-thiocyanate 1:13; e) cobalt(II)-DM-hCAII-thiocyanate 1:1357.

Table 3.2: Magnetic susceptibility tensor anisotropy parameters calculated with the program FANTEN for adducts of cobalt(II)-DM-hCAII and cobalt(II)-WT-hCAII (published data from ref. 4).

	Ligand	pH	$\Delta\chi_{ax}$ (10^{-32} m^3)	$\Delta\chi_{rh}$ (10^{-32} m^3)	Q-factor	Number of PCS	Coordination	Published
DM-hCAII	**No**	**6.0**	3.23 ± 0.01	-1.10 ± 0.02	0.062	169	4	No
DM-hCAII	**No**	**6.3**	3.15 ± 0.02	-0.59 ± 0.02	0.061	161	4	No
WT-hCAII	**No**	**5.8**	2.24±0.02	-1.18 ± 0.02	0.09	165	4	Yes[46]
WT-hCAII	**No**	**6.8**	2.81±0.03	-0.63 ± 0.02	0.08	176	4	Yes[46]
DM-hCAII	**Thiocyanate (1.3-fold)**	**6.3**	6.71 ± 0.05	-2.30 ± 0.03	0.055	157	5	No
DM-hCAII	**Thiocyanate (2-fold)**	**6.3**	7.41 ± 0.05	-2.81 ± 0.04	0.070	150	5	No
DM-hCAII	**Thiocyanate (13-fold)**	**6.3**	7.76 ± 0.04	-2.95 ± 0.05	0.063	141	5	No
DM-hCAII	**Thiocyanate (1357-fold)**	**6.3**	7.69 ± 0.04	-2.95 ± 0.06	0.064	143	5	No
WT-hCAII	**Oxalate**	**5.8**	6.54 ± 0.04	-4.44 ± 0.03	0.07	160	5	Yes[46]
WT-hCAII	**Furosemide**	**6.8**	3.09 ± 0.02	-1.69 ± 0.02	0.07	174	4	Yes[46]

Table 3.3: Magnetic susceptibility tensor anisotropy parameters calculated with the program FANTEN for adducts of cobalt(II)-DM-hCAII using PCS and/or RDC. *Q-factors of PCS/PCS and RDC ******RDC whose NOE have a value equal or higher than 0.75.

	Ligand	pH	$\Delta\chi_{ax}$ (10^{-32} m^3)	$\Delta\chi_{h}$ (10^{-32} m^3)	Q-factor	Number of PCS + RDC
DM-hCAII	**Thiocyanate (1.3-fold)**	**6.3**	7.42	-2.93	0.217	0 + 97
DM-hCAII	**Thiocyanate (1.3-fold)**	**6.3**	7.42	-2.91	0.09/0.217*	150 + 97
DM-hCAII	**Thiocyanate (1.3-fold)**	**6.3**	7.66	-2.98	0.09/0.204*	150 + 83******

Table 3.2 compares the calculated axial and rhombic anisotropies of the cobalt(II)-hCAII, both WT and DM, for the free forms and for different ligand adducts, thiocyanate, oxalate and furosemide using only the PCS data. For thiocyanate, the axial and rhombic anisotropies of the cobalt(II) susceptibility tensor increase with the increasing concentration of the ligand. The largest change is observed from the free enzyme to the 1.3-fold excess thiocyanate, and then these changes become smaller up to 1357-fold excess thiocyanate. For the 1.3-fold thiocyanate sample it might be possible that not all thiocyanate anions are bound to the enzyme. To ensure the stoichiometric binding of thiocyanate to the cobalt(II) ion, its concentration was increased to 2-fold excess, with a significant further increase of the axial and rhombic anisotropies. The 13-fold thiocyanate sample yields a $\Delta\chi$ tensor with higher axiality and slightly higher rhombicity, but for the 1357-fold thiocyanate concentration the $\Delta\chi$ remains constant. Although there are small changes in the absolute values of the anisotropy components for different concentrations of sodium thiocyanate, it is apparent that, for concentrations higher than 2-fold thiocyanate, the $\Delta\chi$ values do not change considerably. Therefore, it can be concluded that the presence of high concentrations of sodium thiocyanate, above 1000-fold excess, do not cause an enough change of $\Delta\chi$ of the cobalt(II) center of the protein to explain the observed EPR differences, meaning that other effects are occurring in the protein that account for these differences.

The data presented in **Table 3.2** shows the anisotropy tensors for the free cobalt(II)-DM-hCAII and cobalt(II)-WT-hCAII at different pH values. A small increase of $\Delta\chi_{ax}$ and a small decrease of $\Delta\chi_{rh}$ from pH 5.8 to 6.8 has been reported for cobalt(II)-WT-hCAII, but those values correspond to tetra-coordination with a small reorientation of the magnetic susceptibility anisotropy tensor[46]. According to the literature[12,48], the number of water molecules for the free enzyme is pH dependent, where at low pH 5-coordination is favored and high pH favors 4-coordination. For the free-cobalt(II)-DM-hCAII at pH 6.3 it is expected that the cobalt(II) ion has a coordination number between four and five, with three histidine ligands complemented by water molecule(s), since the pH is close to neutral. Both $\Delta\chi_{ax}$ and χ_{rh} values decrease slightly in module between pH 6.0 and 6.3, with the $\Delta\chi_{ax}$ values within the range expected for 4-coordinated complexes, again due to a small reorientation of the magnetic susceptibility anisotropy tensor[46]. We can compare the 1:1 adducts formed by different ligands, oxalate, furosemide and thiocyanate with the free cobalt(II)-hCAII protein at the same pH. It can be seen that furosemide provides the lowest axial and rhombic anisotropy increase (pH 6.8), which means the tetra-coordinated geometry is maintained, although with a large anisotropy tensor reorientation[46]. For the oxalate adduct, the $\Delta\chi_{ax}$ and χ_{rh} anisotropy values undergo a high increase to values typical of a penta-coordinated structure and a large anisotropy tensor reorientation[46]. Following the same procedure as for the furosemide and oxalate ligands, the large increase of the $\Delta\chi_{ax}$ and χ_{rh} anisotropy values in module for the 1:1 thiocyanate adduct cobalt(II)-DM-hCAII with the free protein at pH 6.3, leads to the conclusion that thiocyanate forms a penta-coordinated complex at the active site[48,91,108]. The change in the anisotropy values is accompanied by a large reorientation of the magnetic susceptibility tensor when the protein is reacted with thiocyanate, as seen in **Figure 3.12**. At higher concentrations of thiocyanate, up to 1357 fold, the tensor orientation does not significantly change, meaning the coordination of the metal at the active center remains the same. This coordination geometry of the thiocyanate adduct is confirmed by X-ray crystallography, where the crystallographic data for zinc(II)-WT-hCAII (PDB 2CA2[109] and 4YGK[110]) and its T199P/C206S mutant (PDB 1LG6[111]) show that the zinc(II) ion is bound to three protein histidine side-chains, one water molecule and one thiocyanate anion. The coordination geometries for both metal forms of DM-hCAII is confirmed by X-ray-crystallography later in the chapter.

The magnetic anisotropy tensor was also calculated using only the RDC values using FANTEN[73], however RDCs are prone to more errors due to the local mobility of the protein structure. This is verified in **Table 3.3**, where the $\Delta\chi_{ax}$ and $\Delta\chi_{rh}$ values were calculated using only RDC, PCS and RDC and PCS and RDC with a NOE value over 0.75. The NOE values were calculated when performing relaxation experiments, described in section 3.3.5, whose experimental results are shown in **Figure 3.18**. and calculated values are in **Table 3.4**. The NOE value of 0.75 was chosen as the minimum value for the calculations, as it means that the observed coupled nuclei are rigid, and a value below 0.75 means the observed nuclei are more flexible and can introduce errors in the calculated magnetic susceptibility tensor. In **Table 3.3** it is observed that the values for the magnetic susceptibility tensor are similar if they are calculated using PCS alone, RDC alone and combined PCS with all RDC. However, when using a combined PCS and RDC, with a NOE value higher than 0.75, the magnetic susceptibility tensor changes, becoming slightly more axial and slightly more rhombic. Furthermore, the Q-factors for the fits using PCSs together with RDCs becomes much smaller than when using RDCs alone, reflecting that more accurate magnetic susceptibility tensor values are obtained.

3.3.3 CSP NMR studies of zinc(II)-DM-hCAII

To understand which residues of the cobalt(II)-DM-hCAII protein are truly affected by the addition of sodium thiocyanate, and not due to the increasing ionic strength of the protein solution, the chemical shift perturbation (CSP) method[98–100] was applied to the diamagnetic zinc(II)-DM-hCAII analogue. **Figure 3.13** displays the CSP index obtained from combined ^{1}H and ^{15}N CSP values ($\Delta\delta$) for the amide NH moieties of the amino-acid residues along the protein chain protein, at increasing thiocyanate ratios from 1:1.3 to 1:1357. The location of the main secondary structural features of the protein is also shown. The visualization of the affected residues at increasing thiocyanate ratios is identified in the 3D structure of the hCAII (**Appendix A**).

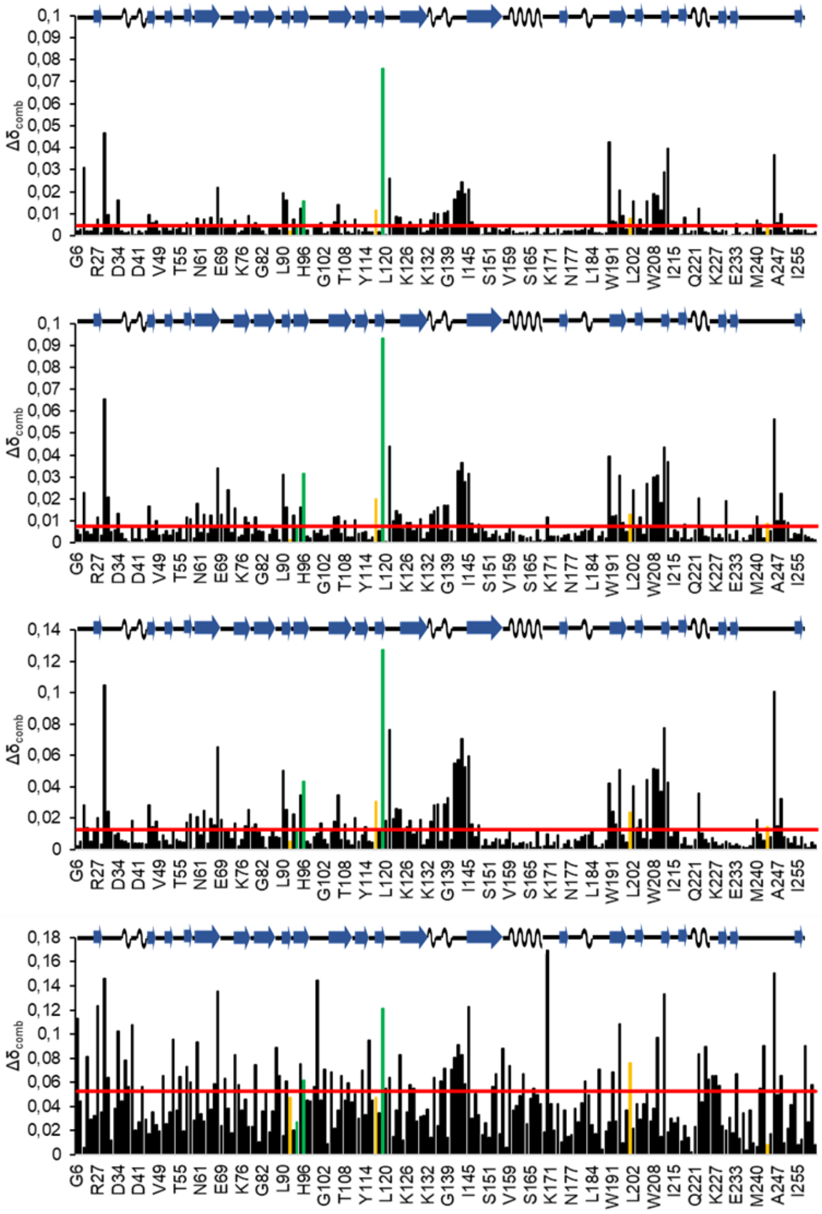
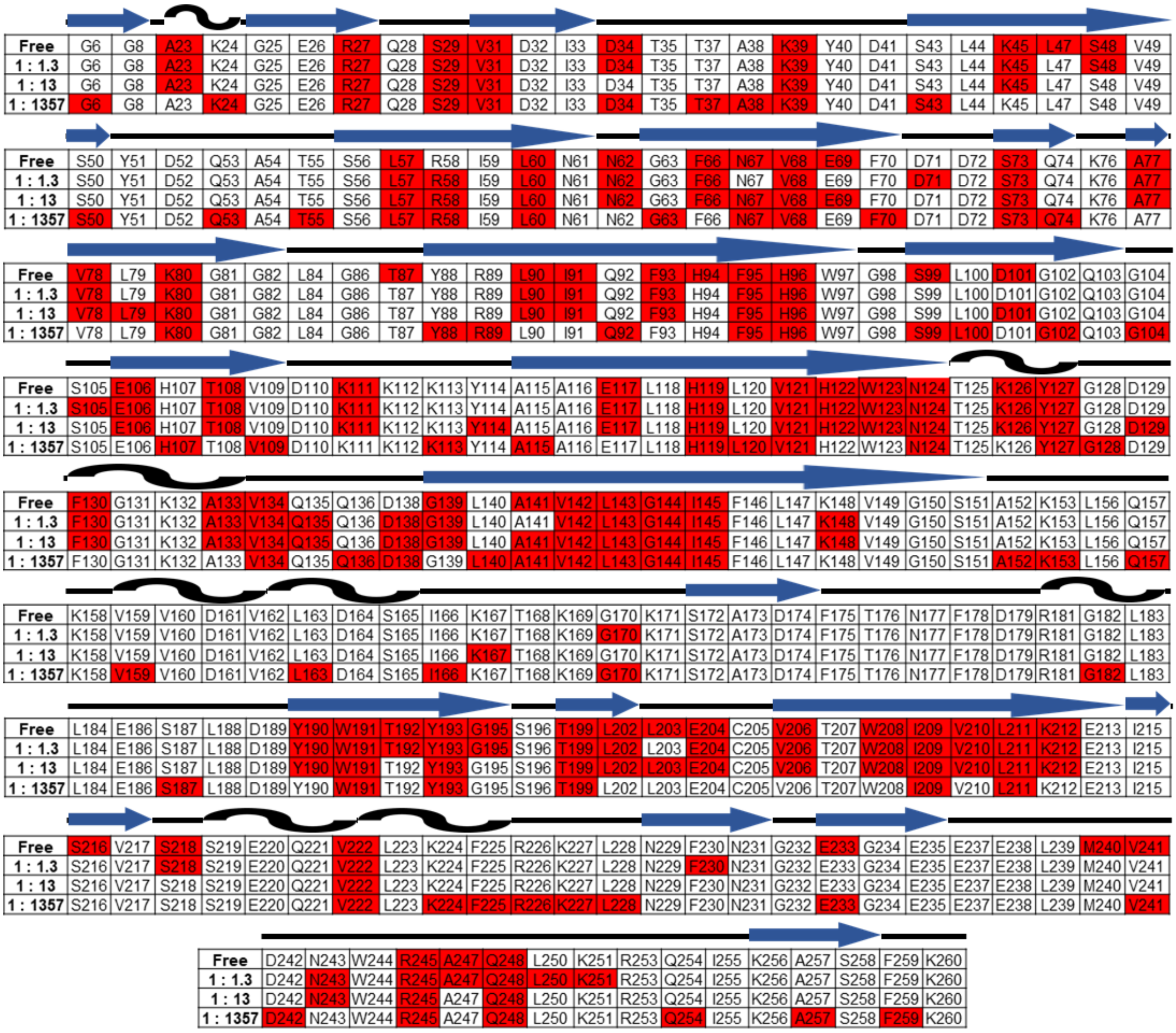


Figure 3.13: Chemical Shift Perturbation analysis for zinc(II)-DM-hCAII at increasing protein:ligand ratios (sodium thiocyanate), 1:1.3, 1:2.7, 1:13 and 1:1357. The green bars represent the residues H94, H96 and H119, which are the primary sphere of coordination of zinc(II) ion. The residues in orange are the second sphere of coordination. The red line is the standard deviation. On top of each plot there are the representations of the secondary structure elements of hCAII.



Free	G6	G8	**A23**	K24	G25	E26	**R27**	Q28	**S29**	**V31**	D32	I33	**D34**	T35	T37	A38	**K39**	Y40	D41	S43	L44	**K45**	**L47**	**S48**	V49
1 : 1.3	G6	G8	**A23**	K24	G25	E26	**R27**	Q28	**S29**	**V31**	D32	I33	**D34**	T35	T37	A38	**K39**	Y40	D41	S43	L44	**K45**	L47	**S48**	V49
1 : 13	G6	G8	**A23**	K24	G25	E26	**R27**	Q28	**S29**	**V31**	D32	I33	D34	T35	T37	A38	**K39**	Y40	D41	S43	L44	**K45**	L47	S48	V49
1 : 1357	**G6**	G8	A23	**K24**	G25	E26	**R27**	Q28	**S29**	**V31**	D32	I33	**D34**	T35	**T37**	**A38**	**K39**	Y40	D41	**S43**	L44	K45	L47	S48	V49

Free	S50	Y51	D52	Q53	A54	T55	S56	**L57**	R58	I59	**L60**	N61	**N62**	G63	**F66**	**N67**	**V68**	**E69**	F70	D71	D72	**S73**	Q74	K76	**A77**
1 : 1.3	S50	Y51	D52	Q53	A54	T55	S56	**L57**	**R58**	I59	**L60**	N61	**N62**	G63	**F66**	N67	**V68**	E69	F70	**D71**	D72	**S73**	Q74	K76	**A77**
1 : 13	S50	Y51	D52	Q53	A54	T55	S56	**L57**	**R58**	I59	**L60**	N61	**N62**	G63	**F66**	**N67**	**V68**	**E69**	F70	D71	D72	**S73**	Q74	K76	**A77**
1 : 1357	**S50**	Y51	D52	**Q53**	A54	**T55**	S56	**L57**	**R58**	I59	**L60**	N61	N62	**G63**	F66	**N67**	**V68**	E69	**F70**	D71	D72	**S73**	**Q74**	K76	A77

Free	**V78**	L79	**K80**	G81	G82	L84	G86	**T87**	Y88	R89	**L90**	**I91**	Q92	**F93**	**H94**	**F95**	**H96**	W97	G98	**S99**	L100	**D101**	G102	Q103	G104
1 : 1.3	**V78**	L79	**K80**	G81	G82	L84	G86	T87	Y88	R89	**L90**	**I91**	Q92	**F93**	H94	**F95**	**H96**	W97	G98	S99	L100	D101	G102	Q103	G104
1 : 13	**V78**	**L79**	**K80**	G81	G82	L84	G86	T87	Y88	R89	**L90**	**I91**	Q92	**F93**	H94	**F95**	**H96**	W97	G98	S99	L100	**D101**	G102	Q103	G104
1 : 1357	V78	L79	**K80**	G81	G82	L84	G86	T87	**Y88**	**R89**	L90	I91	**Q92**	F93	H94	**F95**	**H96**	W97	G98	**S99**	**L100**	D101	**G102**	Q103	**G104**

Free	S105	**E106**	H107	**T108**	V109	D110	**K111**	K112	K113	Y114	A115	A116	**E117**	L118	**H119**	L120	**V121**	**H122**	**W123**	**N124**	T125	**K126**	**Y127**	G128	D129
1 : 1.3	**S105**	**E106**	H107	**T108**	V109	D110	**K111**	K112	K113	Y114	A115	A116	**E117**	L118	**H119**	L120	**V121**	**H122**	**W123**	**N124**	T125	**K126**	**Y127**	G128	D129
1 : 13	S105	**E106**	H107	**T108**	V109	D110	**K111**	K112	K113	**Y114**	A115	A116	**E117**	L118	**H119**	L120	**V121**	**H122**	**W123**	**N124**	T125	**K126**	**Y127**	G128	**D129**
1 : 1357	S105	E106	**H107**	T108	**V109**	D110	K111	K112	**K113**	Y114	**A115**	A116	E117	L118	**H119**	**L120**	**V121**	H122	W123	**N124**	T125	K126	**Y127**	**G128**	D129

Free	**F130**	G131	K132	**A133**	**V134**	Q135	Q136	D138	**G139**	L140	**A141**	**V142**	**L143**	**G144**	**I145**	F146	L147	K148	V149	G150	S151	A152	K153	L156	Q157
1 : 1.3	**F130**	G131	K132	**A133**	**V134**	**Q135**	Q136	**D138**	**G139**	L140	A141	**V142**	**L143**	**G144**	**I145**	F146	L147	**K148**	V149	G150	S151	A152	K153	L156	Q157
1 : 13	**F130**	G131	K132	**A133**	**V134**	**Q135**	Q136	**D138**	**G139**	L140	**A141**	**V142**	**L143**	**G144**	**I145**	F146	L147	**K148**	V149	G150	S151	A152	K153	L156	Q157
1 : 1357	F130	G131	K132	A133	**V134**	Q135	**Q136**	**D138**	G139	**L140**	**A141**	**V142**	**L143**	**G144**	**I145**	F146	L147	K148	V149	G150	S151	**A152**	**K153**	L156	**Q157**

Free	K158	V159	V160	D161	V162	L163	D164	S165	I166	K167	T168	K169	G170	K171	S172	A173	D174	F175	T176	N177	F178	D179	R181	G182	L183
1 : 1.3	K158	V159	V160	D161	V162	L163	D164	S165	I166	K167	T168	K169	**G170**	K171	S172	A173	D174	F175	T176	N177	F178	D179	R181	G182	L183
1 : 13	K158	V159	V160	D161	V162	L163	D164	S165	I166	**K167**	T168	K169	G170	K171	S172	A173	D174	F175	T176	N177	F178	D179	R181	G182	L183
1 : 1357	K158	**V159**	V160	D161	V162	**L163**	D164	S165	**I166**	K167	T168	K169	**G170**	K171	S172	A173	D174	F175	T176	N177	F178	D179	R181	**G182**	L183

Free	L184	E186	S187	L188	D189	**Y190**	**W191**	**T192**	**Y193**	**G195**	S196	**T199**	**L202**	**L203**	**E204**	C205	**V206**	T207	**W208**	**I209**	**V210**	**L211**	**K212**	E213	I215
1 : 1.3	L184	E186	S187	L188	D189	**Y190**	**W191**	**T192**	**Y193**	**G195**	S196	**T199**	**L202**	L203	**E204**	C205	**V206**	T207	**W208**	**I209**	**V210**	**L211**	**K212**	E213	I215
1 : 13	L184	E186	S187	L188	D189	**Y190**	**W191**	T192	**Y193**	G195	S196	**T199**	**L202**	**L203**	**E204**	C205	**V206**	T207	**W208**	**I209**	**V210**	**L211**	**K212**	E213	I215
1 : 1357	L184	E186	**S187**	L188	D189	Y190	**W191**	T192	**Y193**	G195	S196	**T199**	L202	L203	E204	C205	V206	T207	W208	**I209**	V210	**L211**	K212	E213	I215

Free	**S216**	V217	**S218**	S219	E220	Q221	**V222**	L223	K224	F225	R226	K227	L228	N229	F230	N231	G232	**E233**	G234	E235	E237	E238	L239	**M240**	**V241**
1 : 1.3	S216	V217	**S218**	S219	E220	Q221	**V222**	L223	K224	F225	R226	K227	L228	N229	**F230**	N231	G232	E233	G234	E235	E237	E238	L239	M240	V241
1 : 13	S216	V217	S218	S219	E220	Q221	**V222**	L223	K224	F225	R226	K227	L228	N229	F230	N231	G232	E233	G234	E235	E237	E238	L239	M240	V241
1 : 1357	S216	V217	S218	S219	E220	Q221	**V222**	L223	**K224**	**F225**	**R226**	**K227**	**L228**	N229	F230	N231	G232	**E233**	G234	E235	E237	E238	L239	M240	**V241**

Free	D242	N243	W244	**R245**	**A247**	**Q248**	L250	K251	R253	Q254	I255	K256	A257	S258	F259	K260
1 : 1.3	D242	**N243**	W244	**R245**	**A247**	**Q248**	**L250**	**K251**	R253	Q254	I255	K256	A257	S258	F259	K260
1 : 13	D242	**N243**	W244	**R245**	A247	**Q248**	L250	K251	R253	Q254	I255	K256	A257	S258	F259	K260
1 : 1357	**D242**	N243	W244	**R245**	A247	**Q248**	L250	K251	R253	**Q254**	I255	K256	**A257**	S258	**F259**	K260

Figure 3.14: Zinc(II)-DM-hCAII's interacting assigned residues upon addition of different concentrations of sodium thiocyanate. Cells in red represent a perturbation of the residues upon the addition of the ligand. On the top are the representation of the secondary structure elements of hCAII.

The interaction of sodium thiocyanate with the protein is confirmed by the chemical shift perturbation (CSP) analysis of the NMR spectra of zinc(II)-DM-hCAII. From the datasets, some identified interactions reflect the direct perturbation of the residues in the primary sphere of coordination of zinc(II) (His 94, His 96 and His 119) and some neighboring residues. The second sphere of coordination (Gln 92, Glu 117, Thr 199 and Asn 243) and some residues in its vicinity are also perturbed. Although His 64 is not assigned, the residues close to His 64 in the sequence and spatially (the loop and α-helix at the N-terminus) are perturbed. The residue His 64 is responsible for the proton transfer from a catalytic water molecule in the active center to the bulk solvent and it is close to the active center. All these perturbations appear at the first point of the titration (1:1.3) and are common at all points of the titration with sodium thiocyanate, while the number of interactions changes with increasing concentration of sodium thiocyanate. For the protein:ligand 1:1.3 dataset there are 79 interactions, for 1:2.7 there are 74 interactions, for 1:13 there are 77 interactions and for the 1:1357 there are 79 interactions. The decrease in the number of interactions between the ratio 1:1.3 and 1:2.7 might be explained by the transient nature of the interactions and the observation that the interactions become stronger when doubling the ligand concentration. These interactions become even stronger when the ligand concentration is increased to 13-fold excess. However, at 1357-fold ligand concentration excess there are 31 new interactions and 33 fewer interactions. This observation can be explained by the dramatic increase of the ionic strength (from 4.75 mM to 475 mM) that can influence those interactions.

Analyzing the structure carton representations (**Appendix A – Figures 7.6 to 7.9**) together with **Figure 3.14**, some hydrophobic residues, that are on the surface of the protein and solvent exposed, also interact with sodium thiocyanate (V31, A38, L47, L57, L60, F70, Y88, L100, V109, A115, V121, L140, V142, A152, L163, W191, Y193, V222, L228, V241, A257 and F259). Furthermore, two of these hydrophobic residues are in the active cavity of the protein (V121 and V142). Other hydrophobic non-solvent accessible residues are also affected by the presence of sodium thiocyanate (V68, F95L, L120, Y127, V134, A141, L143, I145, V159, I166, I209, L211 and F225,) and this can be explained by the changes in the conformation and/or dynamics of the protein. This evidence can be supported by a study of the interaction of zinc(II)-hCAII with a SCN^- and other anions in the Hofmeister series, using X-ray crystallography, isothermal titration calorimetry and molecular dynamics simulations.[110] This study provided evidence that the interaction of SCN^-, a weakly hydrated anion at the chaotropic end of this series, is favored with weakly hydrated amino-acid side chains, and hydrophobic pockets formed by several nonpolar side chains, such as binding site (I-4) near the mouth of the binding pocket, as well as complementarily shaped hydrophobic pockets on the surface of the protein. This study also suggests that SCN^-, upon forming ion pairs with zinc(II), cause entropically favourable and enthalpically unfavourable rearrangements of water inside the binding pocket, which extend up to 8 Å away from its surface. The binding of chaotropic ions like SCN^- to hydrophobic sites in the protein leads to the displacement of enthalpically and entropically unstable water molecules that may affect the global protein structure and dynamics.

3.3.4 Circular dichroism studies of zinc(II)-DM-hCAII

As mentioned before, the global structure and dynamics of the protein might be affected by the excess amount of thiocyanate, a chaotropic agent that can denature the protein. To test if the zinc(II)-DM-hCAII protein has its global structure affected by the presence of an increasing concentration of sodium thiocyanate, circular dichroism (CD) was used to access the loss of global secondary structure.

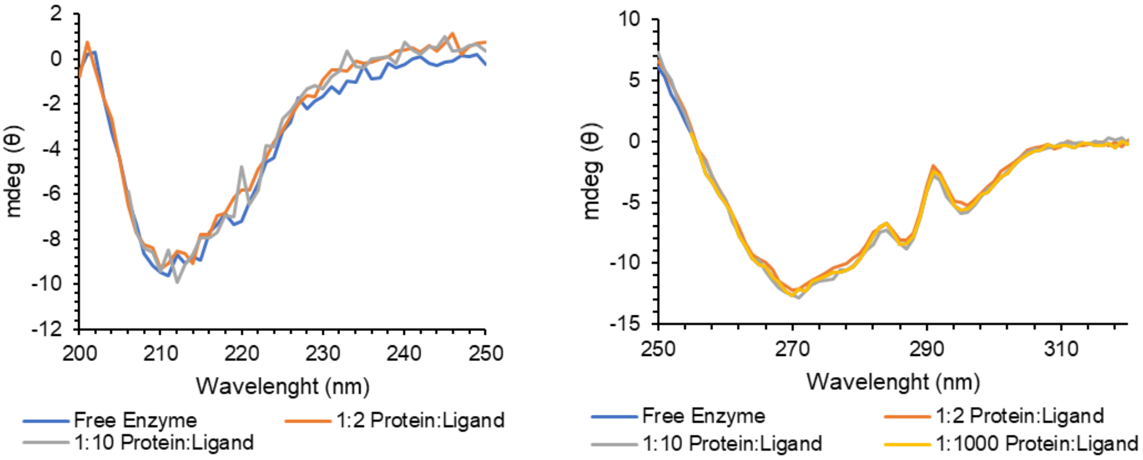


Figure 3.15: Circular dichroism spectra of zinc(II)-WT-hCAII free and bound to increasing concentrations of sodium thiocyanate. Spectra on the left corresponds to the Far-UV range, and those on the right correspond to the Near-UV range.

Circular dichroism is a useful technique to determine the global secondary structure of proteins and is usually applied in the wavelength range from 200 to 250 nm. As sodium thiocyanate absorbs in this range, its possible interference can be problematic depending on its concentration. A sodium thiocyanate concentration higher than 10 times the protein's concentration was found to impact the protein CD spectra, and those data were discarded. Therefore, the extended wavelength range between 200 nm and 320 nm in the near-UV region was used to acquire the CD spectra, which are informative regarding the chemical environment of the aromatic protein residues (histidine, tyrosine and tryptophan). The influence of each of the tryptophan residues present in zinc(II)-hCAII on its CD spectrum has been studied, showing that they had a dominant contribution to the spectrum in the whole region.[112]

Figure 3.15 shows that for any of the tested concentrations of sodium thiocyanate, up to 10-fold excess in the 200-250 nm range, and up to 1000-fold excess in the 250-320 nm range, the CD spectra of zinc(II)-DM-hCAII did not change. This means that the environment of the protein aromatic residues remained unchanged. Many of the protein tryptophan residues are in the hydrophobic core of the protein, indicating that up to a 500 mM concentration of sodium thiocyanate the protein core structure does not change, as probed by CD.

3.3.5 NMR relaxation studies of zinc(II)-DM-hCAII

To corroborate the results described in the previous sections, ^{15}N NMR relaxation studies were used to study the effect of thiocyanate on the overall dynamics of the zinc(II)-DM-hCAII protein.[95–97] This work was also compared to the program HYDRONMR[113], which calculates the R_1 and R_2 parameters using a structure deposited in the PDB databank. This allows us to have a starting point to understand how the dynamics of the system change with increasing concentration of sodium thiocyanate.

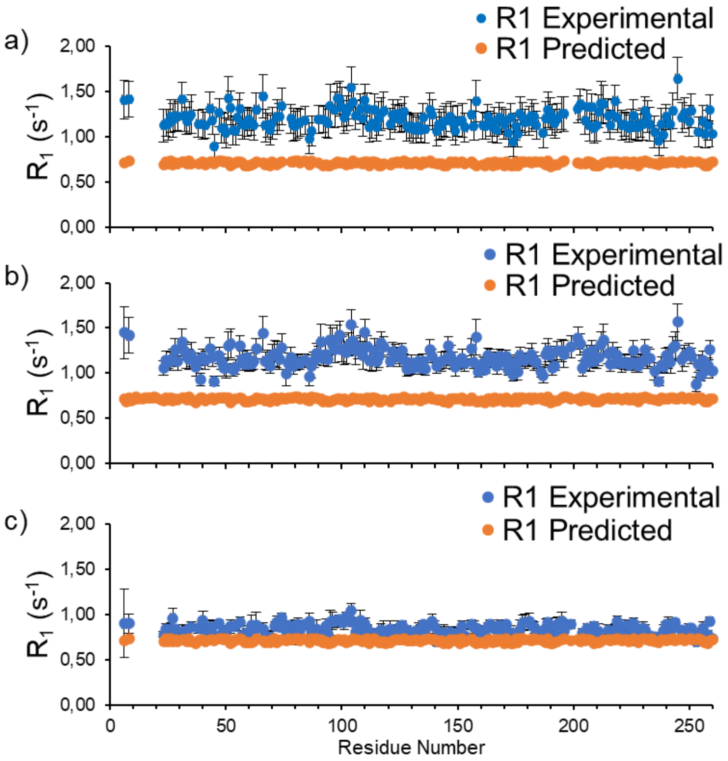


Figure 3.16: Backbone ^{15}N experimental and predicted (HYDRONMR)[113] R_1 vs residue number at different sodium thiocyanate concentrations: a) Free zinc(II)-DM-hCAII; b) 1:1.3 protein:sodium thiocyanate ratio and c) 1:1357 protein:sodium thiocyanate ratio.

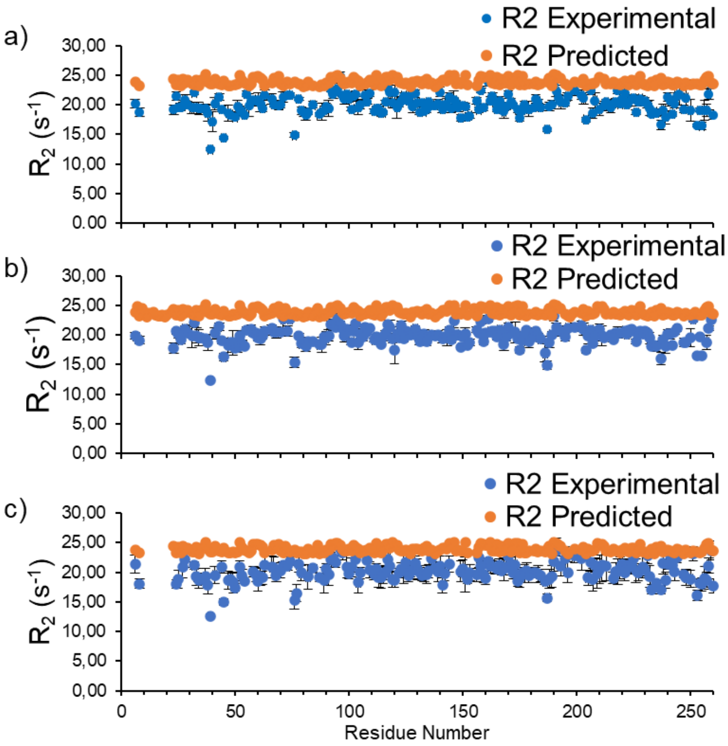


Figure 3.17: Backbone ^{15}N experimental and predicted (HYDRONMR)[113] R_2 vs residue number at different sodium thiocyanate concentrations: a) Free zinc(II)-DM-hCAII; b) 1:1.3 protein:sodium thiocyanate ratio and c) 1:1357 protein:sodium thiocyanate ratio.

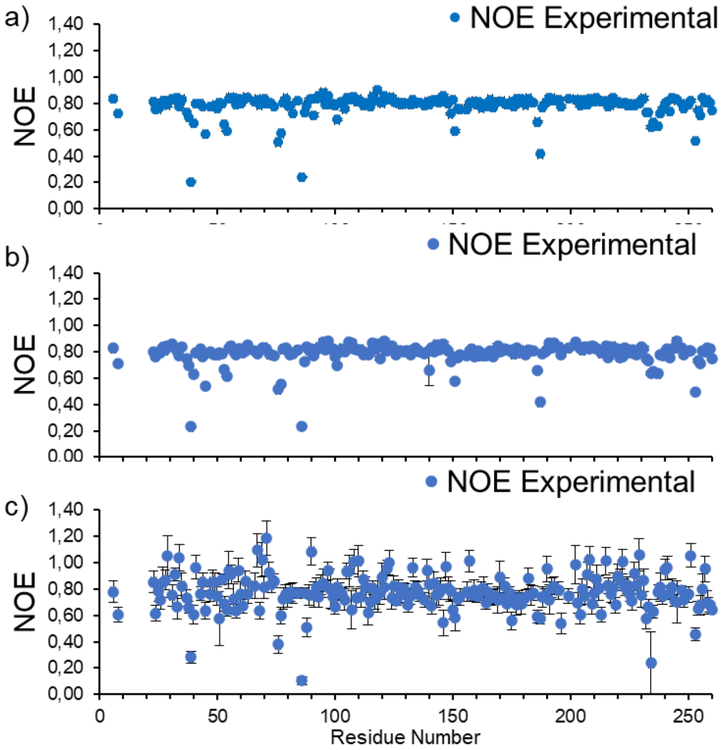


Figure 3.18: Backbone ^{15}N experimental ^{15}N{^{1}H} NOE vs residue number at different sodium thiocyanate concentrations: a) Free zinc(II)-DM-hCAII; b) 1:1.3 protein:sodium thiocyanate ratio and c) 1:1357 protein:sodium thiocyanate ratio.

Table 3.4: Calculated global average ^{15}N relaxation rate parameters R_1 and R_2 with respective standard deviation and correlation time (τ_c) of zinc(II)-DM-hCAII (475 μM) aqueous solutions in the absence and presence of increasing concentrations of sodium thiocyanate.

	Free enzyme	**1:2 ratio**	**1:1357 ratio**
R_1 (s^{-1})	1.2 ± 0.1	1.2 ± 0.1	0.8 ± 0.1
R_2 (s^{-1})	20 ± 2	20 ± 2	20 ± 2
τ_c^{iso} (ns)	10.9 ± 0.5	10.9 ± 0.5	13.2 ± 0.6

Figure 3.16, **Figure 3.17** and **Figure 3.18** show the experimental R_1, R_2 and NOE values obtained for the amide ^{15}N nuclei of the different residues of zinc(II)-DM-hCAII under increasing concentrations of sodium thiocyanate. To determine a t

For all cases, the experimental R_1 and R_2 values are higher than those predicted by HYDRONMR, using a crystal structure (PDB code 3KS3). Furthermore, HYDRONMR considers that the structure is a rigid object, where the local and internal motions are not considered. This means that hCAII behaves as a flexible object in solution rather than a rigid object. At a 1:2 excess concentration of sodium thiocyanate (750 μM), the R_1, R_2, and NOE values do not change appreciably, but at high 1:1000 excess concentration of sodium thiocyanate (475 mM), the R_1 values decrease to values closer to the predicted values, the R_2 values do not change appreciably and the NOE values become more disperse, which could mean that the protein becomes more flexible. **Table 3.4** shows the average R_1 and R_2 values and the calculated isotropic rotational correlation times (τ_c^{iso}) from eqs. (3.2)-(3.4). **Table 3.4** shows that R_2 is independent of the concentration of sodium thiocyanate, and R_1 and τ_c^{iso} change when the concentration of sodium thiocyanate is 475 mM. The value of R_1 decreases, while the value of τ_c^{iso} increases. A lower R_1 indicates that the reorientation of the protein becomes slower in the presence of high excess of thiocyanate. This is reflected in the higher global calculated τ_c^{iso}, which increased from 10.9 ns to 13.2 ns. A higher τ_c^{iso} is indicative that the hydrodynamic radius of the protein increased, thus taking a longer time for a full rotation. Thiocyanate may contribute to these changes because with the increasing concentration of sodium thiocyanate, the viscosity of the medium increases slightly.[114] Furthermore, since sodium thiocyanate is a chaotropic agent, at 475 mM concentration it may remove some structural waters from the protein's surface and thus increasing it apparent hydrodynamic radius.

Taking into consideration the different relaxation techniques used, one can conclude that the overall secondary structure of the proteins does not change, and its hydrophobic core remains the same. Furthermore, the protein hydrodynamic size increases with the addition of 475 mM thiocyanate, taking longer to reorientate. Taken together, the data is indicative that the protein has swollen in size and maintains its general structure.

The relaxation experiments for zinc(II)-DM-hCAII proved to be crucial to understand the effect of sodium thiocyanate on the protein. According to these experiments, sodium thiocyanate affects protein mobility when the thiocyanate concentration is one thousand-fold above protein concentration. In the absence or up to a two-fold excess concentration of thiocyanate, the protein has the same overall R_1, R_2 and τ_c^{iso} values.

3.3.6 X-ray crystallography studies

To confirm the coordination geometry of the thiocyanate adducts of zinc(II) and cobalt(II)-DM-hCAII at a high ligand concentration, we successfully crystallized the two proteins, as it can be seen in **Figure 3.19.**

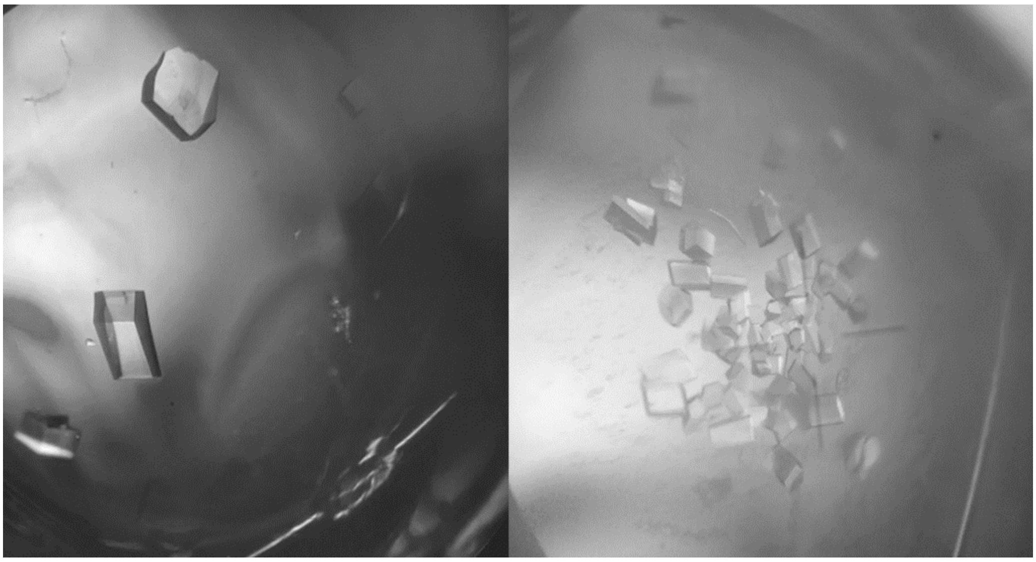


Figure 3.19: Crystals of zinc(II)-DM-hCAII (left) and cobalt(II)-DM-hCAII.

The zinc(II)-DM-hCAII protein formed large single crystals from the simple vapor diffusion technique, whereas for cobalt(II)-DM-hCAII it was necessary to use crystal seeds from zinc(II)-DM-hCAII. The use of a high concentration of seeds led to the formation of many small single crystals, when compared with the zinc(II)-DM-hCAII sample. It was not possible to co-crystallize the proteins with sodium thiocyanate. Instead, the crystals were soaked overnight with 500 mM sodium thiocyanate. This overnight incubation with 500 mM sodium thiocyanate proved to be crucial because longer periods of incubation led to the darkening of the crystals and to a lack of a characteristic protein diffraction pattern. This lack of a diffraction pattern means that the protein was not regularly organized in a crystalline structure, thus confirming the effect of the chaotropic nature of sodium thiocyanate at 500 mM concentration.

The structures of zinc(II)-DM-hCAII-thiocyanate and cobalt(II)-DM-hCAII-thiocyanate were solved using the molecular replacement technique, using the highest diffracting structure of zinc(II)-WT-hCAII (PDB: 3K34) with a maximum resolution of 1.75 Å and 1.46 Å, respectively. The relevant data collection and refinement statistics are shown in **Table 3.5.** The structures of cobalt(II)-DM-hCAII and zinc(II)-DM-hCAII were compared with the highest resolution deposited structure of hCAII bound to thiocyanate (4YGK)[110].

Table 3.5: Data collection and refinement statistics for zinc(II)-DM-hCAII-thiocyanate and cobalt(II)-DM-hCAII-thiocyanate.

	Cobalt(II)-DM-hCAII-thiocyanate	**Zinc(II)-DM-hCAII-thiocyanate**
Wavelength (Å)	1.608	1.541
Resolution range	40.85 - 1.46 (1.52 - 1.46)	23.77 - 1.75 (1.813 - 1.75)
Space group	P 1 21 1	P 1 21 1
Unit cell a b c (Å), α β γ (°)	42.19 41.33 72.53 90 104.47 90	42.09 41.33 72.25 90 104.43 90
Total reflections	102945 (1085)	47244 (4579)
Unique reflections	35299 (749)	24493 (2410)
Multiplicity	2.9 (1.4)	1.9 (1.9)
Completeness (%)	84.11 (17.85)	99.84 (99.75)
Mean I/sigma(I)	22.87 (3.98)	11.44 (2.31)
Wilson B-factor (Å^2)	13.75	10.89
R-merge	0.047 (0.136)	0.071 (0.396)
R-meas	0.056 (0.188)	0.100 (0.560)
R-pim	0.031 (0.129)	0.071 (0.396)
CC1/2	0.997 (0.940)	0.993 (0.659)
CC*	0.999 (0.984)	0.998 (0.891)
Reflections used in refinement	35299 (749)	24468 (2410)
Reflections used for R-free	1825 (35)	1160 (98)
R-work	0.139 (0.237)	0.172 (0.262)
R-free	0.164 (0.268)	0.197 (0.317)
CC(work)	0.973 (0.913)	0.963 (0.830)
CC(free)	0.957 (0.863)	0.945 (0.775)
Number of non-hydrogen atoms	2397	2466
macromolecules	2090	2090
ligands	19	19
solvent	293	362
Protein residues	257	257
RMS (bonds) (Å)	0.023	0.011
RMS (angles) (º)	2.03	1.57
Ramachandran favored (%)	96.47	96.08
Ramachandran allowed (%)	3.53	3.92
Ramachandran outliers (%)	0.00	0.00
Rotamer outliers (%)	0.43	0.43
Clashscore	2.87	6.95
Average B-factor	17.78	14.98
macromolecules	16.25	13.24
ligands	18.72	17.50
solvent	28.66	24.92
Number of TLS groups	1	0

Statistics for the highest-resolution shell are shown in parentheses.

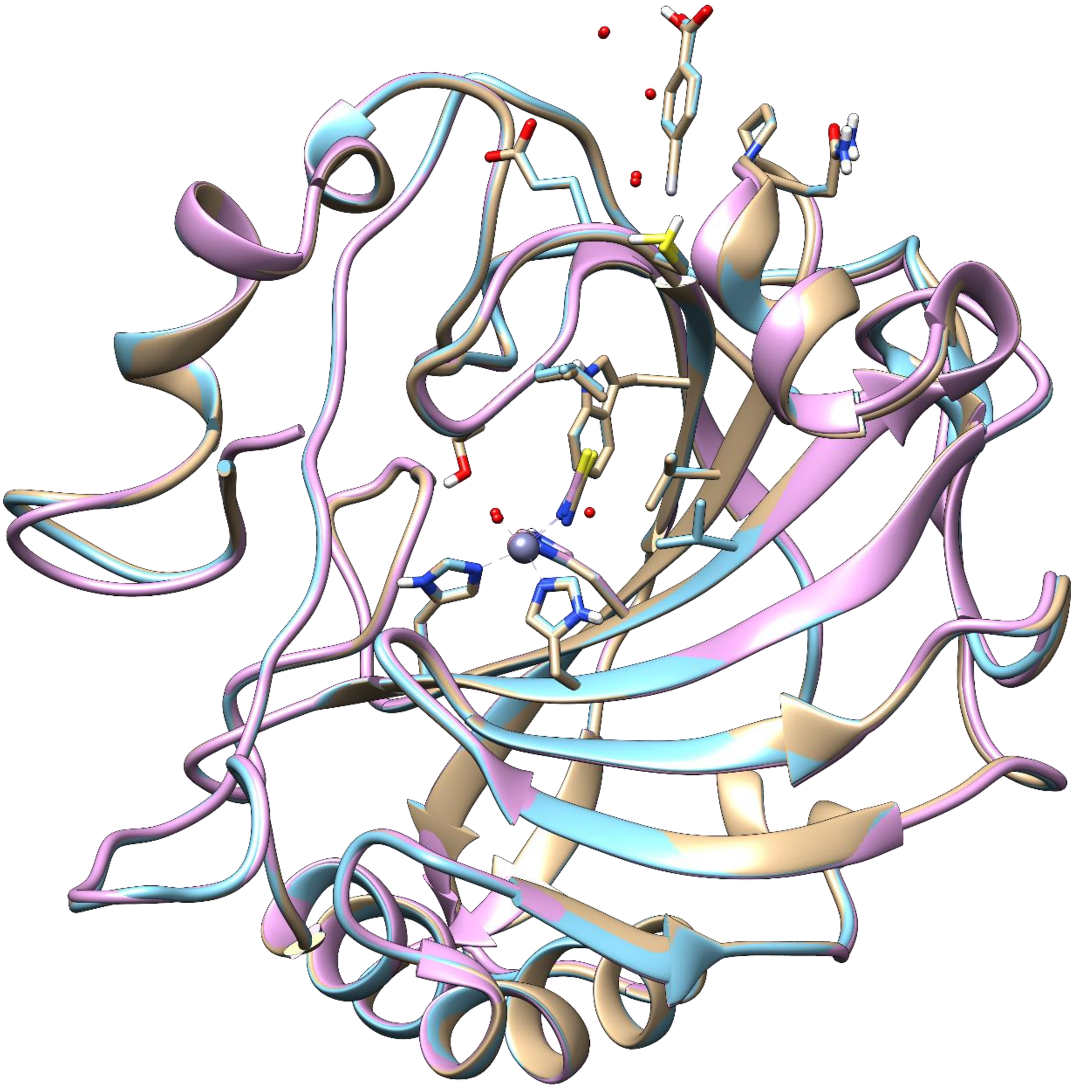


Figure 3.20: Cartoon representation of zinc(II)-WT-hCAII-thiocyanate (purple, PDB code: 4YGK, 1.5 Å); zinc(II)-DM-hCAII-thiocyanate (cream, 1.75 Å) and cobalt(II)-DM-hCAII-thiocyanate (light blue, 1.46 Å)

The solved structures show one protein molecule in the asymmetric unit. For cobalt(II)-DM-hCAII-SCN^- and zinc(II)-DM-hCAII-SCN^- there is a single cobalt(II) ion and a single zinc(II) ion, respectively, and a single mercuric benzoic acid bound to cysteine 204.

The overall structure alignment shows that structures of zinc(II)-WT-hCAII-SCN^-, zinc(II)-DM-hCAII-SCN^- and cobalt(II)-DM-hCAII-SCN^- are superimposable, indicating that they have the same overall structure (**Figure 3.20**). For all proteins, the metal ions are coordinated by five ligands (three histidines, one water molecule and one thiocyanate ion). There are, however, small differences in the active center, where the histidines are at slightly different positions. To determine the regions where there are differences, the three structures were superimposed by their C_α atoms and the global RMSD and the RMSD for each residue were calculated.

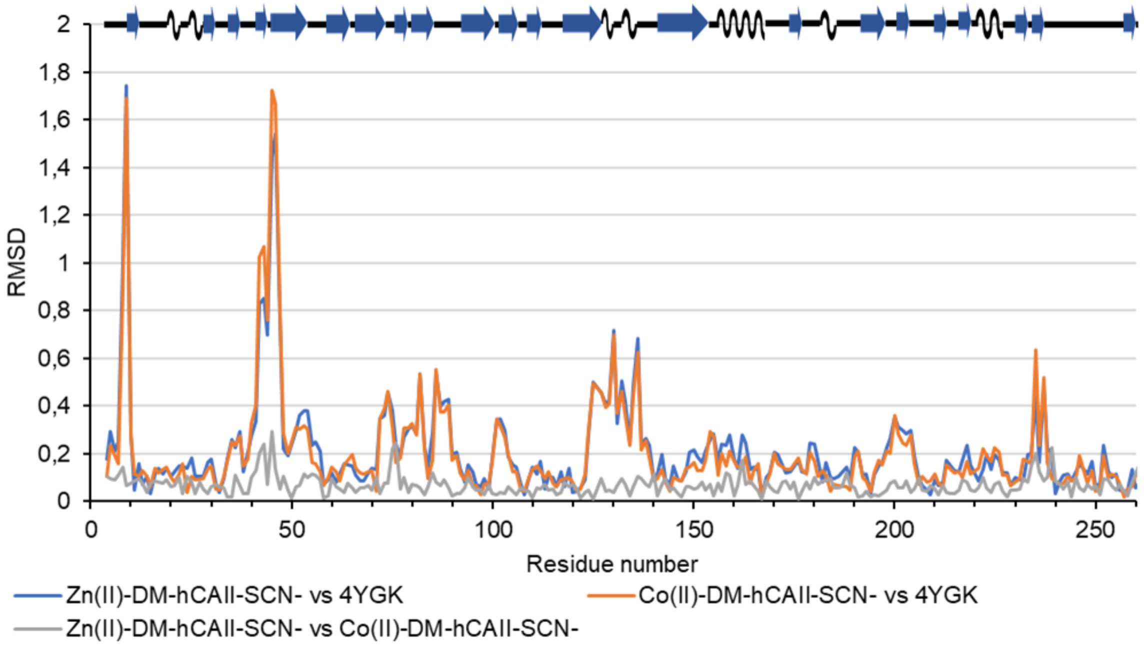


Figure 3.21: C$_\alpha$ RMSD between zinc(II)-DM-hCAII-SCN$^-$ and 4YGK, cobalt(II)-DM-hCAII-SCN$^-$ and 4YGK, and zinc(II)-DM-hCAII-SCN$^-$ and cobalt(II)-DM-hCAII-SCN$^-$.

Table 3.6: Calculated global RSMD between the different hCAII-SCN$^-$ structures.

	Global RMSD (Å)
Zinc(II)-DM-hCAII-SCN$^-$ - 4YGK	0.301
Cobalt(II)-DM-hCAII-SCN$^-$ - 4YGK	0.306
Zinc(II)-DM-hCAII-SCN$^-$ - Cobalt(II)-DM-hCAII-SCN$^-$	0.079

Figure 3.21 shows the C$_\alpha$ RMSD between the cobalt(II)-DM-hCAII, zinc(II)-DM-hCAII and zinc(II)-WT-hCAII SCN$^-$ adducts, and **Table 3.6** shows the global RMSD comparisons. A low RMSD indicates that there are minimal differences between the structures, reflecting that the mutations of DM-hCAII do not affect the overall structure of hCAII. The regions where the RMSD is higher, with a value higher than 1 Å, are loops exposed to the solvent and the residues are 7 and 8 and residues 38 to 42. This higher value reflects a different position in the crystallographic structure, with respect to the 4YGK structure. Regions where the RMSD is higher than 0.4 Å belong to residues that are exposed to the solvent. All these differences might arise from the different crystallization conditions, which imply that different crystal packing forces are in play. The deposited structure 4YGK was crystallized using citrate as a precipitant and in this work the proteins were crystallized using ammonium sulfate. Comparing both metal forms of DM-hCAII-SCN$^-$, the global RMSD is lower and the RMSD at different regions of the protein is also lower, relative to the other comparisons involving Zn-WT-hCAII-SCN, thus reinforcing the conclusion that the crystallization conditions affect the mobility of some local residues.

The incorporation of cobalt(II) ion in the active site is confirmed by NMR due to the extremely different chemical shifts in the ^{1}H-^{15}N-HSQC spectra of the cobalt(II) derivative (**Figure 3.8**). Furthermore, the incorporation of a metal ion in a protein can be confirmed by acquiring X-ray data at the K-edge of an atomic nucleus, which will give rise to an anomalous signal. This anomalous signal arises from the absorption by the atom of some X-ray photons which are re-emitted in the form of fluorescence. Then, some of these re-emitted photons will arrive at the detector with a delay and will encode a non-zero scattering coefficient. With this scattering coefficient, it is possible to calculate single-wavelength anomalous diffraction (SAD) maps that show where the anomalous signal arises from.[115] For the cobalt(II)-DM-hCAII, the K-edge of cobalt is at a wavelength of 1.6083 Å.

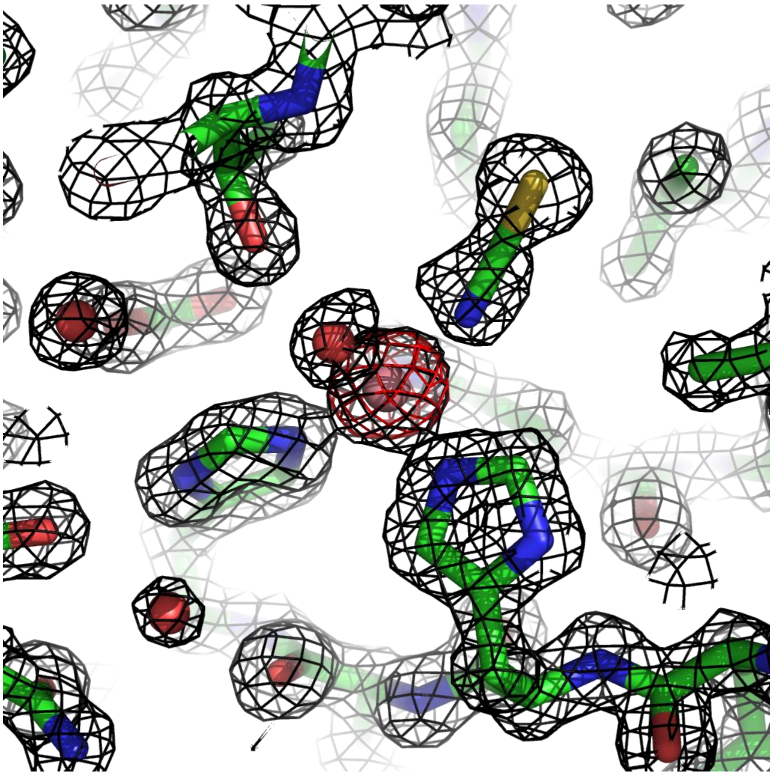


Figure 3.22: Active center of cobalt(II)-DM-hCAII bound to thiocyanate.The black mesh represents the electron density and the red mesh represents the anomalous signal arising from the cobalt(II) ion from DM-hCAII.

The anomalous signal arising from the active center of cobalt(II)-DM-hCAII-SCN$^-$ confirmed the presence of the cobalt(II) ion in the active center of the protein (**Figure 3.22**). Furthermore, as expected, the structure of hCAII does not change with the exchange of the metal ion at the active site.

3.4 Conclusions

The objectives of this work were to clarify the observed effects of high concentrations of sodium thiocyanate on the electronic structure of cobalt(II)-bCA, reflected in changes of their UV-vis and EPR spectra[3,14] (behaving as a type C ligand), and to investigate whether these effects can be exploited using paramagnetic NMR. To fulfill these objectives, a double mutant of hCAII, DM-hCAII, was developed. The mutations in the N-terminal of the protein, H3N and H4N, ensured a single binding site of a cobalt(II) ion in the DM-hCAII protein. To understand the effects of sodium thiocyanate on the chemical environment of the protein, several structural biology techniques were employed, mainly EPR, NMR, CD and X-ray crystallography.

The EPR titration of cobalt(II)-DM-hCAII with thiocyanate at 1:1 and 1:1000-fold excess confirmed the previous study with the cobalt(II)-bCA form[44]. It showed quite rhombic EPR spectra for the free cobalt(II)-DM-hCAII and its 1:1 SCN^- adduct, (1.620; 2.772 and 6.810) reflecting a 5-coordinate geometry around the cobalt(II) ion, typical of a type A ligand. At a 1:1000-fold thiocyanate excess, a more axial EPR spectrum was fitted with three spin systems (4.005; 3.520; and 2.800), resulting from a less distorted 4-coordinate geometry, typical of a type C ligand. Therefore, the study of the effects of thiocyanate excess as a type C ion could be extended to the human protein, cobalt(II)-DM-hCAII.

Samples of cobalt(II)-DM-hCAII and zinc(II)-DM-hCAII were monitored by ^1H-^{15}N-HSQC NMR titrations with increasing concentrations of sodium thiocyanate up to a 1000-fold excess. From these spectra PCS and RDC data were determined and used to calculate the axial ($\Delta\chi_{ax}$) and rhombic ($\Delta\chi_{rh}$) anisotropies of the magnetic susceptibility tensors for the cobalt(II) ion. The analysis showed that the $\Delta\chi$ tensor anisotropies increased substantially up to 1:2-fold thiocyanate excess, ensuring the presence of the 1:1 thiocyanate adduct, whose values are typical of a penta-coordinated structure and a large anisotropy tensor reorientation, as observed for other type A ligands like oxalate[4]. However, only smaller changes in the $\Delta\chi$ tensor anisotropies are seen upon passing 1:2 to a thousand-fold excess, which apparently does not correlate with the drastic changes observed for the EPR spectrum at high excess of thiocyanate.

To understand the type and extent of the effects that sodium thiocyanate has on the protein, further NMR studies, involving CSP index analysis and ^{15}N relaxation experiments, were carried out. The CSP analysis revealed which Zn-DM-hCAII protein residues interact with the thiocyanate anion at different concentrations. At 1000-fold excess concentration of ligand, new interactions involve mostly hydrophobic residues exposed at the surface of the protein and some hydrophobic non-solvent accessible residues, probably reflecting changes in the conformation and/or dynamics of the protein at high ionic strength. The ^{15}N relaxation experiments complemented the CSP results by showing that the hydrodynamic radius of Zn-DM-hCAII increases in the presence of a 1000-fold excess concentration of thiocyanate. The CD experiments reinforced the NMR data where there are not major changes in the protein structure, confirming that up to 10-fold thiocyanate excess the protein maintains it overall secondary structure and up to 1000-fold thiocyanate excess it maintains its hydrophobic core environment. Finally, a comparison of the new X-ray crystal structures of zinc(II)-DM-hCAII-SCN^- and cobalt(II)-DM-hCAII-SCN^- with that of Zn-WT-hCAII-SCN^-[22] showed that, although they have the same overall structure, comparison of their C_α RMSD shows that these are higher than 0.4 Å for residues in solvent exposed loops, which, due to their higher flexibility, might adopt slightly different conformations arising from different crystal packing forces influenced by the different crystallization precipitants used.

Considering the results of all the experiments, it can be concluded that sodium thiocyanate up to 1:1357 protein:ligand ratio, or 475 mM, increases the structural dynamics of the surface of the zinc(II)- and cobalt(II)-hCAII derivatives. The binding of this highly chaotropic ion to a hydrophobic surface displaces water molecules that may affect the global protein structure and dynamics[22]. However, the invariance of the $\Delta\chi$ tensor anisotropies, which are extremely sensitive to structural changes, for concentrations sodium thiocyanate higher than two-fold, show there is not a considerable change in the protein structure.

.

4

Chapter 4. – Investigation of nickel(II)-WT-hCAII

4.1 Introduction

The presence of two or more protein molecules in the asymmetric unit of protein crystals (meaning that they are crystallographically non-equivalent) is common. On occasions, there are clear structural differences between them, such as displacements of secondary structure elements, different conformations of loops, and rotation of domains with respect to one another. More often, the differences are small or almost insignificant. This phenomenon is called non-crystallographic symmetry[116] and it occurs frequently, much more than for small molecules, where it is known as partial symmetry. The causes for the non-crystallographic symmetry can be countless and unpredictable. One might consider there is a driving force that is not understood with our current theoretical and computational tools. This work describes an interesting case of non-crystallographic symmetry in two metalloprotein molecules in crystals of nickel(II)-substituted human Carbonic Anhydrase II (hCAII), which goes unexplained as for most cases of this type.

Nickel (II) has a $3d^8$ configuration with two potentially unpaired electrons, thus yielding a total spin S=1. The exception is nickel(II) with a square planar coordination geometry, where the electrons are paired and therefore it is diamagnetic. In the case of nickel(II)-hCAII, the active center is either five-coordinated or six coordinated, with a square pyramidal geometry or pseudo-octahedral geometry, respectively. When nickel (II) has a five or six coordinated geometry, the ground state is singly degenerate and the energy separation between the ground state and the excited states is large. This leads to an electron spin relaxation time close to 100 ps and a very small magnetic susceptibility tensor anisotropy ($\Delta\chi$) is expected and observed[46,50], meaning that resonances arising from nuclei close to the paramagnetic metal center may not be observed.

To date, several nickel(II)-containing proteins have been studied by NMR in solution due to the paramagnetic properties of nickel(II). Such examples present in the literature are: hCAII[32,43,46,117] (sometimes complemented by UV-vis studies[3-5]) rubredoxin[118], thioredoxin[119], azurin[120], HYPA[121], UreE[122] and CooT.[123] However, proteins containing nickel(II) are far unexplored by solid-state NMR (SSNMR). The recent methodological and technical improvements have brought SSNMR sensitivity and resolution to values comparable to solution NMR. As mentioned in Chapter 3, (SS NMR section), the dipolar coupling mechanism has a huge contribution to the SSNMR spectra causing the broadening of the resonances. To overcome this drawback, higher speed (> 100 kHz) MAS at high magnetic fields can be employed to decrease the dipolar coupling contribution. Another strategy requires a paramagnetic probe that decreases the relaxation time of the affected nuclei that are close to the paramagnetic center and, with a faster relaxation, a larger number of scans with short acquisition times can be employed to increase the signal to noise ratio.

The aim of the work described in this chapter was to produce crystals of hCAII containing the nickel (II) ion, with the double purpose of addressing the theoretical foundation of paramagnetic NMR and its effects on proteins. Furthermore, a second purpose is the optimization of SSNMR strategies for paramagnetic proteins[36]. As the X-ray crystallography study, carried out for nickel(II)-WT-hCAII crystals, presented an interesting example of non-crystallographic symmetry, a preliminary study of crystalline nickel(II)-WT-hCAII was also undertaken by SSNMR to allow the direct comparison between these two structural techniques.

4.2 Materials and methods

4.2.1 Expression and purification of WT-hCAII

Zinc(II)-hCAII was expressed and purified as previously reported[23]. The expression vector, pCAM coding the WT-hCAII sequence, was inserted in competent *Escherichia coli BL21-CodonPlus (DE3)-RIPL* cells using a standard heat shock protocol. The colonies were selected in LB-Agar plates supplemented with ampicillin and chloramphenicol. Overnight pre-inoculum of 10 mL preceded the large growth steps.

Double-labeled ^{13}C-^{15}N-WT-hCAII was expressed using an adapted Marley protocol. The protocol consists in growing overnight the cultures in two litters of LB-media supplemented with the respective selection antibiotics at 37 °C. In the following day, the saturated cultures are transferred from the LB-media to M9 minimal medium, by centrifuging the cultures at 4000 rpm for 15 minutes then resuspending in one litter M9 medium. The M9 medium contains $CaCl_2$, $MgSO_4$, 0.5 mM $ZnSO_4$, as well the respective antibiotics. The culture is left to adapt to the new medium for one hour and to deplete the unlabeled precursors of protein synthesis. Afterwards the expression of ^{13}C^{15}N-labbeled protein is induced with a solution containing 1.2 g/L of ^{15}N-enriched $(NH_4)_2SO_4$ 3g/L, ^{13}C-glucose and 1 mM IPTG. The expression of the protein was halted after 5 hours of expression by centrifuging the cultures at 8000 rpm. Cell pellets were stored at -20 °C. All the previous expression steps were done at 37 °C, 160 rpm.

The cell pellet was resuspended in a final volume of 70 mL of buffer, 20 mM Tris-SO_4, 0.5 mM $ZnSO_4$ pH 8. The cells were sonicated for 30 seconds with a resting period on ice for 3 minutes and this was repeated for a total of 10 times. The lysate was ultracentrifuged at 40000 rpm for 40 minutes and the supernatant was filtered with a 0.45 micra filter.

The lysate was firstly purified in a 5 mL Histrap FF, using an AKTA Prime System, previously equilibrated with 20 mM Tris-SO_4, 0.5 mM $ZnSO_4$ pH 8. The elution of the protein is done with a gradient of 20 column volumes from 0 to 50 % buffer containing 20 mM Tris-SO_4, 500 mM imidazole, 0.5 mM $ZnSO_4$ pH 8. Eluted fractions were analyzed by SDS-PAGE, and the fractions containing hCAII were pooled. Pooled fractions were concentrated using a 10,000 MWCO Amicon centrifugation device to 10 mL.

As a second purification step a 320 mL Superdex 75 pg 26/60 size exclusion column, using a AKTA Prime system, using as running buffer: 50 mM Sodium Phosphate pH 7, running at 2.5 mL/min. Purified protein was pooled and stored at 4 °C.

4.2.2 Demetallation and metalation protocol of WT-hCAII

The nickel (II) derivative was obtained from the zinc (II) enzyme by a demetallation/metalation approach.[2] For demetallation, a solution containing 200 mM sodium phosphate, 50 mM pyridine-2,6-dicarboxylic acid, pH 7, was added to the protein solution and incubated overnight at 4 °C. After demetallation, the apo-hCAII was buffer exchanged to 10 mM HEPES, pH 6.8, using 10,000 MWCO Amicon centrifugation devices. Protein samples were in a water buffered solution 10 mM HEPES, pH 8 with 10 % D_2O, with a protein concentration ranging from 0.3 mM to 1.5 mM.

The paramagnetic nickel (II)-WT-hCAII were prepared by titrating the of apo-WT-hCAII (10 mM HEPES, pH 8), with nickel (II) sulfate solution using mono-dimensional (^{1}H with solvent pre-saturation) and bidimensional spectra (^{1}H-^{15}N HSQC) acquired on Bruker Avance NEO 500 MHz NMR spectrometer (11.7 T, 125.8 MHz ^{13}C Larmor frequency). In these spectra, the disappearance of the resonances of the apo protein and the appearance of the paramagnetic resonances of the metalated protein at different chemical shift values were monitored. The titration was stopped when the resonances that correspond to the residues close to the paramagnetic center fully disappeared and when the new resonances, at a different chemical shift, did not increase in intensity anymore.

4.2.3 Crystallization, X-ray data collection, and structure determination

The purified nickel(II)-WT-hCAII protein sample had a concentration of 15 mg/ml in a solution containing 10 mM HEPES, pH 6.8. Crystallization trials were performed by the sitting drop vapor diffusion method at 20 °C, by mixing an equal volume of the sample and of a solution containing 100 mM HEPES pH 7.5, 2.9 M ammonium sulfate. Bunches of large, thin, and superposed crystals started to grow overnight and grew to final size in a few days. The crystal morphology did not allow to simply cut the protruding edges to get a suitable sample: for this reason, the micro-seeding technique was used to try and grow single crystals.

The best and largest crystal obtained was used to collect a SAD data set at the nickel edge (1.48306 Å) and diffracted up to 1.3 Å resolution at the synchrotron (beamline ID23-1 ESRF, France) with data completeness being good up to 1.45 Å. It belonged to the monoclinic space group P21 with two independent molecules in the asymmetric unit and a solvent content of about 40%. The data set was processed using XDS[124] and scaled using XSCALE[124]. The X-ray structure was solved using a default procedure implemented in the program auto SHARP[125]. Likewise, the model tracing was performed by looping between BUCCANEER[126] and ARP/wARP[127], which, in the end, allowed for a complete chain tracing. Further details on structure solving are described in the Results and in the

Discussion sections. The refinement was then carried out using REFMAC[128] making use of TLS restraints and water molecules were added during refinement. In between refinement cycles, the model was subjected to manual adjustments using COOT[106]. The stereochemical quality of the refined model was assessed using the program MOLPROBITY[129]. The SAD-solved structure of human nickel(II)-WT-hCAII has been deposited at the PDB under the accession code 6H6S. The relevant data collection and refinement statistics are also shown in **Table 4.1**.

4.2.4 Solid-state NMR experiments

Micro-crystals of ^{13}C-^{15}N-nickel(II)-WT-hCAII were obtained by the sitting drop vapor diffusion method at 293 K by mixing an equal volume of the sample (58.5 mg/mL of protein in 10 mM HEPES, pH 6.8) and of a solution containing 100 mM HEPES pH 7.5, 2.9 M ammonium sulfate[35]. SSNMR experiments were performed on the crystalline preparation of nickel(II)-WT-hCAII in 1.3 and 3.2 mm rotors. All the spectra were recorded on a Bruker Avance III HD 800 MHz spectrometer (19 T, 201.2 MHz ^{13}C Larmor frequency), equipped with 3.2 and 1.3 mm MAS probe heads in triple-resonance mode. The inter-scan delay was set to 2.5 and 1 s in ^{13}C-detected and ^{1}H-detected experiments, respectively. The temperature at the stator outlet was set to 260 K and 248 K in ^{13}C-detected and ^{1}H-detected experiments, respectively. Standard ^{13}C-detected SSNMR spectra were acquired at a MAS frequency of 14 kHz using the pulse sequences reported in the literature[130]. The nonselective 90° pulses were set to 2.5 μs at 100 kHz RF-field amplitude (^{1}H), 4.6 μs at 54 kHz RF-field amplitude (^{15}N), and 4 μs at 62.5 kHz RF-field amplitude (^{13}C). Proton decoupling was applied at 100 kHz with a SW$_f$-TPPM sequence[131]. Standard ^{1}H-detected 1D ^{1}H and 2D ^{1}H-^{15}N CP HSQC SSNMR spectra were acquired at a MAS frequency of 65 kHz using the pulse sequences reported in the literature[132,133]. The ^{1}H spectral window for 1D ^{1}H was set to 195 ppm. The nonselective 90° pulses were 1.65 μs (^{1}H) and 4 μs (^{15}N). Proton decoupling was applied at 17 kHz with a SW$_f$-TPPM sequence[134]. All the spectra were processed with the Bruker TopSpin 3.2 software and analyzed with the program CARA.

4.3 Results

4.3.1 X-ray crystallography study

Crystalline nickel (II)-WT-hCAII was prepared, and X-ray crystallography was used to determine its structure. The structure of substituted nickel(II)-WT-hCAII was solved at 1.45 Å resolution and it was deposited in the PDB database under the accession code 6H6S. The structure was determined and solved by Doctor Vito Calderone. The relevant data collection and refinement statistics are shown in **Table 4.1**. The solved structure shows two protein molecules in the asymmetric unit, as other deposited structures of hCAII. In each protein molecule there are two nickel (II) ions present, one in the active center and the other at a secondary binding site at the N-terminal. In addition, it is apparent that in both sites the coordination of the nickel (II) is different. At the secondary binding site, the nickel (II) ion has a square planar geometry in both molecules, whereas two different coordination geometries are present at the active center in the different molecules, a penta-coordinated and an hexa-coordinated one.

Table 4.1: Dara collection and refinement statistics for nickel(II).WT-hCAII.

	Nickel(II)-WT-hCAII
Diffraction source	ID23-1 ESRF
Wavelength (Å)	1.483
Temperature (K)	100
Crystal-to-detector distance (mm)	126.70
Oscillation range (°)	0.150
Total rotation range (°)	360
Space group	$P2_1$
a b c (Å), α β(°)	73.48 41.18 84.04, 109.2
Mosaicity (°)	0.150
Resolution range (Å)	45.34 – 1.45
Total reflections	478304 (54398)
Unique reflections	154486 (22243)
Completeness (%)	93.8 (83.7)
CC1/2	99.9 (93.3)
Anomalous reflections	68173 (6868)
Anomalous multiplicity	2.4
I/σ(I)	12.6 (3.1)
R_{merge} †	0.04 (0.24)
Wilson plot B factor (Å^2)	20.8
N° of sites	4
f’ / f’’ (refined)	-7.3 / 2.8
FOM (SHARP)	0.22
FOM (PARROT)	0.76
R_{cryst} / R_{free}‡ (%)	14.9 / 20.3
Protein atoms	4118
Water molecules	807
Ions	4
RMSD bond lengths (Å)	0.008
RMSD bond angles (º)	1.1

† $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle / \sum_{hkl}\sum_{hkl} I_i(hkl)$, where I_i(*hkl*) is the mean intensity of the *i*th observation of symmetry-related reflections *hkl*.

‡$R_{cryst} = \sum_{hkl}\left||F_{obs}| - |F_{calc}|\right| / \sum_{hkl}|F_{obs}|$, where F_{calc} is the calculated protein structure factor from the atomic model (R_{free} was calculated with a randomly selected 7% of the reflections).

4.3.2 Comparison between the two non-equivalent molecules in the asymmetric unit

A comprehensive analysis of the two protein molecules in the asymmetric unit was carried out under several considerations. Such considerations were their superposition, the atomic interactions (including within the asymmetric unit and with symmetry related molecules), the contact surfaces and the thermal factor behavior. Additionally, the obtained structure (6H6S) was compared with two previously determined structures, a nickel (II)-hCAII structure (1RZE) with one nickel (II) ion per protein and one protein molecule per asymmetric unit[29], zinc(II)-hCAII structure (3MWO), which has two protein molecules in the asymmetric unit and one zinc(II) ion per protein molecule[135].

4.3.3 Protein structure and nickel environment determination

To corroborate the substitution of the zinc(II) ion, the data collection was carried at the nickel edge and the structure determination was performed using the single-wavelength anomalous dispersion (SAD).

The substructure model of nickel(II)-WT-hCAII was extensively built to almost completeness using the high-quality preliminary phases along with the good data resolution obtained. After the structure determination, the presence of a second nickel(II) ion in the structure became clear, which was present at the N-terminal site. The coordination of the second nickel(II) ion is comprises by two residues, His3 (N) and His4 (N and ND) and one water molecule. This study is the first to report a secondary binding site of nickel(II) in hCAII.

4.3.4 Solid state NMR study

The X-ray crystallography showed that nickel(II)-WT-hCAII has two nickel(II) ions per protein molecule, which was a novel discovery. To date, nickel (II) proteins have been used by solution NMR due to their paramagnetic properties, such as hCAII[32,43,46,117], rubredoxin[118], thioredoxin[119] and azurin[120]. However, proteins containing nickel(II) are far unexplored by solid-state NMR. The X-ray crystallography study paved the path to characterize the crystalline nickel(II)-WT-hCAII by SSNMR to allow a direct comparison between the results of these two structural techniques. Here, some preliminary data are presented showing that high quality NMR spectra can be acquired in the solid state for double labeled ^{13}C-^{15}N-nickel(II)-WT-hCAII, as illustrated by **Figure 4.1-4.3**.

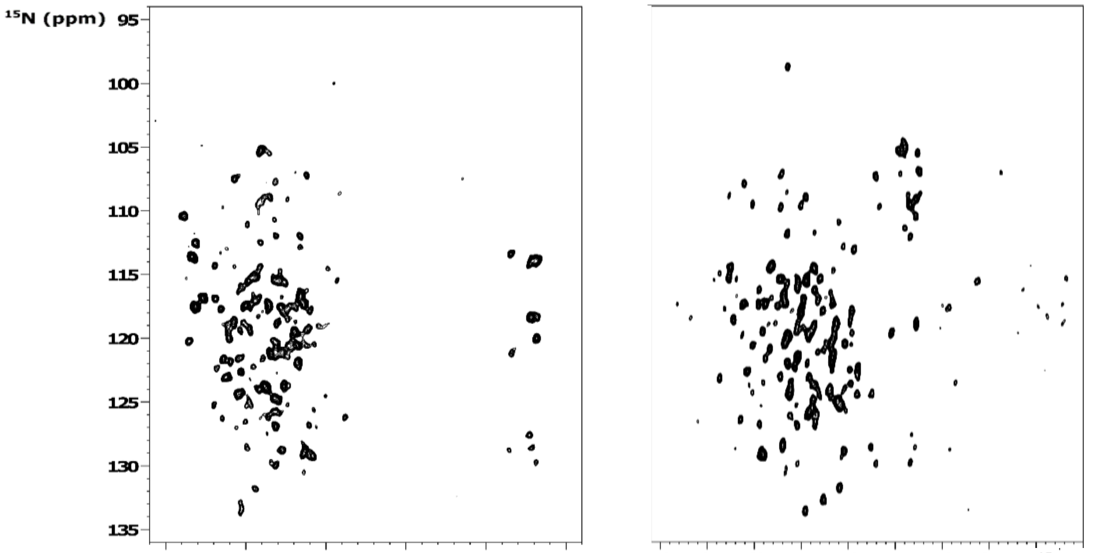


Figure 4.1: SSNMR NCO and NCA correlation spectra acquired on nickel(II)-WT-hCAII. The spectra were acquired at 800 MHz, 260 K at the stator outlet, 14 kHz MAS (adapted from [36]).

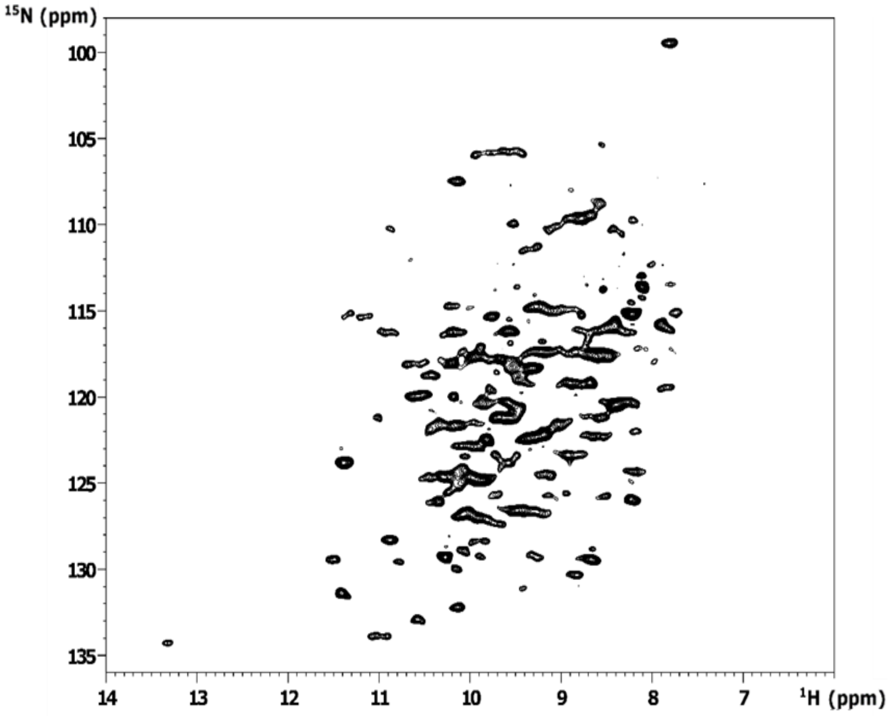


Figure 4.2: SSNMR HN correlation spectrum of nickel(II)-WT-hCAII at 800 MHz, 248 K at the stator outlet, 65 kHz MAS (adapted from [36]).

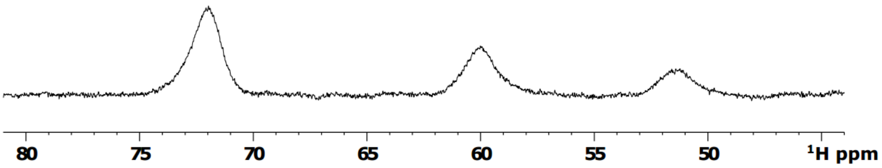


Figure 4.3: Enlargement of the low-field region of the ^{1}H SSNMR spectrum of nickel(II)-WT-hCAII. The spectrum was acquired at 800 MHz, 248 K at the stator outlet, 65 kHz MAS (adapted from [36]).

The acquired spectra have high quality and many resonances are resolved, indicating that the nickel(II) ion induces large hyperfine shifts that can break the dipolar bath even in the absence of protein deuteration. This happens because the presence of the unpaired electrons from the nickel(II) ion interact with the nuclei and change the frequency of the nuclei, thus disrupting the dipolar coupling. The paramagnetic nickel(II) ion induces large contact shifts, due to the covalent character of its bonding to the nitrogen atoms of the histidines' imidazole rings at the active center. This effect is seen in the ^{1}H NMR spectrum, where the ^{1}H resonances from the protons of the histidine's (H94, H96 and H119) imidazole rings in the *meta*-like position relative to the metal-binding nitrogen are in the 50-80 ppm region (**Figure 4.3**). The NCO and NCA spectra are valuable for the protein backbone assignment. A closer inspection of these spectra reveals that there is no doubling of the resonances, meaning that the different coordination geometry in the protein's active center, seen by X-ray crystallography, do not affect the NMR spectra.

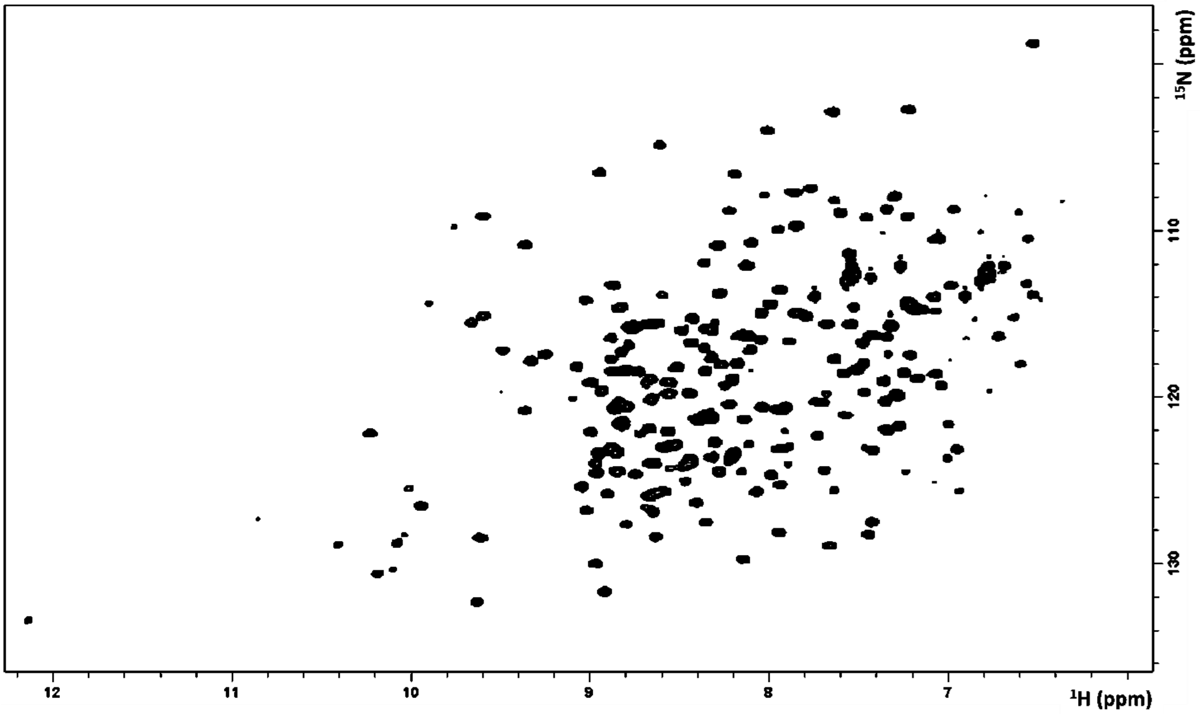


Figure 4.4: ^{1}H-^{15}N HSQC spectrum of nickel(II)-WT-hCAII, in 10 mM HEPES, pH 6.8, recorded at a 500 MHz. Protein concentration is 3300 µM.

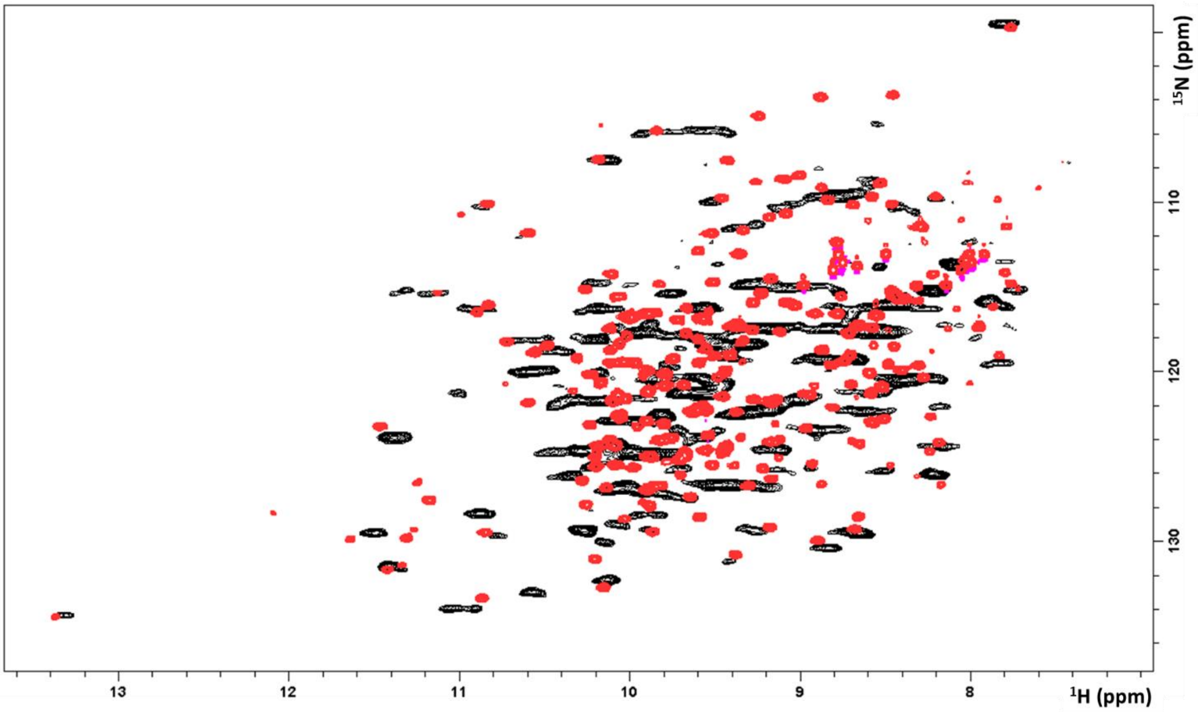


Figure 4.5: Superposition of solid state NMR ^{1}H-^{15}N-HSQC spectrum (black) and solution state ^{1}H-^{15}N-HSQC spectrum (red) of nickel(II)-WT-hCAII.

Using higher MAS speeds, it is possible to acquire a 2D ^{1}H-^{15}N CP HSQC SSNMR spectrum, which is the analog of the classic ^{1}H-^{15}N HSQC solution NMR spectrum (**Figure 4.4**). The superposition of solid state and solution NMR spectra (**Figure 4.5**) shows that both spectra have a similar resonance dispersion, however the solid-state resonances are broadened due to the dipolar coupling, as expected.

Furthermore, nickel(II)-WT-hCAII's solid-state resonances are shifted to high field due to the presence of 2.9 M ammonium sulfate and to crystal packing forces that change the chemical environment of the observed residues.

4.4 Discussion

4.4.1 X-ray crystal structure

Molecular determinants of the asymmetry

Before the present work, in the published X-ray crystal structure of nickel (II) substituted hCAII (1RZE)[29] there is only one protein molecule per asymmetric unit and one nickel (II) ion replacing the native zinc(II) ion, with no metal bound in the N-terminal secondary site.

To try to explain the non-crystallographic symmetry observed in the new structure (6H6S), the two molecules found in the asymmetric unit were superimposed. The molecular contact analysis of the two molecules in the asymmetric unit showed that only 40 % of the residues involved in these contacts are the same, whereas the other 60 % of the residues are different, which shows an asymmetry. Additionally, the residues involved in the molecular contacts between the molecules from the asymmetric unit and the symmetry mates are totally different for the two protein molecules of the asymmetric unit.

B-factors, also known as temperature factors, describe the attenuation of the scattered X-ray caused by thermal motion.[82] Comparison of the B-factors from each of the two molecules, A and B, in the asymmetric unit of the 6H6S structure with those from the 1RZE structure (**Figure 4.6**), shows that there are very small differences between molecules A and B in the B-factor distribution, but they show significantly increased values relative to 1RZE in clustered regions. However, these regions, where the B-factors are higher, are not obviously correlated with the residues that differ in the crystal packing interactions. As for the molecular contacts mentioned above, only the differences can be estimated but not their origin.

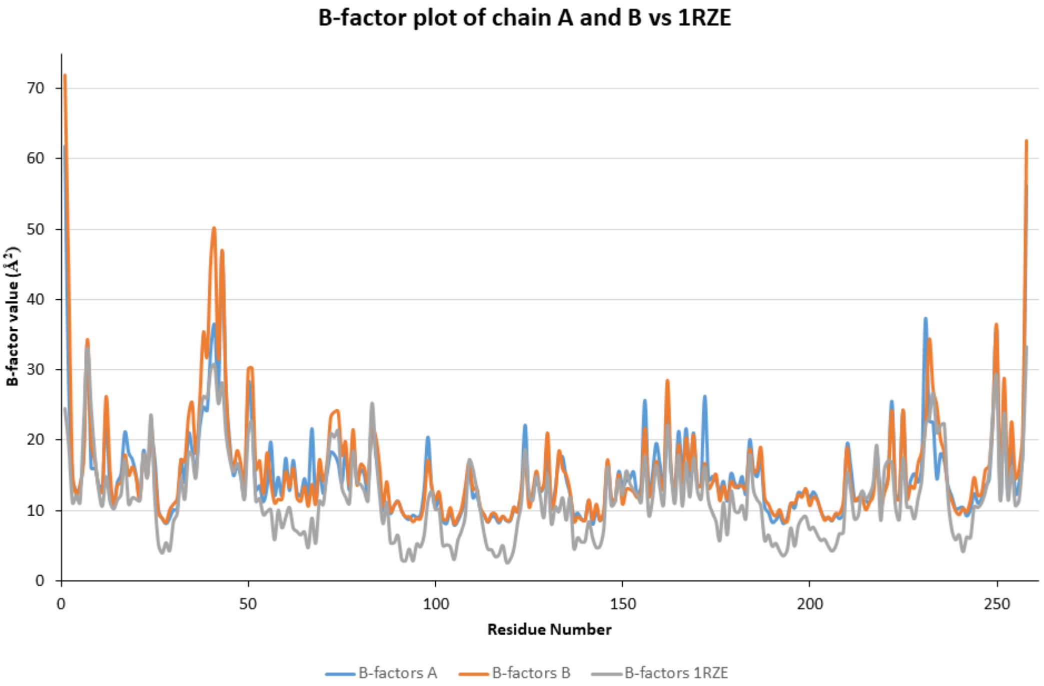


Figure 4.6: Comparison of the B-factor values between molecule A and B of 6H6S vs 1RZE (adapted from [35]).

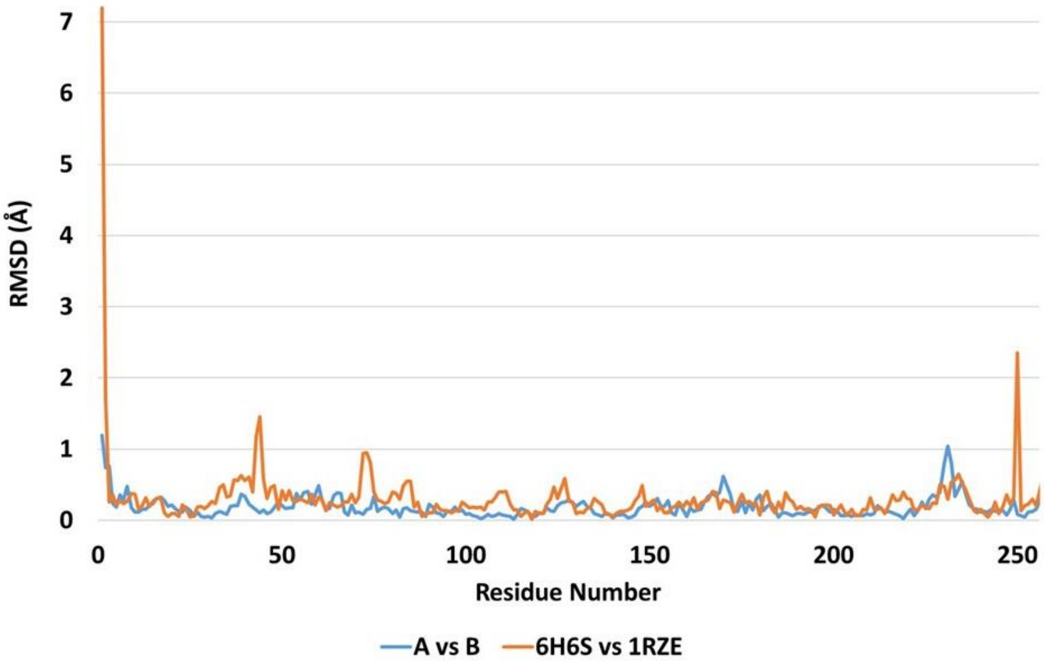


Figure 4.7: C$_\alpha$ RMSD between molecule A and B (blue) in the asymmetric unit and between molecule A of 6H6S and 1RZE(orange) (adapted from [35]).

To try to understand the determinants of the presence of two non-equivalent protein molecules in the asymmetric unit, their structure was superimposed and compared (**Figure 4.7**). This figure shows the RMSD for the C$_\alpha$ atoms between molecules A and B, from 6H6S, and between molecule A of 6H6S and 1RZE. For molecules A and B, from 6H6S, the global RMSD is around 0.18 Å, which is globally very small. In addition, the RMSD is especially small around the active center, due to the structure conservation. The deviation value is higher in two regions where the electron density, and thus, the model position, are not well defined, but they are not related to the active center in an obvious way.

Molecules A and B were superimposed using only the positions of the nickel ion and the coordinating nitrogens from the three histidines of the active center to determine the differences of the coordination geometry in the active site. The superposition was represented in two ways, as atomic positions and as occupied volumes in space filling representations (**Figure 4.8** and **Figure 4.9**, respectively).

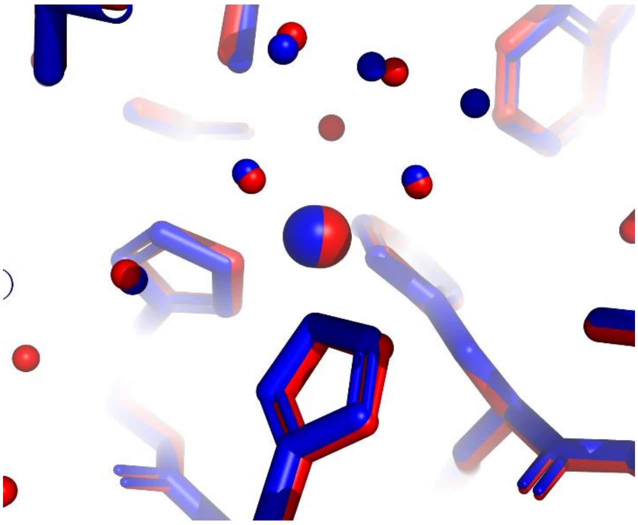


Figure 4.8: Superposition of the active site region between molecule A and B in the asymmetric unit. Penta coordinate in blue and hexacoordinated in red (adapted from [35]).

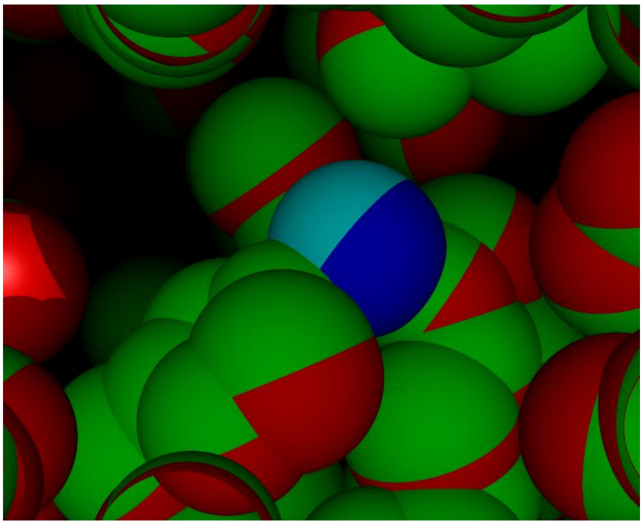


Figure 4.9: Superposition of the active site region of molecule A (red) and B (green) with the respective nickel atoms for A (cyan) and B (blue) (adapted from [35]).

The structure superposition (**Figure 4.8**) confirms that the atomic coordinates' differences in the active site are very small. The obvious difference is the absence of one water molecule coordinating the nickel ion in the penta-coordinated site (molecule B) and the presence of an extra water molecule that is not interacting with the nickel ion. Two more water molecules' positions are modestly shifted in the second coordination sphere between molecules A and B. Analyzing the volumes of the active site's atoms (**Figure 4.9**), it is also visible that there are only minor differences that cannot be easily related to the different arrangements of the water molecules in the active site.

The number of intramolecular hydrogen bonds in the protein molecules is 160 for molecule A and 163 for molecule B, which are very close. Furthermore, the interface areas and the number of interactions occurring between the A and B molecules were analyzed using the software PISA.[136] The difference in interface areas is significant, which are ca. 1850 Å^2 for molecule A and 2100 Å^2 for molecule B.

The nickel(II) ion in nickel(II)-hCAII from 1RZE has an octahedral geometry like the molecule A from 6H6S. Furthermore, the nickel (II) ion's ligands are three histidines, two water molecules and a sulfate ion, that has a partial occupancy. Compared to 6H6S, the sulfate ion is missing in the latter, even when the crystallization conditions are comparable. In the place of the sulfate ion, there is an array of water molecules and this is confirmed by the absence of positive F_0-F_c density in the putative coordinate of the sulfur atom and by the value of B-factor for the oxygen atom of the water molecule in that position relative to atoms in the surrounding environment. Furthermore, the shape of the streak of electron density does not match either HEPES, which was the buffer used in the crystallization, or ethylene glycol, the cryo-protectant.

Both structures have the same space group ($P2_1$), however the cell dimensions are different, in such a way that the resulting cell volume of our structure accommodates two independent molecules instead of one present in 1RZE. The parameters of the unit cell are a= 42.7 Å, b= 41.7 Å, c= 73.0 Å, β= 104.6° for 1RZE and a= 73.5 Å, b= 41.2 Å, c=84.0 Å, β= 109.2° for our structure 6H6S. The cell volume of our structure is about twice that of 1RZE, i.e., the two non-equivalent protein molecules A and B from our structure are crystallographic equivalents in 1RZE and with nearly the same poses in the two crystals. An argument can be made that the higher resolution of our study allows detection of smaller differences between the two molecules in the unit cell, thus doubling the volume of the unit cell compared with 1RZE. However, despite the larger cell dimensions, the volume of the cell is not doubled. For the 6H6S structure the cell volume is 240219 Å^3 and the cell volume for 1RZE is 125785 Å^3. This means that there is a significant difference of 10000 Å^3 missing. This means the solvent contact is different for both structures, 40.4 % for 6H6S vs 43.2 % for 1RZE.

This difference is puzzling, and some suggestions to explain the lattice variation are proposed here. It is known that the protein sequence is the same, the crystallization conditions are similar, and the C_α RMSD between 1RZE and 6H6S is relatively low (0.30 Å). The latter parameter does not show local discrepancies, except for regions where the electron density was weakly defined (**Figure 4.7**). The difference in temperature of the data acquisition of the two datasets could explain this phenomenon. The 1RZE dataset was collected at a diffractometer at room temperature, before the cryo-streaming was implemented[137]. In this work, the dataset was collected at a synchrotron at 100 K, meaning that the large difference of temperatures might play a role in the lattice's contraction. This contraction might have induced a dehydration of the lattice, which might be the link between the two different hydrations of the nickel atoms in the active site[116]. However, simply counting water molecules in both structures is not a reliable dehydration parameter because at higher resolutions a higher number of water molecules are observed due to their clear electron density. Furthermore, in the present case the number of water molecules is lower, therefore this evaluation is meaningless.

The reduced number of water molecules in the lattice of 6H6S might explain the increase in the number of interactions between the protein molecules of the crystal relative to 1RZE. In the 1RZE structure, each protein has 11 hydrogen bonds with its symmetry mates, whereas in 6H6S, the protein molecules A and B establish 14 hydrogen bonds each. Similarly, the interface area between the protein molecules and the symmetry mates is different: around 1600 $Å^2$ for 1RZE and around 1850 $Å^2$ for molecule A and 2100 $Å^2$ for molecule B in 6H6S.

To emphasize the differences between 6H6S and 1RZE structures, both structures were superposed as follows: molecule A, from 6H6S, was superposed with the molecule from 1RZE and the resulting C_α RMSD was calculated. Then, the symmetry of the 1RZE structure was expanded to show its symmetry partners. The 1RZE's symmetry partner closest to molecule B from 6H6S was chosen to calculate the C_α RMSD. As a result, the global RMSD underwent a significant rise from 0.30 Å to 0.38 Å, meaning that, even at a short range, there is a significant shift of the molecules in the lattice between 1RZE and 6H6S (**Figure 4.10**).

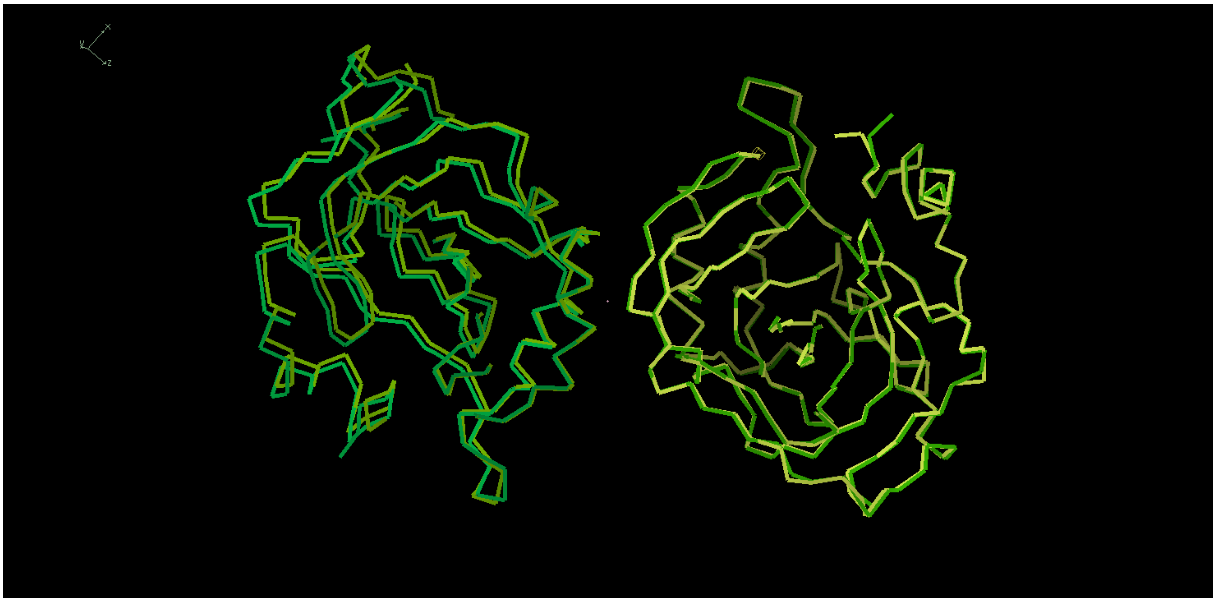


Figure 4.10: Superposition of 6H6S (molecule A and B in green) with 1RZE (molecule A and its symmetry mate in yellow)[35].

This poses the question if the small asymmetry in the crystal is induced by the low temperature, as suggested by the absence of asymmetry in the room-temperature 1RZE structure, or if this asymmetry also exists at room temperature. To answer this question solid-state NMR (SSNMR) can be employed (see **Section 4.4.3**).

There is another interesting comparison to be made with the structure of the native enzyme zinc(II)-hCAII (3MWO, a re-refinement of 3KSI)[135], which crystallized under the same space group and has two molecules per asymmetric unit like 6H6S, with a doubled cell with respect to other hCAIIs. Furthermore, the data for 3MWO was collected at 100 K. The C_α RMSD between the two non-equivalent protein molecules of 3MWO is 0.15 Å whereas the C_α RMSD between 6H6S and 3MWO is 0.17 Å and 0.18 Å for molecule A and B, respectively (**Figure 4.10**).

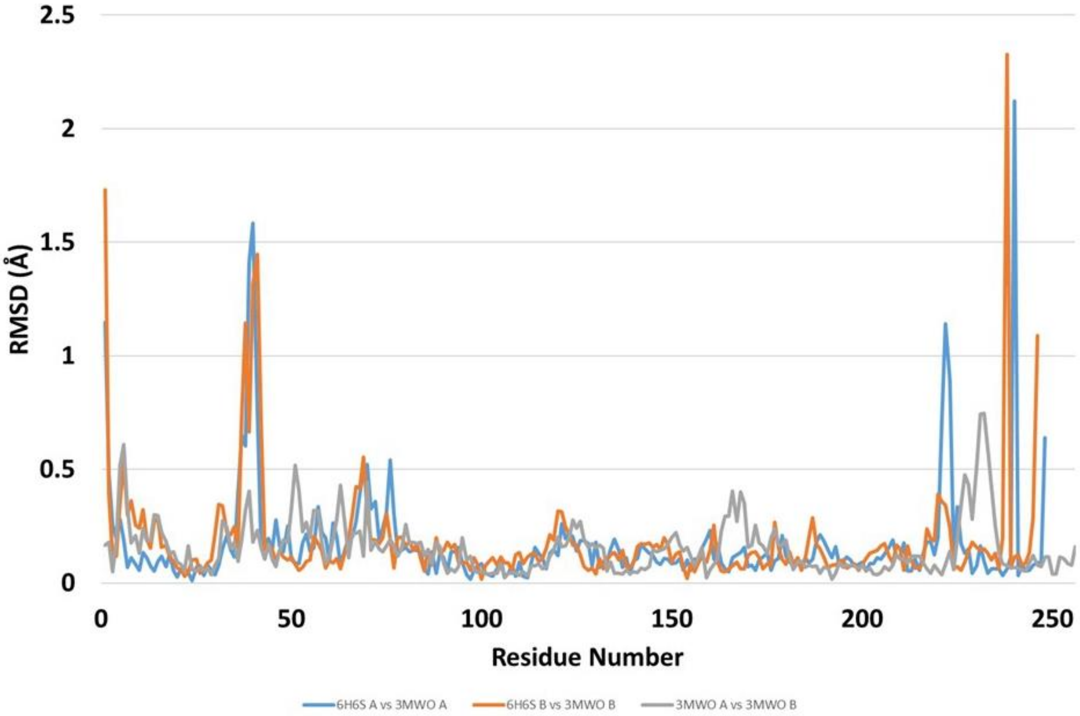


Figure 4.11: C_α RMSD between chains A and B of 6H6S and chains A and B of 3MWO respectively and between chains A and B of 3MWO (adapted from [35]).

Figure 4.11 shows that there is extremely low deviation between the 3MWO and 6H6S structures except in two regions, one close to the N-terminal and the other at the C-terminal. It is worth mentioning that these regions where the RMSD is higher, the N-terminal and C-terminal, are over 20 Å and 7 to 10 Å, respectively, away from the active center. Furthermore, the previously mentioned regions are loops exposed to the solvent, meaning that there is room for flexibility, which is proven by the low electron density in 3MWO at those regions.

4.4.2 Metal coordination facts

After solving the structure, the presence of two nickel atoms per molecule instead of one was obvious. The first nickel atom is present in the expected region, the active center, and the second one is present at the N-terminal site of the molecule. The presence of both nickel atoms, per molecule, was confirmed by using an omitted anomalous difference Fourier map, which indicated one very intense and one weak peak for each molecule.

The two nickel atoms present in the active centers, one per molecule in the asymmetric unit, replace the native zinc atoms at the active center, and they have a full occupancy and a thermal factor of about 10 Å^2. The structure was also refined with partial occupancies for the nickel atoms at the active site, however it provided positive F_0-F_e densities and less plausible thermal factors.

The nickel atom is an anomalous scatterer and at its edge energy it can introduce artifacts on the electron density. This can be avoided by having a high energy dataset; however, it is not the case. To circumvent this, the dataset was refined using two different principles in parallel, and at the end the results were compared. In one approach, the refinement was carried out using the averaged structure factors and, in the other approach, using the separated F^+ and F^-. Both approaches led to roughly the same results regarding the R_{free}/R_{cryst} and geometry and in terms of electron density, thus confirming the coordination environment around the nickel sites. The nickel sites' environments are not equivalent, where one (molecule A) is hexa-coordinated (**Figure 4.12**) and the other one (molecule B) is mostly penta-coordinated (**Figure 4.13**). To exclude artifacts, an omitted map was calculated using the CPP4 software OMIT[138], which minimizes the local phase bias. With the omitted maps, it is safe to conclude that the electron density around the two nickel atoms of the active site is different.

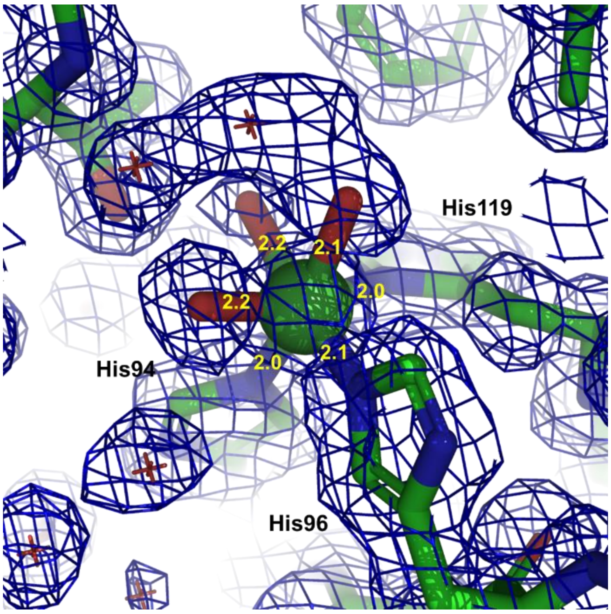


Figure 4.12: Representation of the primary nickel binding site in the first of the two molecules in the asymmetric unit showing the nickel octahedral coordination. The sixth position of the putative octahedral coordination is shown with a significant electron density (contoured at 1 σ level) accounting for it (adapted from [35]).

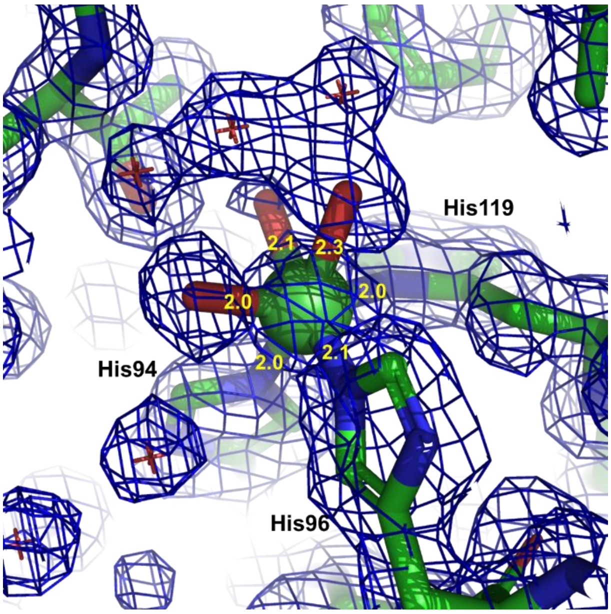


Figure 4.13: Representation of the primary nickel binding site in the second of the two molecules in the asymmetric unit showing the nickel penta coordination. The sixth position of the putative octahedral coordination is shown with a very faint electron density (contoured at 1 σ level) accounting for it (adapted from [35]).

The two nickel atoms at the N-terminal of each molecule in the asymmetric unit have lower intensity in the anomalous difference map. Both nickel atoms are bound to two histidines of the molecule, His 3 and His 4. The coordination is formed by two nitrogens from different backbone residues, one nitrogen from the histidine sidechain and one water molecule (**Figure 4.14**). For both cases, the coordination is square planar, i.e., it is diamagnetic. The lower electron density of the second nickel can be explained by the mobility of the N-terminal, instead of a partial occupancy. This means that the binding of nickel to the secondary site is efficient. The binding of the nickel(II) ion in the secondary site is confirmed by NMR. The NMR experiments, described later in the chapter, show that there is a shift in the residues close to the secondary His 4 and His 64. Both nickel atoms, at the N-terminal, refine with a full occupancy and with thermal factors consistent with protein atoms bound to them. The thermal factors for the N-terminal nickel are three to four times higher than the thermal factors of the nickel atoms bound in the active center. The binding of a second metal at the N-terminal site is not new. Another metal has been reported and it was copper (II) in the structure 1RZC (see section 1.1.3). However, the binding of copper is different from that of nickel whereas the copper is coordinated by His 4 and His 64 and has an occupancy of 0.2.

The 3MWO structure has two molecules in the asymmetric unit and both zinc(II) metal ions have the same tetrahedral coordination geometry, whereas 6H6S has different nickel geometries. Although zinc(II) is less inclined than nickel(II) to adopt different geometries, it makes the observation of different nickel(II) geometries in 6H6S more intriguing, because this is the only difference in the two non-equivalent molecules in the asymmetric unit.

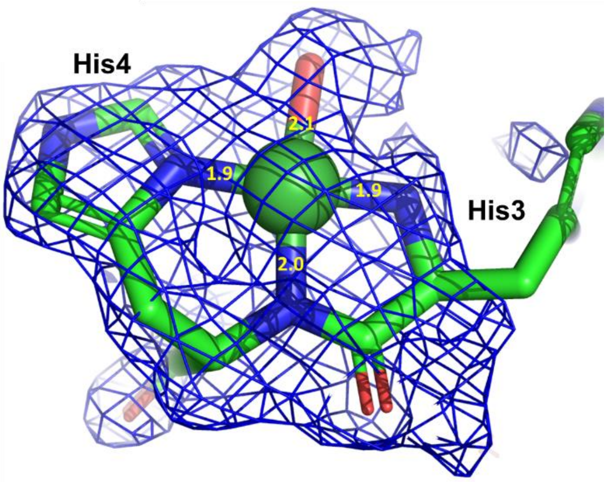


Figure 4.14: Representation of the secondary nickel binding site showing the density (contoured at 1σ level) accounting for the nickel square planar coordination (adapted from [35]).

General considerations on partial symmetry: speculations on the “Jahn-Teller” and the “Butterfly” effects

Despite the very detailed analysis of the two non-equivalent molecules, it is not possible to directly derive the reason for this occurrence. Therefore, the present system is a good example to speculatively discuss the non-crystallographic symmetry phenomenon that is regularly found in protein crystals, which most of the time is unknown and unpredictable. First, it is required to assume that the nickel(II) ion in hCAII at room temperature undergoes a 50/50 equilibrium between a five- and six-coordinated geometry and that this equilibrium is in fast-exchange. Furthermore, it is reasonable to assume that the overall shape of the protein does not change with the nickel(II) coordination number, as it is buried in that active site cavity. If these assumptions are true, then it is expectable that at room temperature the crystal of nickel(II)-hCAII contains a mixture of five- and six-coordinated nickel(II) in the active site and the exchange is also fast, as seen by solid-state NMR (see later). By lowering the temperature, the equilibrium is frozen, however the distribution is not random, but organized in an ordered way, with an A and a B molecule in the asymmetric unit, which repeats itself in the crystal lattice. Compared to a random distribution of states, this repeated pattern amounts to a decrease in entropy. Although there is less entropy, this effect is not detrimental due to the low temperature; however, enthalpy also needs to decrease in order to more than compensate for that effect, due to overall lowering of the free energy. To decrease the enthalpy, the number of intra- and intermolecular interactions must be greater, and this is achieved by creating an asymmetric unit with two inequivalent molecules, which is observed. To further confirm this, a solid-state NMR spectrum of nickel(II)-hCAII should be acquired at 100 K, however to date there are no commercial apparatus capable of performing these measurements at high spinning and at low temperatures.

In inorganic chemistry, the Jahn-Teller effect occurs regularly. This phenomenon states that, under specific conditions, there is a removal of orbital degeneracy by lowering the system's symmetry, which in turn decreases the potential energy of the system. There are many cases where proteins crystallize with more than one molecule in the asymmetric unit, meaning that lowering the symmetry is energetically more beneficial. This also compensates for the lowering of entropy that occurs due to the orderly distribution of different coordination geometries, such as those observed in the nickel(II)-hCAII.

Having said that there is asymmetry in the protein and asymmetry in the nickel coordination sphere and both are related, it is not possible to determine which is the origin of the phenomena. There are many cases of non-crystallographic symmetry in proteins in the literature, and few cases involving metalloproteins. Knowing this, it makes more sense that the protein distortion is the cause for the nickel asymmetry. The small protein distortions might not occur due to a single structural change but to an ensemble of small perturbations that propagate throughout the protein, which at the end affects the metal coordination geometry. This looks like a barely deterministic situation because the relation between the cause and the effect is unknown. To say that a small variation can have a certain effect is different from saying that a slightly different variation can lead to an opposite effect. This happens in the theory of deterministic chaos, which includes the Lorenz attractor, the Mandelbrot set and the fractals. The butterfly effect from the chaos theory states that small changes in initial conditions can lead to large effects on a later state through nonlinear systems. With this, it is speculated that a butterfly effect may exist in the structural features of proteins.

Recapping, lowering the symmetry stabilizes the structure of the whole crystal and small changes related with the lower symmetry are propagated unpredictably through the whole molecule.

4.4.3 Comparison of the X-ray crystal structure with the preliminary SSNMR data

The previous discussion on the 6H6S structure of the nickel(II)-WT-hCAII obtained at 100 K posed the question if the small asymmetry in the crystal is induced by the low temperature, as suggested by the absence of asymmetry in the room-temperature 1RZE structure, or if this asymmetry also exists at room temperature. To answer this question SSNMR can be employed to detect molecular inequivalences. SSNMR has shown to be sensitive to detect small changes in the metal coordination that translate to different chemical shifts[46]. An example that shows this is a 4Fe4S protein, where a modest difference in the asymmetric unit, C_α RMSD = 0.67 Å, is reflected in doubling the resonances of the ligand bound to the protein. These doubled proton resonances of the ligand had a difference of up to 16 ppm[139]. For nickel(II)-WT-hCAII, the recorded proton SSNMR spectrum only had one set of peaks. These peaks correspond to the imidazole rings of the coordinating histidines and these resonances are essentially at the same position as the resonances recorded using solution NMR. The different coordination geometries of nickel(II), five and six coordination, are expected to give different paramagnetic chemical shifts. Since only one set of paramagnetic shifts is observed, then there must be a dynamical equilibrium between these two coordination geometries, as previously suggested using electronic spectroscopy[140]. Although this equilibrium is frozen at low temperatures, the two different nickel(II) environments in the nickel(II)-WT-hCAII interconvert rapidly on the NMR timescale at room temperature, at variance with the case of the 4Fe4S protein. According to literature, [43] (although the rearrangement values in the coordination environment could be larger) it could be expected contact shift values in the 10 ppm range, implying a interconversion frequency larger than 8 kHz. A close inspection of the NCA and NCO spectra (**Figure 4.1**) also does not seem to show doubling of the resonances, although the PCSs are expected to be in the −1 to 0.5 ppm range (i.e.: larger than the linewidth of the ^{13}C resonances).

4.5 Conclusions

In this work, we have investigated the crystal structure of the nickel(II)-substituted hCAII. It is not the first time that the structure of this system has been determined. However, we find that, possibly because of the lower temperature, the present structure shows significant differences relative to the previously solved one. In this case there is a secondary nickel binding site at the very N-terminus of the molecule which has never been observed in nickel-substituted hCAII. Furthermore, there are two independent molecules in the asymmetric unit instead of one, and the nickel(II) ions in the active site of the two molecules show a clearly different coordination geometry. In this case the observation is made striking by the alteration, even within the same crystal, of the coordination geometry of the nickel(II) ions which, to a coordination chemist, would appear as coordinatively stiff objects. This system is an ideal case to discuss the subtleties of the energetic reasons that often favor lowering the symmetry in biomolecular crystals[141]

The small asymmetry observed in the X-ray structure of the nickel(II)-WT-hCAII at 100 K was further studied by SSNMR of the same system at room temperature. This showed that there is a dynamical equilibrium between the two nickel(II) coordination geometries, that is frozen at low temperatures, but is in fast exchange in the NMR timescale at room temperature, in agreement with the absence of asymmetry in the room-temperature 1RZE structure.[10]

5

Chapter 5. – The pseudo-contact shifts of Copper(II)-hCAII

5.1 Introduction

Copper proteins are often involved in electron-transfer processes or catalyze oxidative reactions, and for that purpose they use the two available oxidation states, copper(I) and copper(II). copper(I) is a diamagnetic d^{10} metal ion, with very small and localized effects on the NMR spectra of proteins. However, the presence of the oxidized state copper(II), with a d^9 electronic configuration, has large effects on the NMR spectra of proteins, due to the paramagnetic properties resulting from its unpaired electron. The hyperfine coupling of the magnetic moment associated with this unpaired electron with those of nearby nuclei strongly affects their chemical shifts and relaxation rates. Consequently, a strong increase of their nuclear transverse relaxation rates (R_2), induced by copper(II), produces broad NMR lines, as their NMR signal linewidths are proportional to their R_2 values, which can strongly limit their detectability.[50]

Mononuclear copper centers in proteins are classified into two classes, type 1 centers, typical of blue copper proteins, and type 2 centers.[142] Their different geometries, coordinating atoms, and consequently very different energy separations between their ground and excited electronic states lead to very different electron spin relaxation times, with $T_{1e} \cong 0.1\text{-}0.5$ ns for type 1 copper(II) and $T_{1e} \cong 2\text{-}8$ ns for type 2 copper(II)[50,143]. The relatively fast electron relaxation causes only limited broadening of the NMR spectra of oxidized type 1 blue copper proteins, allowing early assignment of proton NMR spectra of several oxidized proteins of this kind[143–147] and to solve the first solution structure of a ^{15}N-labeled protein (plastocyanin) by NMR using standard 2D and 3D heteronuclear NMR experiments that were adapted to the fast relaxing signals of nuclei in the vicinity of the copper(II) site[148]. However, in the case of type 2 copper(II) proteins, the slower electronic relaxation of the metal ion causes the proton signals of residues close to the paramagnetic center to become undetectably broad, which limits the applicability of the classical ^{1}H NMR based structural methods. Newer techniques based on ^{13}C direct detection provided important progress[143]. As the dipolar contributions to nuclear relaxation depend on the square of the magnetogyric ratio (γ_I) of the observed nucleus and $(\gamma_{1H}/\gamma_{13C})^2 \cong 16$, going from ^{1}H to ^{13}C detection leads to a decrease in relaxation rates by a factor of about 16 occurs. The resulting reduction in sensitivity can be overcome by the use of high field spectrometers, where the increase of R_2 is compensated by a decrease of the longitudinal relaxation rate R_1. 2D ^{13}C-^{13}C correlation experiments were used to detect and assign resonances from most of the nuclei of residues that are usually inaccessible by ^{1}H NMR. Using paramagnetic R_1 enhancements as structural constraints to locate the copper(II) ion within the protein frame, as well as ^{1}H and ^{13}C pseudocontact shifts (PCS's) as structural constraints, the solution structure of copper(II) CopC from *Pseudomonas syringae* was obtained, where information is only missing in a sphere with a 6 Å radius from the copper ion[149].

The aim of the present study is to investigate experimentally why the PCS values induced by a paramagnetic center predicted by the classical semi-empirical (SE)[56–58,150] and quantum chemistry (QC)[49,151–157] theories are not in agreement and which of the two explains better the experimental data. In order to make the experimental comparison of the expected results using SE and QC theories feasible, the S = 1/2 spin system present in copper(II)-hCAII was chosen. The theoretical details of the problem are described in the next section. However, the apo form of hCAII binds two equivalents of copper(II) with high affinity. Besides replacing the physiological zinc(II) site in normal zinc(II)-hCAII, forming a type 2 center, the copper(II) ion also binds at a secondary site at the N-terminal region, where it is coordinated by two Nε2 atoms from His4 and His64, as shown by X-ray diffraction studies[29]. The location and coordination mode of the secondary site in the protein in solution was further studied using paramagnetic NMR techniques and X-ray absorption spectroscopy (EXAFS)[75], while the binding affinities of copper(II) to the two protein sites (K_a values of 5×10^{12} M^{-1} and 1×10^{10} M^{-1}) were determined by isothermal titration calorimetry (ITC) techniques[39]. Therefore, in order to prevent the presence of the secondary copper(II) site, double and triple mutants of hCAII involving mutations of the histidine residues in the secondary site copper(II) binding (DM-hCAII and TM-hCAII) were obtained, and their copper(II) derivatives were used in EPR, NMR and X-ray crystallography studies.

5.2 Semiempirical and quantum chemistry theories of the pseudo-contact shift

The paramagnetic centers of metalloproteins couple with the nuclei present in these macromolecules through dipolar interactions and, when in solution, the dipolar interaction reduces to an isotropic rotational average. This effect was firstly described by Bloembergen and Dickinson10, McConnell and Robertson[57], Marshall[150], and later Kurland and McGarvey[58], and is known as pseudocontact shift (PCS). The PCSs depend on the anisotropy of the molecular paramagnetic susceptibility tensor χ and can be described by a semi-empirical (SE) framework as:

$$\delta^{\mathrm{pc}} = \frac{1}{12\pi r^3}\left[\Delta\chi_{ax}(3\cos^2\theta - 1) + \frac{3}{2}\Delta\chi_{rh}\sin^2\theta\cos2\varphi\right] \qquad \text{(Eq. 5.1)}$$

where δ^{pcs} is the calculated PCS of a nucleus in a molecule, r , θ and φ are the polar coordinates of the observed nuclear spin in the main frame of χ centered at the unpaired electron of the paramagnetic center, and $\Delta\chi_{\mathrm{ax}}$ and $\Delta\chi_{\mathrm{rh}}$ are the axial and rhombic components, respectively, of the χ anisotropy tensor. Eq. 5.1 predicts the shift of the nuclei coupled to the paramagnetic center and it has been extensively used to obtain structural information also on many biological systems, especially in the last two decades, either using transition metal ions[53,59–65] or lanthanides ions.[52,53,160,66–71,158,159]

Recently, the validity of the SE framework was questioned by quantum chemistry work[151–156] The QC theory states that the isotropic average of the nuclear spin-electron spin dipole-dipole interaction in the point dipole and spin Hamiltonian approximations provides the shift[49,157] which is described in Eq. 5.2:

$$\delta^{\mathrm{sd}} = \frac{1}{12\pi r^3}\mathrm{Tr}\left[\frac{\mu_0\mu_{\mathrm{B}}^2}{kT} g_{\mathrm{e}}\mathbf{g}\cdot\langle\mathbf{S}\mathbf{S}^T\rangle\cdot\left(\frac{3\mathbf{r}\mathbf{r}^T}{r^2} - \mathbf{1}\right)\right] \qquad \text{(Eq. 5.2)}$$

To compare Eq. 5.1 and Eq. 5.2, SE theory was rewritten following the formalism of Eq. 5.2 [161]

$$\delta^{\mathrm{pc}} = \frac{1}{12\pi r^3} \mathrm{Tr}\left[\frac{\mu_0 \mu_{\mathrm{B}}^2}{kT} \mathbf{g} \cdot \langle \mathbf{S}\mathbf{S}^T \rangle \cdot \mathbf{g}^T \cdot \left(\frac{3\mathbf{r}\mathbf{r}^T}{r^2} - \mathbf{1}\right)\right], \qquad \text{(Eq. 5.3)}$$

In both equations r is the metal-nucleus vector and $\langle \boldsymbol{SS}^T \rangle$ is the spin dyadic, which for S>1/2 depends on the zero-field splitting tensor $\boldsymbol{D}$. Eq. 5.2 differs from Eq. 5.3 by a missing factor $\boldsymbol{g}^T/g_e$.[161] Eq. 5.2 has been recently used to analyze experimental data. [162–164]

When two theoretical approaches clash, an experimental verification is required, which is presented in this work. For both equations we need to measure $\boldsymbol{g}$ and $\boldsymbol{D}$ tensors, however their measurement is challenging due to different experimental conditions from those in which the PCS are measured.[70] For S>1/2, quantum calculations are required to translate $\boldsymbol{D}$ and $\boldsymbol{g}$ tensors into PCS, for both Eq. 5.2 and Eq. 5.3. Using first principles calculations for a macromolecule containing cobalt(II) and for a small cobalt (II) complex[49,162], the QC equation performs better than the SE equation at predicting the shifts. To avoid the use of quantum calculations, we decided to use an S=1/2 system with high energy excited states, where $\boldsymbol{D}$ is absent, so that $\langle \boldsymbol{SS}^T \rangle = \frac{S(S+1)}{3}\mathbf{1}$, and the PCSs can be expressed in terms of a measurable $\boldsymbol{g}$. With such systems, Eq. 5.2 and Eq. 5.3 reduce to Eq. 5.4 and Eq. 5.5, respectively:

$$\delta^{\mathrm{sd}} = \frac{1}{12\pi r^3} \frac{\mu_0 \mu_{\mathrm{B}}^2 S(S+1) g_e}{3kT} \left[\left(g_z - \frac{g_x + g_y}{2}\right)(3\cos^2\theta - 1) + \frac{3}{2}\left(g_x - g_y\right)\sin^2\theta \cos 2\varphi\right] \qquad \text{(Eq. 5.4)}$$

$$\delta^{\mathrm{pc}} = \frac{1}{12\pi r^3} \frac{\mu_0 \mu_{\mathrm{B}}^2 S(S+1)}{3kT} \left[\left(g_z^2 - \frac{g_x^2 + g_y^2}{2}\right)(3\cos^2\theta - 1) + \frac{3}{2}\left(g_x^2 - g_y^2\right)\sin^2\theta \cos 2\varphi\right] \qquad \text{(Eq. 5.5)}$$

The prediction of the shifts for Eqs. 4 and 5 can vary by a factor larger than 2. For example, in an axial $\boldsymbol{g}$ system, the ratio between the two equations, and therefore the difference in the calculated shift is

$$\frac{\delta^{\mathrm{pc}}}{\delta^{\mathrm{sd}}} = \frac{g_{\parallel}^2 - g_{\perp}^2}{g_{\mathrm{e}}(g_{\parallel} - g_{\perp})} = \frac{g_{\parallel} + g_{\perp}}{g_{\mathrm{e}}} \approx 2\ . \qquad \text{(Eq. 5.6)}$$

With a difference this large we can determine, through the experimental g values, which equations is valid for an S=1/2 system.

To understand which equation stands to the experimental evidence we need a system that has the following requirements: a) has a metal ion where the EPR observables have a direct translation to NMR observables, e.g. a system with an S=1/2 with high energy excited states, such as copper (II); b) a relatively large system where many nuclei are far from the paramagnetic center and the point-dipole approximation holds, e.g. proteins, which have proved to be ideal according to solution and solid-state NMR studies[36,61–63,165]; c) a system where g-values are measured by EPR in solution at room temperature, e.g. the same conditions used to measure PCS from high resolution NMR.

There are some drawbacks in this approach. The magnetic susceptibility anisotropy (and hence the g anisotropy) of copper(II) proteins is relatively small, with typical $\Delta\chi_{\mathrm{ax}}$ values of 6 x 10 $^{-33}$ m^3, which will produce small PCSs from about 8 to 12 Å radius around the paramagnetic center. In addition, due to high energy separation between the ground electronic state and the excited state, the electron spin relaxation is slow, which can cause broadening of resonances in the NMR spectra.

The system that can fulfill the above requirements is the enzyme hCAII. This 30 kDa protein, where the native zinc(II) ion in its active site was replaced with copper(II), was used to acquire NMR and EPR data under the same experimental conditions and the zinc(II)-hCAII provided the diamagnetic reference shifts. The mutations were performed to avoid the binding of copper(II) ions to the secondary binding site.

5.3 Materials and methods

5.3.1 Preparation of the protein mutants and their copper(II) complexes

The wild-type form of hCAII is capable of binding two metal ions at different sites of the molecule. Therefore, to avoid the binding of an extra metal ion at the N-terminal, two constructs of hCAII were designed, the double and triple mutants.

The human carbonic anhydrase II double mutant (H3N, H4N) has the following sequence:

MS**NN**WGYGKHNGPEHWHKDFPIAKGERQSPVDIDTHTAKYDPSLKPLSVSYDQATSLRILN
NGHAFNVEFDDSQDKAVLKGGPLDGTYRLIQFHFHWGSLDGQGSEHTVDKKKYAAELHLVHWNTKY
GDFGKAVQQPDGLAVLGIFLKVGSAKPGLQKVVDVLDSIKTKGKSADFTNFDPRGLLPESLDYWTYP
GSLTTPPLLECVTWIVLKEPISVSSEQVLKFRKLNFNGEGEPEELMVDNWRPAQPLKNRQIKASFK.

The double mutant human carbonic anhydrase II (DM-hCAII) was cloned following a standard PCR protocol. For the H3N and H4N mutations the used primers were: forward: 5’- GA AGG AGA TAT ACC ATG GCC AAT AAC TGG GGG TAC GGC AAA CAC -3’ and reverse: 5’ - GTG TTT GCC GTA CCC CCA GTT ATT GGC CAT GGT ATA TCT CCT TC - 3’.

The human carbonic anhydrase II triple mutant (H3N, H4N, H64N) has the following sequence:

MS**NN**WGYGKHNGPEHWHKDFPIAKGERQSPVDIDTHTAKYDPSLKPLSVSYDQATSLRILN
NG**N**AFNVEFDDSQDKAVLKGGPLDGTYRLIQFHFHWGSLDGQGSEHTVDKKKYAAELHLVHWNTKY
GDFGKAVQQPDGLAVLGIFLKVGSAKPGLQKVVDVLDSIKTKGKSADFTNFDPRGLLPESLDYWTYP
GSLTTPPLLECVTWIVLKEPISVSSEQVLKFRKLNFNGEGEPEELMVDNWRPAQPLKNRQIKASFK

The triple mutant human carbonic anhydrase II (TM-hCAII) was cloned following a standard PCR protocol. For the H3N and H4N mutations the used primers were: forward: 5’- GA AGG AGA TAT ACC ATG GCC AAT AAC TGG GGG TAC GGC AAA CAC -3’ and reverse: GTG TTT GCC GTA CCC CCA GTT ATT GGC CAT GGT ATA TCT CCT TC -3’. For the H64N mutation the primers were: forward: 5’- C CTG AGG ATC CTC AAC AAT GGT AAT GCT TTC AAC GTG GAG TTT G -3’ and reverse. 5’- C AAA CTC CAC GTT GAA AGC ATT ACC ATT GTT GAG GAT CCT CAG G -3’.

The sequence of hCAII’s mutants was confirmed by sequencing. All the expression and purification procedures were the same for both mutants. The plasmid, containing the sequence of the mutant of hCAII, was transformed into *E. coli* BL21-CodonPlus (DE3)-RIPL cells using a standard heat shock protocol. Afterwards, 1 mL of LB was added to the cell culture and incubated at 37 °C, 200 rpm for 1 h 30’. The culture was plated in LB-Agar plates supplemented with ampicillin and chloramphenicol and incubated at 37 °C overnight. The next day, *E. coli* colonies were inoculated in 6 x 10 mL LB media supplemented with ampicillin and chloramphenicol and incubated for 8 hours at 37 °C, 200 rpm. Afterwards, each 10 mL culture was transferred to 1 L LB media (6 liters of LB total) supplemented with ampicillin and chloramphenicol and incubated at 37 °C, 160 rpm overnight. In the following day, cells were harvested at 4000 rpm for 15 minutes. The cells harvested from 2 L of LB were re-suspended in 1

L M9 media, for a total of 3 L of M9, supplemented with $CaCl_2$, $ZnSO_4$, $MgSO_4$, ampicillin, chloramphenicol, 3 mg/L of glucose and 1.2 mg/L $^{15}N\text{-}NH_4SO_4$. Cultures were incubated for 1 hour at 37 °C, 160 rpm. Protein expression was induced with 1 mM IPTG. Protein expression was halted after 5 hours of induction and the cells were harvested at 7500 rpm and stored at -20 °C. The cell paste was re-suspended in 200 mL of 20 mM Tris-SO_4, 500 µM $ZnSO_4$, pH 8. Cells were sonicated 10 times (30 seconds ON, 3 minutes OFF, power 70%) on ice and then ultracentrifuged at 40000 rpm for 40 minutes at 277 K The protein was purified with a two-step purification protocol, which comprises a HisTrap and Size Exclusion Chromatographies.

The copper (II) derivatives were obtained from the zinc(II) adduct by a demetalation/metalation approach. For demetalation, a solution containing 200 mM sodium phosphate, 50 mM pyridine-2,6-dicarboxylic acid, pH 7 was added to the protein solution and incubated overnight at 4 °C. After the demetalation, the apo-hCAII double and triple mutant were buffer exchanged to 10 mM HEPES, pH 5.1 and then to pH 8 using 10,000 MWCO Amicon ultracentrifugation devices. The paramagnetic copper (II)-hCAII double and triple mutants were prepared by titration of apo-hCAII double and triple mutants, respectively (1.5 mM in 10 mM HEPES, pH 8) with copper (II) sulfate solution monitored with 1D ^{1}H and 2D ^{1}H-^{15}N HSQC solution NMR spectra at 298 K. The buffer was then exchanged to 10 mM HEPES pH 6 and solutions of the furosemide, acetazolamide, p-toluenesulfonamide (PTS) and oxalate inhibitors (**Figure 5.1**) were added to the protein solution to reach a final concentration of 2.5 mM of oxalate. The zinc(II) diamagnetic references were prepared following the same protocol.

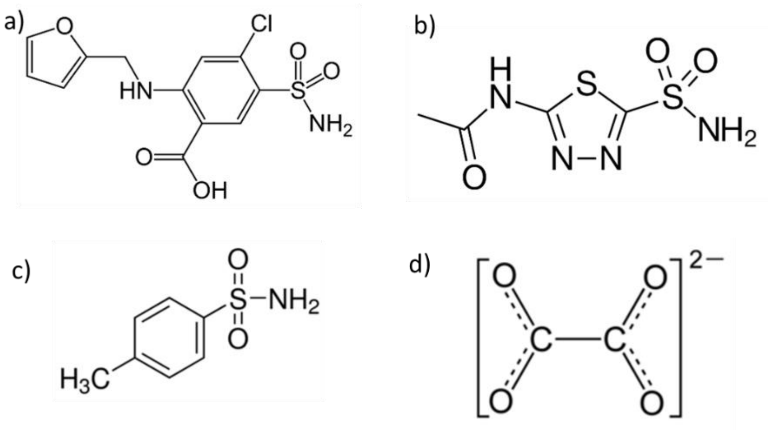


Figure 5.1: Molecular structures of a) furosemide; b) acetazolamide; c) p-toluenesulfonamide and d) oxalate ion.

5.3.2 Spectroscopic measurements and data analysis

NMR measurements

Solution 2D ^{1}H-^{15}N HSQC NMR spectra of the ^{15}N isotopically enriched zinc(II) and copper(II) substituted triple mutant of carbonic anhydrase II, for the evaluation of PCS, were recorded on a Bruker AVANCE III NMR spectrometer, operating at 950 MHz ^{1}H Larmor frequency, equipped with a triple resonance cryo-probe. Protein samples were prepared in a water buffered solution (10 mM HEPES, pH 6) containing 2.5 mM oxalate and with protein concentrations of 1.5 mM. 2D ^{1}H-^{15}N HSQC NMR spectra of the zinc(II) and copper(II) substituted triple mutant of carbonic anhydrase II were acquired at 303, 298, 293 and 288 K. All the spectra were processed with the Bruker TopSpin software package, and analyzed with the program CARA (the program can be downloaded for free from www.nmr.ch).[94] The spectral assignment of the diamagnetic zinc(II) protein was previously published[46], whereas the paramagnetic spectra were easily reassigned, because of the small shifts relative to the diamagnetic reference. The PCSs were calculated from the difference in the values of chemical shifts between the paramagnetic copper(II) and diamagnetic zinc(II) spectra of the protein. The $\Delta\chi$ tensor was evaluated using the program FANTEN[73] on the We-NMR/West-Life web portals from the best fit of 116 experimental PCSs on the protein structure (PDB code: 3KS3).[12]

EPR measurements and fit

The EPR spectra were acquired in solution at room temperature (298 K) on a BRUKER ELEXSYS EPR spectrometer operating at X-band equipped with a super-high sensitivity probehead. The modulation frequency was set to 100 kHz and the microwave frequency was set to 9.843 GHz. Centerfield was set to 330 mT and the total field sweep was 200 mT; 3 scans were taken for each spectrum. The fit of the EPR spectrum was performed using EasySpin.[101]

Crystallization, X-ray data collection, and structure determination

The initial solutions of the metalated forms of copper(II)-DM-hCAII with PTS and acetazolamide adducts, had a concentration close to 0.7 mM, in a solution containing 10 mM HEPES, pH 8. The crystallizations from these solutions were performed using the sitting drop vapor diffusion method by mixing equal amounts of sample volume and of a solution containing 100 mM HEPES pH 7.5, 2.9 M ammonium sulfate and 1 mM 4-hidroxymercuric benzoic acid sodium salt. Small crystals of copper(II)-DM-hCAII-PTS grew within days and copper(II)-DM-hCAII-acetazolamide took more than one month to form. Attempts to crystallize copper(II)-DM-hCAII-furosemide and copper(II)-DM-hCAII-oxalate were unfruitful.

X-ray diffraction data from the crystal of copper(II)-DM-hCAII-PTS and copper(II)-DM-hCAII-acetazolamide was collected at the Chemistry Department “Ugo Schiff”, using a D8 Venture diffractometer equipped with a copper anode, with a maximum resolution of 1.75 Å and 1.50 Å, respectively. The data sets were processed using XDS[124] and scaled using XSCALE.[124] The relevant statistics are shown in **Table 5.1**. The refinement was then carried out using REFMAC.[128]

Analysis of PCSs

To extract the best-fit g-values from the experimental PCSs, and either the relationship from Eq.5.4 and Eq.5.5, the data were parametrically fitted using the equation

$$\delta^{\mathrm{sd}} = \frac{1}{12\pi r^3}\frac{\mu_0\mu_{\mathrm{B}}^2 S(S+1)}{3kT}\left[G_{ax}(3cos^2\theta - 1) + \frac{3}{2}G_{rh}sin^2\theta cos2\varphi\right] \qquad \text{(Eq. 5.7)}$$

against the protein structure. The best fit G_{ax} and G_{rh} values obtained for the system can be used to predict g_z and g_y, if the value of g_x, for instance, is fixed.

5.4 Results and discussion

5.4.1 EPR and X-ray crystal studies of copper(II)-hCAII mutants and their inhibitor derivatives

The comparison of the predicted g_x g_x and g_z values from experimental PCSs using Eq. 5.7 with experimental g values requires the use of a copper(II)-hCAII system with a single copper(II) replacing the physiological zinc(II) site with a single coordination geometry and without binding at the secondary terminal site. This can be evaluated by the obtained EPR spectrum, which should show the expected number of hyperfine lines for a single species.

The EPR spectrum of demetalated WT-hCAII binding to two equivalents of copper(II) (copper(II)/copper(II)- WT-hCAII), both at the traditional zinc(II) site (labeled as CuB) and at the N-terminus of the protein (H3, H4, H64) (labeled as CuA), shows two sets of distinct EPR signals associated with each copper-binding site, corresponding to their distinctive electronic structures.[39] The spectrum corresponds to the absence of magnetic coupling of the two sites, as expected for their large separation of 13.6 Å. The EPR parameters obtained for the CuA site are g = (2.017, 2.067, 2.190) with a hyperfine splitting A_z = 197 G, and for the CuB site, g = (2.029, 2.096, 2.326) and hyperfine splitting of A_z = 127 G.[39]

We started by titrating WT-hCAII with two equivalents of copper(II) salt. The difference spectrum of copper(II)-WT-hCAII (**Figure 5.2 a)**) does not present the four hyperfine lines in the g_z region expected for a single copper (II) ion, as eight hyperfine lines are observed. Comparing the EPR spectrum with the literature (**Figure 5.2 b) B**) there are some similarities in the g_z, where there are present five hyperfine lines at this region. The g_x region is different and the resonances are less resolved than in literature. The differences may arise from the temperature difference, in this work the EPR spectrum was acquired at room temperature and in the literature, it was acquired at 21 K. Since both spectrum are similar, one can conclude that the copper(II) ions are bound in the CuA and CuB sites.

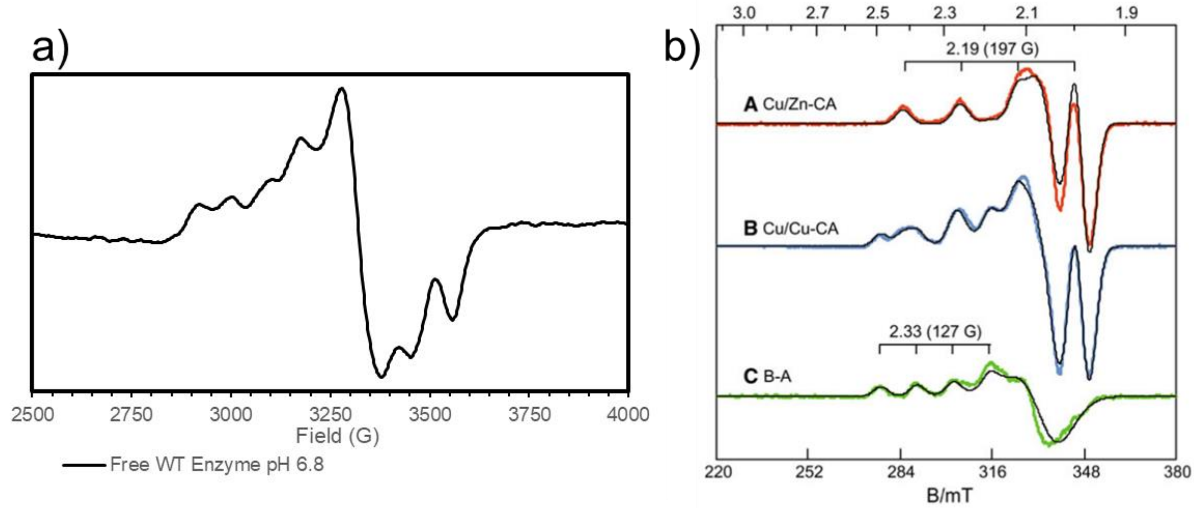


Figure 5.2: a) Normalized room temperature EPR spectrum of 0.7 mM WT-copper(II)-hCAII; b) Low temperature (21 K) EPR spectra of 0.85 mM WT-hCAII reacted with: A) 1.5 equivalents of Zn^{2+} and 0.75 equivalents of Cu^{2+}, B) 1.9 equivalents of Cu^{2+}. Spectrum C represents the difference spectrum between the spectrum A and B. [39]

The double mutant (DM-hCAII) (H3N, H4N) and a triple mutant hCAII (TM-hCAII) (H3N, H4N, H64N), which were produced, were designed to bind the copper(II) ion only at the active center. However, TM-hCAII probably lost its function because H64 is an essential residue to transport protons from the active center to the bulk solvent. Nevertheless, copper(II)-WT-hCAII with copper(II) at the active center also does not present the same activity of the native zinc(II)-WT-hCAII. Therefore, the EPR spectra of the copper(II)-TM-hCAII and copper(II)-DM-hCAII systems were studied next.

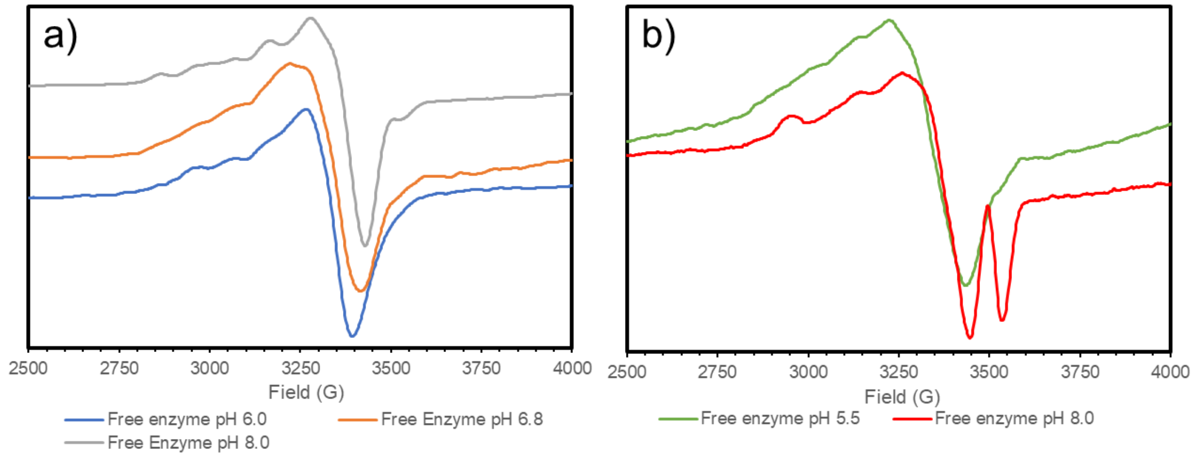


Figure 5.3: Room temperature solution EPR spectra of: a) 0.7 mM copper(II)-TM-hCAII at pH 6.0 (blue), pH 6.8 (orange) and pH 8 (gray) and b) 1.1 mM copper(II)-DM-hCAII at pH 8.0. Both mutants were titrated with one equivalent of copper(II) salt.

The EPR spectra of copper(II)-TM-hCAII show distinct changes with pH (**Figure 5.3 a**). At pH 6, the four expected hyperfine lines in the g_z region are sharp compared with pH 6.8, where they are broadened, and at pH 8.0, where five broadened hyperfine lines can be seen. The first hyperfine line for the spectrum at pH 6.0 is large, compared with the others, suggesting that there is overlap of a second species with different coordination in the active center. The EPR spectra for the different pH values suggest the presence of more than one species with different coordination geometries, making it not possible to accurately measure the g values. The absence of a single coordination geometry in all cases might be a result of fast exchange of water molecules and/or hydroxyl ions at pH 8.0.

A pH dependance is also observed in copper(II)-DM-hCAII (**Figure 5.3 b**). At a pH 8.0, the hyperfine lines in the g_z region are sharper than the pH 5.5 sample, where the hyperfine lines are more broadened. The g_x regions is better resolved at a higher pH value. For copper(II)-DM-hCAII it is also not possible to accurately measure the g values.

As the free copper(II)-HCAII systems do not show a single coordination geometry at different pH values that could provide a single set of g values, their derivatives with known ligand inhibitors were studied next. The ligand derivatives were studied in the following order: furosemide, acetazolamide, p-toluenesulfonamide (PTS) and oxalate. All these ligands have high affinity for the active center of hCAII[11,166–168], therefore we assumed that they have similar affinities for TM-hCAII and DM-hCAII, because the mutations are far from the active center. The inhibitor ligands furosemide, acetazolamide and PTS are sulfonamides and oxalate is a di-carboxylate, whose structures are represented in **Figure 5.1**.

The first ligand derivative to be studied, as stated above, was furosemide, which has an affinity constant for zinc(II)-WT-hCAII close to 3.14 µM.[166] copper(II)-TM-hCAII, and the corresponding DM-hCAII complexes, were reacted with furosemide and their interactions were studied by EPR at different pH values (**Figure 5.4**). The EPR spectra of copper(II)-TM-hCAII in the presence of 2 mM to 4 mM furosemide (2.8 mole ligand/protein ratio) is strongly pH dependent, particularly the observed number of hyperfine lines in the g_z region. All the acquired EPR spectra present more than the four hyperfine lines in that region expected for a copper(II) center with a single stereochemistry, indicating the presence of two different coordination geometries. When the pH is increased, a second coordination geometry becomes more resolved, while at lower pH the hyperfine lines from the two coordination geometries are more overlapped. The same is observed for the copper(II)-DM-hCAII adduct, where the EPR spectra of copper(II)-DM-hCAII have some similarities at pH 8 and pH 10 with those of copper(II)-TM-hCAII. The only difference between both mutants is the presence/absence of H64, which may influence the binding of ligands in the active center of the protein.

Furosemide is a ligand that, according to the published structure deposited with the PDB code **1Z9Y**[169] binds to the metal center of hCAII through a nitrogen and a sulfur atom from the sulfonamide group. In addition, according to the structure, there are no water molecules binding to the zinc(II) metal center. Therefore, chemical exchange of water molecules or hydroxyl ions are, in principle, not responsible for the observed differences in the coordination geometry of the copper(II) center. However, one should be careful with this analysis, as the deposited structure is for zinc(II)-WT-hCAII and not the copper(II) protein, where the coordination geometry might be slightly different. Another reason to change the pH in the EPR study was to test if the protonation of the histidine residues coordinating the copper(II) ion could affect its coordination, however none of the pH values studied provided a single coordination of the copper(II) center in the mutant complexes.

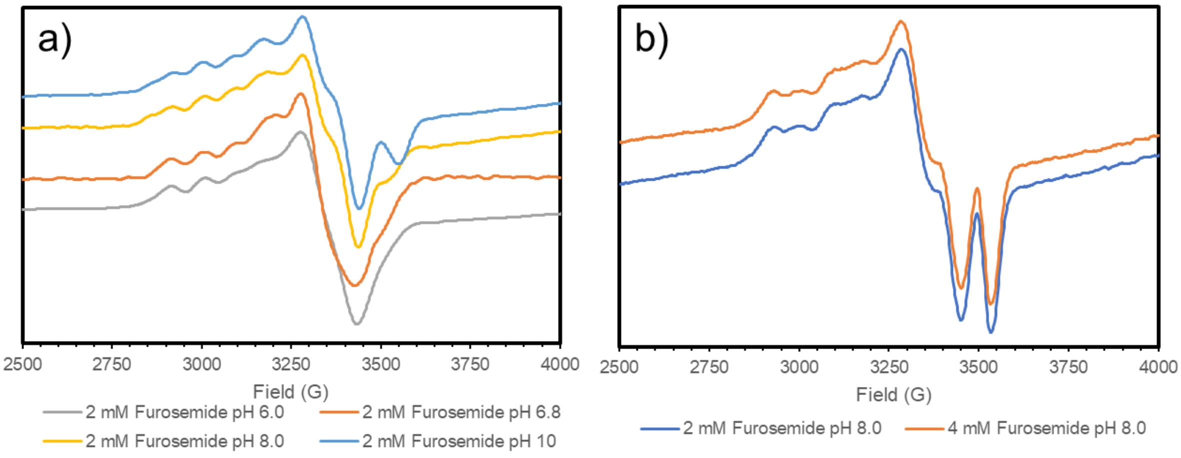


Figure 5.4: Room temperature solution EPR spectra of 0.7 mM copper(II)-TM-hCAII (a) at different pH (6.0; 6.8; 8.0 and 10) and copper(II)-DM-hCAII (b) in the presence of 2 or 4 mM furosemide at different pH values.

The next inhibitor to be studied was acetazolamide, which has a higher binding affinity to zinc(II)-WT-hCAII, with an affinity constant around 0.02 µM.[11] This could solve the problem of multiple coordination geometries for the adducts of copper(II)-hCAII mutants, since the two orders of magnitude stronger binding could provide a tighter coordination geometry. Copper(II)-TM-hCAII and copper(II)-DM-hCAII were reacted with acetazolamide at different pH values and different concentrations, 2 mM and 4 mM (2.8 mole ligand/protein ratio), and EPR spectra were collected (**Figure 5.5**). The acetazolamide adducts of both mutants showed EPR spectra with a varying number of hyperfine lines in the g_z region. For copper(II)-TM-hCAII, seven hyperfine lines were observed at pH 6 and six at pH 8. To make sure that the active center of the protein was fully occupied by the inhibitor, the concentration of acetazolamide was doubled from 2 mM to 4 mM, resulting in the increase of the number of hyperfine lines from six to seven and an increase of their broadening in the EPR spectrum at pH 8. This result is puzzling, as an inhibitor concentration of 2 mM should already be close to or at the saturation point of the protein, however it did not seem to be the case.

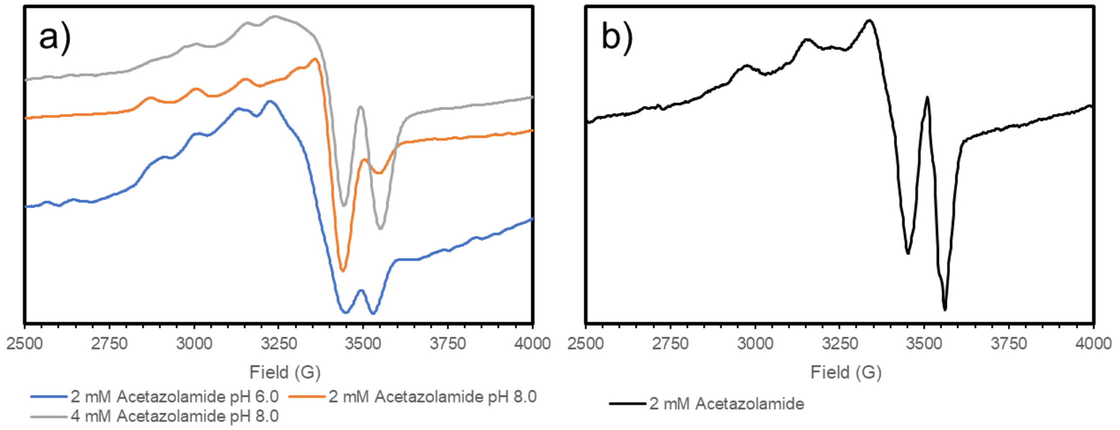


Figure 5.5: Normalized room temperature solution EPR spectra of a) 0.7 mM copper(II)-TM-hCAII and b) 1.1 mM copper(II)-DM-hCAII in the presence of acetazolamide in different conditions: a) blue pH 6.0 in the presence of 2 mM Acetazolamide, orange pH 8.0 in the presence of 2 mM Acetazolamide and gray pH 8.0 in the presence of 4 mM Acetazolamide; b) pH 8.0 in the presence of 4 mM acetazolamide

The next attempt was to study the EPR spectra of the adducts of copper(II)-DM-hCAII and copper(II)-TM-hCAII with PTS. This ligand is a small sulfonamide composed by a sulfonamide group bound to a toluene moiety (**Figure 5.1**) that has an affinity to zinc(II)-WT-hCAII around 2.5 μM.[167] This ligand has poor solubility in water, therefore solid PTS was added to maximize the compound solubilization. The EPR spectra obtained (**Figure 5.6**) show that the binding of PTS provides spectra with five hyperfine lines in the g_z region for both copper(II)-TM-hCAII and copper(II)-DM-hCAII, which are sharper for the first mutant but not for the second, indicating that there is more than one coordination environment in the active center of both systems.

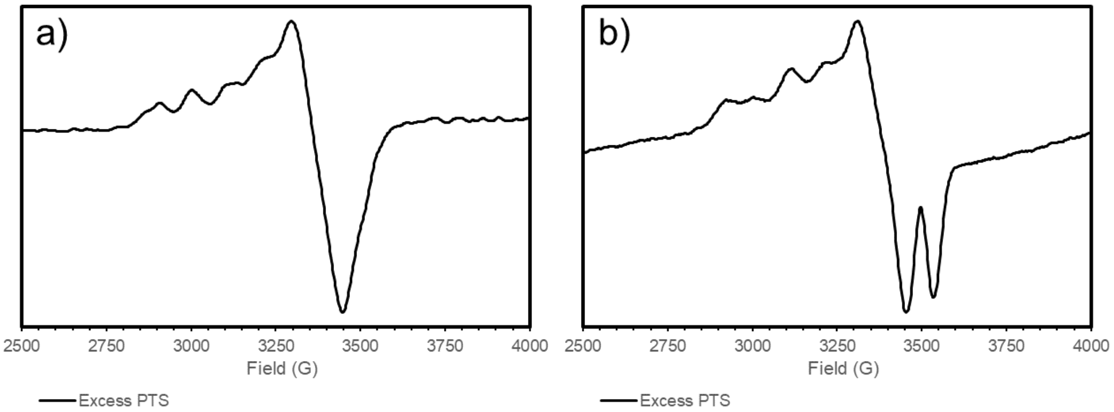


Figure 5.6: Normalized room temperature solution EPR spectra of (a) 0.7 mM copper(II)-TM-hCAII and (b) 1.1 mM copper(II)-DM-hCAII in the presence of excess PTS (pH 8).

In order to understand why sulfonamides do not provide a single coordination geometry in the active site of the copper(II)-hCAII mutants, we attempted to obtain crystals from some of the prepared EPR samples, in order to directly analyze the active center of the copper(II)-DM-hCAII and copper(II)-TM-hCAII adducts. Samples of copper(II)-DM-hCAII adducts with acetazolamide and with PTS gave good crystals for X-ray diffraction studies, but those with furosemide did not crystallize, as well as all the adducts of copper(II)-TM-hCAII. **Figure 5.7 a)** and **b)** show the electron density at the active site of copper(II)-DM-hCAII reacted with PTS and acetazolamide.

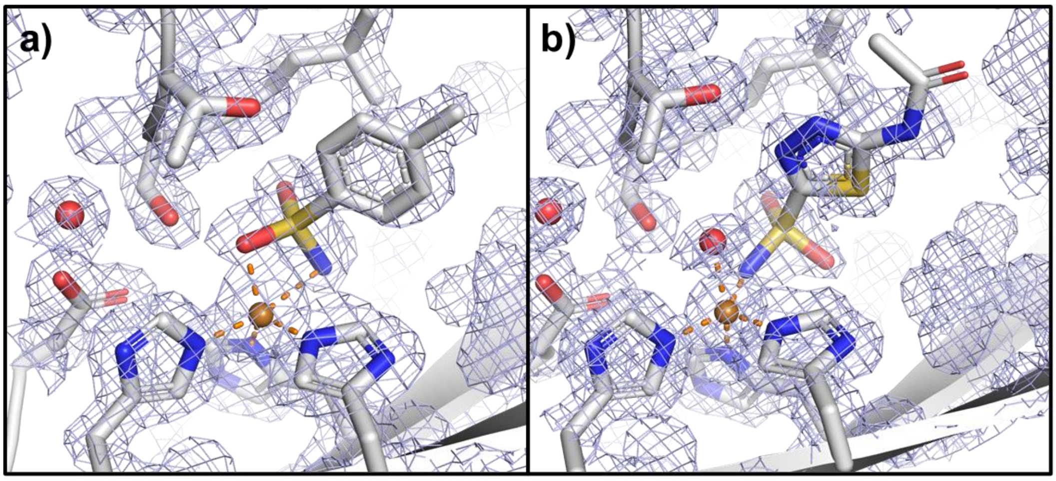


Figure 5.7: Active center of copper(II)-DM-hCAII reacted with: a) PTS and b) Acetazolamide. The electron density is represented in light blue with a contour level of 1 with a contour level of 1 σ. The copper ions are represented as brown sphere and water molecules are represented as red spheres.

The X-ray diffraction data obtained for the complexes of copper(II)-DM-hCAII with acetazolamide and PTS were of good quality and were used to determine their structure, which were solved at 1.50 Å and at 1.75 Å, respectively, by Doctor Vito Calderone. The relevant data collection and refinement statistics are shown in **Table 5.1**. The structure of copper(II)-DM-hCAII reacted with PTS and acetazolamide showed that the copper(II) ion is five coordinated, for both cases. Furthermore, both copper (II) centers are coordinated by His94, His96, His119. For the PTS complex, the remaining coordination is done by one oxygen and one nitrogen from the PTS molecule and for the Acetazolamide complex, the remaining coordination is done by one water molecule and by one nitrogen from the sulfonamide group of Acetazolamide. The analysis of the electron density maps (**Figure 5.7**) at the respective active centers showed that both ligands, PTS and acetazolamide, have parts where the electron density is not well defined, indicating that the ligands have a low occupancy at those coordinates. This can have two causes, either there was not enough ligand available during the crystallization to saturate the binding site, or the ligand is flexible inside the active center. The latter hypothesis seems to be more probable because the ligand was in excess for both samples used in the crystallization. The structure of both ligands shows a ring that is free to rotate inside the protein binding site and can unevenly distribute the unpaired electron around the metal center, thus providing different EPR spectra as well as weak electron densities in the X-ray diffraction maps. Even though the crystallization of the samples for copper(II)-TM-hCAII in the presence of PTS and acetazolamide was not successful, we assume that the explanation given above about the rotation of the ligands in the active site is still valid for those systems.

Table 5.1: Data collection and refinement statistics for copper(II)-DM-hCAII adducts of PTS and acetazolamide and PTS

	copper(II)-DM-hCAII-PTS	**copper(II)-DM-hCAII- acetazolamide**
Wavelength (Å)	1.541	1.541
Resolution range (Å)	39.85 - 1.75 (1.81 - 1.75)	24.5 - 1.5 (1.55 - 1.50)
Space group	P 1 21 1	P 1 21 1
Unit cell (a, b, c) (Å) (α, β, γ) (°)	42.13 41.34 72.5 90 104.43 90	41.94 41.28 72.04 90 104.25 90
Total reflections	44424 (2668)	62762 (5316)
Unique reflections	24197 (2013)	37968 (3748)
Multiplicity	1.8 (1.3)	1.7 (1.4)
Completeness (%)	98.11 (82.17)	98.55 (98.14)
Mean I/sigma(I)	8.55 (1.36)	26.49 (6.82)
Wilson B-factor	13.53	9.11
R-merge	0.09261 (0.4547)	0.035 (0.110)
R-meas	0.131 (0.643)	0.049 (0.155)
R-pim	0.09261 (0.4547)	0.035 (0.110)
CC1/2	0.988 (0.499)	0.997 (0.898)
CC*	0.997 (0.816)	0.999 (0.973)
Reflections used in refinement	24180 (2010)	37964 (3748)
Reflections used for R-free	1210 (100)	1897 (187)
R-work	0.149 (0.237)	0.130 (0.155)
R-free	0.209 (0.300)	0.188 (0.244)
CC(work)	0.958 (0.631)	0.970 (0.886)
CC(free)	0.951 (0.529)	0.940 (0.750)
Number of non-hydrogen atoms	2300	2580
macromolecules	2047	2047
ligands	33	37
solvent	220	496
Protein residues	257	257
RMS(bonds) (Å)	0.016	0.169
RMS(angles) (°)	1.89	5.28
Ramachandran favored (%)	96.08	94.90
Ramachandran allowed (%)	3.92	4.71
Ramachandran outliers (%)	0.00	0.39
Rotamer outliers (%)	1.80	0.90
Clashscore	3.43	8.56
Average B-factor	17.89	16.42
macromolecules	16.42	12.50
ligands	26.51	29.98
solvent	30.28	31.57

Previous studies[170,171], show that the binding of the sulfonamides to the active center of either copper substituted bovine and human carbonic anhydrase change the T_1^{-1} of the water in the active center, thus indicating the sulfonamide displaces one water molecule of the active center. The same studies also show that oxalate abolishes the T_1^{-1} of the water in the active center, thus indicating that oxalate displaces both water molecules at the active center. Since oxalate is capable of displacing both water molecules it generates a complex with a five-coordination. However, sulfonamides since are only capable of only displacing one water molecule meaning the fifth position is still available to bind with a water molecule or a hydroxyl ion. Furthermore, the water exchange in the fifth position of the copper(II) center might cause the extra hyperfine lines in the EPR spectra.

Knowing that the presence of intramolecular rotation of the ligands in the active site of the copper(II)-hCAII mutants is problematic for the EPR spectra, oxalate was chosen as the next known inhibitor of hCAII. Oxalate is a small bidentate di-carboxylic acid that has an affinity constant for cobalt(II)-WT-hCAII close 10 μM.[168] This ligand might provide a single coordination geometry for the active site of the copper(II)-hCAII mutants.[172]

Both copper(II)-TM-hCAII and copper(II)-DM-hCAII mutants were reacted with oxalate and their solutions were studied by EPR (**Figure 5.8**). copper(II)-TM-hCAII-oxalate provides an EPR spectrum with the expected number of four hyperfine lines in the g_z region, which does not occur for copper(II)-DM-hCAII-oxalate. Both samples have similar concentrations (1.5 mM and 1.1 mM, respectively) are in the same buffer conditions (10 mM HEPES, pH 6.0) and the ligand: protein ratio is approximately 1.6. The key difference between these mutants is the mutation in H64, present in the TM but not in the DM mutant. This key residue (H64) is involved in the proton transfer process from the active site to the bulk solvent. There is a water network between these two locations and H64 might influence the water positions of the water network, thus giving a different environment around the coordination sphere of copper(II) in the active center. Increasing the concentration of oxalate in copper(II)-DM-hCAII did not result in any spectral changes. Changing the pH to 8.0 then to 5.5 did not led to a spectrum with four hyperfine lines. Moreover, increasing the pH led to an increase in intensity of the hyperfine line in the g_x region and decreasing the pH 5.5 led to an increase in the number of hyperfine lines to 5 and some of these hyperfine lines were broadened.

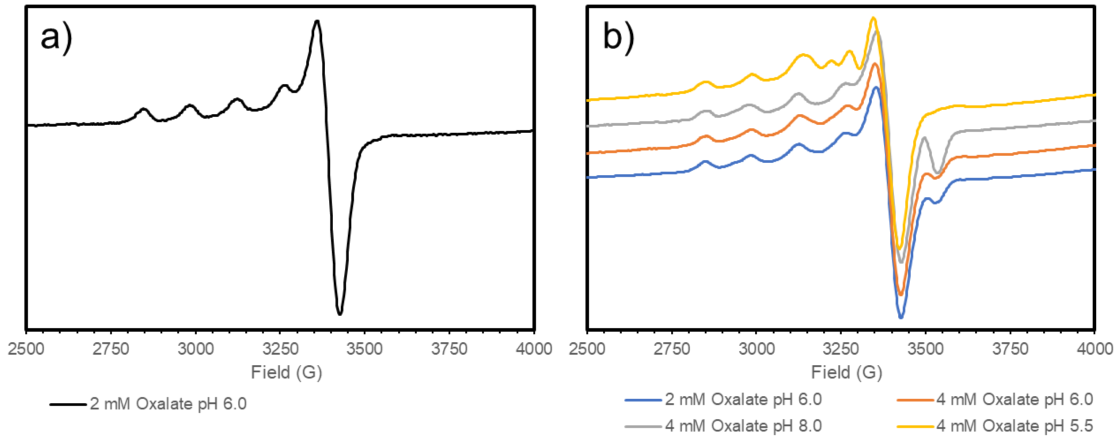


Figure 5.8: Normalized room temperature solution EPR spectra of: a) 1.5 mM copper(II)-TM-hCAII pH 6.0 in the presence of 2 mM oxalate pH 6.0; b) 1.5 mM copper(II)-TM-hCAII pH 6.0 in the presence of 2.0 mM oxalate (blue), pH 6.0 in the presence of 4 mM oxalate, pH 8.0 in the presence of 4 mM oxalate (gray) and pH 5.5 in the presence of 4 mM oxalate (yellow).

The EPR spectrum of copper(II)-hCAII-TM-oxalate was fitted to a single of copper(II) geometry, with calculated g-values $g_x = 2.064$, $g_y = 2.082$ and $g_z = 2.307$, and hyperfine splittings $A_x = 0.113$ G, $A_y = 13.609$ G and $A_z = 135.15$ G.

5.4.2. Acquisition and analysis of PCS data

In order to obtain the experimental PCS values, NMR data were acquired for some of the samples of the zinc(II) and copper(II)-TM-hCAII adducts studied by EPR, in the same experimental conditions. The ^{1}H-^{15}N HSQC NMR data acquired for zinc(II) and copper(II)-TM-hCAII furosemide adducts (**Figure 5.9**) at 298 K have a very high quality and are well resolved. The spectra are different due to the presence of paramagnetic copper(II) in the active center of copper(II)-TM-hCAII. The differences in the chemical shifts for both metal forms are not very different, therefore the copper(II)-TM-hCAII spectrum could be assigned from the zinc(II)-TM-hCAII assignment. This small difference in the chemical shifts (the PCS values) is expected due to the small magnetic susceptibility anisotropy of copper(II). However, these PCS values were not used further, as the furosemide adduct of copper(II)-TM-hCAII does not have a single coordination geometry, as shown by EPR.

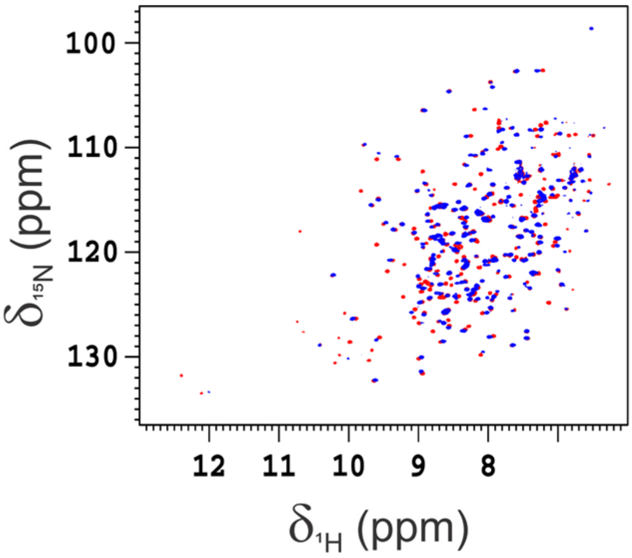


Figure 5.9: Superimposed 2D ^{1}H-^{15}N HSQC spectra of the 1.1 mM zinc(II) (red) and copper(II) (blue) substituted TM-hCAII in the presence of 2 mM furosemide, acquired at 298 K on a 500 MHz Bruker Avance III spectrometer equipped with a triple-resonance cryo-probe.

After confirmation from the EPR data for the copper(II)-TM-hCAII-oxalate adduct that its active site had a single coordination species, 2D ^{1}H-^{15}N HSQC NMR spectra were acquired for copper(II)-TM-hCAII and zinc(II)-TM-hCAII under the same conditions as the EPR measurements (**Figure 5.10**) in order to compare the *g* values of the copper(II) protein obtained from the experimental PCS values using Eq. 5.7 with those from the EPR spectrum.

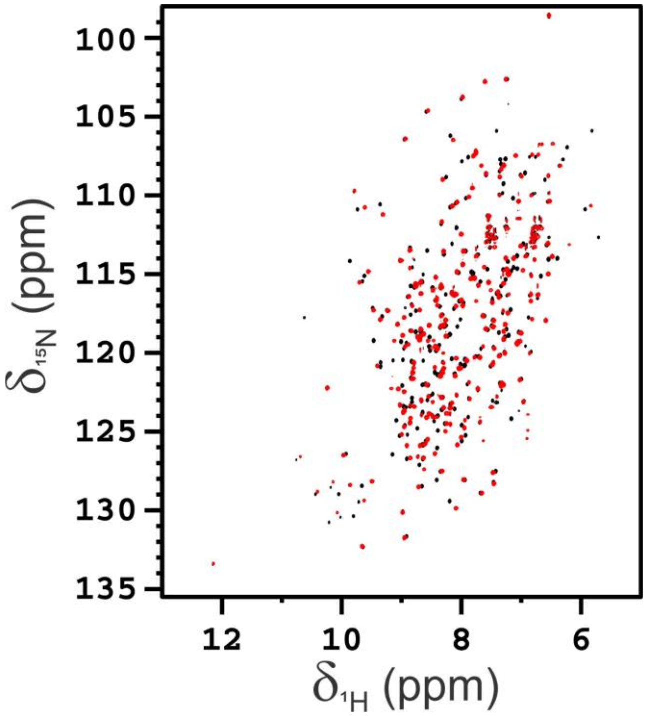


Figure 5.10: Superimposed 2D ^{1}H-^{15}N HSQC NMR spectra of the 1.5 mM zinc(II) (black) and copper(II) (red) substituted TM-hCAII in the presence of 2 mM oxalate, acquired at 298 K on a Bruker Avance III NMR spectrometer, operating at 950 MHZ ^{1}H Larmor frequency.

The acquired NMR data for both metal forms are of high quality and the dispersion of the resonances allows an accurate measurement of the amide proton chemical shifts in both proteins, which are not very different in the copper(II) and zinc(II) proteins. This allows the assignment of the amide protons of the copper(II) protein from that of the zinc(II)protein. The differences in their chemical shifts are again small due to the small magnetic susceptibility anisotropy arising from the copper(II) center. In addition, some resonances from the copper(II) protein are less intense than in the zinc(II) form, which arises from the increased nuclear R_2 relaxation (increased broadening) caused by the presence of the unpaired electron of copper(II).

The fit for the EPR spectrum of copper(II)-TM-hCAII provides g values $g_x = 2.064$, $g_y = 2.082$ and $g_z = 2.307$. The observed PCSs were calculated by subtracting the chemical shifts of the amide protons of paramagnetic copper(II)-TM-hCAII from the corresponding ones in the diamagnetic reference, zinc(II)-TM-hCAII. The measured PCSs were then compared with the calculated PCSs using Eq. 5.4 and Eq. 5.5 and the results are shown in **Figure 5.11**.

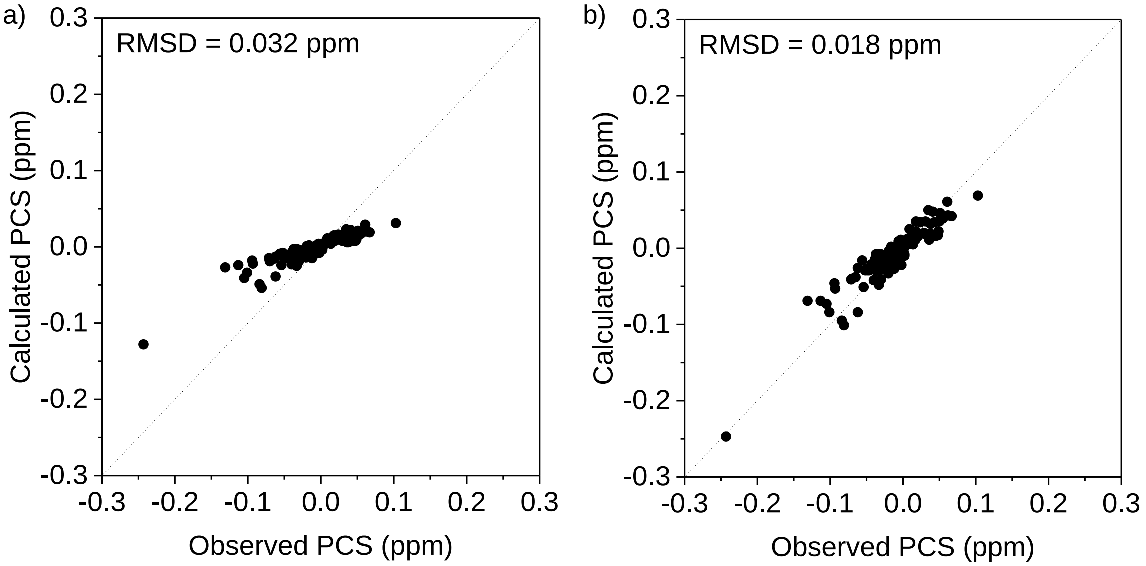


Figure 5.11: Agreement between experimental PCSs and values calculated with Eq. 5.4 (a) or Eq. 5.5 (b), as obtained from the EPR g-values and a 3-parameters fit (the three Euler angles defining the main frame of the g tensor) against the protein structure of human carbonic anhydrase II (PDB code 3KS3[12]).

The plots clearly demonstrate that the PCSs calculated with Eq. 5.5 are in reasonably good agreement with the experimental data, whereas those calculated with Eq. 5.4 are in strong disagreement. To date, and to our knowledge, this is the first time that both equations are compared using only experimental data and without employing theoretical estimations of parameters.

To confirm this finding, the PCSs obtained previously for the copper(II) bound forms of the type-1 copper proteins pea plastocyanin[145] and *Anabaena variabilis* plastocyanin[72] (with the cadmium derivative taken as a diamagnetic reference) were also reanalyzed. In the first case, the analysis was limited to the experimental values measured for protein backbone protons (NH and H_α), the positions of which are more reliable than those of sidechain nuclei. In the second case, the PCSs were taken from the project report by Robert Dagil (http://dagil.dk/bach.pdf), due to the unavailability of the supporting information of the respective publication.[72] In these two cases, Eq 5 also provides strikingly better agreement between predicted and experimental data than Eq. 5.4 (**Figure 5.12**).

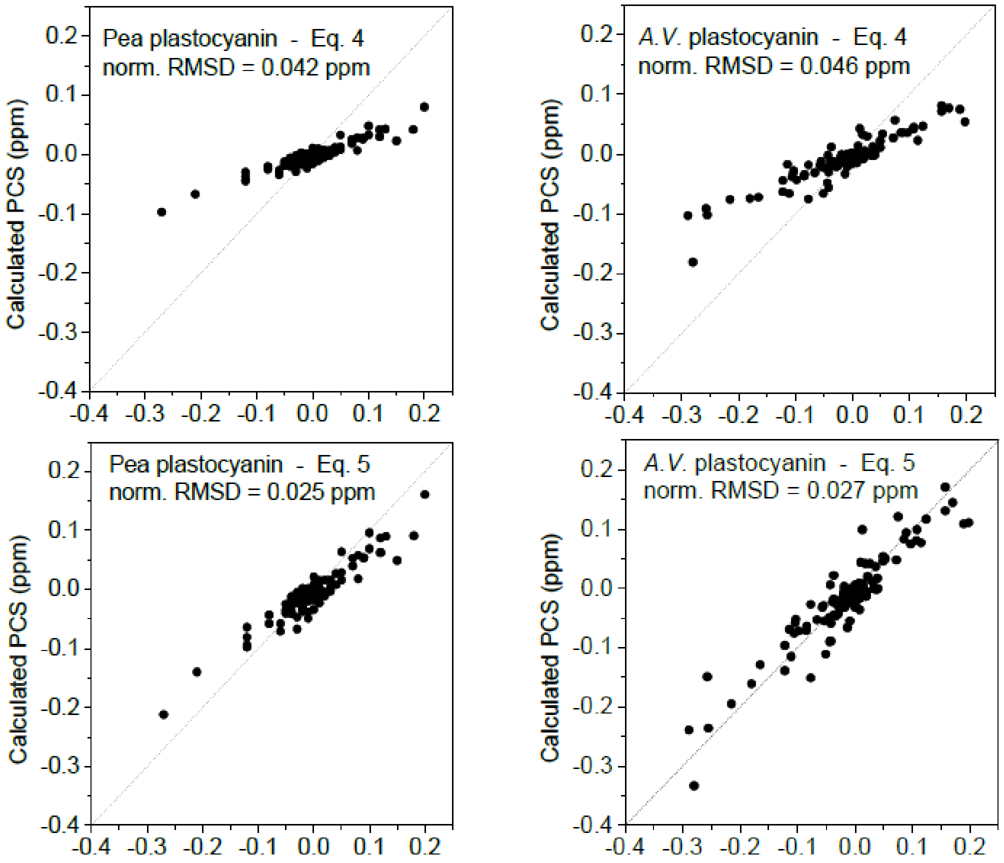


Figure 5.12: Agreement between experimental PCSs and values calculated with Eq. 5.4 or Eq. 5.5, as obtained from the EPR g-values and a 3-parameters fit (the three Euler angles defining the main frame of the g tensor) against the protein structures (PDB codes 9PCY[173] and 2GYM[174], respectively).

It is also useful to compare the experimental g values obtained from EPR with those that can be obtained as best-fit parameters of the experimental PCS data sets (**Figure 5.13**) against the three protein structures (**Figure 5.14**) using Eq. 5.4 and Eq. 5.5 (**Table 5.2**). The best-fit G_{ax} and G_{rh} values obtained for the three systems are also reported in **Table 5.2**. The best-fit G_{ax} and G_{rh} values can be used to predict g_z and g_y, if the value of g_x, for instance, is fixed. Using the SE relationship (Eq. 5.5), for all systems g_z is found to be close to the EPR values and in the typical range for tetragonal copper(II) of 2.2-2.4, reported in the literature. On the contrary, using the QC relationship (Eq. 5.4), g_z values as large as 2.5-2.8, far away from the observed EPR values, are obtained, thus confirming the inaccuracy of this equation.

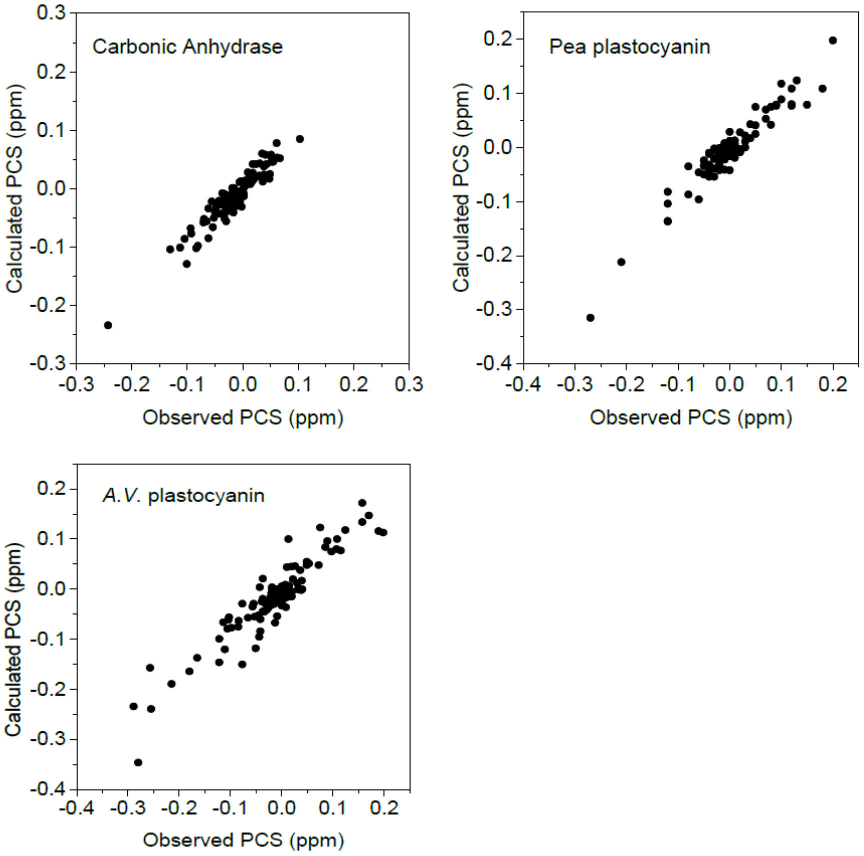


Figure 5.13: Agreement between experimental and calculated PCS, as obtained from a 5-parameter fit (G_{ax}, G_{rh}, and three Euler angles defining the main frame of the $\boldsymbol{g}$ tensor) against the protein structures.

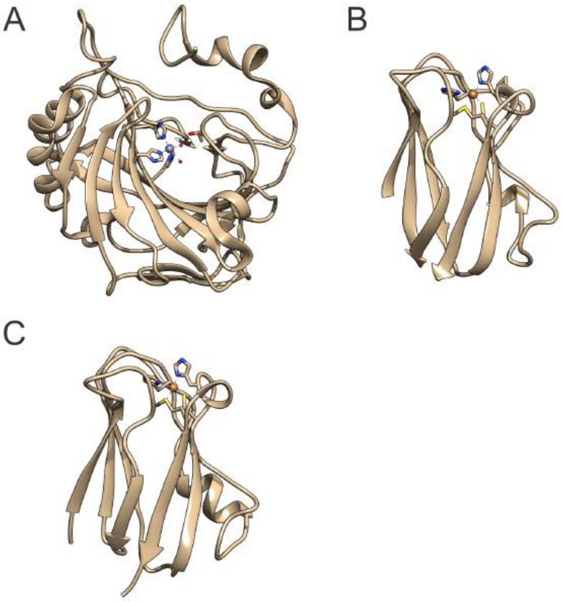


Figure 5.14: Protein structures used for the evaluation of the agreement between experimental and calculated PCS: (A) carbonic anhydrase II (PDB: 3KS3)[12], (B) Pea plastocyanin (PDB: 9PCY)[173], (C) A.V. plastocyanin (PDB:2GIM).[174]

Table 5.2: EPR and PCS-derived g tensor parameters for the three investigated copper(II) proteins

Protein	**copper(II)-TM-hCAII-oxalate**	**Pea plastocyanin**[1]	**Anabaene variabilis plastocyanin**[3]
PDB	3KS3	9PCY	2GYM
g_x, g_y, g_z (experimental, from EPR)	2.064, 2.082, 2.306	2.042, 2.059, 2.226[2]	2.042, 2.059, 2.226
g_x, g_y, g_z from PCS and Eq. 5.5	(2.05)[4], 2.12±0.01, 2.37±0.01	(2.04)[4], 2.10±0.01, 2.28±0.02	(2.04)[4], 2.06±0.01, 2.23±0.01
g_x, g_y, g_z from PCS and Eq. 5.4	(2.05)[4], 2.20±0.02, 2.76±0.02	(2.04)[4], 2.16±0.03, 2.56±0.04	(2.04)[4], 2.08±0.02, 2.45±0.02
G_{ax}, G_{rh} (experimental from PCS)	1.26±0.02, -0.29±0.03	0.91±0.05, -0.25±0.06	0.77±0.03, -0.08±0.03
$\boldsymbol{g_e\left(g_z-\frac{g_x+g_y}{2}\right), g_e(g_x-g_y)}$, from experimental EPR values	0.47, -0.04	0.35, -0.03	0.35, -0.03
$\boldsymbol{g_z^2-\frac{g_x^2+g_y^2}{2}, g_x^2+g_y^2}$, from experimental EPR values	1.02, -0.08	0.75, -0.07	0.75, -0.07

[1] From ref. [145]5.

[2] For spinach plastocyanin at T=100 K, from ref [175].

[3] From ref. [72].

[4] The value of g_x was fixed.

Table 5.2 shows the comparison of the g-values, both calculated and measured, for three different proteins. As expected, the calculated g-values for the different proteins are different using different equations. If we use Eq. 5.5, we obtain a g_z value between 2.2-2.4, which are the expected values for a tetragonal copper(II) geometry, and if we use Eq. 5.4 we obtain g-values between 2.5-2.8. Knowing this, we conclude that Eq. 5.5 is inaccurate in predicting the g-values for copper(II) systems.

A further analysis of the g-values obtained using Eq. 5.5 shows that their best fit values are, in all cases, slightly larger than the EPR observed g-values. The origin of this small discrepancy can be understood if the equation for PCS is considered in its most general form of Eq. 5.1. For the present copper(II) systems, Eq. 5.5 is a good approximation to Eq. 5.1, except for neglecting the so-called temperature independent paramagnetism (TIP) whose origin comes from second order perturbation theory.[176–178]

For copper(II) systems, *g*-values can be translated into the axial anisotropy of the magnetic susceptibility tensor ($\Delta\chi_{ax}$), through the following relationship:

$$\Delta\chi_{ax}=\frac{\mu_0\mu_B^2S(S+1)}{3kT}\left(g_z^2-\frac{g_x^2+g_y^2}{2}\right) \qquad \text{(Eq. 5.8)}$$

The comparison between the $\Delta\chi_{ax}$ values calculated as the best fit parameters from the fit of the PCSs using Eq. 5.1 and those calculated from calculated from the experimental EPR *g* values is shown in **Table 5.3**.

Table 5.3: $\Delta\chi_{ax}$ values (in 10^{-32} m^3) calculated from the EPR *g* values and obtained from the fit of the PCSs, and *g* values calculated from the PCS and measured through the EPR spectra

	Carbonic anhydrase II-oxalate	**Pea plastocyanin**	**A.v. plastocyanin**
$\Delta\chi_{ax}$ from the *g* values	0.67±0.06	0.50	0.50
$\Delta\chi_{ax}$ from the PCSs[a]	0.76±0.09	0.63±0.07	0.51±0.02
g_x, g_y, g_z (experimental, from EPR)	2.064, 2.082, 2.306	2.042, 2.059, 2.226[b]	2.042, 2.059, 2.226[b]
g_x, g_y, g_z from PCSs and Eq. 5.5	(2.05)[c], 2.12±0.01, 2.37±0.01	(2.04)[c], 2.10±0.01, 2.28±0.02	(2.04)[c], 2.06±0.01, 2.23±0.01

[a] The rhombic anisotropy is fixed to the same ratio between rhombic and axial anisotropy of the semi-empirical $\Delta\boldsymbol{\chi}$
[b] For spinach plastocyanin at T=100 K[175].
[c] The value of g_x has been fixed

The $\Delta\chi_{ax}$ values calculated directly from Eq. 5.1 are slightly larger value than those calculated through the EPR g values, which can be accounted for by the presence of a small (ca. 12 %) TIP contribution. In addition, for the other two copper(II) proteins, a small TIP contribution is also present (ca. 9 % and 2 %, respectively). The presence of TIP can be confirmed through the measurement of the PCSs of copper(II)-TM-hCAII-oxalate at different temperatures and fitting them using Eq. 5.1 to determine the corresponding $\Delta\chi_{ax}$ values. **Figure 5.15** shows the 1/T dependence of the $\Delta\chi_{ax}$ values. The data show a relatively good linear dependence in the limited range of temperatures allowed by the system and extrapolate to a few percent at infinite temperature, also consistent with the presence of a small TIP.

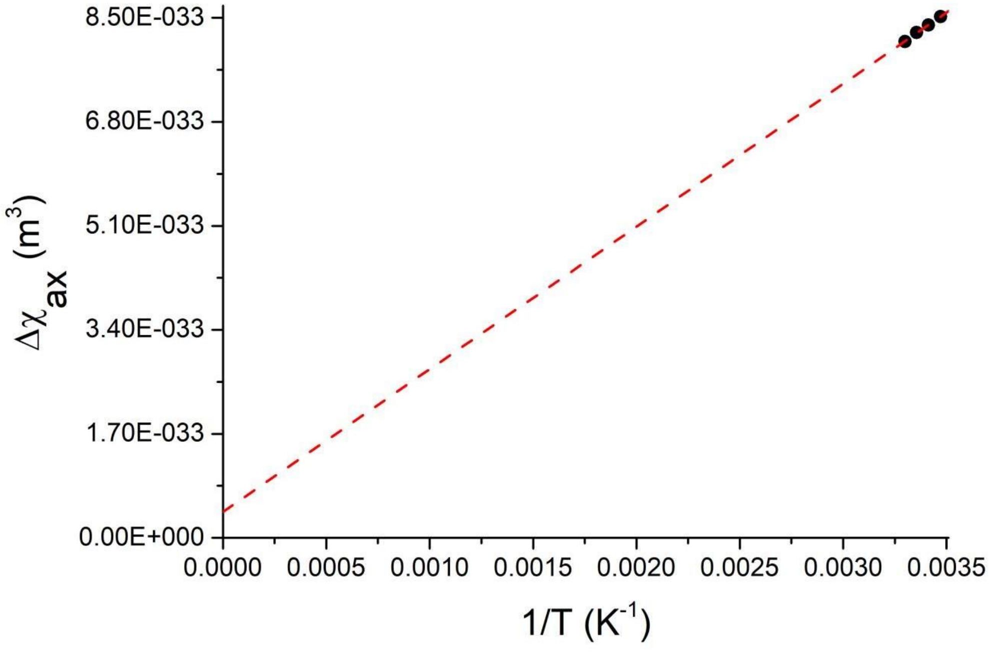


Figure 5.15: Plot of the 1/T dependence of the $\Delta\chi_{ax}$ value. The $\Delta\chi_{ax}$ values were determined with the program FANTEN from the fit of experimental PCSs, measured at 288, 293, 298 and 303 K, against the crystal structure of WT-hCAII (PDB code: 3KS3[12]).

5.5 Conclusions and perspectives

In this work, we developed a copper(II)-protein system, copper(II)-TM-hCAII-oxalate, that allowed to show that the obtained experimental PCS values are in much better agreement with those predicted by the classical semi-empirical (SE) theory than by quantum chemistry (QC) theories. Double and triple mutants of human Carbonic Anhydrase II were developed obtain the PCS values arising from a single copper (II) ion, avoiding the binding of a interaction of a second copper(II) ion at the N-terminal site of the protein. To understand the nature of the observed paramagnetic phenomena, a combination of structural biology techniques were used, namely NMR, X-ray crystallography and EPR. This work revealed that the sample construct, sample treatment, sample condition and type of inhibitor present were crucial to assure a single coordination geometry of copper(II) in the protein. For example, the copper(II)-TM-hCAII-furosemide adduct provided high quality ^{15}N-^{1}H-HSQC NMR data which allowed measurement of many PCS values and to calculate its magnetic susceptibility tensor. However, solution EPR data showed that its active center copper(II) ion had more than one coordination geometry. Similar attempts with acetazolamide and PTS adducts proved unfruitful in terms of EPR data that showed the presence of several copper(II) coordination geometries. Nevertheless, in combination with X-ray crystal structural studies, EPR helped choose oxalate, a smaller bidentate and rigid ligand, to obtain the ideal system for our objective. EPR and NMR are techniques with different timescales, which is an essential feature for experimental complementarity. Using EPR we determined there were several coordination geometries using the different ligands. However if we used only NMR we would have wrongly assumed the coordination geometry of copper(II) in the active site, due to the presence of a single set of resonances.

To understand the different results for the different ligands, X-ray crystallography was also used. Although the physical conditions of a crystal in the data acquisition of this technique, are different from those in EPR and NMR experiments, X-ray crystallography can prove to be a useful ally in understanding differences in stereochemistry. The crystals' datasets were acquired at very low temperature, therefore the motion is slowed down. One also must take into consideration the crystal packing forces. Nonetheless, we could conclude that local movements in the protein active site could affect the acquired data, due to the low electron density at the ligand region in the active site.

The oxalate ligand proved to be the best of the four used inhibitor ligands because it gave rise to a stable single coordination geometry at the active center, although only for the copper(II)-TM-hCAII-oxalate construct. This was not the case for the corresponding DM-hCAII construct, as it is possible that residue H64's rotations affect the water network close to the metal ion center. Therefore, the copper(II)-TM-hCAII-oxalate system could be used to acquire EPR and ^{15}N-^{1}H-HSQC NMR data in solution in the same conditions and use the experimental PCSs and g values to compare them with predicted values from SE and QC theories, using Eqs. 5 and 4, respectively.

In summary, the present analysis shows that the PCS predicted with the SE theory, described by eq 5, are in good agreement with the *g* anisotropy values obtained through EPR measurements. In contrast, the PCS values predicted using Eq. 5.4, obtained from the QC derivation of the spin dipolar part of the HFC, are not in agreement with the experimental *g* values outside of any possible experimental error. This confirms the validity of the SE theory for the analysis of the PCS. The equation derived from QC theory seems unable to reproduce the paramagnetic shifts, probably because it does not take into account all contributions that one should consider when a paramagnetic metal ion is substituted for a diamagnetic analogue. In particular, the dipole−dipole interaction between the magnetic moments of the electron and nucleus involves the spin-only part of the electron magnetic moment, thus excluding the orbital component[161].

The present results are important for the development of paramagnetic NMR applied to biological systems[28,29,30–32,34,38,39,41,42,71] that needs a reliable theoretical framework for the interpretation of experimental data while maintaining the possibility of investigation of the electronic structure for different metal centers using *ab-initio* calculations. The conclusions drawn here apply not only to solution NMR but also to solid state NMR studies, including oriented lipid bilayers and membrane proteins, to in-cell NMR studies etc. Having obtained superior agreement to the experimental data using the SE framework with respect to the QC framework, this work endorses the SE view that a nucleus that is far away from the metal senses coupling with both the spin magnetic moment and the orbital magnetic moment of the electron, i.e., with the full magnetic moment of the metal center, in turn proportional to the magnetic susceptibility.[84] This poses a firm point of reference on the nature of the magnetic interaction between nuclei and unpaired electron(s).

This view was recently confirmed by the development of a more rigorous QC approach to determine the hyperfine shift. In the QC approach, the calculation of the hyperfine shielding tensor involves two contributions to the hyperfine coupling tensor **A,** the spin-dipolar ($\mathbf{A}_{sd}$) and Fermi contact ($\mathbf{A}_{con}$) contributions. The part of the hyperfine shielding tensor due to A_{sd} was identified with the spin-dipolar (SD) shielding tensor. Then, defining a non-symmetric tensor, χ', it can assume the place traditionally reserved for χ in the SE theory.[49] When spin-orbit coupling is considered in the QC framework, an additional term A_{so} is introduced in **A** which cannot be assigned to either the contact or the SD part of the hyperfine shift. Hence, strictly speaking, the comparison of calculated contact and spin-dipolar parts of the hyperfine shift to the experimentally distinguished contact and pseudocontact shifts is only meaningful as long as A_{so} is negligible. In the opposite case, only the total hyperfine shift should be compared to the experiment. The non-symmetric tensor χ' expression was derived while neglecting A_{so}.[161]

As the spin−orbit coupling is responsible for the orbital contributions to both the ***g*** tensor and the **A** tensor, the QC approach leading to χ' incorporates this orbital contribution in the first but not the second tensor, i.e., in a very unsymmetrical way. A complete QC theory of the hyperfine shielding tensor including all of the contributions (not only the SD one) to the magnetic hyperfine shielding and to the **A** tensor was recently published. After inclusion of previously neglected terms in the full Hamiltonian, the SE equations can be deduced from a rigorous QC-based treatment.[180] The inclusion of further terms in the rigorous QC treatment has demonstrated that the pseudocontact shift is indeed dependent on the magnetic susceptibility tensor, providing a definitive proof of the SE Kurland−McGarvey Eq. 5.1, which led to using the magnetic susceptibility tensor computed with the *ab-initio* state averaged complete active space self-consistent field (SACASSCF) theory in structural applications based on PCSs to cobalt(II) protein and a nickel(II) complex.[87,181]

6 General Conclusions

The hCAII is a robust system that can coordinate many different metal ions without changing its structure. This is a crucial feature to study the paramagnetic properties that arise from many transition metal ions and its effects of protein systems. This feature allowed to address the different problems, posed throughout the thesis, in the study of paramagnetic effects.

In this thesis we were able to successfully produce hCAII and its mutants and create different metal derivatives that were required to answer many different questions. Furthermore, the mutant preparation of hCAII allowed us to create a protein platform with an important characteristic that WT-hCAII does not have, the capability of binding a single metal ion. This feature proved crucial to understand the origin of the pseudocontact shifts (see Chapter 5). The usage of different structural biology techniques also gave a more comprehensive picture of the different systems, proving the point that integrated structural biology is the future of research.

The used metal systems present different features that are dependent on the used metal ion. In this thesis we explored two sources of structural information, the PCS and the paramagnetic RDC. For the measured PCS, the cobalt system provided the largest PCS and copper the lowest, as expected.[50] The paramagnetic RDCs were only measured for the cobalt(II)-DM-hCAII. For the copper and nickel systems the measurement of the paramagnetic RDC was not done. For copper the typical magnetic susceptibility tensor does not produce RDC larger than 1 Hz, therefore they are not measurable.[50] For the explored nickel system only the octahedral coordination geometry can provide measurable RDC, however only at magnetic fields close to 1000 MHz or higher. This could be interesting to explore in CERM, since they have the new 1.2 GHz spectrometer, meaning new information on paramagnetic nickel probes can be exploited and complemented to literature.

The acquisition of PCS and paramagnetic RDC can be used as structural restraints to further refine the protein structure giving a more accurate result and further validation on many structures determined by X-ray crystallography. [182] Furthermore, these observables can also be used into quantum calculations that will give us a better understanding of the paramagnetic systems.

With this developed work and analysis, I hope that it can be applied to future structure biology restrains to better understand the structure and function of proteins.

7 References

(1) Tripp, B. C.; Smith, K.; Ferry, J. G. Carbonic Anhydrase: New Insights for an Ancient Enzyme. *J. Biol. Chem.* **2001**, *276* (52), 48615–48618. https://doi.org/10.1074/jbc.R100045200.

(2) Lipton, A. S.; Heck, R. W.; Ellis, P. D. Zinc Solid-State NMR Spectroscopy of Human Carbonic Anhydrase: Implications for the Enzymatic Mechanism. *J. Am. Chem. Soc.* **2004**, *126* (14), 4735–4739. https://doi.org/10.1021/ja0305609.

(3) Bertini, I.; Canti, G.; Luchinat, C.; Borghi, E. Investigation of the System Copper(II) Carbonic Anhydrase and HCO3-1/CO2. *J. Inorg. Biochem.* **1983**, *18* (3), 221–229. https://doi.org/10.1016/0162-0134(83)85004-1.

(4) Lindskog, S. Structure and Mechanism of Carbonic Anhydrase. *Pharmacol. Ther.* **1997**, *74* (1), 1–20. https://doi.org/10.1016/S0163-7258(96)00198-2.

(5) Banci, L.; Bertini, I.; Luchinat, C.; Moratal, J. M. The Mechanism of Action of Carbonic Anhydrase. In *Enzymatic and Model Carboxylation and Reduction Reactions for Carbon Dioxide Utilization*; Springer Netherlands: Dordrecht, 1990; pp 181–197. https://doi.org/10.1007/978-94-009-0663-1_10.

(6) Sethi, K. K.; Vullo, D.; Verma, S. M.; Tanç, M.; Carta, F.; Supuran, C. T. Carbonic Anhydrase Inhibitors: Synthesis and Inhibition of the Human Carbonic Anhydrase Isoforms I, II, VII, IX and XII with Benzene Sulfonamides Incorporating 4,5,6,7-Tetrabromophthalimide Moiety. *Bioorg. Med. Chem.* **2013**, *21* (19), 5973–5982. https://doi.org/10.1016/j.bmc.2013.07.044.

(7) Aggarwal, M.; Boone, C. D.; Kondeti, B.; McKenna, R. Structural Annotation of Human Carbonic Anhydrases. *J. Enzyme Inhib. Med. Chem.* **2013**, *28* (2), 267–277. https://doi.org/10.3109/14756366.2012.737323.

(8) Whittington, D. A.; Waheed, A.; Ulmasov, B.; Shah, G. N.; Grubb, J. H.; Sly, W. S.; Christianson, D. W. Crystal Structure of the Dimeric Extracellular Domain of Human Carbonic Anhydrase XII, a Bitopic Membrane Protein Overexpressed in Certain Cancer Tumor Cells. *Proc. Natl. Acad. Sci.* **2001**, *98* (17), 9545–9550. https://doi.org/10.1073/pnas.161301298.

(9) Mishra, C. B.; Tiwari, M.; Supuran, C. T. Progress in the Development of Human Carbonic Anhydrase Inhibitors and Their Pharmacological Applications: Where Are We Today? *Med. Res. Rev.* **2020**, *40* (6), 2485–2565. https://doi.org/10.1002/med.21713.

(10) Bryk, A. H.; Wiśniewski, J. R. Quantitative Analysis of Human Red Blood Cell Proteome. *J. Proteome Res.* **2017**, *16* (8), 2752–2761. https://doi.org/10.1021/acs.jproteome.7b00025.

(11) Bozdag, M.; Ferraroni, M.; Ward, C.; Carta, F.; Bua, S.; Angeli, A.; Langdon, S. P.; Kunkler, I. H.; Al-Tamimi, A. M. S.; Supuran, C. T. Carbonic Anhydrase Inhibitors Based on Sorafenib

Scaffold: Design, Synthesis, Crystallographic Investigation and Effects on Primary Breast Cancer Cells. *Eur. J. Med. Chem.* **2019**, *182*, 111600. https://doi.org/10.1016/j.ejmech.2019.111600.

(12) Avvaru, B. S.; Kim, C. U.; Sippel, K. H.; Gruner, S. M.; Agbandje-McKenna, M.; Silverman, D. N.; McKenna, R. A Short, Strong Hydrogen Bond in the Active Site of Human Carbonic Anhydrase II. *Biochemistry* **2010**, *49* (2), 249–251. https://doi.org/10.1021/bi902007b.

(13) Elder, I.; Fisher, Z.; Laipis, P. J.; Tu, C.; McKenna, R.; Silverman, D. N. Structural and Kinetic Analysis of Proton Shuttle Residues in the Active Site of Human Carbonic Anhydrase III. *Proteins Struct. Funct. Genet.* **2007**, *68* (1), 337–343. https://doi.org/10.1002/prot.21403.

(14) Vernier, W.; Chong, W.; Rewolinski, D.; Greasley, S.; Pauly, T.; Shaw, M.; Dinh, D.; Ferre, R. A.; Nukui, S.; Ornelas, M.; Reyner, E. Thioether Benzenesulfonamide Inhibitors of Carbonic Anhydrases II and IV: Structure-Based Drug Design, Synthesis, and Biological Evaluation. *Bioorganic Med. Chem.* **2010**, *18* (9), 3307–3319. https://doi.org/10.1016/j.bmc.2010.03.014.

(15) Jude, K. M.; Wright, S. K.; Tu, C.; Silverman, D. N.; Viola, R. E.; Christianson, D. W. Crystal Structure of F65A/Y131C-Methylimidazole Carbonic Anhydrase V Reveals Architectural Features of an Engineered Proton Shuttle. *Biochemistry* **2002**, *41* (8), 2485–2491. https://doi.org/10.1021/bi015808q.

(16) Pilka, E. S.; Kochan, G.; Oppermann, U.; Yue, W. W. Crystal Structure of the Secretory Isozyme of Mammalian Carbonic Anhydrases CA VI: Implications for Biological Assembly and Inhibitor Development. *Biochem. Biophys. Res. Commun.* **2012**, *419* (3), 485–489. https://doi.org/10.1016/j.bbrc.2012.02.038.

(17) Buemi, M. R.; Di Fiore, A.; De Luca, L.; Angeli, A.; Mancuso, F.; Ferro, S.; Monti, S. M.; Buonanno, M.; Russo, E.; De Sarro, G.; De Simone, G.; Supuran, C. T.; Gitto, R. Exploring Structural Properties of Potent Human Carbonic Anhydrase Inhibitors Bearing a 4-(Cycloalkylamino-1-Carbonyl)Benzenesulfonamide Moiety. *Eur. J. Med. Chem.* **2019**, *163*, 443–452. https://doi.org/10.1016/j.ejmech.2018.11.073.

(18) Picaud, S. S.; Muniz, J. R. C.; Kramm, A.; Pilka, E. S.; Kochan, G.; Oppermann, U.; Yue, W. W. Crystal Structure of Human Carbonic Anhydrase-Related Protein VIII Reveals the Basis for Catalytic Silencing. *Proteins Struct. Funct. Bioinforma.* **2009**, *76* (2), 507–511. https://doi.org/10.1002/prot.22411.

(19) Koruza, K.; Lafumat, B.; Nyblom, M.; Mahon, B. P.; Knecht, W.; McKenna, R.; Fisher, S. Z. Structural Comparison of Protiated, H/D-Exchanged and Deuterated Human Carbonic Anhydrase IX. *Acta Crystallogr. Sect. D Struct. Biol.* **2019**, *75* (10), 895–903. https://doi.org/10.1107/S2059798319010027.

(20) RCSB PDB - 5MSA: Crystal structure of human carbonic anhydrase isozyme XII with 2,3,5,6-Tetrafluoro-4-(propylthio)benzenesulfonamide https://www.rcsb.org/structure/5MSA (accessed Jul 1, 2021).

(21) RCSB PDB - 5OGJ: Crystal structure of human carbonic anhydrase isozyme XIII with 2-(Cyclooctylamino)-3,5,6-trifluoro-4-[(2-hydroxyethyl)thio]benzenesulfonamide https://www.rcsb.org/structure/5OGJ (accessed Jul 1, 2021).

(22) La Regina, G.; Coluccia, A.; Famiglini, V.; Pelliccia, S.; Monti, L.; Vullo, D.; Nuti, E.; Alterio, V.; De Simone, G.; Monti, S. M.; Pan, P.; Parkkila, S.; Supuran, C. T.; Rossello, A.; Silvestri, R. Discovery of 1,1′-Biphenyl-4-Sulfonamides as a New Class of Potent and Selective Carbonic Anhydrase XIV Inhibitors. *J. Med. Chem.* **2015**, *58* (21), 8564–8572. https://doi.org/10.1021/acs.jmedchem.5b01144.

(23) Cerofolini, L.; Giuntini, S.; Louka, A.; Ravera, E.; Fragai, M.; Luchinat, C. High-Resolution Solid-State NMR Characterization of Ligand Binding to a Protein Immobilized in a Silica Matrix. *J. Phys. Chem. B* **2017**, *121* (34), 8094–8101. https://doi.org/10.1021/acs.jpcb.7b05679.

(24) Putnam, R. W. Intracellular PH Regulation. In *Cell Physiology Source Book*; Elsevier, 2012; pp 303–321. https://doi.org/10.1016/B978-0-12-387738-3.00017-2.

(25) Chiche, J.; Ilc, K.; Laferrière, J.; Trottier, E.; Dayan, F.; Mazure, N. M.; Brahimi-Horn, M. C.; Pouysségur, J. Hypoxia-Inducible Carbonic Anhydrase IX and XII Promote Tumor Cell Growth by Counteracting Acidosis through the Regulation of the Intracellular PH. *Cancer Res.* **2009**, *69* (1), 358–368. https://doi.org/10.1158/0008-5472.CAN-08-2470.

(26) Guyton, A. C.; Hall, J. E. *Textbook of Medical Phisiology 11th Edition*; 2011.

(27) Richmond, C. M.; Leventer, R.; Ryan, M. M.; Delatycki, M. B. Cerebellar Ataxia with Normal Intellect Associated with a Homozygous Truncating Variant in CA8. *Clin. Genet.* **2020**, *97* (3), 516–520. https://doi.org/10.1111/cge.13666.

(28) Türkmen, S.; Guo, G.; Garshasbi, M.; Hoffmann, K.; Alshalah, A. J.; Mischung, C.; Kuss, A.; Humphrey, N.; Mundlos, S.; Robinson, P. N. CA8 Mutations Cause a Novel Syndrome Characterized by Ataxia and Mild Mental Retardation with Predisposition to Quadrupedal Gait. *PLoS Genet.* **2009**, *5* (5), e1000487. https://doi.org/10.1371/journal.pgen.1000487.

(29) Hakansson, K.; Wehnert, A.; Liljas, A. X-Ray Analysis of Metal-Substituted Human Carbonic Anhydrase II Derivatives. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **1994**, *50* (1), 93–100. https://doi.org/10.1107/S0907444993008790.

(30) Coleman, J. E. Mechanism of Action of Carbonic Anhydrase. *J. Biol. Chem.* **1967**, *242* (22), 5212–5219. https://doi.org/10.1016/S0021-9258(18)99413-5.

(31) Avvaru, B. S.; Arenas, D. J.; Tu, C.; Tanner, D. B.; McKenna, R.; Silverman, D. N. Comparison of Solution and Crystal Properties of Co(II)–Substituted Human Carbonic Anhydrase II. *Arch. Biochem. Biophys.* **2010**, *502* (1), 53–59. https://doi.org/10.1016/j.abb.2010.07.010.

(32) Bertini, I.; Borghi, E.; Luchinat, C.; Monnanni, R. Nickel Carbonic Anhydrase: A Re-Examination of the Electronic Spectra with the Help of CD Spectra. *Inorganica Chim. Acta* **1982**, *67* (C), 99–102. https://doi.org/10.1016/S0020-1693(00)85049-5.

(33) Bertini, I.; Canti, G.; Luchinat, C.; Scozzafava, A. Characterization of Cobalt(II) Bovine Carbonic Anhydrase and of Its Derivatives. *J. Am. Chem. Soc.* **1978**, *100* (15), 4873–4877. https://doi.org/10.1021/ja00483a038.

(34) Bertini, I.; Luchinat, C. The Structure of Cobalt(II)-substituted Carbonic Anhydrase and Its Implications for the Catalytic Mechanism of the Enzyme. *Ann. N. Y. Acad. Sci.* **1984**, *429* (1), 89–98. https://doi.org/10.1111/j.1749-6632.1984.tb12318.x.

(35) Silva, J. M.; Giuntini, S.; Cerofolini, L.; Geraldes, C. F. G. C.; Macedo, A. L.; Ravera, E.; Fragai,

M.; Luchinat, C.; Calderone, V. Non-Crystallographic Symmetry in Proteins: Jahn–Teller-like and Butterfly-like Effects? *JBIC J. Biol. Inorg. Chem.* **2019**, *24* (1), 91–101. https://doi.org/10.1007/s00775-018-1630-0.

(36) Malanho Silva, J.; Cerofolini, L.; Giuntini, S.; Calderone, V.; Geraldes, C. F. G. C.; Macedo, A. L.; Parigi, G.; Fragai, M.; Ravera, E.; Luchinat, C. Metal Centers in Biomolecular Solid-State NMR. *J. Struct. Biol.* **2019**, *206* (1), 99–109. https://doi.org/10.1016/j.jsb.2018.11.013.

(37) Okrasa, K.; Kazlauskas, R. J. Manganese-Substituted Carbonic Anhydrase as a New Peroxidase. *Chem. Eur. J.* **2006**, *12* (6), 1587–1596. https://doi.org/10.1002/chem.200501413.

(38) Jing, Q.; Okrasa, K.; Kazlauskas, R. J. Manganese-Substituted α-Carbonic Anhydrase as an Enantioselective Peroxidase. *Top. Organomet. Chem.* **2009**, *25* (41 mm), 45–61. https://doi.org/10.1007/978-3-540-87757-8_3.

(39) Song, H.; Weitz, A. C.; Hendrich, M. P.; Lewis, E. A.; Emerson, J. P. Building Reactive Copper Centers in Human Carbonic Anhydrase II. *J. Biol. Inorg. Chem.* **2013**, *18* (6), 595–598. https://doi.org/10.1007/s00775-013-1009-1.

(40) Andring, J. T.; Kim, C. U.; McKenna, R. Structure and Mechanism of Copper–Carbonic Anhydrase II: A Nitrite Reductase. *IUCrJ* **2020**, *7* (2), 287–293. https://doi.org/10.1107/S2052252520000986.

(41) Jing, Q.; Okrasa, K.; Kazlauskas, R. J. Stereoselective Hydrogenation of Olefins Using Rhodium-Substituted Carbonic Anhydrase - A New Reductase. *Chem. Eur. J.* **2009**, *15* (6), 1370–1376. https://doi.org/10.1002/chem.200801673.

(42) Piazzetta, P.; Marino, T.; Russo, N.; Salahub, D. R. Direct Hydrogenation of Carbon Dioxide by an Artificial Reductase Obtained by Substituting Rhodium for Zinc in the Carbonic Anhydrase Catalytic Center. A Mechanistic Study. *ACS Catal.* **2015**, *5* (9), 5397–5409. https://doi.org/10.1021/acscatal.5b00185.

(43) Moratal, J. M.; Jiménez, H. R.; Castells, J.; Salgado, J.; Martínez-Ferrer, M.-J.; Donaire, A. 1H NMR and UV-Vis Spectroscopic Characterization of Sulfonamide Complexes of Nickek(II)-Carbonic Anhydrase. Resonance Assignments Based on NOE Effects. *J. Inorg. Biochem.* **1992**, *45* (4), 231–243. https://doi.org/10.1016/0162-0134(92)84012-C.

(44) Bencini, A.; Bertini, I.; Canti, G.; Gatteschi, D.; Luchinat, C. The Epr Spectra of the Inhibitor Derivatives of Cobalt Carbonic Anhydrase. *J. Inorg. Biochem.* **1981**, *14* (1), 81–93. https://doi.org/10.1016/S0162-0134(00)80016-1.

(45) Nitsche, C.; Otting, G. Pseudocontact Shifts in Biomolecular NMR Using Paramagnetic Metal Tags. *Prog. Nucl. Magn. Reson. Spectrosc.* **2017**, *98–99*, 20–49. https://doi.org/10.1016/j.pnmrs.2016.11.001.

(46) Cerofolini, L.; Staderini, T.; Giuntini, S.; Ravera, E.; Fragai, M.; Parigi, G.; Pierattelli, R.; Luchinat, C. Long-Range Paramagnetic NMR Data Can Provide a Closer Look on Metal Coordination in Metalloproteins. *J. Biol. Inorg. Chem.* **2018**, *23* (1), 71–80. https://doi.org/10.1007/s00775-017-1511-y.

(47) Rubio, S.; Borrás, C.; Alzuet, G.; Borrás, J. Comparison of the Interaction of Cobalt Bovine Carbonic Anhydrase II with Acetazolamide and Methazolamide and the Reaction of Apoenzyme

with Cobalt(II) Complexes of Acetazolamide and Methazolamide: Spectrophotometric Study. *Biochem. Mol. Biol. Educ.* **2003**, *31* (1), 28–33. https://doi.org/10.1002/bmb.2003.494031010153.

(48) Bertini, I.; Luchinat, C. Cobalt(II) as a Probe of the Structure and Function of Carbonic Anhydrase. *Acc. Chem. Res.* **1983**, *16* (8), 272–279. https://doi.org/10.1021/ar00092a002.

(49) Benda, L.; Mareš, J.; Ravera, E.; Parigi, G.; Luchinat, C.; Kaupp, M.; Vaara, J.; Benda, L.; Vaara, J. Pseudo-Contact NMR Shifts over the Paramagnetic Metalloprotein CoMMP-12 from First Principles. *Angew. Chemie Int. Ed.* **2016**, *55* (47), 14713–14717. https://doi.org/10.1002/anie.201608829.

(50) Bertini, I.; Luchinat, C.; Parigi, G.; Ravera, E. *NMR of Paramagnetic Molecules. Applications to Metallobiomolecules and Models.*; 2016.

(51) Gil, V. M. S.; Geraldes, C. F. G. C. *Ressonância Magnética Nuclear. Fundamentos, Métodos e Aplicações*, 2nd ed.; Fundação Calouste Gulbenkian: Lisboa, 2002.

(52) Suturina, E. A.; Kuprov, I. Pseudocontact Shifts from Mobile Spin Labels. *Phys. Chem. Chem. Phys.* **2016**, *18* (38), 26412–26422. https://doi.org/10.1039/c6cp05437d.

(53) Schmitz, C.; Stanton-Cook, M. J.; Su, X.-C.; Otting, G.; Huber, T. Numbat: An Interactive Software Tool for Fitting Δχ-Tensors to Molecular Coordinates Using Pseudocontact Shifts. *J. Biomol. NMR* **2008**, *41* (3), 179–189. https://doi.org/10.1007/s10858-008-9249-z.

(54) Rinaldelli, M.; Ravera, E.; Calderone, V.; Parigi, G.; Murshudov, G. N.; Luchinat, C. Simultaneous Use of Solution NMR and X-Ray Data in REFMAC5 for Joint Refinement/Detection of Structural Differences. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2014**, *70* (4), 958–967. https://doi.org/10.1107/S1399004713034160.

(55) Clore, G. M. Practical Aspects of Paramagnetic Relaxation Enhancement in Biological Macromolecules. In *Methods in Enzymology*; Academic Press Inc., 2015; Vol. 564, pp 485–497. https://doi.org/10.1016/bs.mie.2015.06.032.

(56) Bloembergen, N.; Dickinson, W. C. On the Shift of the Nuclear Magnetic Resonance in Paramagnetic Solutions. *Phys. Rev.* **1950**, *79* (1), 179–180. https://doi.org/10.1103/PhysRev.79.179.

(57) McConnell, H. M.; Robertson, R. E. Isotropic Nuclear Resonance Shifts. *J. Chem. Phys.* **1958**, *29* (6), 1361–1365. https://doi.org/10.1063/1.1744723.

(58) Kurland, R. J.; McGarvey, B. R. Isotropic NMR Shifts in Transition Metal Complexes: The Calculation of the Fermi Contact and Pseudocontact Terms. *J. Magn. Reson.* **1970**, *2* (3), 286–301. https://doi.org/10.1016/0022-2364(70)90100-9.

(59) Schmitz, C.; Vernon, R.; Otting, G.; Baker, D.; Huber, T. Protein Structure Determination from Pseudocontact Shifts Using ROSETTA. *J. Mol. Biol.* **2012**, *416* (5), 668–677. https://doi.org/10.1016/j.jmb.2011.12.056.

(60) Wang, X.; Srisailam, S.; Yee, A. A.; Lemak, A.; Arrowsmith, C.; Prestegard, J. H.; Tian, F. Domain-Domain Motions in Proteins from Time-Modulated Pseudocontact Shifts. *J. Biomol. NMR* **2007**, *39* (1), 53–61. https://doi.org/10.1007/s10858-007-9174-6.

(61) McDermott, A.; Polenova, T. Solid State NMR: New Tools for Insight into Enzyme Function. *Curr.*

Opin. Struct. Biol. **2007**, *17* (5), 617–622. https://doi.org/10.1016/j.sbi.2007.10.001.

(62) Parthasarathy, S.; Nishiyama, Y.; Ishii, Y. Sensitivity and Resolution Enhanced Solid-State NMR for Paramagnetic Systems and Biomolecules under Very Fast Magic Angle Spinning. *Acc. Chem. Res.* **2013**, *46* (9), 2127–2135. https://doi.org/10.1021/ar4000482.

(63) Yamamoto, K.; Dürr, U. H. N.; Xu, J.; Im, S. C.; Waskell, L.; Ramamoorthy, A. Dynamic Interaction between Membrane-Bound Full-Length Cytochrome P450 and Cytochrome B5 Observed by Solid-State NMR Spectroscopy. *Sci. Rep.* **2013**, *3*, 2538. https://doi.org/10.1038/srep02538.

(64) Yamamoto, K.; Xu, J.; Kawulka, K. E.; Vederas, J. C.; Ramamoorthy, A. Use of a Copper-Chelated Lipid Speeds up NMR Measurements from Membrane Proteins. *J. Am. Chem. Soc.* **2010**, *132* (20), 6929–6931. https://doi.org/10.1021/ja102103n.

(65) Yamamoto, K.; Vivekanandan, S.; Ramamoorthy, A. Fast NMR Data Acquisition from Bicelles Containing a Membrane-Associated Peptide at Natural-Abundance. *J. Phys. Chem. B* **2011**, *115* (43), 12448–12455. https://doi.org/10.1021/jp2076098.

(66) Keizers, P. H. J.; Saragliadis, A.; Hiruma, Y.; Overhand, M.; Ubbink, M. Design, Synthesis, and Evaluation of a Lanthanide Chelating Protein Probe: CLaNP-5 Yields Predictable Paramagnetic Effects Independent of Environment. *J. Am. Chem. Soc.* **2008**, *130* (44), 14802–14812. https://doi.org/10.1021/ja8054832.

(67) Charnock, G. T. P.; Kuprov, I. A Partial Differential Equation for Pseudocontact Shift. *Phys. Chem. Chem. Phys.* **2014**, *16* (37), 20184–20189. https://doi.org/10.1039/c4cp03106g.

(68) Bertini, I.; Kursula, P.; Luchinat, C.; Parigi, G.; Vahokoski, J.; Wilmanns, M.; Yuan, J. Accurate Solution Structures of Proteins from X-Ray Data and a Minimal Set of NMR Data: Calmodulin-Peptide Complexes as Examples. *J. Am. Chem. Soc.* **2009**, *131* (14), 5134–5144. https://doi.org/10.1021/ja8080764.

(69) Canales, A.; Mallagaray, A.; Pérez-Castells, J.; Boos, I.; Unverzagt, C.; André, S.; Gabius, H. J.; Cañada, F. J.; Jiménez-Barbero, J. Breaking Pseudo-Symmetry in Multiantennary Complex N-Glycans Using Lanthanide-Binding Tags and NMR Pseudo-Contact Shifts. *Angew. Chemie - Int. Ed.* **2013**, *52* (51), 13789–13793. https://doi.org/10.1002/anie.201307845.

(70) Pintacuda, G.; John, M.; Su, X. C.; Otting, G. NMR Structure Determination of Protein - Ligand Complexes by Lanthanide Labeling. *Acc. Chem. Res.* **2007**, *40* (3), 206–212. https://doi.org/10.1021/ar050087z.

(71) Suturina, E. A.; Mason, K.; Geraldes, C. F. G. C.; Kuprov, I.; Parker, D. Beyond Bleaney’s Theory: Experimental and Theoretical Analysis of Periodic Trends in Lanthanide-Induced Chemical Shift. *Angew. Chemie Int. Ed.* **2017**, *56* (40), 12215–12218. https://doi.org/10.1002/anie.201706931.

(72) Jensen, M. R.; Hansen, D. F.; Ayna, U.; Dagil, R.; Hass, M. A. S.; Christensen, H. E. M.; Led, J. J. On the Use of Pseudocontact Shifts in the Structure Determination of Metalloproteins. *Magn. Reson. Chem.* **2006**, *44* (3 SPEC. ISS.), 294–301. https://doi.org/10.1002/mrc.1771.

(73) Rinaldelli, M.; Carlon, A.; Ravera, E.; Parigi, G.; Luchinat, C. FANTEN: A New Web-Based Interface for the Analysis of Magnetic Anisotropy-Induced NMR Data. *J. Biomol. NMR* **2015**.

https://doi.org/10.1007/s10858-014-9877-4.

(74) Salgado, J.; Jiménez, H. R.; Moratal, J. M.; Kroes, S.; Warmerdam, G. C. M.; Canters, G. W. Paramagnetic Cobalt and Nickel Derivatives of Alcaligenes Denitrificans Azurin and Its M121Q Mutant. A 1 H NMR Study †. *Biochemistry* **1996**, *35* (6), 1810–1819. https://doi.org/10.1021/bi951748a.

(75) Nettles, W. L.; Song, H.; Farquhar, E. R.; Fitzkee, N. C.; Emerson, J. P. Characterization of the Copper(II) Binding Sites in Human Carbonic Anhydrase II. *Inorg. Chem.* **2015**, *54* (12), 5671–5680. https://doi.org/10.1021/acs.inorgchem.5b00057.

(76) Duer, M. J. Solid-State NMR Spectroscopy : Principles and Applications. **2002**, 567.

(77) Pell, A. J.; Pintacuda, G.; Grey, C. P. Paramagnetic NMR in Solution and the Solid State. *Prog. Nucl. Magn. Reson. Spectrosc.* **2019**, *111*, 1–271. https://doi.org/10.1016/j.pnmrs.2018.05.001.

(78) Drago, R. S. *Physical Methods in Chemistry*; Saunders, P. W. B., Ed.; 1977.

(79) Carrington, A.; McLachlan, A. D. Introduction to Magnetic Resonance : With Applications to Chemistry and Chemical Physics. **1979**, 266.

(80) Palmer, G. The Electron Paramagnetic Resonance of Metalloproteins. *Biochem. Soc. Trans.* **1985**, *13* (3), 548–560. https://doi.org/10.1042/bst0130548.

(81) Krzystek, J.; Ozarowski, A.; Telser, J. Multi-Frequency, High-Field EPR as a Powerful Tool to Accurately Determine Zero-Field Splitting in High-Spin Transition Metal Coordination Complexes. *Coord. Chem. Rev.* **2006**, *250* (17–18), 2308–2324. https://doi.org/10.1016/j.ccr.2006.03.016.

(82) Rhodes, G. *Crystallography Made Crystal Clear*, 3rd editio.; Elsevier, 2006. https://doi.org/10.1016/B978-0-12-587073-3.X5000-4.

(83) Benvenuti, M.; Mangani, S. Crystallization of Soluble Proteins in Vapor Diffusion for X-Ray Crystallography. *Nat. Protoc.* **2007**, *2* (7), 1633–1651. https://doi.org/10.1038/nprot.2007.198.

(84) Cerofolini, L.; Silva, J. M.; Ravera, E.; Romanelli, M.; Geraldes, C. F. G. C.; Macedo, A. L.; Fragai, M.; Parigi, G.; Luchinat, C. How Do Nuclei Couple to the Magnetic Moment of a Paramagnetic Center? A New Theory at the Gauntlet of the Experiments. *J. Phys. Chem. Lett.* **2019**, *10* (13), 3610–3614. https://doi.org/10.1021/acs.jpclett.9b01128.

(85) Bertini, L. B.; Bertini, I. B.; Luchinat, C.; Scozzafava, A. Nuclear Relaxation in the Magnetic Coupled System Cu2Co 2SOD: Histidine-44 Is Detached upon Anion Binding. *J. Am. Chem. Soc.* **1987**, *109* (8), 2328–2334. https://doi.org/10.1021/JA00242A014.

(86) Banci, L.; Bertini, I.; Luchinat, C.; Piccioli, M. Transient versus Steady State NOE in Paramagnetic Molecules Cu2Co2SOD as an Example. *FEBS Lett.* **1990**, *272* (1–2), 175–180. https://doi.org/10.1016/0014-5793(90)80477-Z.

(87) Ravera, E.; Gigli, L.; Suturina, E. A.; Calderone, V.; Fragai, M.; Parigi, G.; Luchinat, C. A High-Resolution View of the Coordination Environment in a Paramagnetic Metalloprotein from Its Magnetic Properties. *Angew. Chemie Int. Ed.* **2021**, *60* (27), 14960–14966. https://doi.org/10.1002/ANIE.202101149.

(88) Balayssac, S.; Bertini, I.; Bhaumik, A.; Lelli, M.; Luchinat, C. Paramagnetic Shifts in Solid-State NMR of Proteins to Elicit Structural Information. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105* (45),

17284–17289. https://doi.org/10.1073/pnas.0708460105.

(89) Bertini, I.; Emsley, L.; Lelli, M.; Luchinat, C.; Mao, J.; Pintacuda, G. Ultrafast MAS Solid-State NMR Permits Extensive 13 C and 1 H Detection in Paramagnetic Metalloproteins. *J. Am. Chem. Soc.* **2010**, *132* (16), 5558–5559. https://doi.org/10.1021/ja100398q.

(90) Banci, L.; Dugad, L. B.; La Mar, G. N.; Keating, K. A.; Luchinat, C.; Pierattelli, R. 1H Nuclear Magnetic Resonance Investigation of Cobalt(II) Substituted Carbonic Anhydrase. *Biophys. J.* **1992**, *63* (2), 530–543. https://doi.org/10.1016/S0006-3495(92)81607-7.

(91) Bertini, I.; Jonsson, B. H. B. H.; Luchinat, C.; Pierattelli, R.; Vila, A. J. A. J.; Plerattelli, R.; Vila, A. J. A. J. Strategies of Signal Assignments in Paramagnetic Metalloproteins. An NMR Investigation of the Thiocyanate Adduct of the Cobalt(II)-Substituted Human Carbonic Anhydrase II. *J. Magn. Reson. Ser. B* **1994**, *104* (3), 230–239. https://doi.org/10.1006/jmrb.1994.1080.

(92) Bertini, I.; Dei, A.; Luchinat, C.; Monnanni, R. Acid-Base Properties of Cobalt(II)-Substituted Carbonic Anhydrases. *Inorg. Chem.* **1985**, *24* (3), 301–303. https://doi.org/10.1021/ic00197a012.

(93) Bertini, I.; Canti, G.; Luchinat, C. Water in the Coordination Sphere of Metallocarbonic Anhydrases: A Solvent Proton Longitudinal Relaxation Study at Several Frequencies. *Inorganica Chim. Acta* **1981**, *56* (C), 99–107. https://doi.org/10.1016/S0020-1693(00)88554-0.

(94) Keller, R. *The Computer Aided Resonance Assignment Tutorial*; 2004.

(95) Cavanagh, J. Protein NMR Spectroscopy : Principles and Practice. **2007**, 885.

(96) Gairí, M.; Dyachenko, A.; González, M. T.; Feliz, M.; Pons, M.; Giralt, E. An Optimized Method for 15N R1 Relaxation Rate Measurements in Non-Deuterated Proteins. *J. Biomol. NMR* **2015**, *62* (2), 209–220. https://doi.org/10.1007/s10858-015-9937-4.

(97) Blake-Hall, J.; Walker, O.; Fushman, D. Characterization of the Overall Rotational Diffusion of a Protein From ¹⁵N Relaxation Measurements and Hydrodynamic Calculations. In *Protein NMR Techniques*; Humana Press: New Jersey, 2004; Vol. 278, pp 139–160. https://doi.org/10.1385/1-59259-809-9:139.

(98) Williamson, M. P. Using Chemical Shift Perturbation to Characterise Ligand Binding. *Prog. Nucl. Magn. Reson. Spectrosc.* **2013**, *73*, 1–16. https://doi.org/10.1016/J.PNMRS.2013.02.001.

(99) Viegas, A.; Sardinha, J.; Freire, F.; Duarte, D. F.; Carvalho, A. L.; Fontes, C. M. G. A.; Rom˜ao, M. J.; Rom˜ao, R.; Macedo, A. L.; Cabrita, E. J. Solution Structure, Dynamics and Binding Studies of a Family 11 Carbohydrate-Binding Module from Clostridium Thermocellum (CtCBM11). *Biochem. J* **2013**, *451*, 289–300. https://doi.org/10.1042/BJ20120627.

(100) Schumann, F. H.; Hubert, A. E.; Ae, R.; Maurer, T.; Wolfram, A. E.; Ae, G.; Neidig, K.-P.; Hans, A. E.; Kalbitzer, R. Combined Chemical Shift Changes and Amino Acid Specific Chemical Shift Mapping of Protein-Protein Interactions. https://doi.org/10.1007/s10858-007-9197-z.

(101) Stoll, S.; Schweiger, A. EasySpin, a Comprehensive Software Package for Spectral Simulation and Analysis in EPR. *J. Magn. Reson.* **2006**, *178* (1), 42–55. https://doi.org/10.1016/j.jmr.2005.08.013.

(102) Pilbrow, J. R. Effective g Values for S = 3/2 and S = 5/2. *J. Magn. Reson.* **1978**, *31* (3), 479–490.

https://doi.org/10.1016/S0022-2364(78)80015-8.

(103) Banci, L.; Bencini, A.; Benelli, C.; Gatteschi, D.; Zanchini, C. Spectral-Structural Correlations in High-Spin Cobalt(II) Complexes. *Struct. versus Spec. Prop.* **2007**, 37–86. https://doi.org/10.1007/BFB0111296.

(104) DOC-M86-EXX242 PROTEUM3 Software User Manual. *Bruker AXS Inc.* **2017**.

(105) Liebschner, D.; Afonine, P. V.; Baker, M. L.; Bunkoczi, G.; Chen, V. B.; Croll, T. I.; Hintze, B.; Hung, L. W.; Jain, S.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R. D.; Poon, B. K.; Prisant, M. G.; Read, R. J.; Richardson, J. S.; Richardson, D. C.; Sammito, M. D.; Sobolev, O. V.; Stockwell, D. H.; Terwilliger, T. C.; Urzhumtsev, A. G.; Videau, L. L.; Williams, C. J.; Adams, P. D. Macromolecular Structure Determination Using X-Rays, Neutrons and Electrons: Recent Developments in Phenix. *Acta Crystallogr. Sect. D Struct. Biol.* **2019**, *75*, 861–877. https://doi.org/10.1107/S2059798319011471.

(106) Emsley, P.; Cowtan, K. Coot : Model-Building Tools for Molecular Graphics. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2004**, *60* (12), 2126–2132. https://doi.org/10.1107/S0907444904019158.

(107) Misochko, E. Y.; Akimov, A. V.; Korchagin, D. V.; Nehrkorn, J.; Ozerov, M.; Palii, A. V.; Clemente-Juan, J. M.; Aldoshin, S. M. Purely Spectroscopic Determination of the Spin Hamiltonian Parameters in High-Spin Six-Coordinated Cobalt(II) Complexes with Large Zero-Field Splitting. *Inorg. Chem.* **2019**, *58* (24), 16434–16444. https://doi.org/10.1021/ACS.INORGCHEM.9B02195/SUPPL_FILE/IC9B02195_SI_001.PDF.

(108) Bertini, I.; Luchinat, C. The Reaction Pathways of Zinc Enzymes and Related Biological Catalysts. *Bioinorg. Chem.* **1994**, 37–106.

(109) Eriksson, A. E.; Kylsten, P. M.; Jones, T. A.; Liljas, A. Crystallographic Studies of Inhibitor Binding Sites in Human Carbonic Anhydrase II: A Pentacoordinated Binding of the SCN− Ion to the Zinc at High PH. *Proteins Struct. Funct. Bioinforma.* **1988**, *4* (4), 283–293. https://doi.org/10.1002/PROT.340040407.

(110) Fox, J. M.; Kang, K.; Sherman, W.; Héroux, A.; Sastry, G. M.; Baghbanzadeh, M.; Lockett, M. R.; Whitesides, G. M. Interactions between Hofmeister Anions and the Binding Pocket of a Protein. *J. Am. Chem. Soc.* **2015**, *137* (11), 3859–3866. https://doi.org/10.1021/JACS.5B00187.

(111) Shenghua Huang, §; Björn Sjöblom, ∥; A. Elisabeth Sauer-Eriksson, § and; Bengt-Harald Jonsson*, ⊥. Organization of an Efficient Carbonic Anhydrase: Implications for the Mechanism Based on Structure−Function Studies of a T199P/C206S Mutant†,‡. *Biochemistry* **2002**, *41* (24), 7628–7635. https://doi.org/10.1021/BI020053O.

(112) Freskgaard, P.-O.; Maartensson, L.-G.; Jonasson, P.; Jonsson, B.-H.; Carlsson, U. Assignment of the Contribution of the Tryptophan Residues to the Circular Dichroism Spectrum of Human Carbonic Anhydrase II. *Biochemistry* **1994**, *33* (47), 14281–14288. https://doi.org/10.1021/bi00251a041.

(113) García de la Torre, J.; Huertas, M. .; Carrasco, B. HYDRONMR: Prediction of NMR Relaxation of Globular Proteins from Atomic-Level Structures and Hydrodynamic Calculations. *J. Magn. Reson.* **2000**, *147* (1), 138–146. https://doi.org/10.1006/jmre.2000.2170.

(114) Janz, G. J.; Oliver, B. G.; Lakshminarayanan, G. R.; Mayer, G. E. Electrical Conductance, Diffusion, Viscosity, and Density of Sodium Nitrate, Sodium Perchlorate, and Sodium Thiocyanate in Concentrated Aqueous Solutions. *J. Phys. Chem.* **1970**, *74* (6), 1285–1289. https://doi.org/10.1021/j100701a022.

(115) Bunkóczi, G.; McCoy, A. J.; Echols, N.; Grosse-Kunstleve, R. W.; Adams, P. D.; Holton, J. M.; Read, R. J.; Terwilliger, T. C. Macromolecular X-Ray Structure Determination Using Weak, Single-Wavelength Anomalous Data. *Nat. Methods* **2015**, *12* (2), 127–130. https://doi.org/10.1038/nmeth.3212.

(116) Fichtner, K. Non-Space-Group Symmetry in Crystallography. *Comput. Math. with Appl.* **1986**, *12* (3–4), 751–762. https://doi.org/10.1016/0898-1221(86)90421-9.

(117) Moratal, J. M.; Martínez-Ferrer, M.-J.; Donaire, A.; Castells, J.; Salgado, J.; Jiménez, H. R. Spectroscopic Studies of Nickel(II) Carbonic Anhydrase and Its Adducts with Inorganic Anions. *J. Chem. Soc., Dalt. Trans.* **1991**, No. 12, 3393–3399. https://doi.org/10.1039/DT9910003393.

(118) Goodfellow, B. J.; Duarte, I. C. N.; MacEdo, A. L.; Volkman, B. F.; Nunes, S. G.; Moura, I.; Markley, J. L.; Moura, J. J. G. An NMR Structural Study of Nickel-Substituted Rubredoxin. *J. Biol. Inorg. Chem.* **2010**, *15* (3), 409–420. https://doi.org/10.1007/s00775-009-0613-6.

(119) Jensen, M. R.; Led, J. J. Metal-Protein Interactions: Structure Information from Ni 2+-Induced Pseudocontact Shifts in a Native Nonmetalloprotein. *Biochemistry* **2006**, *45* (29), 8782–8787. https://doi.org/10.1021/bi0604431.

(120) Donaire, A.; Salgado, J.; Moratal, J.-M. Determination of the Magnetic Axes of Cobalt(II) and Nickel(II) Azurins from 1 H NMR Data: Influence of the Metal and Axial Ligands on the Origin of Magnetic Anisotropy in Blue Copper Proteins †. *Biochemistry* **1998**, *37* (24), 8659–8673. https://doi.org/10.1021/bi971974f.

(121) Spronk, C. A. E. M.; Żerko, S.; Michał Górka, ·; Koźmiński, W.; Bardiaux, B.; Zambelli, B.; Musiani, F.; Piccioli, M.; Priyanka Basak, ·; Faith, ·; Blum, C.; Johnson, R. C.; Hu, H.; Merrell, · D Scott; Maroney, M.; Ciurli, S. Structure and Dynamics of Helicobacter Pylori Nickel-Chaperone HypA: An Integrated Approach Using NMR Spectroscopy, Functional Assays and Computational Tools. *JBIC J. Biol. Inorg. Chem.* **2018**, *23* (3), 1309–1330. https://doi.org/10.1007/s00775-018-1616-y.

(122) Merloni, A.; Dobrovolska, O.; Zambelli, B.; Agostini, F.; Bazzani, M.; Musiani, F.; Ciurli, S. Molecular Landscape of the Interaction between the Urease Accessory Proteins UreE and UreG. *Biochim. Biophys. Acta - Proteins Proteomics* **2014**, *1844* (9), 1662–1674. https://doi.org/10.1016/J.BBAPAP.2014.06.016.

(123) Alfano, M.; Veronesi, G.; Musiani, F.; Zambelli, B.; Signor, L.; Proux, O.; Rovezzi, M.; Ciurli, S.; Cavazza, C. A Solvent-Exposed Cysteine Forms a Peculiar NiII-Binding Site in the Metallochaperone CooT from Rhodospirillum Rubrum. *Chem. - A Eur. J.* **2019**, *25* (67), 15351–15360. https://doi.org/10.1002/CHEM.201903492.

(124) Kabsch, W. XDS. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66* (2), 125–132. https://doi.org/10.1107/S0907444909047337.

(125) Vonrhein, C.; Blanc, E.; Roversi, P.; Bricogne, G. Automated Structure Solution With

AutoSHARP. In *Macromolecular Crystallography Protocols, Volume 2*; Humana Press: New Jersey; Vol. 364, pp 215–230. https://doi.org/10.1385/1-59745-266-1:215.

(126) Cowtan, K. The Buccaneer Software for Automated Model Building. 1. Tracing Protein Chains. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2006**, *62* (9), 1002–1011. https://doi.org/10.1107/S0907444906022116.

(127) Langer, G.; Cohen, S. X.; Lamzin, V. S.; Perrakis, A. Automated Macromolecular Model Building for X-Ray Crystallography Using ARP/WARP Version 7. *Nat. Protoc.* **2008**, *3* (7), 1171–1179. https://doi.org/10.1038/nprot.2008.91.

(128) Murshudov, G. N.; Skubák, P.; Lebedev, A. A.; Pannu, N. S.; Steiner, R. A.; Nicholls, R. A.; Winn, M. D.; Long, F.; Vagin, A. A. REFMAC 5 for the Refinement of Macromolecular Crystal Structures. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2011**, *67* (4), 355–367. https://doi.org/10.1107/S0907444911001314.

(129) Chen, V. B.; Arendall, W. B.; Headd, J. J.; Keedy, D. A.; Immormino, R. M.; Kapral, G. J.; Murray, L. W.; Richardson, J. S.; Richardson, D. C. MolProbity: All-Atom Structure Validation for Macromolecular Crystallography. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66* (1), 12–21. https://doi.org/10.1107/S0907444909042073.

(130) Schuetz, A.; Wasmer, C.; Habenstein, B.; Verel, R.; Greenwald, J.; Riek, R.; Böckmann, A.; Meier, B. H. Protocols for the Sequential Solid-State NMR Spectroscopic Assignment of a Uniformly Labeled 25 KDa Protein: HET-s(1-227). *ChemBioChem* **2010**, *11* (11), 1543–1551. https://doi.org/10.1002/cbic.201000124.

(131) Thakur, R. S.; Kurur, N. D.; Madhu, P. K. Swept-Frequency Two-Pulse Phase Modulation for Heteronuclear Dipolar Decoupling in Solid-State NMR. *Chem. Phys. Lett.* **2006**, *426* (4–6), 459–463. https://doi.org/10.1016/j.cplett.2006.06.007.

(132) Knight, M. J.; Pell, A. J.; Bertini, I.; Felli, I. C.; Gonnelli, L.; Pierattelli, R.; Hermann, T.; Emsley, L.; Pintacuda, G. Structure and Backbone Dynamics of a Microcrystalline Metalloprotein by Solid-State NMR. *Proc. Natl. Acad. Sci. U.S.A 109*, 10811095–11100.

(133) Zhou, D. H.; Shah, G.; Cormos, M.; Mullen, C.; Sandoz, D.; Rienstra, C. M. Protondetected Solid-State NMR Spectroscopy of Fully Protonated Proteins at 40 KHz Magicangle Spinning. *J. Am. Chem. Soc 129*, 11791–11801. https://doi.org/10.1021/ja073462m.

(134) Laage, S.; Sachleben, J.; Steuernagel, S.; Pierattelli, R.; Pintacuda, G.; Emsley, L. Fast Acquisition of Multi-Dimensional Spectra in Solid-State NMR Enabled by Ultra-Fast MAS. *J. Magn. Reson 196*, 133–141.

(135) Robbins, A. H.; Domsic, J. F.; Agbandje-Mckenna, M.; McKenna, R. Emerging from Pseudo-Symmetry: The Redetermination of Human Carbonic Anhydrase II in Monoclinic P21 with a Doubled a Axis. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66* (8), 950–952. https://doi.org/10.1107/S0907444910023723.

(136) Krissinel, E.; Henrick, K. Inference of Macromolecular Assemblies from Crystalline State. *J. Mol. Biol.* **2007**, *372* (3), 774–797. https://doi.org/10.1016/j.jmb.2007.05.022.

(137) Håkansson, K.; Carlsson, M.; Svensson, L. A.; Liljas, A. Structure of Native and Apo Carbonic Anhydrase II and Structure of Some of Its Anion-Ligand Complexes. *J. Mol. Biol.* **1992**, *227* (4),

1192–1204. https://doi.org/10.1016/0022-2836(92)90531-N.

(138) Vellieux, F. M. D.; Dijkstra, B. W. Computation of Bhat's OMIT Maps with Different Coefficients. *J. Appl. Crystallogr.* **1997**, *30* (3), 396–399. https://doi.org/10.1107/S0021889896012551.

(139) Bertarello, A.; Schubeis, T.; Fuccio, C.; Ravera, E.; Fragai, M.; Parigi, G.; Emsley, L.; Pintacuda, G.; Luchinat, C. Paramagnetic Properties of a Crystalline Iron-Sulfur Protein by Magic-Angle Spinning NMR Spectroscopy. *Inorg. Chem.* **2017**, *56* (11), 6624–6629. https://doi.org/10.1021/acs.inorgchem.7b00674.

(140) Bertini, I.; Borghi, E.; Luchinat, C. Characterization of Nickel(II) Bovine Carbonic Anhydrase and Its Inhibitor Derivatives. *Bioinorg. Chem.* **1978**, *9* (6), 495–504. https://doi.org/10.1016/S0006-3061(00)80133-7.

(141) Silva, J. M.; Giuntini, S.; Cerofolini, L.; Geraldes, C. F. G. C.; Macedo, A. L.; Ravera, E.; Fragai, M.; Luchinat, C.; Calderone, V. Non-Crystallographic Symmetry in Proteins: Jahn–Teller-like and Butterfly-like Effects? *JBIC J. Biol. Inorg. Chem.* **2019**, *24* (1), 91–101. https://doi.org/10.1007/s00775-018-1630-0.

(142) Randalí, D. W.; Gamelin, D. R.; Lacroixédward, L. B.; Solomon, L. I. Electronic Structure Contributions to Electron Transfer in Blue Cu and Cu A. *JBIC* **2000**, *5*, 16–19.

(143) Bertini, I.; Pierattelli, R. Copper(II) Proteins Are Amenable for NMR Investigations*. *Pure Appl. Chem* **2004**, *76* (2), 321–333.

(144) Kalverda, A. P.; Salgado, J.; Dennison, C.; Canters, G. W. Analysis of the Paramagnetic Copper(II) Site of Amicyanin by 1 H NMR Spectroscopy †. *Biochemistry* **1996**, *35* (9), 3085–3092. https://doi.org/10.1021/bi9518508.

(145) Ubbink, M.; Lian, L. Y.; Modi, S.; Evans, P. A.; Bendall, D. S. Analysis of the 1H-NMR Chemical Shifts of Cu(I)-, Cu(II)- and Cd-Substituted Pea Plastocyanin. Metal-Dependent Differences in the Hydrogen-Bond Network around the Copper Site. *Eur. J. Biochem.* **1996**, *242* (1), 132–147. https://doi.org/10.1111/j.1432-1033.1996.0132r.x.

(146) Vila, A. J.; Ramirez, B. E.; Bilio, A. J. Di; Mizoguchi, T. J.; Richards, J. H.; Gray, H. B. Paramagnetic NMR Spectroscopy of Cobalt(II) and Copper(II) Derivatives of Pseudomonas Aeruginosa His46Asp Azurin. **1997**.

(147) Bertini, I.; Ciurli, S.; Dikiy, A.; Gasanov, R.; Luchinat, C.; Martini, G.; Safarov, N. High-Field NMR Studies of Oxidized Blue Copper Proteins: The Case of Spinach Plastocyanin. **1999**. https://doi.org/10.1021/ja983833m.

(148) Bertini, I.; Ciurli, S.; Dikiy, A.; Fernàndez, C. O.; Luchinat, C.; Safarov, N.; Shumilin, S.; Vila, A. J. The First Solution Structure of a Paramagnetic Copper(II) Protein: The Case of Oxidized Plastocyanin from the Cyanobacterium Synechocystis PCC6803. **2001**. https://doi.org/10.1021/ja0033685.

(149) Arnesano, F.; Banci, L.; Bertini, I.; Felli, I. C.; Luchinat, C.; Thompsett, A. R. A Strategy for the NMR Characterization of Type II Copper(II) Proteins: The Case of the Copper Trafficking Protein CopC from Pseudomonas Syringae. **2003**. https://doi.org/10.1021/ja034112c.

(150) Marshall, W. *In Paramagnetic Resonance*; Low, W., Ed.; Academic Press: New York.

(151) Arbuznikov, A. V.; Vaara, J.; Kaupp, M. Relativistic Spin-Orbit Effects on Hyperfine Coupling

Tensors by Density-Functional Theory. *J. Chem. Phys.* **2004**, *120* (5), 2127–2139. https://doi.org/10.1063/1.1636720.

(152) Pennanen, T. O.; Vaara, J. Nuclear Magnetic Resonance Chemical Shift in an Arbitrary Electronic Spin State. *Phys. Rev. Lett.* **2008**, *100* (13), 133002–4. https://doi.org/10.1103/PhysRevLett.100.133002.

(153) Van Den Heuvel, W.; Soncini, A. NMR Chemical Shift as Analytical Derivative of the Helmholtz Free Energy. *J. Chem. Phys.* **2013**, *138* (5), 54113. https://doi.org/10.1063/1.4789398.

(154) Martin, B.; Autschbach, J. Temperature Dependence of Contact and Dipolar NMR Chemical Shifts in Paramagnetic Molecules. *J. Chem. Phys.* **2015**, *142* (5), 54108. https://doi.org/10.1063/1.4906318.

(155) Vaara, J.; Awais Rouf, S.; Mares, í. Magnetic Couplings in the Chemical Shift of Paramagnetic NMR. **2015**. https://doi.org/10.1021/acs.jctc.5b00656.

(156) Neese, F. Quantum Chemistry and EPR Parameters. *eMagRes* **2017**, *6* (1), 1–22. https://doi.org/10.1002/9780470034590.emrstm1505.

(157) Mondal, A.; Gaultois, M. W.; Pell, A. J.; Iannuzzi, M.; Grey, C. P.; Hutter, J.; Kaupp, M. Large-Scale Computation of Nuclear Magnetic Resonance Shifts for Paramagnetic Solids Using CP2K. *J. Chem. Theory Comput.* **2018**, *14* (1), 377–394. https://doi.org/10.1021/acs.jctc.7b00991.

(158) Prosser, R. S.; Volkov, V. B.; Shiyanovskaya, I. V. Novel Chelate-Induced Magnetic Alignment of Biological Membranes. *Biophys. J.* **1998**, *75* (5), 2163–2169. https://doi.org/10.1016/S0006-3495(98)77659-3.

(159) Prosser, R. S.; Volkov, V. B.; Shiyanovskaya, I. V. Solid-State NMR Studies of Magnetically Aligned Phospholipid Membranes: Taming Lanthanides for Membrane Protein Studies. *Biochem. Cell Biol.* **1998**, *76* (2–3), 443–451. https://doi.org/10.1139/o98-058.

(160) Ye, L.; Van Eps, N.; Li, X.; Ernst, O. P.; Prosser, R. S. Utilizing Tagged Paramagnetic Shift Reagents to Monitor Protein Dynamics by NMR. *Biochim. Biophys. Acta - Proteins Proteomics* **2017**, *1865* (11), 1555–1563. https://doi.org/10.1016/j.bbapap.2017.09.011.

(161) Parigi, G.; Benda, L.; Ravera, E.; Romanelli, M.; Luchinat, C. Pseudocontact Shifts and Paramagnetic Susceptibility in Semiempirical and Quantum Chemistry Theories. *J. Chem. Phys.* **2019**, *150* (14). https://doi.org/10.1063/1.5037428.

(162) Mareš, J.; Vaara, J. Ab Initio Paramagnetic NMR Shifts via Point-Dipole Approximation in a Large Magnetic-Anisotropy Co(II) Complex. *Phys. Chem. Chem. Phys.* **2018**, *20* (35), 22547–22555. https://doi.org/10.1039/C8CP04123G.

(163) Chyba, J.; Novák, M.; Munzarová, P.; Novotný, J.; Marek, R. Through-Space Paramagnetic NMR Effects in Host-Guest Complexes: Potential Ruthenium(III) Metallodrugs with Macrocyclic Carriers. *Inorg. Chem.* **2018**, *57* (15), 8735–8747. https://doi.org/10.1021/acs.inorgchem.7b03233.

(164) Srb, P.; Svoboda, M.; Benda, L.; Lepšík, M.; Tarábek, J.; Šícha, V.; Grüner, B.; Grantz-Šašková, K.; Brynda, J.; Řezáčová, P.; Konvalinka, J.; Veverka, V. Capturing a Dynamically Interacting Inhibitor by Paramagnetic NMR Spectroscopy. *Phys. Chem. Chem. Phys.* **2019**, *21* (10), 5661–5673. https://doi.org/10.1039/c9cp00416e.

(165) Ravera, E.; Takis, P. G.; Fragai, M.; Parigi, G.; Luchinat, C. NMR Spectroscopy and Metal Ions in Life Sciences. *Eur. J. Inorg. Chem.* **2018**, *2018* (44), 4752–4770. https://doi.org/10.1002/ejic.201800875.

(166) Navon, G.; Shani, J.; Panige1, R.; Schoenberglc, S. Application of the Carbonic Anhydrase Inhibitory Effect of Furosemide to the Study of Furosemide Release from Two of Its Diuretic Derivatives. *J. Med. Chem.* **1975**, *18* (1).

(167) Rivière-Baudet, M.; Supuran, C. T.; Scozzafava, A.; Briganti, F.; Baz, F. El; Maarouf, Z. Ben; Riviere, P. INTERACTION OF ISOZYMES I AND II OF CARBONIC ANHYDRASE WITH Ge(IV) AND Sb(Lll) DERIVATIVES. *Main Gr. Met. Chem.* **1997**, *20* (10), 641–648. https://doi.org/10.1515/MGMC.1997.20.10.641.

(168) Bertini, I.; Luchinat, C.; Scozzafava, A. Binding Affinity of Bicarboxylate Ions for Cobalt(II) Bovine Carbonic Anhydrase. *Bioinorg. Chem.* **1978**, *9* (2), 93–100. https://doi.org/10.1016/S0006-3061(00)80283-5.

(169) Honndorf, V. S.; Heine, A.; Klebe, G.; Supuran, C. T. RCSB PDB - 1Z9Y: carbonic anhydrase II in complex with furosemide as sulfonamide inhibitor https://www.rcsb.org/structure/1z9y (accessed Aug 21, 2021). https://doi.org/10.2210/pdb1z9y/pdb.

(170) Bertini, I.; Canti, G.; Luchinat, C.; Scozzafava, A. Spectroscopic Investigation of Copper(II) Bovine Carbonic Anhydrase and Its Inhibitor Derivatives. *J. Chem. Soc. Dalt. Trans.* **1978**, No. 10, 1269. https://doi.org/10.1039/dt9780001269.

(171) Bertini, I.; Luchinat, C.; Monnanni, R.; Scozzafava, A. Different Behavior of Sulfonamides with Respect to Copper-Substituted Bovine and Human Carbonic Anhydrases. *J. Inorg. Biochem.* **1982**, *16* (2), 155–160. https://doi.org/10.1016/S0162-0134(00)80223-8.

(172) Morpurgo, L.; Desideri, A.; Rigo, A.; Viglino, P.; Rotilio, G. Reaction of N,N-Diethyldithiocarbamate and Other Bidentate Ligands with Zn, Co and Cu Bovine Carbonic Anhydrases. Inhibition of the Enzyme Activity and Evidence for Stable Ternary Enzyme-Metal-Ligand Complexes. *Biochim. Biophys. Acta (BBA)/Protein Struct. Mol.* **1983**, *746* (3), 168–175. https://doi.org/10.1016/0167-4838(83)90071-7.

(173) Moore, J. M.; Lepre, C. A.; Gippert, G. P.; Chazin, W. J.; Case, D. A.; Wright, P. E. High-Resolution Solution Structure of Reduced French Bean Plastocyanin and Comparison with the Crystal Structure of Poplar Plastocyanin. *I. Mol. Bid* **1991**, *221*, 533–555.

(174) Schmidt, L.; Christensen, H. E. M.; Harris, P. Structure of Plastocyanin from the Cyanobacterium Anabaena Variabilis. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2006**, *62* (9), 1022–1029. https://doi.org/10.1107/S0907444906023638.

(175) Penfield, K. W.; Gewirth, A. A.; Solomon, E. I. Electronic Structure and Bonding of the Blue Copper Site in Plastocyanin. *J. Am. Chem. Soc.* **1985**, *107* (15), 4519–4529. https://doi.org/10.1021/ja00301a024.

(176) Boča, R. Theoretical Foundations of Molecular Magnetism. *Coordination Chemistry Reviews*. 2002, p 83. https://doi.org/10.1016/s0010-8545(02)00003-6.

(177) Walder, B. J.; Patterson, A. M.; Baltisberger, J. H.; Grandinetti, P. J. Hydrogen Motional Disorder in Crystalline Iron Group Chloride Dihydrates. *J. Chem. Phys.* **2018**, *149* (8), 84503.

https://doi.org/10.1063/1.5037151.

(178) Vleck, J. H. Van. The Theory of Electric and Magnetic Susceptibilities. *Math. Gaz.* **1934**, *18* (231), 328. https://doi.org/10.2307/3605487.

(179) Ravera, E.; Parigi, G.; Luchinat, C. What Are the Methodological and Theoretical Prospects for Paramagnetic NMR in Structural Biology? A Glimpse into the Crystal Ball. *J. Magn. Reson.* **2019**, *306*, 173–179. https://doi.org/10.1016/j.jmr.2019.07.027.

(180) Lang, L.; Ravera, E.; Parigi, G.; Luchinat, C.; Neese, F. Solution of a Puzzle: High-Level Quantum-Chemical Treatment of Pseudocontact Chemical Shifts Confirms Classic Semiempirical Theory. *J. Phys. Chem. Lett.* **2020**, *11* (20), 8735–8744. https://doi.org/10.1021/acs.jpclett.0c02462.

(181) Ravera, E.; Gigli, L.; Czarniecki, B.; Lang, L.; Kümmerle, R.; Parigi, G.; Piccioli, M.; Neese, F.; Luchinat, C. A Quantum Chemistry View on Two Archetypical Paramagnetic Pentacoordinate Nickel(II) Complexes Offers a Fresh Look on Their NMR Spectra. *Inorg. Chem.* **2021**, *60* (3), 2068–2075. https://doi.org/10.1021/acs.inorgchem.0c03635.

(182) Carlon, A.; Ravera, E.; Andrałojć, W.; Parigi, G.; Murshudov, G. N. G. N.; Luchinat, C. How to Tackle Protein Structural Data from Solution and Solid State: An Integrated Approach. *Prog. Nucl. Magn. Reson. Spectrosc.* **2016**, *92–93*, 54–70. https://doi.org/10.1016/j.pnmrs.2016.01.001.

8

Appendix A: Cobalt Chapter

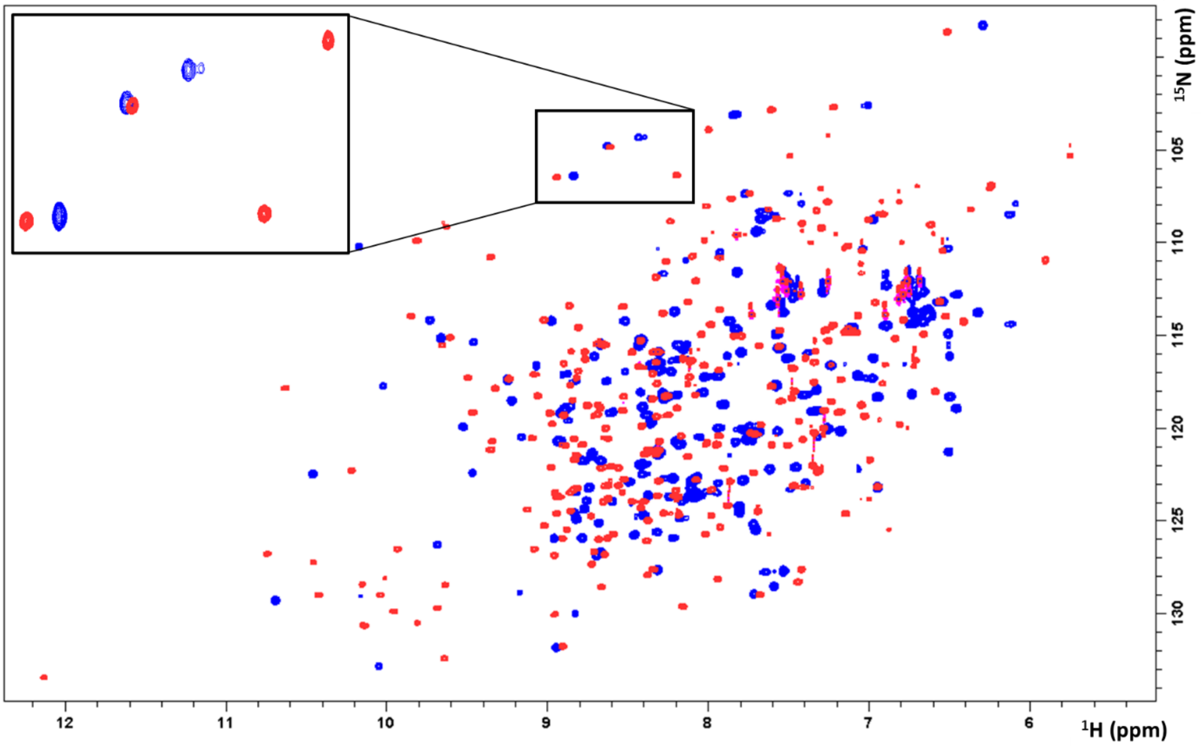


Figure 8.1: ^{1}H-^{15}N HSQC spectra of cobalt(II)-DM-hCAII (red) and zinc(II)-DM-hCAII (blue) in the presence of 1000 µM sodium thiocyanate, in 10 mM HEPES, pH 6.3, 1.2 GHz. Protein concentration is 500 µM.

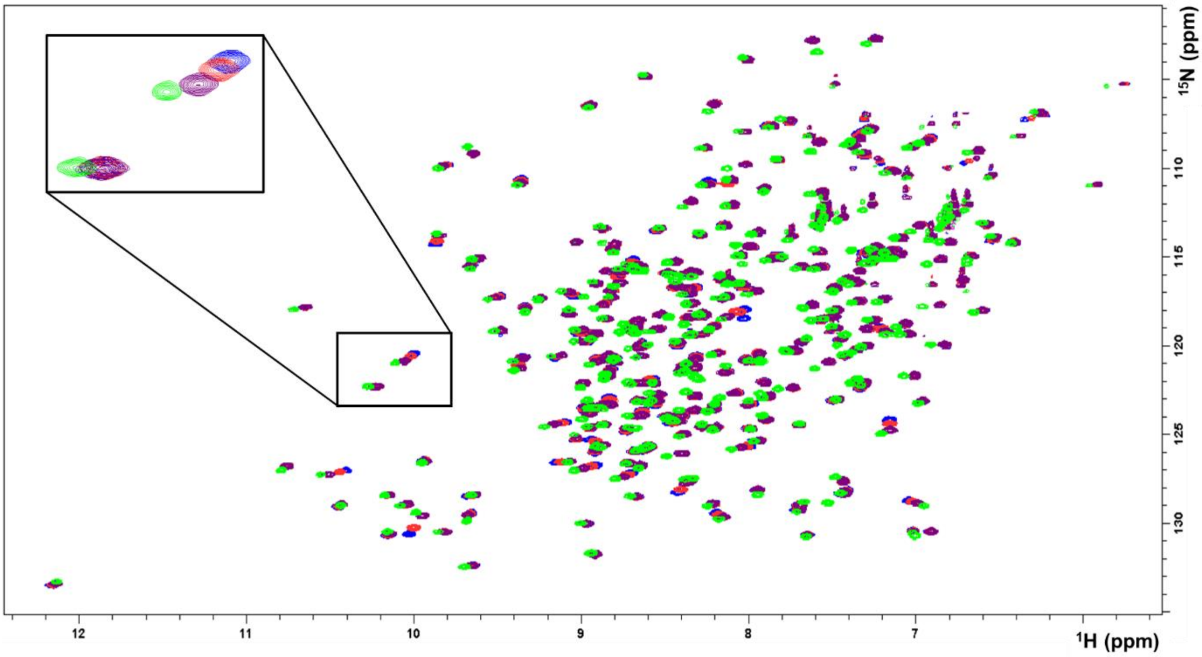


Figure 8.2: ^{1}H-^{15}N HSQC spectra of zinc(II)-DM-hCAII titration with sodium thiocyanate, in 10 mM HEPES, pH 6.3. (Blue - free enzyme; Red - protein:ligand ratio 1:1.3; purple - protein:ligand ratio 1:13; green – protein:ligand ratio 1:1357). Spectra were recorded at 500 MHz. Protein concentration is 350 µM.

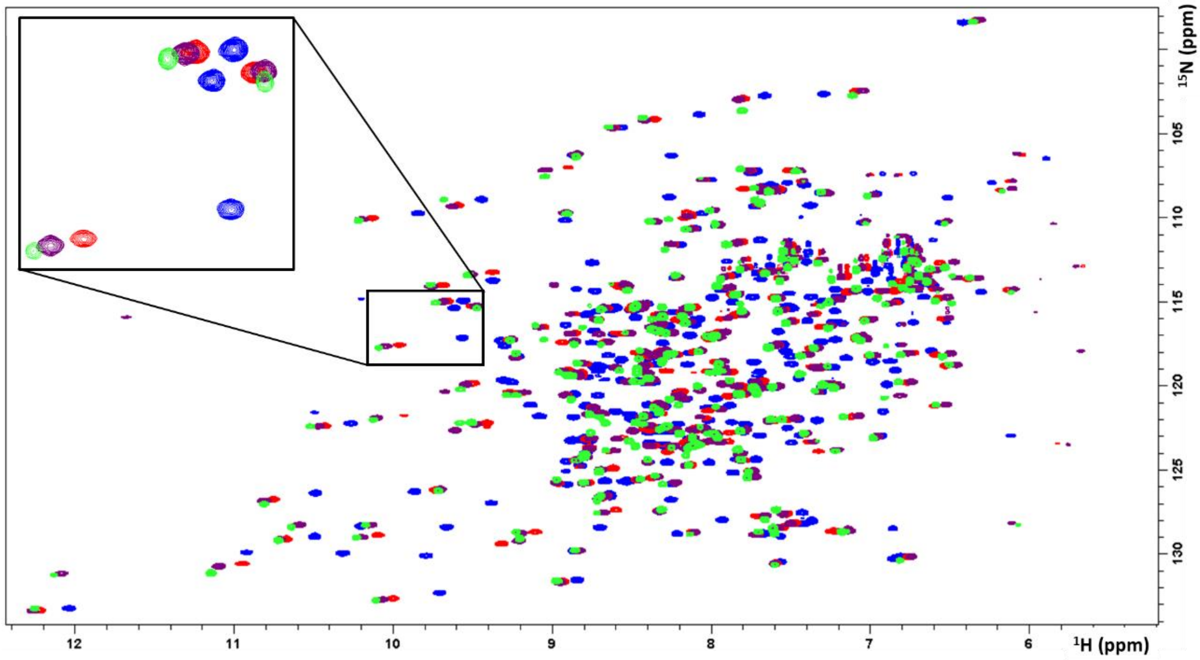


Figure 8.3: ^{1}H-^{15}N HSQC spectra of cobalt(II)-DM-hCAII titration with sodium thiocyanate, in 10 mM HEPES, pH 6.3. (Blue - free enzyme; Red - protein:ligand ratio 1:1.3; purple - protein:ligand ratio 1:13; green – protein:ligand ratio 1:1357). Spectra were recorded at 500 MHz. Protein concentration is 350 µM.

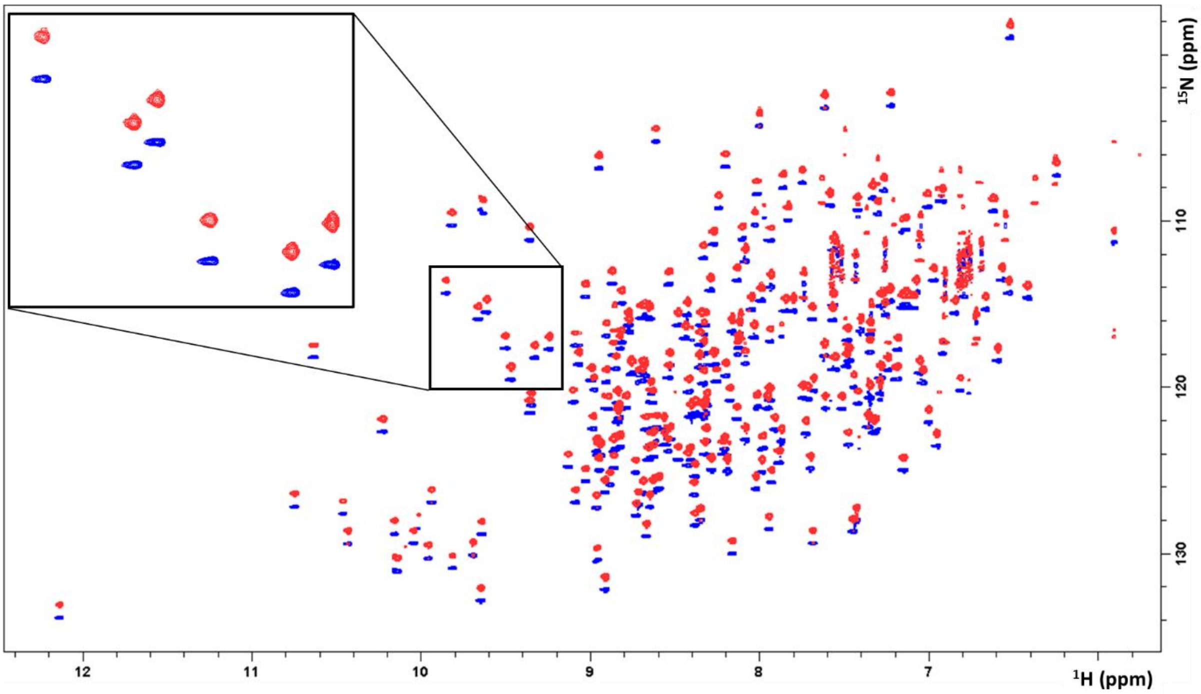


Figure 8.4: Downfield (blue) and upfield (red) doublet components of ^{1}H-^{15}N-HSQC-IPAP spectrum of zinc(II)-DM-hCAII in the presence of 1000 µM sodium thiocyanate, in 10 mM HEPES, pH 6.3, recorded 1.2 GHz. Protein concentration is 500 µM.

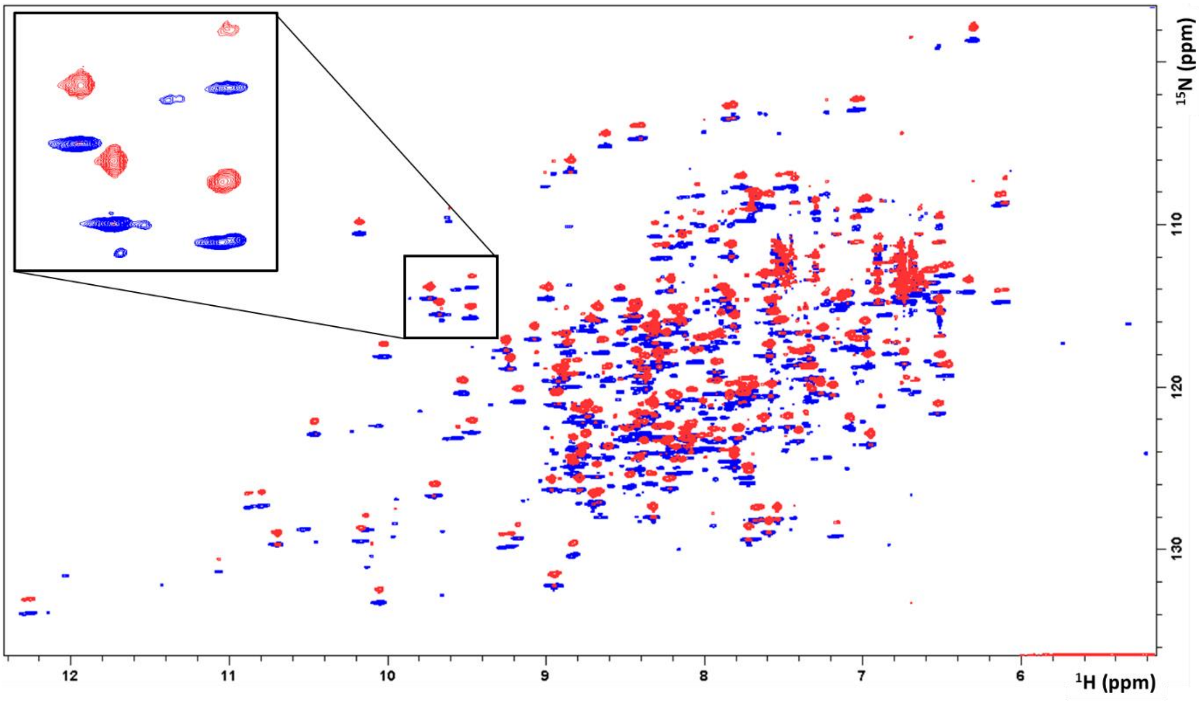


Figure 8.5: Downfield (blue) and upfield (red) doublet components of ^{1}H-^{15}N-HSQC-IPAP spectrum of cobalt(II)-DM-hCAII) in the presence of 1000 µM sodium thiocyanate, in 10 mM HEPES, pH 6.3, recorded at the 1.2GHz Protein concentration is 500 µM.

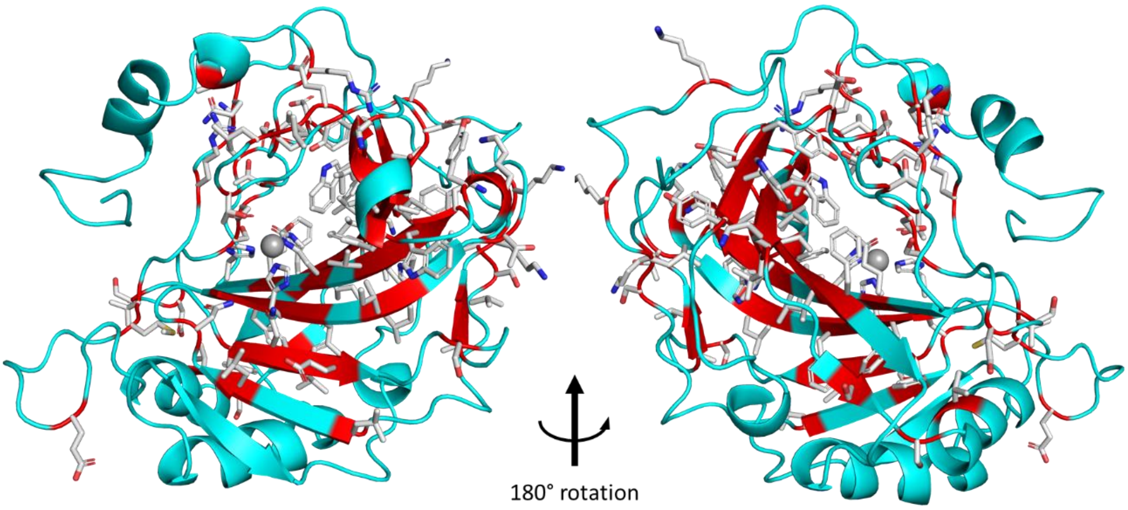


Figure 8.6: Cartoon representation of zinc(II)-DM-hCAII with a protein:ligand ratio of 1:1.3. The residues in red are the most affected by the ligand.

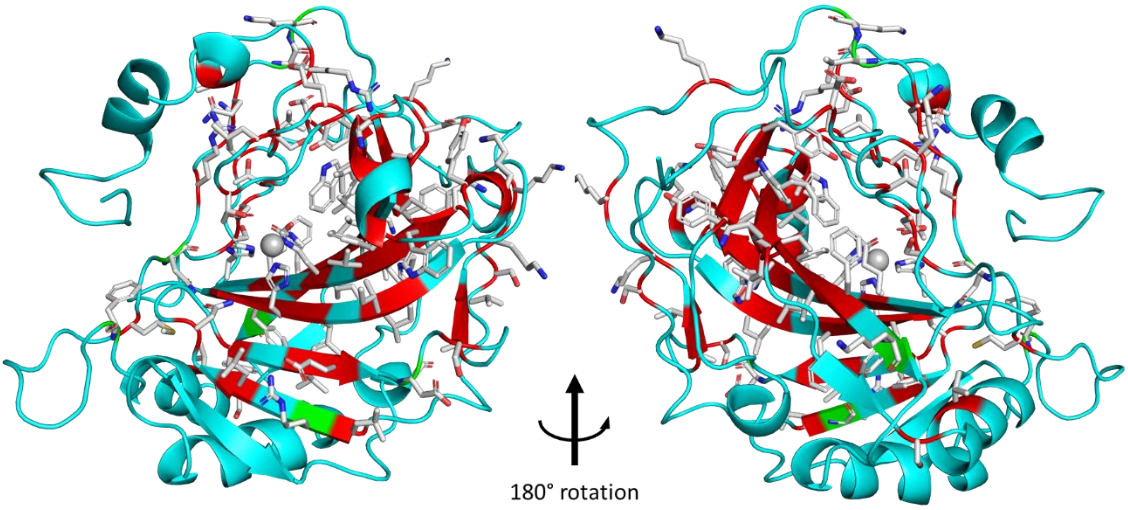


Figure 8.7: Cartoon representation of zinc(II)-DM-hCAII with a protein:ligand ratio of 1:2.7. The residues in red are the most affected by the ligand from the previous addition and are maintained. The residues in light green are the new interactions with sodium thiocyanate.

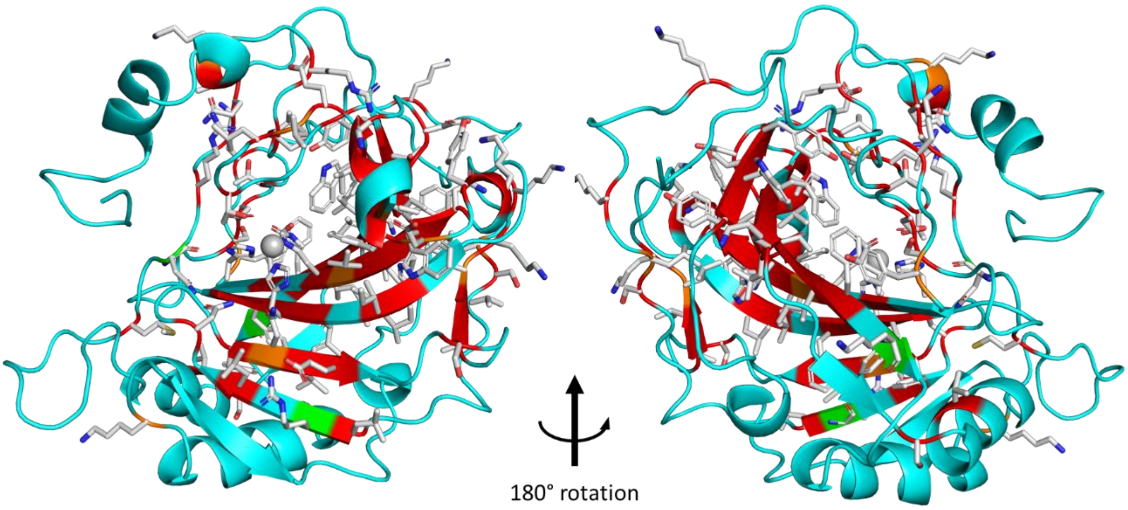


Figure 8.8: Cartoon representation of zinc(II)-DM-hCAII with a protein:ligand ratio of 1:13. The residues in red and green are the most affected by the ligand from the previous additions and are maintained. The residues in orange are the new interactions with sodium thiocyanate.

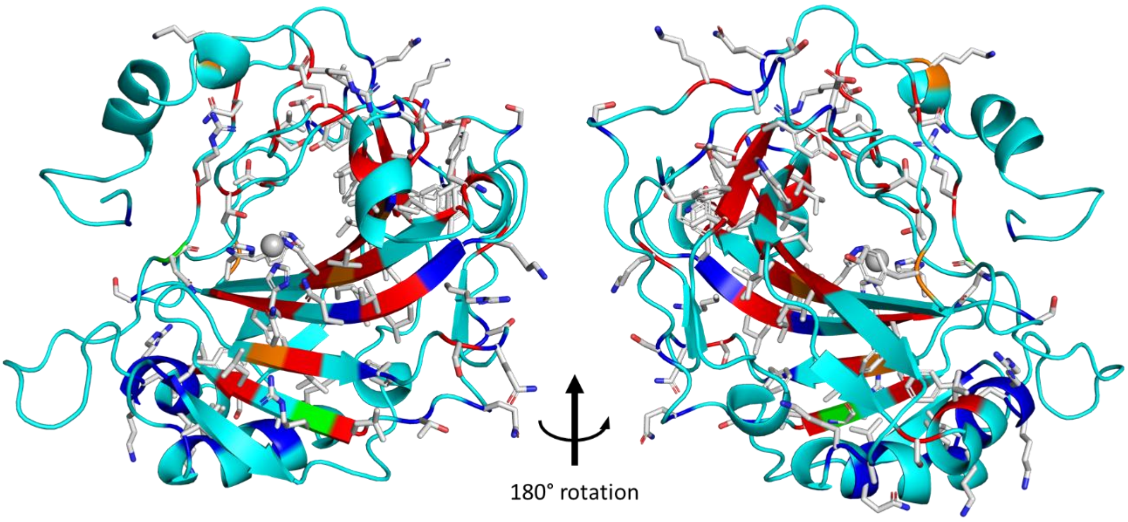


Figure 8.9: Cartoon representation of zinc(II)-DM-hCAII with a protein:ligand ratio of 1:1357. The residues in red, green and orange are the most affected by the ligand from the previous additions and are maintained. The residues in blue are the new interactions with the sodium thiocyanate.

Table 8.1: PCS values (in ppm) used in the plot regarding the free cobalt(II)-DM-hCAII and cobalt(II)-DM-hCAII-thiocyanate 1:1.3.

Free cobalt(II)-DM-hCAII				**Cobalt(II)-DM-hCAII-thiocyanate 1:1.3**			
Residue #			Measured PCS	Residue #			Measured PCS
6	G	H	0,04	6	G	H	0,112
8	G	H	-0,137	8	G	H	-0,12
24	K	H	0,03	24	K	H	-0,329
25	G	H	0,006	25	G	H	-0,459
26	E	H	-0,041	26	E	H	-0,399
27	R	H	-0,086	31	V	H	-0,201
28	Q	H	-0,174	33	I	H	0,433
31	V	H	-0,347	34	D	H	0,129
32	D	H	-0,149	35	T	H	0,199
33	I	H	0,029	37	T	H	0,075
34	D	H	-0,016	38	A	H	0,091
35	T	H	0,03	39	K	H	0,025
37	T	H	-0,006	40	Y	H	0,045
38	A	H	-0,005	41	D	H	0,024
39	K	H	-0,027	43	S	H	-0,011
40	Y	H	-0,01	44	L	H	-0,008
41	D	H	-0,019	45	K	H	-0,075
43	S	H	-0,029	47	L	H	0,044
44	L	H	-0,033	48	S	H	-0,055
45	K	H	-0,054	49	V	H	-0,024
47	L	H	-0,015	50	S	H	-0,085
48	S	H	-0,063	51	Y	H	-0,092
49	V	H	-0,03	52	D	H	-0,094
50	S	H	-0,076	53	Q	H	-0,094
51	Y	H	-0,037	54	A	H	-0,141
52	D	H	-0,052	55	T	H	-0,165
53	Q	H	-0,049	56	S	H	-0,208
54	A	H	-0,059	57	L	H	-0,199
55	T	H	-0,084	58	R	H	-0,206
56	S	H	-0,076	61	N	H	-0,206
57	L	H	-0,108	62	N	H	-0,09
58	R	H	-0,12	63	G	H	-0,055
59	I	H	-0,124	69	E	H	-0,369
60	L	H	-0,248	70	F	H	-0,131
61	N	H	-0,173	71	D	H	-0,114
62	N	H	-0,286	72	D	H	0,011
63	G	H	-0,207	73	S	H	0,009
69	E	H	-0,207	74	Q	H	-0,037
70	F	H	-0,192	76	K	H	-0,077
72	D	H	-0,085	77	A	H	-0,122

74	Q	H	-0,068	78	V	H	-0,169
76	K	H	-0,079	79	L	H	-0,174
77	A	H	-0,106	80	K	H	-0,123
78	V	H	-0,119	81	G	H	-0,108
80	K	H	-0,108	82	G	H	-0,042
81	G	H	-0,094	84	L	H	-0,137
82	G	H	-0,063	86	G	H	-0,12
86	G	H	-0,094	87	T	H	-0,121
87	T	H	-0,104	88	Y	H	-0,173
88	Y	H	-0,161	89	R	H	-0,146
89	R	H	-0,199	99	S	H	-0,101
99	S	H	-0,156	100	L	H	0,027
100	L	H	-0,052	101	D	H	0,082
101	D	H	0,034	102	G	H	0,177
102	G	H	0,043	103	Q	H	0,298
103	Q	H	0,057	104	G	H	0,698
104	G	H	0,05	109	V	H	0,713
105	S	H	0,352	110	D	H	0,345
108	T	H	0,067	111	K	H	0,313
109	V	H	0,166	112	K	H	0,421
110	D	H	0,053	114	Y	H	1,274
111	K	H	0,031	115	A	H	0,796
112	K	H	0,082	116	A	H	1,281
113	K	H	0,107	123	W	H	-0,211
114	Y	H	0,379	124	N	H	-0,158
115	A	H	0,364	126	K	H	-0,095
116	A	H	0,656	127	Y	H	-0,016
123	W	H	-0,296	128	G	H	0,001
124	N	H	-0,19	129	D	H	0,074
125	T	H	-0,124	130	F	H	0,218
126	K	H	-0,086	131	G	H	0,348
127	Y	H	-0,062	132	K	H	0,21
128	G	H	-0,037	133	A	H	0,132
129	D	H	-0,019	134	V	H	0,227
130	F	H	-0,013	135	Q	H	0,085
131	G	H	0,069	136	Q	H	-0,082
132	K	H	0,044	138	D	H	-0,243
133	A	H	0,009	146	F	H	1,833
134	V	H	0,072	147	L	H	1,749
135	Q	H	0,103	148	K	H	0,577
136	Q	H	0,005	149	V	H	0,334
138	D	H	-0,096	150	G	H	0,155
147	L	H	0,957	151	S	H	0,121

148	K	H	0,382	152	A	H	0,015
149	V	H	0,198	153	K	H	-0,023
150	G	H	0,143	156	L	H	-0,175
151	S	H	0,113	157	Q	H	-0,206
152	A	H	0,078	161	D	H	-0,319
153	K	H	0,121	162	V	H	-0,365
156	L	H	0,093	163	L	H	-0,416
157	Q	H	0,048	164	D	H	-0,317
159	V	H	-0,009	165	S	H	-0,277
161	D	H	-0,028	166	I	H	-0,325
162	V	H	-0,065	167	K	H	-0,255
163	L	H	-0,092	168	T	H	-0,202
164	D	H	-0,086	169	K	H	-0,105
165	S	H	-0,088	170	G	H	0,038
166	I	H	-0,122	171	K	H	-0,058
167	K	H	-0,113	172	S	H	-0,078
168	T	H	-0,117	173	A	H	-0,193
169	K	H	-0,12	174	D	H	-0,131
170	G	H	-0,079	175	F	H	-0,181
172	S	H	-0,068	176	T	H	-0,19
174	D	H	-0,059	177	N	H	-0,163
175	F	H	-0,073	178	F	H	-0,237
176	T	H	-0,058	179	D	H	-0,239
177	N	H	-0,052	182	G	H	-0,15
178	F	H	-0,055	183	L	H	-0,11
179	D	H	-0,019	184	L	H	0,031
181	R	H	0,041	186	E	H	0,23
182	G	H	0,049	187	S	H	0,248
183	L	H	0,124	188	L	H	0,179
184	L	H	0,158	189	D	H	0,241
186	E	H	0,151	190	Y	H	0,196
187	S	H	0,13	191	W	H	0,456
188	L	H	0,075	192	T	H	0,036
189	D	H	0,091	193	Y	H	-0,289
190	Y	H	0,026	207	T	H	-1,062
191	W	H	0,068	208	W	H	-0,953
192	T	H	-0,073	210	V	H	0,434
193	Y	H	-0,253	212	K	H	0,579
207	T	H	-0,555	213	E	H	0,567
209	I	H	-0,456	215	I	H	0,81
210	V	H	-0,013	216	S	H	0,273
211	L	H	0,466	217	V	H	0,23
212	K	H	0,189	218	S	H	-0,008

213	E	H	0,231	219	S	H	0,02
215	I	H	0,5	221	Q	H	-0,109
217	V	H	0,228	222	V	H	-0,149
218	S	H	0,134	223	L	H	-0,173
219	S	H	0,073	224	K	H	-0,288
221	Q	H	0,053	225	F	H	-0,477
222	V	H	0,069	226	R	H	-0,449
223	L	H	0,012	227	K	H	-0,416
224	K	H	-0,028	228	L	H	-0,519
225	F	H	-0,073	229	N	H	-0,387
226	R	H	-0,12	230	F	H	-0,288
227	K	H	-0,184	231	N	H	-0,154
228	L	H	-0,205	232	G	H	-0,004
229	N	H	-0,209	233	E	H	-0,018
230	F	H	-0,209	234	G	H	-0,04
231	N	H	-0,138	235	E	H	-0,031
232	G	H	-0,053	237	E	H	-0,082
233	E	H	-0,048	238	E	H	-0,133
234	G	H	-0,04	239	L	H	-0,214
235	E	H	-0,043	241	V	H	-0,537
237	E	H	-0,063	242	D	H	-0,497
238	E	H	-0,108	250	L	H	-0,26
240	M	H	-0,277	251	K	H	-0,249
241	V	H	-0,373	253	R	H	-0,249
242	D	H	-0,396	254	Q	H	-0,214
248	Q	H	-0,294	255	I	H	-0,078
251	K	H	-0,088	256	K	H	-0,053
253	R	H	-0,093	257	A	H	0,043
254	Q	H	-0,1	258	S	H	0,141
255	I	H	-0,071	259	F	H	0,128
256	K	H	-0,084	260	K	H	0,098
257	A	H	-0,032				
258	S	H	0,005				
259	F	H	0,014				
260	K	H	0,019				

Table 8.2: PCS values (in ppm) used in the plot regarding the cobalt(II)-DM-hCAII-thiocyanate 1:2 and cobalt(II)-DM-hCAII-thiocyanate 1:13.

Cobalt(II)-DM-hCAII-thiocyanate 1:2				**Cobalt(II)-DM-hCAII-thiocyanate 1:13**			
Residue #			Measured PCS	Residue #			Measured PCS
8	G	H	-0,128	6	G	H	0,131
32	D	H	0,019	8	G	H	-0,117
33	I	H	0,508	24	K	H	-0,423
34	D	H	0,144	25	G	H	-0,565
35	T	H	0,238	26	E	H	-0,479
37	T	H	0,092	33	I	H	0,521
38	A	H	0,112	34	D	H	0,155
39	K	H	0,029	35	T	H	0,24
40	Y	H	0,053	37	T	H	0,096
41	D	H	0,036	38	A	H	0,113
43	S	H	-0,007	39	K	H	0,03
44	L	H	-0,001	40	Y	H	0,058
45	K	H	-0,072	41	D	H	0,035
47	L	H	0,052	43	S	H	-0,006
48	S	H	-0,055	44	L	H	-0,003
49	V	H	-0,027	45	K	H	-0,064
50	S	H	-0,09	47	L	H	0,059
51	Y	H	-0,106	48	S	H	-0,046
53	Q	H	-0,107	49	V	H	-0,021
54	A	H	-0,158	50	S	H	-0,09
55	T	H	-0,178	51	Y	H	-0,107
56	S	H	-0,246	52	D	H	-0,101
57	L	H	-0,215	53	Q	H	-0,11
58	R	H	-0,241	54	A	H	-0,156
59	I	H	-0,298	55	T	H	-0,181
61	N	H	-0,209	56	S	H	-0,236
62	N	H	0,037	63	G	H	-0,019
69	E	H	-0,434	71	D	H	-0,113
70	F	H	-0,119	72	D	H	0,033
71	D	H	-0,115	73	S	H	0,018
72	D	H	0,032	74	Q	H	-0,028
74	Q	H	-0,03	76	K	H	-0,079
76	K	H	-0,074	77	A	H	-0,131
77	A	H	-0,13	78	V	H	-0,166
78	V	H	-0,141	79	L	H	-0,176
80	K	H	-0,105	80	K	H	-0,247
81	G	H	-0,105	81	G	H	-0,106
82	G	H	-0,032	82	G	H	-0,031
84	L	H	-0,14	84	L	H	-0,142
86	G	H	-0,125	86	G	H	-0,13

87	T	H	-0,078	87	T	H	-0,125
88	Y	H	-0,153	88	Y	H	-0,173
89	R	H	-0,106	89	R	H	-0,041
90	L	H	-0,18	99	S	H	-0,087
99	S	H	-0,029	100	L	H	0,033
100	L	H	0,079	101	D	H	0,086
102	G	H	0,236	102	G	H	0,211
103	Q	H	0,358	103	Q	H	0,353
108	T	H	1,008	104	G	H	0,83
109	V	H	0,815	109	V	H	0,838
110	D	H	0,404	110	D	H	0,412
111	K	H	0,356	111	K	H	0,367
112	K	H	0,489	112	K	H	0,495
113	K	H	0,569	114	Y	H	1,467
115	A	H	0,864	123	W	H	-0,179
124	N	H	-0,088	124	N	H	-0,146
125	T	H	-0,084	126	K	H	-0,091
126	K	H	-0,069	127	Y	H	-0,016
127	Y	H	0,009	128	G	H	0,015
128	G	H	0,034	129	D	H	0,101
129	D	H	0,099	131	G	H	0,413
130	F	H	0,31	132	K	H	0,246
131	G	H	0,443	133	A	H	0,167
132	K	H	0,271	134	V	H	0,265
133	A	H	0,207	135	Q	H	0,099
134	V	H	0,255	147	L	H	1,918
135	Q	H	0,133	148	K	H	0,631
136	Q	H	-0,059	149	V	H	0,371
138	D	H	-0,248	150	G	H	0,157
139	G	H	-0,332	151	S	H	0,124
148	K	H	0,622	152	A	H	0,007
149	V	H	0,387	153	K	H	-0,05
150	G	H	0,159	156	L	H	-0,23
151	S	H	0,124	157	Q	H	-0,264
152	A	H	0,01	164	D	H	-0,366
153	K	H	-0,053	165	S	H	-0,322
156	L	H	-0,226	166	I	H	-0,37
157	Q	H	-0,251	167	K	H	-0,295
159	V	H	-0,516	168	T	H	-0,221
160	V	H	-0,587	169	K	H	-0,098
161	D	H	-0,368	171	K	H	-0,054
162	V	H	-0,423	172	S	H	-0,079
163	L	H	-0,477	173	A	H	-0,213

164	D	H	-0,357	174	D	H	-0,145
165	S	H	-0,315	175	F	H	-0,204
166	I	H	-0,361	176	T	H	-0,216
167	K	H	-0,281	177	N	H	-0,192
168	T	H	-0,221	178	F	H	-0,28
169	K	H	-0,097	179	D	H	-0,283
170	G	H	0,059	182	G	H	-0,189
171	K	H	-0,057	183	L	H	-0,168
172	S	H	-0,079	184	L	H	0,011
173	A	H	-0,207	186	E	H	0,249
174	D	H	-0,141	187	S	H	0,275
175	F	H	-0,197	188	L	H	0,203
176	T	H	-0,208	189	D	H	0,278
177	N	H	-0,186	190	Y	H	0,231
178	F	H	-0,293	192	T	H	0,07
179	D	H	-0,3	193	Y	H	-0,322
182	G	H	-0,18	208	W	H	-1,011
183	L	H	-0,199	209	I	H	-0,076
184	L	H	-0,018	210	V	H	0,559
186	E	H	0,193	212	K	H	0,663
187	S	H	0,269	213	E	H	0,647
188	L	H	0,194	215	I	H	0,876
189	D	H	0,266	216	S	H	0,283
190	Y	H	0,248	217	V	H	0,233
191	W	H	0,526	218	S	H	-0,037
192	T	H	0,059	219	S	H	0,008
193	Y	H	-0,305	221	Q	H	-0,142
204	E	H	-0,553	222	V	H	-0,176
210	V	H	0,524	223	L	H	-0,21
212	K	H	0,617	224	K	H	-0,343
213	E	H	0,63	225	F	H	-0,57
215	I	H	0,861	226	R	H	-0,523
216	S	H	0,279	227	K	H	-0,484
217	V	H	0,234	228	L	H	-0,592
218	S	H	-0,017	229	N	H	-0,429
219	S	H	0,009	230	F	H	-0,305
220	E	H	-0,052	231	N	H	-0,157
221	Q	H	-0,134	232	G	H	0,009
222	V	H	-0,178	233	E	H	-0,02
223	L	H	-0,178	234	G	H	-0,041
224	K	H	-0,328	235	E	H	-0,027
225	F	H	-0,546	237	E	H	-0,087
226	R	H	-0,474	238	E	H	-0,141

227	K	H	-0,461	239	L	H	-0,226
228	L	H	-0,576	241	V	H	-0,572
229	N	H	-0,421	242	D	H	-0,526
230	F	H	-0,297	247	A	H	-0,094
231	N	H	-0,158	248	Q	H	-0,752
232	G	H	0,009	250	L	H	-0,288
233	E	H	-0,018	251	K	H	-0,286
234	G	H	-0,045	253	R	H	-0,286
235	E	H	-0,024	254	Q	H	-0,237
237	E	H	-0,087	255	I	H	-0,079
238	E	H	-0,119	256	K	H	-0,043
239	L	H	-0,224	257	A	H	0,061
241	V	H	-0,513	258	S	H	0,172
247	A	H	-0,022	259	F	H	0,151
250	L	H	-0,308	260	K	H	0,115
251	K	H	-0,28				
253	R	H	-0,287				
254	Q	H	-0,238				
255	I	H	-0,08				
256	K	H	-0,046				
257	A	H	0,058				
258	S	H	0,168				
259	F	H	0,145				
260	K	H	0,112				

Table 8.3: PCS values (in ppm) used in the plot regarding the cobalt(II)-DM-hCAII-thiocyanate 1:357.

Cobalt(II)-DM-hCAII-thiocyanate 1:1357			
Residue #			Measured PCS
6	G	H	0,14
8	G	H	-0,124
24	K	H	-0,458
25	G	H	-0,557
26	E	H	-0,475
31	V	H	-0,172
33	I	H	0,522
34	D	H	0,088
35	T	H	0,241
37	T	H	0,096
38	A	H	0,111
39	K	H	0,027
40	Y	H	0,057
41	D	H	0,032
44	L	H	-0,007

45	K	H	-0,061
47	L	H	0,06
48	S	H	-0,049
49	V	H	-0,024
50	S	H	-0,098
51	Y	H	-0,11
52	D	H	-0,045
53	Q	H	-0,113
54	A	H	-0,158
55	T	H	-0,185
56	S	H	-0,236
57	L	H	-0,232
58	R	H	-0,238
59	I	H	-0,34
63	G	H	-0,002
70	F	H	-0,131
71	D	H	-0,115
72	D	H	0,028
74	Q	H	-0,035
76	K	H	-0,083
77	A	H	-0,132
81	G	H	-0,109
82	G	H	-0,038
86	G	H	-0,138
87	T	H	-0,132
88	Y	H	-0,186
89	R	H	-0,078
99	S	H	-0,107
100	L	H	0,016
101	D	H	0,022
102	G	H	0,215
103	Q	H	0,334
104	G	H	0,798
105	S	H	1,617
109	V	H	0,841
110	D	H	0,41
111	K	H	0,36
112	K	H	0,489
114	Y	H	1,444
115	A	H	0,822
116	A	H	1,39
123	W	H	-0,204
124	N	H	-0,146

126	K	H	-0,099
127	Y	H	-0,031
128	G	H	0,009
129	D	H	0,09
130	F	H	0,262
131	G	H	0,395
132	K	H	0,229
133	A	H	0,151
134	V	H	0,245
135	Q	H	0,086
136	Q	H	-0,114
138	D	H	-0,306
147	L	H	1,898
148	K	H	0,629
150	G	H	0,152
151	S	H	0,117
153	K	H	-0,013
156	L	H	-0,23
157	Q	H	-0,265
162	V	H	-0,441
163	L	H	-0,493
164	D	H	-0,365
165	S	H	-0,319
166	I	H	-0,37
167	K	H	-0,293
168	T	H	-0,221
169	K	H	-0,094
170	G	H	0,069
172	S	H	-0,076
173	A	H	-0,206
175	F	H	-0,204
176	T	H	-0,213
177	N	H	-0,194
178	F	H	-0,281
179	D	H	-0,284
182	G	H	-0,185
183	L	H	-0,163
184	L	H	0,02
186	E	H	0,245
187	S	H	0,274
188	L	H	0,202
189	D	H	0,28
190	Y	H	0,26

191	W	H	0,531
192	T	H	0,077
193	Y	H	-0,329
207	T	H	-1,195
208	W	H	-1,003
209	I	H	-0,08
210	V	H	0,571
212	K	H	0,654
213	E	H	0,639
215	I	H	0,869
216	S	H	0,287
217	V	H	0,228
219	S	H	0,005
221	Q	H	-0,144
222	V	H	-0,295
223	L	H	-0,284
224	K	H	-0,346
225	F	H	-0,576
226	R	H	-0,529
227	K	H	-0,491
228	L	H	-0,599
229	N	H	-0,431
230	F	H	-0,301
231	N	H	-0,16
232	G	H	0,01
234	G	H	-0,052
235	E	H	-0,026
237	E	H	-0,086
238	E	H	-0,148
239	L	H	-0,23
241	V	H	-0,589
242	D	H	-0,604
250	L	H	-0,282
251	K	H	-0,28
253	R	H	-0,286
254	Q	H	-0,227
255	I	H	-0,077
256	K	H	-0,042
257	A	H	0,061
258	S	H	0,17
259	F	H	0,15
260	K	H	0,115

Table 8.4: Experimental RDC values (in Hertz) used in the plot regarding the cobalt(II)-DM-hCAII-thiocyanate 1:2 with and without NOE data.

RDC			**RDC with no NOE under 0.75**		
Residue #		RDC (Hz)	Residue #		RDC (Hz)
33	I	-12,65	33	I	-12,65
34	D	0,97	34	D	0,97
35	T	-2,8	35	T	-2,8
38	A	-2,31	41	D	-10,46
40	Y	10,22	43	S	-4,87
41	D	-10,46	44	L	7,06
43	S	-4,87	47	L	9,73
44	L	7,06	48	S	4,5
45	K	10,95	49	V	12,53
47	L	9,73	50	S	1,95
48	S	4,5	51	Y	15,94
49	V	12,53	56	S	2,68
50	S	1,95	57	L	-9,37
51	Y	15,94	59	I	-4,75
53	Q	-9,98	70	F	-12,53
54	A	14,24	72	D	8,4
56	S	2,68	74	Q	0,97
57	L	-9,37	80	K	7,3
59	I	-4,75	81	G	6,21
70	F	-12,53	84	L	-1,58
72	D	8,4	99	S	-10,83
74	Q	0,97	100	L	-9,25
80	K	7,3	102	G	-3,77
81	G	6,21	109	V	-7,91
82	G	4,75	110	D	-2,8
84	L	-1,58	111	K	4,38
86	G	4,26	112	K	-3,29
99	S	-10,83	113	K	1,95
100	L	-9,25	126	K	13,99
102	G	-3,77	133	A	-13,38
109	V	-7,91	134	V	-1,1
110	D	-2,8	136	Q	-10,59
111	K	4,38	139	G	-5,96
112	K	-3,29	148	K	-6,08
113	K	1,95	150	G	10,46
126	K	13,99	153	K	2,8
133	A	-13,38	156	L	13,26
134	V	-1,1	157	Q	6,94
136	Q	-10,59	163	L	15,94
139	G	-5,96	164	D	4,99

148	K	-6,08	165	S	6,08
149	V	-4,99	166	I	14,6
150	G	10,46	167	K	-10,22
151	S	-3,04	168	T	-2,92
153	K	2,8	169	K	-3,16
156	L	13,26	171	K	4,26
157	Q	6,94	172	S	2,56
163	L	15,94	173	A	-4,38
164	D	4,99	176	T	-9,86
165	S	6,08	178	F	8,52
166	I	14,6	182	G	10,46
167	K	-10,22	188	L	1,22
168	T	-2,92	189	D	1,46
169	K	-3,16	190	Y	-7,06
171	K	4,26	191	W	-2,68
172	S	2,56	193	Y	-3,16
173	A	-4,38	204	E	4,99
176	T	-9,86	210	V	-5,84
178	F	8,52	212	K	-5,72
182	G	10,46	213	E	0,73
187	S	-2,19	216	S	-4,14
188	L	1,22	217	V	-2,19
189	D	1,46	218	S	4,75
190	Y	-7,06	219	S	15,33
191	W	-2,68	221	Q	13,38
193	Y	-3,16	222	V	16,31
204	E	4,99	223	L	10,34
210	V	-5,84	224	K	6,08
212	K	-5,72	225	F	14,6
213	E	0,73	226	R	9,61
216	S	-4,14	227	K	-4,02
217	V	-2,19	228	L	4,62
218	S	4,75	229	N	-0,73
219	S	15,33	230	F	0,61
221	Q	13,38	231	N	4,99
222	V	16,31	250	L	8,76
223	L	10,34	251	K	0,97
224	K	6,08	255	I	5,11
225	F	14,6	256	K	9,61
226	R	9,61	257	A	-4,99
227	K	-4,02	258	S	-4,99
228	L	4,62	259	F	-1,58
229	N	-0,73	260	K	-4,99

230	F	0,61			
231	N	4,99			
232	G	-2,92			
233	E	-1,83			
234	G	6,69			
238	E	4,14			
250	L	8,76			
251	K	0,97			
255	I	5,11			
256	K	9,61			
257	A	-4,99			
258	S	-4,99			
259	F	-1,58			
260	K	-4,99			

Table 8.5: Chemical shift list (in ppm) of cobalt(II)-DM-hCAII-free and cobalt(II)-DM-hCAII-thiocyanate 1:1.3.

Free cobalt(II)-DM-hCAII				**Cobalt(II)-DM-hCAII-thiocyanate 1:1.3**			
Residue #		1H (ppm)	15N (ppm)	Residue #		1H (ppm)	15N (ppm)
6	G	8.054	109.799	6	G	8.124	109.914
8	G	8.097	110.667	8	G	8.116	110.717
24	K	7.194	114.806	24	K	6.837	114.389
25	G	8.258	108.814	25	G	7.793	108.353
26	E	9.289	117.663	26	E	8.930	117.455
27	R	8.705	120.614	31	V	6.036	106.309
28	Q	7.506	115.306	32	D	8.215	118.188
31	V	5.882	106.529	33	I	9.099	128.739
32	D	8.128	118.084	34	D	7.547	128.453
33	I	8.691	128.451	35	T	10.423	122.420
34	D	7.417	128.261	37	T	7.643	108.460
35	T	10.256	122.270	38	A	7.520	127.646
37	T	7.561	108.407	39	K	8.093	122.629
38	A	7.423	127.596	40	Y	8.686	126.725
39	K	8.036	122.572	41	D	7.710	128.904
40	Y	8.631	126.638	43	S	8.416	115.183
41	D	7.669	128.853	44	L	6.948	123.059
43	S	8.397	115.167	45	K	7.405	122.894
44	L	6.921	123.027	47	L	8.956	125.834
45	K	7.435	122.983	48	S	8.364	119.687
47	L	8.899	125.910	49	V	8.319	127.494
48	S	8.360	119.857	50	S	8.229	122.647
49	V	8.314	127.498	51	Y	8.852	124.467
50	S	8.237	122.756	52	D	8.739	121.322
51	Y	8.904	124.529	53	Q	7.886	114.200

52	D	8.781	121.236	54	A	7.208	120.056
53	Q	7.928	114.323	55	T	9.183	120.454
54	A	7.292	120.102	56	S	9.723	126.223
55	T	9.260	120.516	57	L	8.658	117.432
56	S	9.855	126.332	58	R	6.924	114.402
57	L	8.744	117.482	59	I	8.734	121.585
58	R	7.007	114.436	61	N	8.174	120.779
59	I	8.859	121.915	62	N	7.892	123.339
60	L	8.379	123.440	63	G	9.583	109.327
61	N	8.203	120.759	70	F	8.402	118.834
62	N	7.689	123.016	71	D	8.238	116.888
63	G	9.434	108.970	72	D	8.926	131.703
68	V	8.397	121.762	74	Q	7.577	117.614
69	E	8.183	123.593	76	K	7.811	124.106
70	F	8.338	118.785	77	A	8.084	123.526
71	D	8.262	116.912	78	V	8.167	115.689
72	D	8.834	131.578	79	L	8.757	123.707
73	S	8.778	115.805	80	K	8.185	120.395
74	Q	7.547	117.592	81	G	8.840	106.269
76	K	7.807	124.057	82	G	6.971	108.683
77	A	8.096	123.409	84	L	7.576	120.079
78	V	8.226	115.664	86	G	7.698	109.280
79	L	8.785	123.159	87	T	8.315	116.414
80	K	8.200	120.423	88	Y	8.220	125.807
81	G	8.857	106.339	89	R	8.380	122.470
82	G	6.952	108.632	99	S	8.238	111.619
84	L	7.590	120.141	100	L	7.312	118.965
86	G	7.721	109.388	101	D	8.904	120.581
87	T	8.335	116.550	102	G	7.792	102.976
88	Y	8.236	125.879	103	Q	8.102	115.381
89	R	8.325	122.526	104	G	8.895	107.065
99	S	8.187	111.667	105	S	8.681	109.475
100	L	7.229	118.951	109	V	8.275	119.149
101	D	8.853	120.534	110	D	9.992	132.686
102	G	7.659	102.797	111	K	10.124	110.071
103	Q	7.860	115.062	112	K	8.610	124.963
104	G	8.244	106.369	114	Y	9.402	122.491
105	S	7.623	108.107	115	A	8.310	121.696
108	T	7.367	109.316	116	A	9.367	113.292
109	V	7.728	118.649	118	L	9.308	129.432
110	D	9.700	132.364	123	W	8.788	118.732
111	K	9.835	109.781	124	N	8.810	119.318
112	K	8.269	124.588	125	T	7.946	116.879

113	K	8.011	122.069	126	K	7.779	122.663
114	Y	8.503	121.763	127	Y	7.541	115.512
115	A	7.878	121.233	128	G	7.750	107.245
116	A	8.743	112.723	129	D	7.479	116.206
118	L	9.783	130.151	130	F	7.668	119.766
123	W	8.709	118.902	131	G	8.350	104.195
124	N	8.770	119.532	132	K	7.889	119.928
125	T	8.006	117.068	133	A	8.123	124.760
126	K	7.792	122.774	134	V	7.562	108.376
127	Y	7.489	115.424	135	Q	6.618	113.898
128	G	7.714	107.284	136	Q	7.406	118.190
129	D	7.390	116.162	138	D	7.596	114.805
130	F	7.444	119.602	140	L	8.477	115.309
131	G	8.070	103.921	147	L	9.924	121.817
132	K	7.725	119.780	148	K	9.490	119.910
133	A	8.007	124.619	149	V	8.730	121.701
134	V	7.410	108.107	150	G	9.223	118.358
135	Q	6.643	113.975	151	S	8.396	124.571
136	Q	7.494	118.282	152	A	8.329	121.150
138	D	7.736	114.767	153	K	8.632	123.928
140	L	8.597	115.413	156	L	7.304	116.540
147	L	9.128	121.188	157	Q	7.522	122.053
148	K	9.299	119.705	158	K	8.491	115.490
149	V	8.594	121.567	161	D	8.008	117.157
150	G	9.207	118.304	162	V	6.712	113.587
151	S	8.389	124.568	163	L	6.586	121.243
152	A	8.392	121.177	164	D	7.841	115.838
153	K	8.779	124.025	165	S	7.656	113.278
156	L	7.574	116.827	166	I	7.022	116.105
157	Q	7.776	122.370	167	K	7.076	117.123
158	K	8.631	115.503	168	T	6.321	98.307
160	V	7.185	114.302	169	K	7.317	120.410
161	D	8.299	117.484	170	G	8.852	116.899
162	V	7.011	113.820	171	K	7.680	120.149
163	L	6.910	121.577	172	S	8.206	113.599
164	D	8.071	116.077	173	A	8.768	123.203
165	S	7.844	113.444	174	D	8.429	121.923
166	I	7.224	116.267	175	F	7.750	125.120
167	K	7.214	117.288	176	T	7.854	114.626
168	T	6.406	98.433	177	N	9.495	115.306
169	K	7.301	120.392	178	F	7.962	118.622
170	G	8.732	116.762	179	D	7.702	127.786
171	K	7.659	120.098	181	R	8.299	118.288

172	S	8.216	113.639	182	G	7.085	102.536
173	A	8.871	123.303	183	L	7.071	117.288
174	D	8.501	121.988	184	L	6.603	113.173
175	F	7.856	125.201	186	E	8.347	116.541
176	T	7.987	114.802	187	S	7.803	111.497
177	N	9.608	115.411	188	L	8.768	125.817
178	F	8.142	118.785	189	D	7.598	122.124
179	D	7.924	128.016	190	Y	8.892	119.111
181	R	8.399	118.421	191	W	9.950	117.612
182	G	7.287	102.706	192	T	9.652	115.029
183	L	7.305	117.432	193	Y	7.721	125.429
184	L	6.734	113.315	207	T	7.322	123.937
186	E	8.267	116.478	208	W	7.221	128.684
187	S	7.686	111.385	209	I	8.591	127.421
188	L	8.665	125.699	210	V	10.088	128.932
189	D	7.449	122.007	211	L	10.358	128.083
190	Y	8.706	119.139	212	K	7.881	122.843
191	W	9.556	117.209	213	E	9.405	122.248
192	T	9.548	114.983	215	I	9.677	114.070
193	Y	7.737	125.387	216	S	8.302	116.850
195	G	8.911	110.206	217	V	8.713	116.087
205	C	7.214	114.297	218	S	8.334	115.825
207	T	7.831	124.404	219	S	9.262	117.292
208	W	7.558	128.989	220	E	8.097	118.018
209	I	8.245	126.787	221	Q	7.363	117.808
210	V	9.652	128.440	222	V	7.138	114.459
211	L	9.375	127.021	223	L	7.778	120.482
212	K	7.516	122.563	224	K	6.556	114.825
213	E	9.071	121.848	225	F	6.576	118.201
215	I	9.363	113.777	226	R	6.172	108.414
216	S	8.245	116.849	227	K	6.373	113.703
217	V	8.711	116.077	228	L	6.517	118.854
218	S	8.469	115.974	229	N	8.460	118.055
219	S	9.315	117.359	230	F	8.523	116.926
220	E	8.288	118.272	231	N	7.945	110.472
221	Q	7.527	117.961	232	G	8.608	104.707
222	V	7.368	114.755	233	E	8.310	121.234
223	L	7.965	120.815	234	G	8.982	114.075
224	K	6.816	115.019	235	E	7.257	119.933
225	F	6.980	118.631	237	E	8.114	123.140
226	R	6.503	108.844	238	E	8.828	129.851
227	K	6.606	114.046	240	M	8.385	122.662
228	L	6.832	119.101	241	V	6.124	114.323

229	N	8.640	118.228	242	D	6.589	114.299
230	F	8.604	116.986	248	Q	8.011	121.508
231	N	7.961	110.494	251	K	8.560	114.168
232	G	8.559	104.698	253	R	6.993	118.276
233	E	8.285	121.257	254	Q	8.309	117.993
234	G	8.982	114.052	255	I	8.782	124.274
235	E	7.245	119.898	256	K	8.662	126.583
237	E	8.134	123.184	257	A	8.361	123.594
238	E	8.853	129.857	258	S	9.045	116.557
239	L	8.527	125.731	259	F	6.722	118.071
241	V	6.293	114.462	260	K	7.790	124.477
242	D	6.689	114.357				
247	A	7.682	120.069				
248	Q	8.358	121.780				
250	L	8.516	125.778				
251	K	8.724	114.410				
253	R	7.148	118.394				
254	Q	8.425	118.121				
255	I	8.791	124.295				
256	K	8.633	126.499				
257	A	8.288	123.592				
258	S	8.908	116.434				
259	F	6.607	117.966				
260	K	7.711	124.423				

Table 8.6: Chemical shift list (in ppm) of cobalt(II)-DM-hCAII-1:13 and cobalt(II)-DM-hCAII-thiocyanate 1:1357.

Cobalt(II)-DM-hCAII-thiocyanate 1:13				**Cobalt(II)-DM-hCAII-thiocyanate 1:1357**			
Residue #		^{1}H (ppm)	^{15}N (ppm)	Residue #		^{1}H (ppm)	^{15}N (ppm)
6	G	8.143	109.956	6	G	8.265	110.134
8	G	8.121	110.722	8	G	8.154	110.726
24	K	6.754	114.268	24	K	6.784	114.209
25	G	7.685	108.270	25	G	7.724	108.285
26	E	8.848	117.406	26	E	8.857	117.585
31	V	6.067	106.254	31	V	6.119	106.293
32	D	8.237	118.223	32	D	8.263	118.259
33	I	9.192	128.798	33	I	9.222	128.750
34	D	7.578	128.483	34	D	7.608	128.489
35	T	10.461	122.446	35	T	10.511	122.457
37	T	7.665	108.499	37	T	7.717	109.023
38	A	7.543	127.648	38	A	7.589	127.409
39	K	8.111	122.667	39	K	8.194	122.903
40	Y	8.700	126.740	40	Y	8.711	126.836
41	D	7.719	128.899	41	D	7.698	128.732

44	L	6.954	123.062	43	S	8.469	115.272
45	K	7.398	122.873	44	L	6.975	123.141
47	L	8.969	125.809	45	K	7.413	122.891
48	S	8.367	119.667	47	L	8.965	125.624
49	V	8.320	127.488	48	S	8.387	119.626
50	S	8.227	122.615	49	V	8.306	127.416
51	Y	8.837	124.451	50	S	8.274	122.448
52	D	8.730	121.351	51	Y	8.855	124.473
53	Q	7.878	114.170	52	D	8.759	121.735
54	A	7.191	120.050	53	Q	7.959	114.127
55	T	9.166	120.443	54	A	7.225	120.169
56	S	9.694	126.209	55	T	9.219	120.583
57	L	8.639	117.424	56	S	9.710	126.284
58	R	6.906	114.416	57	L	8.688	117.604
59	I	8.673	121.680	58	R	6.947	114.837
61	N	8.165	120.811	59	I	8.658	122.236
62	N	7.936	123.412	61	N	8.199	121.184
63	G	9.618	109.379	62	N	7.957	123.143
70	F	8.417	118.853	63	G	9.674	108.964
71	D	8.234	116.883	70	F	8.458	118.990
72	D	8.946	131.730	71	D	8.279	116.916
74	Q	7.585	117.626	72	D	8.964	131.651
76	K	7.813	124.116	74	Q	7.636	117.755
77	A	8.084	123.516	76	K	7.831	123.873
78	V	8.154	115.658	77	A	8.113	123.609
79	L	8.756	123.794	78	V	8.174	115.910
80	K	8.062	120.299	79	L	8.770	124.051
81	G	8.836	106.259	80	K	8.081	120.063
82	G	6.976	108.693	81	G	8.844	106.409
84	L	7.572	120.069	82	G	7.012	108.756
86	G	7.692	109.256	84	L	7.614	120.140
87	T	8.307	116.346	86	G	7.693	109.174
88	Y	8.217	125.795	87	T	8.299	115.920
89	R	8.488	122.736	88	Y	8.298	125.893
99	S	8.249	111.617	89	R	8.511	123.321
100	L	7.324	118.979	99	S	8.285	111.938
101	D	8.914	120.596	100	L	7.437	119.347
102	G	7.821	103.027	101	D	8.875	120.872
103	Q	8.153	115.441	102	G	7.801	103.682
104	G	9.031	107.215	103	Q	8.129	115.382
105	S	8.906	109.749	104	G	9.039	107.583
109	V	8.400	119.256	105	S	8.902	109.806
110	D	10.060	132.764	109	V	8.462	119.269

111	K	10.190	110.136	110	D	10.100	132.809
112	K	8.686	125.030	111	K	10.211	110.235
114	Y	9.603	122.694	112	K	8.700	125.134
115	A	8.387	121.861	114	Y	9.575	122.236
116	A	9.505	113.419	115	A	8.424	121.863
123	W	8.808	118.701	116	A	9.522	113.458
124	N	8.839	119.330	123	W	8.832	118.703
125	T	7.907	116.845	124	N	8.897	119.429
126	K	7.776	122.643	125	T	8.017	117.096
127	Y	7.553	115.524	126	K	7.814	122.752
128	G	7.760	107.233	127	Y	7.576	115.622
129	D	7.498	116.216	128	G	7.812	107.148
130	F	7.716	119.804	129	D	7.526	116.337
131	G	8.412	104.248	130	F	7.730	120.173
132	K	7.922	119.936	131	G	8.425	104.111
133	A	8.148	124.789	132	K	7.879	119.741
134	V	7.596	108.440	133	A	8.136	124.762
135	Q	6.618	113.864	134	V	7.632	108.490
136	Q	7.385	118.163	135	Q	6.634	113.859
138	D	7.565	114.820	136	Q	7.434	118.257
140	L	8.432	115.258	138	D	7.591	114.873
146	F	12.081	131.207	140	L	8.473	115.367
147	L	10.097	121.961	146	F	12.129	131.305
148	K	9.533	119.961	147	L	10.118	122.065
149	V	8.765	121.746	148	K	9.566	120.253
150	G	9.227	118.361	149	V	8.697	121.905
151	S	8.398	124.564	150	G	9.222	118.098
152	A	8.321	121.145	151	S	8.394	124.419
153	K	8.602	123.904	152	A	8.356	121.360
156	L	7.245	116.483	153	K	8.656	123.876
157	Q	7.465	121.980	156	L	7.260	116.612
158	K	8.428	115.389	157	Q	7.551	122.039
161	D	8.008	117.087	158	K	8.304	115.369
162	V	6.638	113.521	161	D	7.982	117.126
163	L	6.512	121.158	162	V	6.684	113.473
164	D	7.790	115.782	163	L	6.576	121.181
165	S	7.614	113.239	164	D	7.806	115.617
166	I	6.977	116.069	165	S	7.658	113.368
167	K	7.045	117.074	166	I	7.030	116.183
168	T	6.301	98.278	167	K	7.083	117.188
169	K	7.322	120.402	168	T	6.343	98.391
170	G	8.883	116.940	169	K	7.344	120.337
171	K	7.686	120.148	170	G	9.046	117.266

172	S	8.204	113.588	171	K	7.716	120.311
173	A	8.744	123.168	172	S	8.219	113.474
174	D	8.415	121.893	173	A	8.751	122.825
175	F	7.728	125.108	174	D	8.469	122.232
176	T	7.825	114.589	175	F	7.761	125.166
177	N	9.468	115.277	176	T	7.847	114.634
178	F	7.922	118.591	177	N	9.466	115.423
179	D	7.655	127.737	178	F	7.949	118.763
181	R	8.272	118.258	179	D	7.656	127.991
182	G	7.040	102.494	181	R	8.304	118.626
183	L	7.018	117.255	182	G	7.106	102.789
184	L	6.575	113.135	183	L	7.062	117.361
186	E	8.367	116.560	184	L	6.618	113.097
187	S	7.832	111.529	186	E	8.369	116.700
188	L	8.790	125.849	187	S	7.900	111.635
189	D	7.631	122.143	188	L	8.795	125.806
190	Y	8.931	119.156	189	D	7.622	122.078
191	W	10.039	117.706	190	Y	8.950	119.226
192	T	9.678	115.025	191	W	10.085	117.773
193	Y	7.718	125.448	192	T	9.725	115.105
207	T	7.203	123.840	193	Y	7.768	125.541
208	W	7.135	128.653	207	T	7.216	123.948
209	I	8.670	127.565	208	W	7.170	128.715
210	V	10.188	129.039	209	I	8.712	127.603
211	L	10.579	128.322	210	V	10.223	129.086
212	K	7.962	122.890	211	L	10.637	128.441
213	E	9.479	122.331	212	K	7.976	122.930
215	I	9.750	114.128	213	E	9.500	122.207
216	S	8.311	116.843	215	I	9.756	114.029
217	V	8.713	116.080	216	S	8.338	116.762
218	S	8.305	115.797	217	V	8.701	116.137
219	S	9.248	117.281	218	S	8.369	116.046
220	E	8.055	117.954	219	S	9.258	117.341
221	Q	7.326	117.769	220	E	8.076	118.007
222	V	7.088	114.397	221	Q	7.353	117.785
223	L	7.737	120.410	222	V	7.086	114.246
224	K	6.502	114.775	223	L	7.712	120.513
225	F	6.485	118.114	224	K	6.583	114.994
226	R	6.097	108.313	225	F	6.537	118.297
227	K	6.310	113.644	226	R	6.159	108.468
228	L	6.446	118.800	227	K	6.364	113.621
229	N	8.417	118.016	228	L	6.495	118.799
230	F	8.503	116.905	229	N	8.441	117.968

231	N	7.940	110.463	230	F	8.532	116.885
232	G	8.620	104.716	231	N	7.969	110.435
233	E	8.310	121.251	232	G	8.638	104.662
234	G	8.982	114.085	233	E	8.324	121.624
235	E	7.261	119.942	234	G	8.982	114.108
237	E	8.109	123.138	235	E	7.296	119.978
238	E	8.822	129.850	237	E	8.114	123.118
240	M	8.382	122.405	238	E	8.850	129.830
241	V	6.086	114.285	240	M	8.620	123.296
242	D	6.562	114.299	241	V	6.103	114.482
247	A	7.833	120.077	242	D	6.567	114.490
248	Q	7.932	121.444	247	A	7.823	120.519
251	K	8.521	114.118	248	Q	7.965	121.284
253	R	6.958	118.244	251	K	8.523	114.374
254	Q	8.281	117.968	253	R	6.996	118.221
255	I	8.781	124.262	254	Q	8.246	118.120
256	K	8.667	126.575	255	I	8.793	124.188
257	A	8.380	123.619	256	K	8.687	126.623
258	S	9.076	116.568	257	A	8.471	123.621
259	F	6.748	118.106	258	S	9.098	116.461
260	K	7.806	124.470	259	F	6.800	118.193
				260	K	7.807	124.434

Table 8.7: Chemical shift list (in ppm) of zinc(II)-DM-hCAII-free and zinc(II)-DM-hCAII-thiocyanate 1:1.3.

Free zinc(II)-DM-hCAII				**Zinc(II)-DM-hCAII 1:1.3**			
Residue #		1H (ppm)	15N (ppm)	Residue #		1H (ppm)	15N (ppm)
6	G	8.014	109.739	6	G	8.012	109.738
8	G	8.234	110.825	8	G	8.236	110.845
23	A	8.572	123.019	23	A	8.548	123.215
24	K	7.164	114.837	24	K	7.166	114.828
25	G	8.252	108.817	25	G	8.252	108.798
26	E	9.330	117.714	26	E	9.329	117.752
27	R	8.791	120.679	27	R	8.789	120.752
28	Q	7.680	115.463	28	Q	7.679	115.454
29	S	8.020	117.897	29	S	8.065	118.017
31	V	6.229	106.936	31	V	6.237	106.885
32	D	8.277	118.264	32	D	8.274	118.230
33	I	8.662	128.465	33	I	8.666	128.457
34	D	7.433	128.239	34	D	7.418	128.184
35	T	10.226	122.250	35	T	10.224	122.244
37	T	7.567	108.430	37	T	7.568	108.433
38	A	7.428	127.585	38	A	7.429	127.573
39	K	8.063	122.605	39	K	8.068	122.610

40	Y	8.641	126.673	40	Y	8.641	126.675
41	D	7.688	128.889	41	D	7.686	128.897
43	S	8.426	115.193	43	S	8.427	115.187
44	L	6.954	123.071	44	L	6.956	123.069
45	K	7.489	123.042	45	K	7.480	123.013
47	L	8.914	125.942	47	L	8.912	125.887
48	S	8.423	119.909	48	S	8.419	119.854
49	V	8.344	127.541	49	V	8.343	127.525
50	S	8.313	122.822	50	S	8.314	122.786
51	Y	8.941	124.571	51	Y	8.944	124.562
52	D	8.833	120.511	52	D	8.833	120.491
53	Q	7.977	114.373	53	Q	7.980	114.355
54	A	7.351	120.140	54	A	7.349	120.162
55	T	9.344	120.581	55	T	9.348	120.588
56	S	9.931	126.411	56	S	9.931	126.425
57	L	8.852	117.566	57	L	8.857	117.597
58	R	7.127	114.528	58	R	7.130	114.554
59	I	8.983	122.041	59	I	8.983	122.041
60	L	8.627	123.603	60	L	8.630	123.675
61	N	8.376	121.048	61	N	8.380	121.067
62	N	7.975	123.215	62	N	7.982	123.245
63	G	9.641	109.115	63	G	9.638	109.127
66	F	7.147	110.045	66	F	7.139	110.055
67	N	8.896	120.649	67	N	8.896	120.699
68	V	8.568	121.816	68	V	8.588	121.900
69	E	8.390	123.904	69	E	8.398	123.907
70	F	8.530	118.952	70	F	8.533	118.964
71	D	8.356	117.034	71	D	8.352	117.017
72	D	8.919	131.680	72	D	8.915	131.681
73	S	8.776	115.819	73	S	8.769	115.816
74	Q	7.615	117.645	74	Q	7.614	117.648
76	K	7.886	124.104	76	K	7.888	124.111
77	A	8.202	123.488	77	A	8.206	123.519
78	V	8.345	115.836	78	V	8.336	115.835
79	L	8.928	123.429	79	L	8.931	123.453
80	K	8.308	120.534	80	K	8.308	120.590
81	G	8.951	106.427	81	G	8.948	106.409
82	G	7.015	108.694	82	G	7.013	108.696
84	L	7.713	120.240	84	L	7.713	120.235
86	G	7.815	109.488	86	G	7.818	109.472
87	T	8.439	116.664	87	T	8.436	116.630
88	Y	8.397	125.999	88	Y	8.393	125.997
89	R	8.524	122.708	89	R	8.526	122.693

90	L	8.460	124.293	90	L	8.441	124.263
91	I	8.852	123.295	91	I	8.840	123.185
92	Q	7.255	115.209	92	Q	7.257	115.215
93	F	8.546	113.440	93	F	8.540	113.399
94	H	8.141	113.088	94	H	8.138	113.123
95	F	9.038	117.112	95	F	9.049	117.057
96	H	8.686	115.121	96	H	8.689	115.275
97	W	9.472	119.120	97	W	9.474	119.109
98	G	8.017	107.885	98	G	8.017	107.885
99	S	8.343	111.827	99	S	8.339	111.803
100	L	7.281	118.976	100	L	7.285	118.973
101	D	8.819	121.264	101	D	8.822	121.313
102	G	7.616	102.747	102	G	7.615	102.749
103	Q	7.803	114.951	103	Q	7.804	114.964
104	G	8.194	106.249	104	G	8.197	106.281
105	S	7.271	107.677	105	S	7.265	107.694
106	E	8.239	119.140	106	E	8.251	119.213
107	H	10.646	117.778	107	H	10.645	117.777
108	T	7.300	109.243	108	T	7.301	109.176
109	V	7.562	118.501	109	V	7.562	118.485
110	D	9.647	132.313	110	D	9.647	132.320
111	K	9.804	109.753	111	K	9.811	109.775
112	K	8.187	124.537	112	K	8.189	124.548
113	K	7.904	121.960	113	K	7.903	121.972
114	Y	8.124	121.381	114	Y	8.128	121.401
115	A	7.514	120.856	115	A	7.514	120.840
116	A	8.087	112.041	116	A	8.086	112.026
117	E	9.372	120.943	117	E	9.368	121.050
118	L	9.819	130.394	118	L	9.818	130.407
119	H	8.973	124.936	119	H	8.908	125.330
120	L	8.962	123.637	120	L	8.962	123.632
121	V	9.140	126.508	121	V	9.114	126.493
122	H	8.491	123.923	122	H	8.486	123.921
123	W	9.005	119.249	123	W	8.999	119.186
124	N	8.960	119.714	124	N	8.968	119.700
125	T	8.130	117.193	125	T	8.126	117.177
126	K	7.878	122.856	126	K	7.874	122.824
127	Y	7.551	115.490	127	Y	7.557	115.489
128	G	7.751	107.314	128	G	7.749	107.284
129	D	7.409	116.202	129	D	7.405	116.184
130	F	7.457	119.609	130	F	7.450	119.589
131	G	8.001	103.843	131	G	8.002	103.835
132	K	7.681	119.745	132	K	7.679	119.735

133	A	7.998	124.589	133	A	7.991	124.606
134	V	7.338	108.005	134	V	7.335	108.104
135	Q	6.540	114.016	135	Q	6.533	113.945
136	Q	7.489	118.261	136	Q	7.488	118.288
138	D	7.832	114.856	138	D	7.839	114.929
139	G	7.885	107.554	139	G	7.874	107.555
140	L	8.844	115.727	140	L	8.849	115.739
141	A	8.828	122.859	141	A	8.825	123.023
142	V	7.153	124.127	142	V	7.151	124.330
143	L	8.417	128.209	143	L	8.400	128.033
144	G	9.867	114.212	144	G	9.858	114.045
145	I	9.094	124.249	145	I	9.115	124.273
146	F	10.050	128.926	146	F	10.044	128.905
147	L	8.171	120.347	147	L	8.175	120.355
148	K	8.917	119.291	148	K	8.913	119.274
149	V	8.396	121.353	149	V	8.396	121.350
150	G	9.064	118.164	150	G	9.068	118.169
151	S	8.276	124.482	151	S	8.275	124.486
152	A	8.314	121.090	152	A	8.314	121.126
153	K	8.658	123.911	153	K	8.655	123.900
156	L	7.481	116.716	156	L	7.479	116.703
157	Q	7.728	122.331	157	Q	7.728	122.316
158	K	8.662	115.475	158	K	8.659	115.471
159	V	6.963	113.126	159	V	6.967	113.134
160	V	7.229	114.355	160	V	7.231	114.354
161	D	8.327	117.534	161	D	8.327	117.543
162	V	7.076	113.847	162	V	7.077	113.855
163	L	7.002	121.664	163	L	7.002	121.651
164	D	8.157	116.181	164	D	8.158	116.170
165	S	7.932	113.521	165	S	7.933	113.519
166	I	7.346	116.366	166	I	7.347	116.357
167	K	7.327	117.421	167	K	7.331	117.416
168	T	6.523	98.575	168	T	6.523	98.574
169	K	7.421	120.532	169	K	7.422	120.536
170	G	8.811	116.836	170	G	8.814	116.849
171	K	7.738	120.206	171	K	7.738	120.199
172	S	8.284	113.732	172	S	8.284	113.721
173	A	8.964	123.424	173	A	8.961	123.426
174	D	8.560	122.061	174	D	8.560	122.054
175	F	7.929	125.259	175	F	7.931	125.257
176	T	8.045	114.879	176	T	8.044	114.860
177	N	9.660	115.452	177	N	9.658	115.453
178	F	8.197	118.816	178	F	8.199	118.833

179	D	7.943	128.060	179	D	7.941	128.050
181	R	8.358	118.382	181	R	8.357	118.361
182	G	7.238	102.634	182	G	7.235	102.625
183	L	7.181	117.262	183	L	7.181	117.286
184	L	6.576	113.130	184	L	6.572	113.113
186	E	8.116	116.311	186	E	8.117	116.307
187	S	7.556	111.282	187	S	7.555	111.288
188	L	8.590	125.615	188	L	8.589	125.617
189	D	7.358	121.933	189	D	7.357	121.927
190	Y	8.680	118.687	190	Y	8.696	119.079
191	W	9.488	117.151	191	W	9.494	117.179
192	T	9.621	115.094	192	T	9.616	115.060
193	Y	7.990	125.572	193	Y	8.010	125.616
195	G	8.750	110.107	195	G	8.742	110.060
196	S	8.635	119.957	196	S	8.637	119.936
199	T	6.907	108.241	199	T	6.913	108.293
202	L	9.154	120.611	202	L	9.139	120.562
203	L	5.918	110.868	203	L	5.913	110.863
204	E	8.727	124.782	204	E	8.727	124.707
205	C	7.569	114.674	205	C	7.567	114.652
206	V	7.210	118.912	206	V	7.197	119.003
207	T	8.386	124.919	207	T	8.384	124.915
208	W	8.191	129.364	208	W	8.174	129.452
209	I	8.701	127.101	209	I	8.717	127.188
210	V	9.665	128.425	210	V	9.654	128.392
211	L	8.909	126.658	211	L	8.937	126.735
212	K	7.327	122.348	212	K	7.302	122.041
213	E	8.840	121.597	213	E	8.838	121.608
215	I	8.863	113.345	215	I	8.867	113.334
216	S	8.029	116.641	216	S	8.029	116.593
217	V	8.483	115.843	217	V	8.483	115.844
218	S	8.335	115.788	218	S	8.342	115.831
219	S	9.242	117.270	219	S	9.242	117.278
220	E	8.323	118.327	220	E	8.325	118.330
221	Q	7.474	117.925	221	Q	7.472	117.924
222	V	7.299	114.688	222	V	7.287	114.655
223	L	7.953	120.830	223	L	7.951	120.793
224	K	6.844	115.053	224	K	6.844	115.070
225	F	7.053	118.699	225	F	7.053	118.687
226	R	6.623	109.008	226	R	6.621	108.978
227	K	6.790	114.104	227	K	6.789	114.087
228	L	7.037	119.279	228	L	7.036	119.283
229	N	8.849	118.431	229	N	8.847	118.442

230	F	8.813	117.203	230	F	8.811	117.202
231	N	8.099	110.631	231	N	8.099	110.631
232	G	8.612	104.768	232	G	8.612	104.762
233	E	8.333	121.291	233	E	8.328	121.268
234	G	9.022	114.077	234	G	9.022	114.090
235	E	7.288	119.935	235	E	7.288	119.940
237	E	8.197	123.249	237	E	8.196	123.246
238	E	8.961	129.959	238	E	8.961	129.964
239	L	8.678	125.904	239	L	8.675	125.893
240	M	8.664	123.032	240	M	8.657	123.027
241	V	6.666	114.848	241	V	6.661	114.864
242	D	7.085	114.730	242	D	7.086	114.744
243	N	8.346	119.339	243	N	8.342	119.346
244	W	6.413	114.101	244	W	6.411	114.113
245	R	10.405	126.989	245	R	10.441	127.069
247	A	7.940	120.320	247	A	7.935	120.351
248	Q	8.652	122.024	248	Q	8.662	122.037
250	L	8.629	125.824	250	L	8.627	125.821
251	K	8.812	114.505	251	K	8.809	114.480
253	R	7.241	118.482	253	R	7.242	118.491
254	Q	8.525	118.205	254	Q	8.523	118.189
255	I	8.862	124.377	255	I	8.860	124.388
256	K	8.717	126.537	256	K	8.715	126.554
257	A	8.320	123.594	257	A	8.318	123.580
258	S	8.903	116.437	258	S	8.904	116.437
259	F	6.593	117.945	259	F	6.594	117.946
260	K	7.692	124.411	260	K	7.692	124.396

Table 8.8: Chemical shift list (in ppm) of zinc(II)-DM-hCAII-thiocyanate 1:13 and zinc(II)-DM-hCAII-thiocyanate 1:1357.

Zinc(II)-DM-hCAII 1:10				**Zinc(II)-DM-hCAII 1:1000**			
Residue #		1H (ppm)	15N (ppm)	Residue #		1H (ppm)	15N (ppm)
6	G	8.012	109.743	6	G	8.125	109.956
8	G	8.238	110.859	8	G	8.278	110.874
23	A	8.544	123.042	23	A	8.578	123.012
24	K	7.177	114.872	24	K	7.242	115.053
25	G	8.250	108.772	25	G	8.281	108.807
26	E	9.327	117.826	26	E	9.332	118.032
27	R	8.787	120.877	27	R	8.668	120.615
28	Q	7.677	115.463	28	Q	7.713	115.591
29	S	8.117	118.295	29	S	8.155	118.450
31	V	6.249	106.800	31	V	6.291	106.777
32	D	8.270	118.162	32	D	8.270	118.171
33	I	8.671	128.469	33	I	8.700	128.408

34	D	7.423	128.219	34	D	7.520	128.774
35	T	10.221	122.231	35	T	10.270	122.222
37	T	7.569	108.476	37	T	7.621	108.998
38	A	7.430	127.550	38	A	7.478	127.321
39	K	8.081	122.635	39	K	8.167	122.869
40	Y	8.642	126.709	40	Y	8.654	126.830
41	D	7.684	128.895	41	D	7.666	128.732
43	S	8.430	115.194	43	S	8.482	115.272
44	L	6.957	123.074	44	L	6.982	123.160
45	K	7.462	122.967	45	K	7.474	122.961
47	L	8.910	125.803	47	L	8.905	125.598
48	S	8.413	119.761	48	S	8.436	119.701
49	V	8.341	127.498	49	V	8.330	127.405
50	S	8.317	122.740	50	S	8.372	122.538
51	Y	8.944	124.527	51	Y	8.965	124.500
52	D	8.831	120.491	52	D	8.804	120.712
53	Q	7.988	114.321	53	Q	8.072	114.259
54	A	7.347	120.190	54	A	7.383	120.322
55	T	9.347	120.624	55	T	9.404	120.824
56	S	9.930	126.454	56	S	9.946	126.531
57	L	8.867	117.644	57	L	8.920	117.834
58	R	7.145	114.656	58	R	7.185	114.672
59	I	8.983	122.040	59	I	8.998	122.282
60	L	8.634	123.798	60	L	8.716	123.883
61	N	8.381	121.125	61	N	8.378	121.384
62	N	7.999	123.275	62	N	7.998	123.065
63	G	9.637	109.112	63	G	9.676	108.711
66	F	7.128	110.090	66	F	7.181	110.207
67	N	8.896	120.805	67	N	8.943	120.993
68	V	8.630	122.014	68	V	8.700	122.105
69	E	8.409	123.887	69	E	8.412	123.993
70	F	8.541	118.995	70	F	8.589	119.177
71	D	8.347	116.959	71	D	8.394	117.042
72	D	8.913	131.709	72	D	8.936	131.632
73	S	8.760	115.788	73	S	8.804	116.597
74	Q	7.613	117.655	74	Q	7.671	117.791
76	K	7.892	124.118	76	K	7.914	123.872
77	A	8.215	123.561	77	A	8.245	123.643
78	V	8.320	115.849	78	V	8.335	116.045
79	L	8.932	123.562	79	L	8.951	123.472
80	K	8.309	120.692	80	K	8.370	120.944
81	G	8.942	106.372	81	G	8.953	106.530
82	G	7.007	108.696	82	G	7.050	108.771

84	L	7.714	120.235	84	L	7.765	120.279
86	G	7.822	109.452	86	G	7.831	109.400
87	T	8.432	116.575	87	T	8.431	116.310
88	Y	8.390	126.029	88	Y	8.484	126.172
89	R	8.529	122.659	89	R	8.589	122.621
90	L	8.411	124.185	90	L	8.451	124.169
91	I	8.827	123.269	91	I	8.911	123.443
92	Q	7.260	115.210	92	Q	7.301	115.305
93	F	8.527	113.321	93	F	8.532	113.300
94	H	8.133	113.184	94	H	8.160	113.283
95	F	9.068	116.944	95	F	9.109	116.868
96	H	8.708	115.492	96	H	8.711	115.682
97	W	9.476	119.092	97	W	9.517	119.105
98	G	8.014	107.888	98	G	8.061	107.857
99	S	8.336	111.788	99	S	8.392	112.107
100	L	7.291	118.966	100	L	7.421	119.319
101	D	8.828	121.400	101	D	8.853	121.555
102	G	7.610	102.762	102	G	7.586	103.389
103	Q	7.800	114.967	103	Q	7.795	114.915
104	G	8.201	106.333	104	G	8.241	106.744
105	S	7.254	107.723	105	S	7.285	107.837
106	E	8.270	119.288	106	E	8.273	119.269
107	H	10.640	117.765	107	H	10.711	117.870
108	T	7.303	109.084	108	T	7.339	109.026
109	V	7.562	118.448	109	V	7.621	118.495
110	D	9.648	132.339	110	D	9.690	132.389
111	K	9.823	109.811	111	K	9.851	109.924
112	K	8.191	124.565	112	K	8.211	124.710
113	K	7.899	122.035	113	K	7.924	122.631
114	Y	8.136	121.463	114	Y	8.131	121.062
115	A	7.518	120.827	115	A	7.602	121.215
116	A	8.085	112.004	116	A	8.132	112.067
117	E	9.362	121.229	117	E	9.399	121.334
118	L	9.808	130.419	118	L	9.853	130.421
119	H	8.860	125.521	119	H	8.882	125.731
120	L	8.947	123.737	120	L	8.993	124.086
121	V	9.064	126.455	121	V	9.076	126.494
122	H	8.472	123.884	122	H	8.481	124.036
123	W	8.987	119.063	123	W	9.036	119.050
124	N	8.985	119.690	124	N	9.043	119.701
125	T	8.118	117.168	125	T	8.142	117.196
126	K	7.867	122.770	126	K	7.913	122.862
127	Y	7.569	115.509	127	Y	7.607	115.619

128	G	7.745	107.240	128	G	7.803	107.144
129	D	7.397	116.154	129	D	7.436	116.274
130	F	7.439	119.557	130	F	7.468	119.902
131	G	7.999	103.822	131	G	8.030	103.711
132	K	7.676	119.711	132	K	7.650	119.529
133	A	7.981	124.645	133	A	7.985	124.637
134	V	7.331	108.279	134	V	7.387	108.415
135	Q	6.519	113.820	135	Q	6.548	113.787
136	Q	7.487	118.304	136	Q	7.548	118.405
138	D	7.854	115.044	138	D	7.897	115.147
139	G	7.852	107.559	139	G	7.873	107.623
140	L	8.850	115.742	140	L	8.774	115.611
141	A	8.863	123.279	141	A	8.877	123.496
142	V	7.149	124.700	142	V	7.202	124.892
143	L	8.368	127.702	143	L	8.380	127.468
144	G	9.844	113.739	144	G	9.860	113.631
145	I	9.153	124.331	145	I	9.214	124.496
146	F	10.035	128.871	146	F	10.080	128.948
147	L	8.179	120.345	147	L	8.220	120.442
148	K	8.902	119.253	148	K	8.937	119.554
149	V	8.394	121.356	149	V	8.406	121.484
150	G	9.070	118.163	150	G	9.070	117.908
151	S	8.274	124.483	151	S	8.277	124.340
152	A	8.314	121.132	152	A	8.314	121.655
153	K	8.652	123.882	153	K	8.669	124.574
156	L	7.475	116.720	156	L	7.490	116.865
157	Q	7.729	122.305	157	Q	7.816	122.362
158	K	8.656	115.461	158	K	8.656	115.480
159	V	6.973	113.163	159	V	7.026	113.510
160	V	7.233	114.360	160	V	7.264	114.466
161	D	8.325	117.546	161	D	8.367	117.572
162	V	7.078	113.880	162	V	7.125	113.827
163	L	7.004	121.622	163	L	7.069	121.638
164	D	8.156	116.153	164	D	8.171	115.969
165	S	7.936	113.511	165	S	7.977	113.637
166	I	7.347	116.360	166	I	7.400	116.477
167	K	7.340	117.428	167	K	7.376	117.511
168	T	6.522	98.566	168	T	6.564	98.659
169	K	7.420	120.513	169	K	7.438	120.450
170	G	8.820	116.871	170	G	8.977	117.180
171	K	7.740	120.199	171	K	7.780	120.224
172	S	8.283	113.690	172	S	8.295	113.564
173	A	8.957	123.370	173	A	8.957	123.009

174	D	8.560	122.056	174	D	8.569	122.014
175	F	7.932	125.270	175	F	7.965	125.320
176	T	8.041	114.839	176	T	8.060	114.873
177	N	9.660	115.463	177	N	9.660	115.617
178	F	8.202	118.854	178	F	8.230	119.007
179	D	7.938	128.063	179	D	7.940	128.330
181	R	8.351	118.352	181	R	8.383	118.725
182	G	7.229	102.629	182	G	7.291	102.926
183	L	7.186	117.337	183	L	7.225	117.482
184	L	6.564	113.091	184	L	6.598	113.025
186	E	8.118	116.309	186	E	8.124	116.461
187	S	7.557	111.280	187	S	7.626	111.391
188	L	8.587	125.623	188	L	8.593	125.589
189	D	7.353	121.914	189	D	7.342	121.828
190	Y	8.700	119.055	190	Y	8.690	118.937
191	W	9.510	117.243	191	W	9.554	117.332
192	T	9.608	115.000	192	T	9.648	115.031
193	Y	8.040	125.669	193	Y	8.097	125.760
195	G	8.742	110.060	195	G	8.742	110.060
196	S	8.635	119.914	196	S	8.599	119.896
199	T	6.925	108.390	199	T	6.978	108.501
202	L	9.116	120.467	202	L	9.174	120.531
203	L	5.905	110.834	203	L	5.960	110.913
204	E	8.724	124.593	204	E	8.711	124.537
205	C	7.566	114.617	205	C	7.608	114.512
206	V	7.172	119.145	206	V	7.179	119.310
207	T	8.380	124.915	207	T	8.411	125.043
208	W	8.146	129.612	208	W	8.173	129.698
209	I	8.746	127.339	209	I	8.792	127.444
210	V	9.629	128.348	210	V	9.652	128.352
211	L	8.984	126.843	211	L	9.038	126.987
212	K	7.299	122.028	212	K	7.322	122.172
213	E	8.832	121.610	213	E	8.861	121.813
215	I	8.874	113.317	215	I	8.887	113.238
216	S	8.028	116.520	216	S	8.051	116.425
217	V	8.480	115.836	217	V	8.473	115.894
218	S	8.342	115.802	218	S	8.357	115.888
219	S	9.240	117.280	219	S	9.253	117.354
220	E	8.325	118.330	220	E	8.325	118.330
221	Q	7.468	117.929	221	Q	7.497	117.953
222	V	7.264	114.618	222	V	7.381	114.546
223	L	7.947	120.748	223	L	7.996	120.768
224	K	6.845	115.111	224	K	6.929	115.335

225	F	7.055	118.666	225	F	7.113	118.871
226	R	6.620	108.944	226	R	6.688	109.090
227	K	6.794	114.082	227	K	6.855	114.065
228	L	7.038	119.295	228	L	7.094	119.295
229	N	8.846	118.433	229	N	8.872	118.380
230	F	8.808	117.199	230	F	8.833	117.193
231	N	8.097	110.597	231	N	8.129	110.554
232	G	8.611	104.753	232	G	8.628	104.688
233	E	8.330	121.282	233	E	8.294	121.837
234	G	9.023	114.096	234	G	9.034	114.134
235	E	7.288	119.950	235	E	7.322	120.002
237	E	8.196	123.252	237	E	8.200	123.231
238	E	8.963	129.956	238	E	8.998	129.945
239	L	8.671	125.884	239	L	8.669	125.874
240	M	8.645	123.021	240	M	8.662	123.096
241	V	6.658	114.898	241	V	6.692	115.331
242	D	7.088	114.777	242	D	7.171	115.003
243	N	8.332	119.337	243	N	8.353	119.377
244	W	6.406	114.138	244	W	6.429	114.153
245	R	10.504	127.179	245	R	10.554	127.190
247	A	7.927	120.386	247	A	7.928	120.802
248	Q	8.684	122.076	248	Q	8.717	122.000
250	L	8.621	125.810	250	L	8.630	125.729
251	K	8.807	114.459	251	K	8.803	114.701
253	R	7.244	118.504	253	R	7.282	118.479
254	Q	8.518	118.171	254	Q	8.473	118.332
255	I	8.860	124.384	255	I	8.870	124.398
256	K	8.710	126.568	256	K	8.729	126.573
257	A	8.319	123.566	257	A	8.410	123.565
258	S	8.904	116.433	258	S	8.928	116.340
259	F	6.597	117.955	259	F	6.650	118.044
260	K	7.691	124.386	260	K	7.692	124.329

9

Appendix B: Copper Chapter

Table 9.1: PCS values (in ppm) used in the plot regarding the copper(II)-TM-hCAII bound to oxalate.

Residue number	Residue name	Residue Atom	PCS (ppm)
25	GLY	H	0.055
26	GLU	H	0.029
33	ILE	H	0.05
35	THR	H	0.005
39	LYS	H	0.001
40	TYR	H	-0.015
41	ASP	H	-0.033
43	SER	H	-0.001
44	LEU	H	0.002
47	LEU	H	-0.01
48	SER	H	-0.029
49	VAL	H	-0.03
51	TYR	H	0.008
52	ASP	H	0.006
54	ALA	H	0.009
55	THR	H	0.052
56	SER	H	0.041
57	LEU	H	0.103
59	ILE	H	0.061
71	ASP	H	0.035
72	ASP	H	0.031
74	GLN	H	0.017
76	LYS	H	0.048
80	LYS	H	-0.021
81	GLY	H	-0.021
82	GLY	H	-0.025
84	LEU	H	-0.028
86	GLY	H	-0.005
87	THR	H	-0.022
102	GLY	H	-0.01
103	GLN	H	-0.029

104	GLY	H	-0.056
110	ASP	H	0.011
114	TYR	H	-0.01
115	ALA	H	-0.033
116	ALA	H	-0.084
123	TRP	H	-0.03
125	THR	H	-0.038
126	LYS	H	-0.027
127	TYR	H	-0.012
128	GLY	H	-0.008
129	ASP	H	-0.029
130	PHE	H	-0.019
131	GLY	H	-0.026
132	LYS	H	-0.043
133	ALA	H	-0.05
134	VAL	H	-0.054
135	GLN	H	-0.07
136	GLN	H	-0.03
139	GLY	H	-0.093
145	ILE	H	-0.243
147	LEU	H	-0.081
148	LYS	H	-0.04
149	VAL	H	-0.048
151	SER	H	-0.033
152	ALA	H	-0.02
156	LEU	H	-0.006
157	GLN	H	-0.003
159	VAL	H	0.038
160	VAL	H	0.029
161	ASP	H	0.012
162	VAL	H	0.049
163	LEU	H	0.02
165	SER	H	0.036
166	ILE	H	0.017
167	LYS	H	-0.016
168	THR	H	-0.005
170	GLY	H	-0.019
171	LYS	H	0.014
175	PHE	H	0.067
176	THR	H	0.024
177	ASN	H	0.043
178	PHE	H	0.062
179	ASP	H	0.018

182	GLY	H	0.018
183	LEU	H	0.038
186	GLU	H	-0.024
187	SER	H	-0.027
188	LEU	H	-0.031
191	TRP	H	-0.02
195	GLY	H	-0.052
203	LEU	H	-0.105
208	TRP	H	-0.113
211	LEU	H	-0.062
213	GLU	H	-0.036
216	SER	H	-0.043
217	VAL	H	-0.012
219	SER	H	-0.018
220	GLU	H	-0.022
221	GLN	H	-0.02
222	VAL	H	-0.062
225	PHE	H	-0.046
226	ARG	H	-0.065
227	LYS	H	-0.017
229	ASN	H	-0.026
230	PHE	H	-0.071
231	ASN	H	-0.032
232	GLY	H	-0.035
233	GLU	H	-0.019
235	GLU	H	-0.011
237	GLU	H	-0.032
238	GLU	H	-0.002
239	LEU	H	-0.05
240	MET	H	-0.094
241	VAL	H	-0.131
242	ASP	H	-0.101
250	LEU	H	0.051
251	LYS	H	0.038
253	ARG	H	0.045
254	GLN	H	0.006
255	ILE	H	0.001
256	LYS	H	0.002
257	ALA	H	-0.037
258	SER	H	-0.029
259	PHE	H	-0.015
260	LYS	H	-0.017

2022

JOSÉ PEDRO MALANHO DA SILVA

Ultra-high resolution structure determination of transition metal substituted human carbonic anhydrase 2 - inhibitor complexes

UNIVERSIDADE NOVA DE LISBOA
OMNIS CIVITAS CONTRA SE DIVISA NON STABIT