Computational and EPR studies of the conformational dynamics of the Aurora-A kinase protein activation loop

A thesis submitted to The University of Manchester for the
degree of a Doctor of Philosophy in the Faculty of
Engineering and Physical Sciences

2017

Maria Grazia Concilio

School of Chemistry

Table of contents

7. Paper four: Detection of Ligand-induced Conformational Changes in the Activation Loop of Aurora-A Kinase by PELDOR Spectroscopy ………………………………...145

List of Figures

List of abbreviations

Abbreviation	
ADP	adenosine diphosphate
APE	Ala–Pro–Glu
ATP	Adenosine triphosphate
CPU	central processing unit
CW	continuous-wave
DBI	Davies-Bouldin Index
DEER	double electron-electron resonance
DFG	Asp–Phe–Gly
DFT	density functional theory
DNA	deoxyribonucleic acid
DQC	double-quantum coherence
EFG	electric field gradient
ELDOR	electron-electron double resonance
ENDOR	electron nuclear double resonance
EPR	electron paramagnetic resonance
ESE	electron spin echo
ESEEM	electron spin echo envelope modulation
FRET	Förster resonance energy transfer
FP	Fokker-Planck
GC	Gauge correction
GGA	generalized gradient approximation
GHz	gigahertz
GIAO	gauge including atomic orbital
GPU	graphic processing unit
GTO	Gaussian type orbitals
HF	Hartree-Fock
HYSCORE	hyperfine sublevel correlation
ICA	integrated computational approaches
iRED	isotropic reorientational eigenmode
LSDA	local spin density approximation
MD	molecular dynamics
MHz	megahertz
MMM	modelling of Macromolecular

Abbreviation	
MOMD	microscopic-order-macroscopic-disorder
MW	micro-wave
NMR	nuclear magnetic resonance
NVE	microcanonical ensemble
ODE	ordinary differential equation
OZ	orbital Zeeman
PCA	principal component analysis
pFS	pseudo F-Statistic
PDB	protein data back
PDF	probability density function
PDS	pulsed dipolar spectroscopy
PELDOR	pulsed electron–electron double resonance
PME	particle mesh Ewald
PT	parallel tempering
QM	quantum mechanics
RESP	restrained electrostatic potential
RNA	ribonucleic acid
RMC	relativistic mass correction
RMSD	root mean square deviation
RMSFs	root mean square fluctuations
SANDER	simulated Annealing with NMR-derived Energy Restraints
SDSL	site-directed spin labelling
SIFTER	single frequency technique for refocusing dipolar couplings
SLE	stochastic Liouville equation
SNR	signal-to-noise ratio
SOC	spin-orbit coupling
SPC/E	extended simple point charge
SRLS	slowly relaxing local structure
ST	simulated tempering
STO	Slater-type orbital
TIP3P	transferable Intermolecular Potential 3 Point
TPX2	microtubule-associated protein

List of symbols

Symbols	
$a_{m,l}$	time-dependent amplitudes
A	hyperfine -tensor with components A_{xx}, A_{yy} and A_{zz}
α	fine structure constant
B_0	external magnetic field
$\vec{B}$	magnetic field vector with B_{xx} , B_{yy} and B_{zz}
$B(t)$	background in the DEER signal as function of the time (t)
$C_{m,l}$	Time-correlation function
γ	gyromagnetic ratio
D	dipolar tensor
$\hat{D}$	dipolar tensor
D_R	rotational diffusion constant
$D_{km}^{(l)}$	Wigner matrix of rank l and indices k and m
$\varDelta$	Modulation depth
$\varDelta E$	Energy difference
$\varDelta g^{GC}$	gauge correction term
$\varDelta g^{OZ/SOC}$	orbital Zeeman and spin-orbit coupling correction term
$\varDelta g^{RMC}$	relativistic mass correction term
-e	charge of the electron
E	energy
ϵ_{ij}	Lennard-Jones parameter
$F(t)$	Factor form in the DEER signal as function of the time (t)
g	g-tensor with components g_{xx}, g_{yy} and g_{zz}
g_e	correction factor
g_{iso}	isotropic g-factor
g_n	nuclear g-factor
h	Planck constant
$\hbar$	reduced Planck constant
$\hat{H}$	Hamiltonian
$\hat{\hat{H}}$	superoperator
$Y_{l,m}(\theta,\varphi)$	Spherical harmonics as function of the angles θ and φ
l	angular momentum quantum number
I	current

Symbols	
I_0	echo intensity in absence of the dipolar interactions
$\hat{I}$	nuclear angular momentum operator with components $\hat{I}_{xx}$, $\hat{I}_{yy}$ and $\hat{I}_{zz}$
J	coupling constant
$\vec{L}$	angular momentum vector
$\hat{L}$	angular momentum operator
$\hat{\hat{L}}(\vec{x},t)$	Liouvillian as function of space points $\vec{x}$ and time
k	Boltzmann constant
$\hat{\hat{R}}$	kinetic superoperator
$K(t,r)$	Kernel function of time (t) and distance (r)
n_1, n_2	spin populations two states
m	mass
m_e	electron mass
m_l	magnetic quantum number of the angular momentum
m_s	magnetic quantum number for the spin angular momentum
μ	reduced mass
μ_B	Bohr magneton constant
μ_l	orbital magnetic dipole moment
μ_J	total magnetic moment
μ_N	nuclear magneton
μ_s	spin magnetic dipole moment
μ_0	vacuum permeability
ω_s	Larmor frequency
ω_{AB}	dipolar coupling frequency
$P(r)$	distance distribution as function of the distance (r)
$\mathrm{P}(\vec{x},\hat{\rho},t)$	probability density as function of space points $\vec{x}$ in different spin states $\hat{\rho}$ and time
$\vec{p}$	momentum vector
q	charge
$\hat{\hat{Q}}_{km}$	25 operators term
ψ	wave-function
φ_μ , φ_υ	basis set in the $\varDelta g^{RMC}$ and $\varDelta g^{GC}$
∇	gradient

Symbols	
r	distance between charges q
$\hat{\hat{R}}$	relaxation superoperator
$\vec{r}$	position vector
$\hat{\rho}(t)$	density matrix as function of the time
s	spin angular momentum quantum number
$S(t)$	simulated time-domain signal in function of the time (t)
S^2	order parameter
$\vec{S}$	Spin angular momentum vector
$\hat{S}$	Spin angular momentum operator with components $\hat{S}_{xx}$, $\hat{S}_{yy}$ and $\hat{S}_{zz}$
σ	shielding constant
σ_{ij}	the Lennard-Jones parameter
t	time
t_{max}	maximum dipolar evolution time
T	temperature
T_2	transverse relaxation time
$\hat{T}$	kinetic energy operator
τ_R	rotational correlation time
τ_m	rotational correlation time of the loop motion
$V(r)$	electrostatic potential
$V(t)$	DEER signal in function of the time (t)
ν	frequency
$\vec{v}$	velocity vector
Z_{eff}^{n}	The effective nuclear charge of atom n

Abstract

Aurora-A is a miotic serine/threonine protein kinase that controls many cellular pathways, its function is implicated in a wide range of diseases including cancer and its activity is tightly regulated by changes in conformation of a conserved region known as the activation loop. The structure of the Aurora-A has been extensively studied by X-ray crystallography and in many cases the position of the activation loop in solution is unknown. Therefore, studies of Aurora-A's activation through characterisation of its conformations in solution are important to enhance understanding of molecular processes related to diseases and to support the discovery of small molecule kinase inhibitors.

In this project, continuous-wave (CW) and pulsed electron paramagnetic resonance (EPR) techniques were employed to characterize and resolve the conformational dynamics of the activation loop of Aurora-A kinase. The protein used for this study was spin labelled with the (1-oxyl-2,2,5,5-tetramethyl-pyrroline-3-methyl) methane-thiosulfonate spin label (MTSL) introduced *via* site-directed mutagenesis within the activation loop.

Computational methods, including quantum mechanical (QM) calculations based on density functional theory (DFT) and classical molecular dynamics (MD) simulations were carried out to determine conformational states in which both the activation loop and spin label were engaged, as well as structural fluctuations, order parameters and rotational correlation times related to the motion of the full spin-labelled system. The theoretical data obtained were used for the interpretation of the room temperature CW EPR spectra of the spin labelled Aurora-A kinase. Comparisons made between simulated and experimental spectra revealed that the motion of the protein and spin label occurred on a comparable timescale and ranged between 0.1 ns and 10 ns, indicating that the dynamics of the Aurora-A was dominated by events occurring on long timescales. A comparison between MD data related to un-phosphorylated and phosphorylated activation loops revealed that the latter was endowed by less conformational movement. In both the cases, the region that showed the most conformational dynamics was that localized between residues 282 and 294, which includes residues 287 and 288 that are crucial for the catalytic activity of Aurora-A kinase.

Pulsed electron–electron double resonance (PELDOR) measurements were carried out to determine the average distances and distance distributions between spin-label pairs. Broad distance distributions and multiple populations were identified, indicating high flexibility of the structure. Variations in the distance distributions were observed upon addition of different

protein kinase inhibitors and binding partners, which were attributed to differences in conformational space accessible to the activation loop that were found to be in good agreement with previous studies using X-ray crystallography. This study provided evidence for structural adaptations that could be used for drug design and a methodological approach that has potential to characterize inhibitors in development.

Declaration

No portion of the work referred to in this thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Copyright Statement

The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the “Copyright”) and s/he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.

Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made **only** in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.

The ownership of certain Copyright, patents, designs, trademarks and other intellectual property (the “Intellectual Property”) and any reproductions of copyright works in the thesis, for example graphs and tables (“Reproductions”), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.

Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see http://documents.manchester.ac.uk/DocuInfo.aspx?DocID=487), in any relevant Thesis restriction declarations deposited in the University Library, The University Library’s regulations (see http://www.manchester.ac.uk/library/aboutus/regulations) and in The University’s policy on Presentation of Theses.

Acknowledgments

I would like to thanks to my advisor, Dr. Alistair J. Fielding for giving me the opportunity to undertake this postgraduate research program and for useful discussions, as well as Bruker Ltd. for funding.

I would like to thanks to Prof. Richard Bayliss and Dr. Selena G. Burgess for having provided samples of the protein studied in this project.

I would like to thanks to Prof. David Collison and Dr. Alexandra Simperler for useful comments on the papers produced during this project, as well as to my office mates and colleagues, especially Dr. Fabrizio Moro, Dr. Amgalanbaatar Baldansuren, Dr. Marie-Emmanuelle Boulon, Graham Heaven and Mike Hollas for useful discussions.

I would like to thanks Prof Klaus Muller-dethlefs for the interesting discussion about *ab intio* calculations, national and international politics at lunch time.

An enormous thanks to Dr. Ilya Kuprov for his advice and support during the preparation of my EPSRC postdoctoral fellowship, as well as to Prof. Gareth Morris and his research associate Dr. Mohammadali Foroozandeh, for their support and comments on the proposal.

To Prof. R. E. L.

When there is a problem….
do not look for reasons or people, work out the solution!

1. Preface

1.1 Project aims

The overall aim was to provide a detailed characterization of the conformational states and dynamics in solution of the phosphorylated and un-phosphorylated activation loops of Aurora-A kinase using computational methods and EPR spectroscopy in order to enhance understanding of the biological function of this enzyme. This project provided a number of challenges and was focused on three main goals:

1. The biological systems that are mostly studied using EPR spectroscopy are ribonucleic acid (RNA) and deoxyribonucleic acid (DNA) fragments (1) and rigid protein domains, such as β-sheets (2) and α-helices (3); a limited number of EPR studies have been performed on loops. (4) The weakly structured nature of loops, associated to their flexibility, make them difficult to study with spectroscopic techniques, especially with nuclear magnetic resonance (NMR) spectroscopy. Therefore, a goal was to provide a characterization of the dynamics of the spin-labelled activation loop of Aurora-A using CW EPR spectroscopy.
2. A number of previous works (5-7) reported that exhaustive sampling of the space accessible to the conformations of protein secondary structure and of the rotamers of the MTSL spin label through classical MD simulations was not a straightforward task. Therefore, an ambitious goal was related to the employment of long MD simulations to accomplish significant sampling.
3. Dipolar spectroscopy has been successfully employed to study kinases, (8) however, this approach has never been applied to kinase activation loops and in this work distance distributions and their variation upon addition of different protein kinase small molecule inhibitors and protein binding partners were characterized.

Results obtained in this project are described in four papers, Chapter 1 and 2 provide the theoretical background and the description of the methods employed to accomplish these goals, respectively.

1.2 Rationale for the alternative format and outline of the thesis

This thesis is presented in an alternative format in which each chapter is written in the style typically required for journal publication, papers 1 and 4 (Chapters 4 and 7) are published peer-reviewed articles, while the publication of the original research described in papers 2 and 3 (Chapters 5 and 6) is forthcoming. As mentioned in the presentation of the Thesis policy of the University of Manchester, the choice of writing a thesis in alternative format presents several benefits that are:

1) It reduces the conflict of interest between the exigence of publishing results before the end of the PhD project and the need to complete the thesis itself, since both can be achieved simultaneously.
2) It avoids the redundancy of rewriting articles for journal publication after the examination of the thesis, end also eliminates the risk of self-plagiarism that might arise from such rewriting.
3) It encourages the publication of results during the Ph.D. project and greatly facilitates publication afterwards and the examination of the thesis itself.

In this thesis work, chapter two provides a detailed introduction about kinase proteins and theoretical background related to computational and experimental work described in the papers and chapter three provides a description of the experimental and theoretical methodologies employed in this project. Further details are given in the introduction of each paper, a brief outline of chapters four to seven (papers one to four) is given below:

Paper one: “Density functional theory studies of MTSL nitroxide side chain conformations attached to an activation loop”.

M.G.Concilio, A.J. Fielding, R. W. Bayliss, S. G. Burgess *Theor. Chem. Acc.* **2016**, *135*, 1-6.

This paper describes a characterization of the conformations of MTSL attached to a fragment extracted from the activation loop of Aurora-A kinase using a QM method, rooted on DFT. The work was performed in order to determine the rotamer population and evaluate the Gibbs energy of rotamers in order to establish starting structures of MTSL for MD studies.

Contribution of the authors: R. W. Bayliss, A. J. Fielding and M. G. Concilio designed the research. S. G. Burgess made the MTSL-spin labelled Aurora-A used for EPR spectroscopy. M. G. Concilio designed and organized the work described in this paper, performed the spectroscopic experiments, the MD and DFT calculations, calculated the theoretical EPR spectra and analysed the data. M. G. Concilio wrote the paper with contributions from A. J. Fielding, S. G. Burgess and R. W. Bayliss.

Paper two: "MD and EPR studies of the structure and dynamics of the MTSL spin-labelled activation loop of the Aurora-A kinase".
M. G. Concilio, A. J. Fielding, R. W. Bayliss, S. G. Burgess.

This paper describes the characterization of the conformational dynamics of the Aurora-A kinase's activation loop using clustering, principal component (PC) and isotropic reorientational eigenmode (iRED) analyses to determine conformational states, structural fluctuations, order parameters and rotational correlation times of the motion of the activation loop and of MTSL attached to the activation loop. Theoretical data obtained were then used for the interpretation of the CW EPR spectra of the spin-labelled protein that were used to establish the timescale of the dynamics of the system.

Contribution of the authors: R. W. Bayliss, A. J. Fielding and M. G. Concilio designed the research. S. G. Burgess made the MTSL-spin labelled Aurora-A used for EPR spectroscopy. M. G. Concilio designed and organized the work described in this paper, performed the spectroscopic experiments, the MD and DFT calculations, calculated the theoretical EPR spectra and analysed the data. M. G. Concilio wrote the paper with contributions from A. J. Fielding, S. G. Burgess and R. W. Bayliss.

Paper three: "Characterization of conformational states and dynamics of phosphorylated activation loops of the Aurora-A kinase".
M. G. Concilio, A. J. Fielding, R. W. Bayliss, S. G. Burgess.

This paper reports a combined study employing both MD simulations and CW EPR spectroscopy to characterize conformational states and dynamics of phosphorylated and un-phosphorylated activation loops of Aurora-A kinase in solution. This work was performed in order to validate results obtained in previous work using X-ray crystallography that showed that phosphorylated activation loops are more rigid than un-phosphorylated activation loops.

Contribution of the authors: R. W. Bayliss, A. J. Fielding and M. G. Concilio designed the research. S. G. Burgess made the MTSL-spin labelled Aurora-A used for EPR spectroscopy. M. G. Concilio designed and organized the work described in this paper, performed the spectroscopic experiments, the MD and DFT calculations, calculated the theoretical EPR spectra and analysed the data. M. G. Concilio wrote the paper with contributions from A. J. Fielding, S. G. Burgess and R. W. Bayliss.

Paper four: "Detection of Ligand-induced Conformational Changes in the Activation Loop of Aurora-A Kinase by PELDOR Spectroscopy".

G. Burgess, M. G. Concilio, R. W. Bayliss, S., A. J. Fielding, *ChemistryOpen* **2016**, *5*, 531 – 534.

This paper describes a study of the conformational states of Aurora-A kinase using PELDOR to measure distances and distance distributions between pairs of MTSL spin label attached at positions T284 and T288 (activation loop) and E170 (helix), and their variation upon the addition of different protein kinase inhibitors and binding partners.

Contribution of the authors: R. W. Bayliss, A. J. Fielding and M. G. Concilio designed the research. S. G. Burgess made the MTSL-spin labelled Aurora-A used for EPR spectroscopy and designed the inhibitors. M. G. Concilio performed the experimental PELDOR measurements, the analysis of the data and contributed to the writing of the paper.

2. Introduction

2.1 Types and function of the Kinase proteins

Protein kinases are enzymes that catalyse the transfer of phosphate from adenosine triphosphate (ATP) to an acceptor residue, commonly serine, threonine or tyrosine, within a substrate protein. This function regulates many cellular pathways, including metabolism, transcription (step of gene expression, in which a segment of DNA is copied into RNA), cytoskeletal (basic structural framework of the cell) rearrangement and apoptosis (process of the cell death). Kinase proteins also have function in embryonic development and in the nervous and immune system. Abnormal phosphorylation causes many human diseases, including cancer. (9-12) In the last decade, pharmaceutical companies and academic institutions have reported the development of Aurora kinase inhibitors that induces a mitotic arrest leading to the death of the cell (Figure 2.1.1) (12).

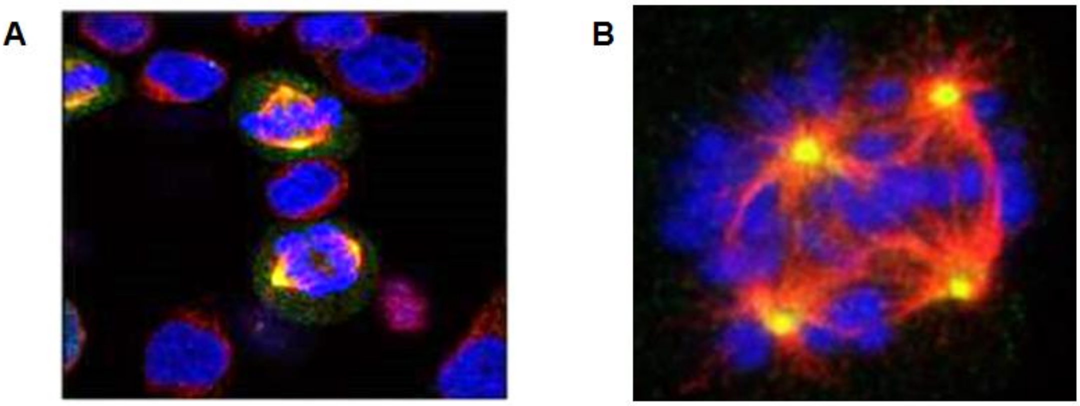


Figure 2.1.1: Aurora-A kinase involved in the interphase step of the mitosis process. (A) Untreated system, (B) process arrested by the inhibitor. Picture adapted from (12).

The first Aurora kinase was discovered in *Drosophila* (9) and because its mutations resulted in a failure of centrosome separation, leading to the formation of monopolar spindles, it was given the name “Aurora” reminiscent of the Aurora Borealis. (10) The evolutionarily conserved Aurora family, comprises Aurora-A, B, and C in humans. Aurora-A kinases represent a family of serine/threonine kinase proteins crucial for the cell cycle control and regulate multiple aspects of mitotic progression (process of the cell cycle in which a cell duplicates itself into two genetically identical daughters), controlling several mitotic events of centrosome (microtubules involved in the mammalian cell division) separation, bipolar spindle formation (mitotic phase in which centrosomes move to opposite sites in the cell forming a bipolar array of microtubules), chromosome segregation (a step in cell reproduction in which chromosomes pair off with their similar homologous) and cytokinesis (cytoplasmic division of a cell at the end of mitosis).

Human Aurora-A is overexpressed in a wide range of human cancers including breast, colorectal, ovarian, and glioma.(12, 13) Aurora-B kinase functions in a multi-protein mitotic complex, which monitors and regulates chromosome segregation. (14) Aurora-C kinase is a chromosomal passenger protein that can complement Aurora-B kinase function in mitotic cells. (12) The deregulation of Aurora kinase activity can result in mitotic abnormality and genetic instability, leading to defects in centrosome function, spindle assembly, chromosome alignment, and cytokinesis.[4] Both the expression level and the activity of Aurora kinases are found to be up-regulated in many kinds of human cancers. Mechanistic studies of kinase activation have been undertaken to build an understanding of the molecular processes that underpin diseases and to facilitate the discovery of small-molecule kinase inhibitors and anti-cancer drugs. (15, 16)

2.2 Aurora-A kinase's activity and inhibition

The catalytic activity of the Aurora-A has been widely studied through X-ray crystallography and kinetic measurements. (17, 18) Aurora-A's activity is tightly regulated by changes in conformation of a conserved region known as the activation loop, its sequence is shown in Figure 2.2.1.

ASP-PHE-GLY-TRP-SER-VAL- HIS-ALA-PRO-SER-SER-ARG-ARG-TPO-TPO-LEU-ALA-GLY-THR-LEU-ASP-TYR-LEU-PRO-PRO-GLU

274-275-276-277-278-279-280-281-282- 283-284-285- 286 -287 -288-289-290-291-292-293-294- 295-296 -297- 298-299

Figure 2.2.1: Sequence of the phosphorylated activation loop of the Aurora-A kinase (B), showing the 26 amino acids present: Aspartate (ASP), phenylalanine (PHE), glycine (GLY), tryptophan (TRP), serine (SER), valine (VAL), histidine (HIS), alanine (ALA), proline (PRO), arginine (ARG), threonine (THR), leucine (LEU), tyrosine (TYR), glutamate (GLU). TPO (in red) represents the phosphorylated threonine residues at position 287 and 288.

In order to observe catalysis, protein kinases have to be activated first, this process occurs when a phosphate from ATP is transferred to two adjacent threonine residues (pT287 and pT288) within the activation loop. The consequent changes in conformation upon phosphorylation lead to specific arrangements of the full protein that can favour or prevent the substrate binding in the active site of the protein (green region in Figure 2.2.2), with a consequent increase or reduction in the catalytic efficiency of the Aurora-A kinase.

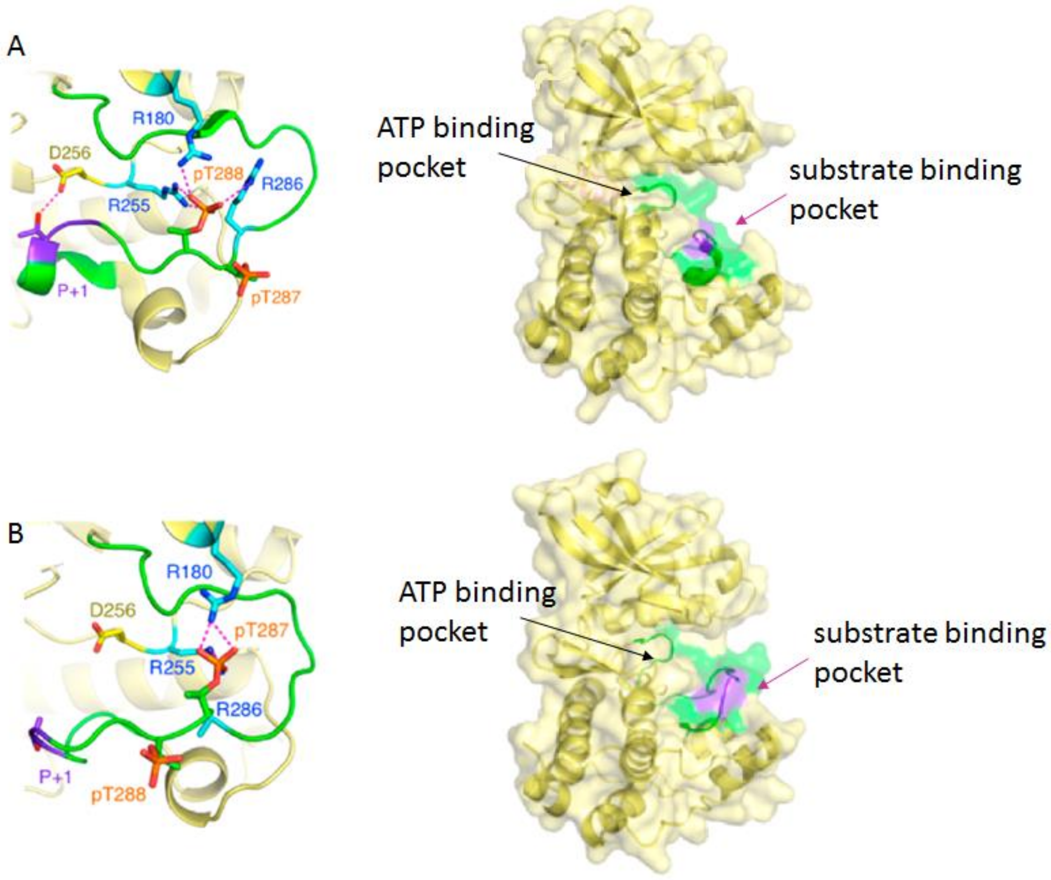


Figure 2.2.2: (A) Active conformation of the activation loop (in green, on the left) extracted from the X-ray crystal structure of Aurora-A (full structure with PDB: 1OL5, on the right). (B) Inactive conformation of the activation loop extracted from the X-ray crystal structure of the Aurora-A (full structure with PDB: 4BN1, on the right).[1] The region highlighted in purple (P+1) contains the residues: aspartate, phenylalanine, and glycine.

In particular, there is a dynamic region within the activation loop that contains aspartate (ASP), phenylalanine (PHE), glycine (GLY), called DFG (purple region in Figure 2.2.2), that regulates catalysis and contributes to ATP binding. The Aurora-A kinase exists in active and inactive states related to the conformations of the activation loop, Figure 2.2.2A shows an active conformation in which the phospho-threonine at position 288 is rotated inward and binds the arginine pocket formed by residues R180, R255 and R286. The P+1 loop interacts with a nearby aspartate, D256, which arranges the activation loop in a suitable conformation for protein substrate binding, highlighted in purple in the surface representation. This conformational state is also called DFG-in conformation. The ASP

---

[1] Picture adapted from reference [17]

residue is oriented toward the ATP binding pocket and is able to coordinate the magnesium ion to bind the *β*- and *γ*-phosphate groups of the ATP.

In the inactive conformation of the activation loop (Figure 2.2.2B), the phospho-T287 interacts with the residues R180 and R255 in the arginine pocket, and there are no interactions between the P+1 and D256, in this case the activation loop is arranged such that substrate binding is obstructed. In the conformational state, the DFG motif is flipped outward, such that the ASP no longer coordinates the magnesium for the catalytic site.

Therefore, the Asp-Phe-Gly (DFG) motif plays an important role in the regulation of kinase activity, structure-based drug design is performed to design compounds able to interact with the DFG motif. Aurora-A inhibitors are classified as either Type I, which bind to the kinase in its active state, or Type II, which bind to an inactive DFG-out conformation of the kinase and occupy an additional hydrophobic pocket within the active site. In this work, the effect of two main families, the imidazo [1,2-a] pyrazine-based (19) and imidazo [4,5-b] pyridine-based (20) inhibitors of inhibitors was studied. The chemical structures are shown in Figure 2.2.3.

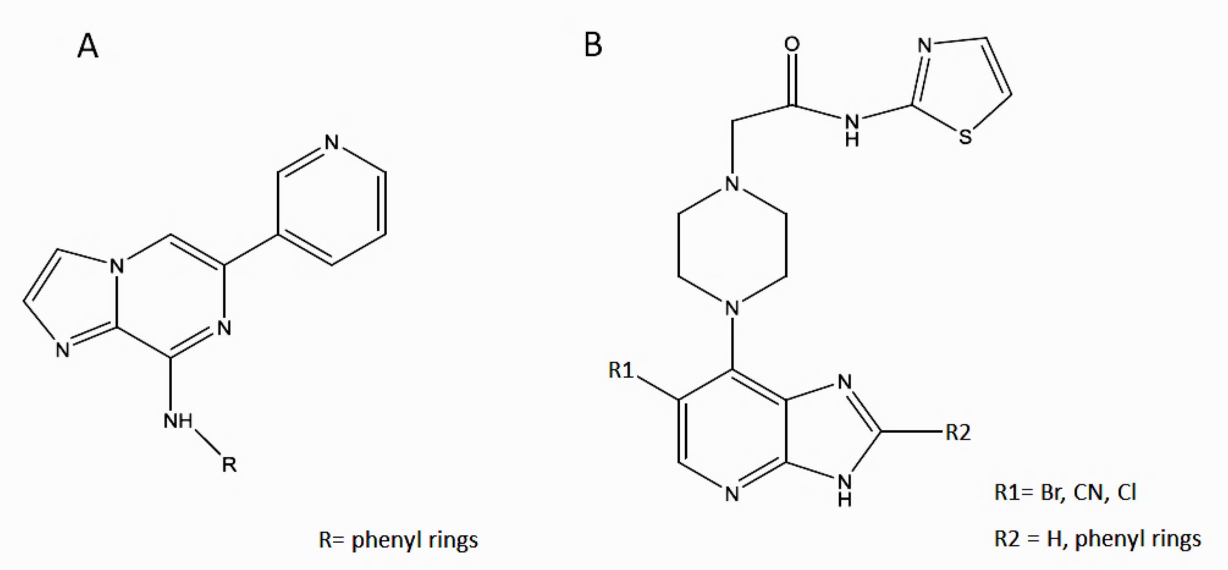


Figure 2.2.3: Chemical structure of the pyrazine-based (A) and imidazo pyridine-based (B) inhibitors.

Both these families act by binding the ATP pocket, forming hydrogen bonding between the N and NH groups and the residues of the Aurora-A. This stop the catalytic activity without affecting the activation loop. Variations in the chemical structure of the inhibitors can be obtained changing the R, R1 and R2 groups for a better affinity to specific pockets. Aurora-A inhibition represents an important intervention strategy for combating proliferative diseases such as cancer, where a central role in mitosis and frequent overexpression make them highly attractive drug targets (15-17, 19, 20). Although no specific Aurora inhibitors have yet reached the market, they have been widely appraised in cancer models and in early stage clinical trials.

Mechanisms of activation and inhibition have been already widely studied in previous work, through kinetic measurements and X-ray crystallography, which provided information about the activity of the protein in the presence and absence of phosphorylation and binding patterns, such as TPX2, used to provide specificity in the catalytic process. However, there are few studies of conformational dynamics of the activation loops in solution, NMR studies for this particular protein, performed by our collaborators (Dr. G. Burgess, Prof. R. W. Bayliss, University of Leeds), provided poor results probably due to the weakly structured nature of the activation loops, this justified the employment of EPR and MD simulations that were used, in this project. The dynamics of the activation loop can be evaluated through the reduction in mobility on phosphorylation and in the variation of the distance distributions (between two MTSL spin labels inserted in two positions within the protein) upon the addition of the inhibitors and binding patters that dock the protein in specific positions (detailed descriptions are given in Chapters 5, 6 and 7).

2.3 History of the electron paramagnetic resonance (EPR) spectroscopy

One of the major spectroscopic techniques that rely on the electron spin, is the electron paramagnetic resonance (EPR) spectroscopy that requires the presence of an unpaired electron in the system and is the only method for the direct detection of paramagnetic species. Thus, it is one of the most powerful tools for the interpretation of the structural properties and dynamics processes of molecules in the condensed phase. The first EPR experiment was performed by Evgeny Zavoisky at Kazan State University in 1944 (21) and after this date a number of improvements were made through the development of more sophisticated hardware and experimental protocols. This lead to new applications spanning across a wide range of areas from quality control to molecular research in fields, such as materials research, structural biology and quantum physics. In particular, EPR experiments have provided valuable information pertaining to metalloprotein structures, photochemistry (22) and photolysis studies (23).

Pulsed techniques (described later in Methodologies), such as electron nuclear double resonance (ENDOR), electron spin echo envelope modulation (ESEEM) and hyperfine sublevel correlation (HYSCORE) are employed to study the geometry of ligands in metal-complexes through measurements of hyperfine couplings to coordinating magnetic nuclei. A very common pulsed EPR technique is double electron-electron resonance (DEER), also known as PELDOR or pulsed electron-electron double resonance (ELDOR) that uses two separate microwave frequencies to measure the dipolar interactions between spin pairs, also and double-quantum coherence (DQC) sequences have been developed for this purpose. (24-28) These techniques have been used to investigate structural changes and conformational flexibility of spin-labeled proteins, protein complexes and membrane proteins (29, 30), and to study inter-qubit interactions in supramolecular arrays of molecular electron spin qubits (31).

From a methodological point of view, EPR has a higher sensitivity at given spin concentration as compared to NMR due to the much larger magnetic moment of electron spins that exceeds that of the proton by a factor of 660. However, the development of the EPR technique has been slower than that of NMR due to the requirement of higher frequencies and shorter relaxation time scales. A recent major improvement of EPR equipment is the use of arbitrary wave generators (AWGs) incorporated into pulse EPR spectrometers (32) for the use of shaped microwave pulses (chirp pulses). Chirp pulses can provide ultra-wide band and uniform excitations to overcome several limitations such as the narrow excitation bandwidth in PELDOR experiments (33) and improve the data quality in DQC and in single frequency technique for refocusing dipolar couplings (SIFTER). (34) Many improvements have been made also in CW EPR spectroscopy to increases the signal-to-noise ratio (SNR) without distorting the signal through employing wavelet transforms (35) and of loop-gap resonators (for CW and pulsed experiments) for frequencies below the X-band (36).

2.4 Angular momentum in quantum mechanics

The following two paragraphs describe the basic theoretical background related to the angular momentum and the magnetic dipole moment of the electron for a correct interpretation of the lines in the EPR spectra. The angular momentum quantum number (*l*) is a very important property in magnetic resonance. Experimental evidence that suggested that particles are characterized by this number was given by hydrogen fine structure and the Stern-Gerlach’s experiment. In classical mechanics (CM), the angular momentum vector is associated with rotational motion and is defined as:

$$\vec{L} = \vec{r} \times \vec{p} = \mathrm{m}|\vec{r} \times \vec{v}| \qquad \text{(Eq. 2.4.1)}$$

where $\vec{r}$ is position vector and $\vec{p} = m\vec{v}$ is the momentum vector of a particle with mass (m) and velocity ($\vec{v}$), while the quantum mechanical angular mometum can be defined as a linear Hermitian[2] angular momentum:

$$\vec{L} = -i\hbar\vec{r}\times\vec{\nabla} \qquad \text{(Eq. 2.4.2)}$$

In both the cases, it is possible to define the total angular momentum operator ($\hat{L}^2$) and its components $\hat{L}_x$, $\hat{L}_y$ and $\hat{L}_z$:

$$\hat{L}^2 = -\hbar^2\left[\frac{\partial^2}{\partial\theta^2} + \cot\theta\frac{\partial}{\partial\theta} + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}\right] \qquad \text{(Eq. 2.4.3)}$$

$$\hat{L}_x = -i\hbar\left[y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right],\quad \hat{L}_y = -i\hbar\left[z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}\right],\quad \hat{L}_z = -i\hbar\left[x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right] \qquad \text{(Eq. 2.4.4-6)}$$

Instead of working with the $\hat{L}_x$ and $\hat{L}_y$, in general, it is easier to define new operators that are given by a linear combinations of $\hat{L}_x$ and $\hat{L}_y$, defined as ladder operators that include raising and lowering operators:

$$\hat{L}_+ = \hat{L}_x + i\hat{L}_y \text{ and } \hat{L}_- = \hat{L}_x - i\hat{L}_y \qquad \text{(Eq. 2.4.7-8)}$$

Angular momentum operators $\hat{L}^2$ and $\hat{L}_z$ have simultaneous eigenfunctions in spherical coordinates that are the spherical harmonics $Y_{l,m}(\theta,\varphi)$:

$$\begin{cases} \hat{L}^2 Y_{l,m}(\theta,\varphi) = l(l+1)\hbar\, Y_{l,m}(\theta,\varphi) \\ \hat{L}_z Y_{l,m}(\theta,\varphi) = m_l\hbar\, Y_{l,m}(\theta,\varphi) \end{cases} \qquad \text{(Eq. 2.4.9)}$$

Eigenvalues of $\hat{L}^2$ and $\hat{L}_z$ are $l(l+1)\hbar$ and $m_l\hbar$, respectively. It is known from quantum mechanics (QM) that *l* has to be an integer positive number and its values are confined between *–1* and *+1*, limiting the choice of m_l. Therefore, in QM angular momentum is "quantized", meaning it comes in discrete amounts, as opposed to the classical case where angular momentum is a continuous variable. This property is applied also to the orbital angular momentum and to the spin angular momentum (s) for an atomic system. The

---

[2] An Hermitian operator is an operator that has real eigenvalues and is composed by orthogonal eigenfunctions.

relationship between the magnitude of the angular momentum and its projection along any direction in space is often visualized in terms of a vector model. In this model, the total angular momentum (indicated with $\vec{J}$) is composed of the orbital angular momentum ($\vec{L}$) and by the spin angular momentum (in general indicated with $\vec{S}$ for electrons and $\vec{I}$ for nuclei), both $\vec{S}$ and $\vec{L}$ process about the direction of the total angular momentum $\vec{J}$ (Figure 2.4.1):

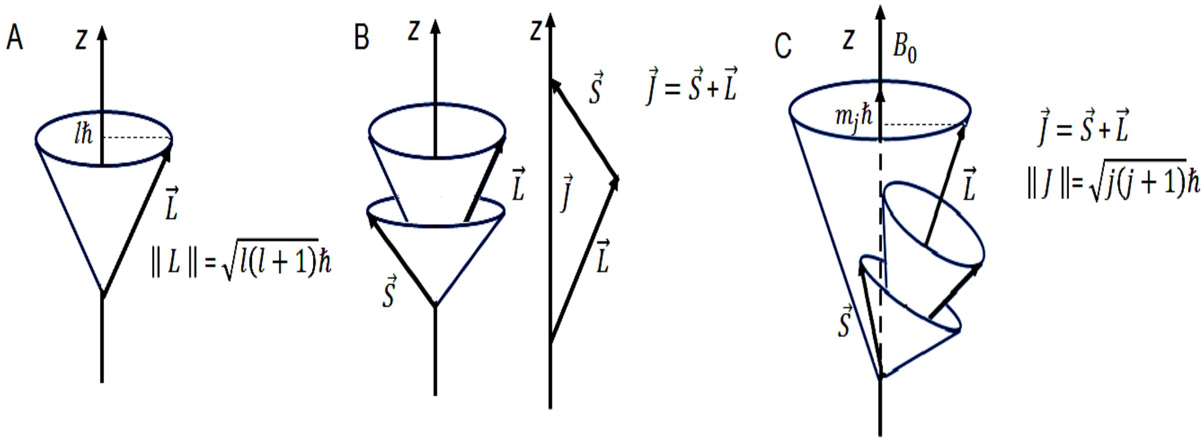


Figure 2.4.1: (A) Vector model for the orbital angular momentum ($\vec{L}$) of an atomic electron, $\|L\|$ represents the magnitude of the vector. (B) Vector model for the total angular momentum ($\vec{J}$) of an atomic electron composed by the orbital angular momentum ($\vec{L}$) and by the spin angular momentum ($\vec{S}$). (C) Total angular momentum in a magnetic field B_0.

Figure 2.4.1 describes the precession of one single electron or multiple electrons for which the spin and orbital angular momenta have been combined to produce composite angular momenta S and L, respectively. Once orbital and spin angular momenta have been combined, the resulting total angular momentum can be visualized as recessing about the external applied magnetic field (B_0).

In the absence of a magnetic field, the $2s + 1$ energy states are degenerate, if a magnetic field is applied, the degeneracy is broken. Considering the case of a $s = \frac{1}{2}$, two energy levels are obtained and in accordance to $m_s = \pm\frac{1}{2}$, the quantized projections along the z-axis are also defined as:

$$\hat{L}_z = -i\hbar \frac{\partial}{\partial\varphi} = i\hbar\, m_s \qquad \text{(Eq. 2.4.10)}$$

and are shown in Figure 2.4.2:

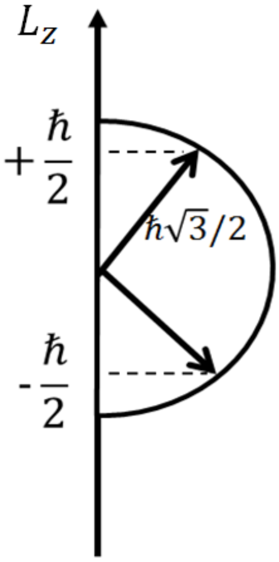


Figure 2.4.2: The orientations of angular momentum along the *z*-axis for *s* = 1/2 for the two permitted values of m_s.

The splitting between the energy states is called the *Zeeman splitting*. Transitions between energy levels represented a further proof of the existence of the spin angular momentum and are at the basis of all the magnetic resonance techniques.

2.5 Orbital and spin magnetic dipole moment of the electron

The magnetic moment of an electron is an intrinsic property that arises from the rotation about its own axis and/or around a circular orbit. Considering an electron with charge (-*e*) and mass (m_e) moving along a circular orbit of radius (*r*) with velocity (v), the orbital magnetic dipole moment of the electron (μ_s) is given by:

$$\mu_l = I\,A = \frac{-ev}{2\pi r}\,\pi r^2 = \frac{-evr}{2} \qquad \text{(Eq. 2.5.1)}$$

where $I = \frac{-e}{T}$ is the current, T and $A = \pi r^2$ are the period and the area of the orbit, respectively. Dividing each term of Eq. 2.5.1 by $\vec{L} = m_e|\vec{r} \times \vec{v}|$ gives:

$$\frac{\mu_l}{\vec{L}} = \frac{-ev\,r}{2m_e v\,r} = \frac{-e}{2m_e} \Rightarrow \mu_l = \gamma\vec{L} \qquad \text{(Eq. 2.5.2)}$$

The quantity $\frac{-e}{2m_e}$ is called the gyromagnetic ratio (γ) of the rotating particle and indicates that a rotating particle is always accompanied by spin angular momentum vector ($\vec{L}$). As shown above (Eq. 2.4.14), the angular momentum is quantized and can assume values that are integer multiples of $\hbar$ along the *z*-axis. The product of the gyromagnetic ratio and $\hbar$ is called the Bohr magneton constant and is defined as $\mu_B = \frac{-e\hbar}{2m_e} = 9.274 \text{ x } 10^{-24}$ A m^2. Rearranging equation Eq. 2.5.1, it is obtained:

$$\mu_l = \frac{-e}{2m_e}\,\vec{L} = \frac{-\mathrm{e}\hbar}{2m_e}\,\sqrt{l(l+1)} = \mu_B\sqrt{l(l+1)} \qquad \text{(Eq. 2.5.3)}$$

The presence of a magnetic moment for a non-zero angular momentum has important consequences for the behaviour of an atom in presence of an external magnetic field (B_0), since it corresponds to non-zero energy of a magnetic dipole. Therefore the energy (*E*) of the interaction between the magnetic moment and the magnetic field along the *z*-axis is given by:

$$E = -B_0\,\mu_l = B_0\,\frac{-e\hat{L}_z}{2m_e} = B_0\,\frac{-e\hbar m_l}{2m_e} \Rightarrow E = \hbar\omega_l m_l \qquad \text{(Eq. 2.5.4)}$$

where the *Larmor* frequency is $\omega_l = \frac{-eB_0}{2m_e}$. Equation 2.5.4 shows that the magnetic energy of an electron depends on the magnetic quantum number m_l and therefore it is quantized. The total energy of the electron in presence of an external magnetic field is given by the magnetic energy plus the energy it had in the absence of an applied field (E_0), consequently the energy of the state *n*+1 is given by:

$$E_{n+1} = E_n + \hbar\omega_l m_l \qquad \text{(Eq. 2.5.5)}$$

As explained above, when a magnetic field is applied the energy states are no longer degenerate and different levels (E_n) are obtained, this is the principle in which magnetic resonance techniques are based. Electrons not only possess the orbital magnetic moment due to orbital motion but also possess magnetic moment due to spin. The spin magnetic moment (μ_S) for a charged particle (of charge *e*) is defined as:

$$\mu_S = \frac{-e}{2m_e}\,\vec{S} = \frac{-e\hbar}{2m_e} = g_e\,\mu_B\sqrt{s(s+1)} \qquad \text{(Eq. 2.5.6)}$$

where s is referred to the spin angular momentum and $g_e \sim 2.0023$ is a correction factor derived from the motion of the electron in the magnetic field. It is a dimensionless quantity and was discovered by Pauli due to the discrepancy between the observed and predicted Zeeman Effect. Therefore, the total magnetic moment is given by:

$$\mu_J = \mu_l + \mu_s = \mu_B\sqrt{l(l+1)} + g_e\mu_B\sqrt{s(s+1)} \qquad \text{(Eq. 2.5.7)}$$

EPR spectroscopy is based on the interaction between an unpaired electron and an external magnetic field and the nucleus of the atom itself. The electron is characterized by the spin quantum number $s = \frac{1}{2}$ and $m_s = \pm\frac{1}{2}$. The energy of the interaction is equal to $E = g_e\ \mu_B m_s$ and the $\varDelta E$ is given by:

$$\varDelta E = h\boldsymbol{\nu} = g_e\,\mu_B\,B_0 \qquad \text{(Eq. 2.5.8)}$$

where is the energy difference, h is the Planck constant and ν_{mw} is the microwave frequency. In a magnetic field B_0, there are two energy states for the electron (Figure 2.5.2):

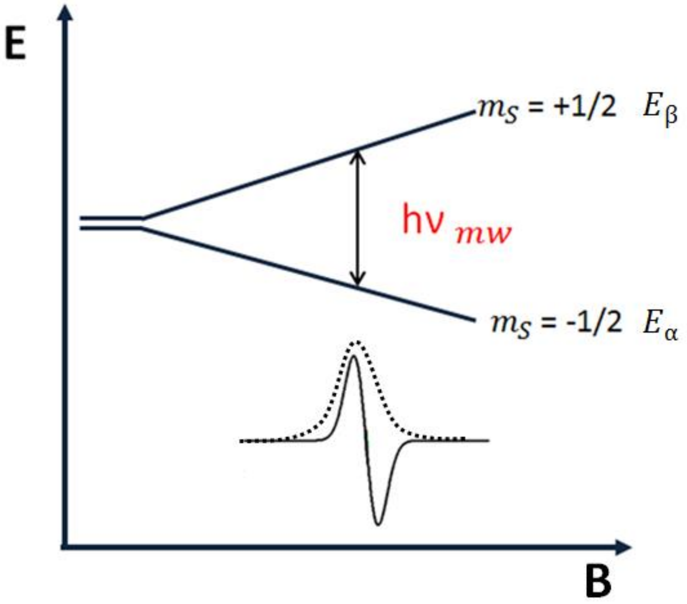


Figure 2.5.2: Representation of the energy states for an electron with spin quantum number s $= \frac{1}{2}$ and magnetic quantum number $m_s = \pm\frac{1}{2}$. Spins flip with the selection rule $\varDelta m_s = \pm 1$, induced by electromagnetic radiation of a certain frequency. The transition between energy levels gives rise to an absorption spectrum (dotted line) that is converted first derivative due to phase sensitive detection with modulation.

The energy levels are expressed by $E_\alpha = +\frac{1}{2} g_e \mu_B B_0$ and $E_\beta = -\frac{1}{2} g_e \mu_{B,} B_0$ and, in this case, only one transition is observed since there are only two energy levels. The spin population is split between two Zeeman levels in accordance to the Maxwell-Boltzmann law:

$$\frac{n_1}{n_2} = e^{\frac{-\Delta E}{kT}} \qquad \text{(Eq. 2.5.9)}$$

where k is the Boltzmann constant, T is the temperature, n_1 and n_2 are the spin populations characterized by $m_s = +\frac{1}{2}$ and $m_s = -\frac{1}{2}$, respectively. At 298 K, at about 3000 G the ratio $\frac{n_1}{n_2}$ is equal to 0.9986, indicating that the populations of the two Zeeman levels are almost equal, but the slight excess in the lower level gives rise to a net absorption. If the electron spin interacts with a non-zero nuclear spin, further transitions are allowed, this kind of interaction is called hyperfine splitting, and the number of lines observed in the EPR spectra (N) is given by:

$$N = (2n_1 I_1 + 1)\,(2n_2 I_2 + 1)\,\ldots\,(2n_n I_n + 1) \qquad \text{(Eq. 2.5.10)}$$

where n represents the number of equivalent nuclei in the system. In this case an electron with spin $s = \frac{1}{2}$ coupled with a nucleus with $I = 1$, six energy levels are observed due to m_I = 0, ±1 and three transitions are allowed (Figure 2.5.3):

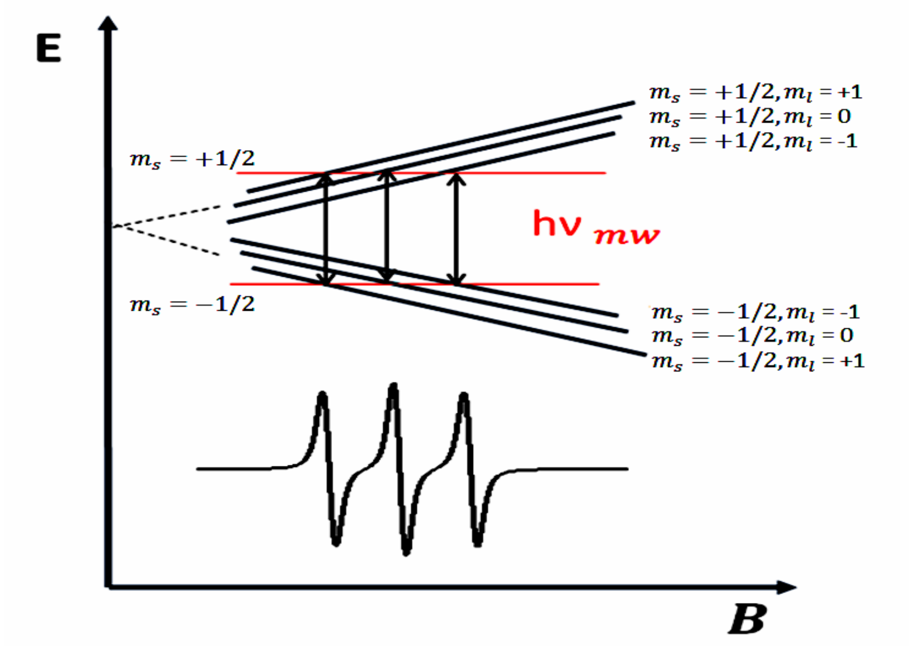


Figure 2.5.3: Representation of the energy states for an electron with spin quantum number s = ½ coupled to a nucleus with nuclear spin number, I = 1, the additional selection rule is $\Delta m_I = 0$.

In this case, the energy of the unpaired electron will depend on the interaction with the magnetic field and nucleus and is defined as:

$$\mathrm{E}(m_{s,}\, m_I) = v_e m_s - v_n m_I + h\, A\, m_s m_I \qquad \text{(Eq. 2.5.11)}$$

The constant A is the hyperfine coupling constant. For simplicity the electron and nuclear Zeeman terms are expressed in frequency units that correspond to $v_e = \frac{g_e\,\mu_B B_0}{h}$ and $v_n = \frac{g_n\,\mu_N\,B_0}{h}$, respectively, where $g_n$is the nuclear g-factor and μ_N is the nuclear magneton. These three transitions give rise to three absorption peaks at different magnetic field positions and are separated by A. These frequencies are accessed in hyperfine techniques such as ENDOR and ESEEM, and are extremely important for measuring very small and unresolved hyperfine couplings.

2.6 Spin Interaction Hamiltonians

Spins interact with their environment and with each other through their associated magnetic moments. Within non-relativistic quantum theory, for the description of a quantum mechanical system, the corresponding interaction Hamiltonians are obtained by replacing magnetic moment vectors μ with the corresponding spin operators (Eq. 2.5.2). In EPR spectroscopy, the energy of the unpaired electron depends on several contributions and the full spin Hamiltonian is given by:

$$\hat{H} = \hat{H}_{EZI} + \hat{H}_{HFI} + \hat{H}_{EX} + \hat{H}_{NZI} + \hat{H}_{ZFI} + \hat{H}_{NQI} + \hat{H}_{NNI} \qquad \text{(Eq. 2.6.1)}$$

where $\hat{H}_{EZI}$ is the spin Hamiltonian representing the electron Zeeman interaction, $\hat{H}_{HFI}$ represents the hyperfine interaction between the electron spin and nuclear spins that is composed by the *Fermi contact* term ($\hat{H}_{FC}$) and electron-nuclear dipole-dipole coupling term ($\hat{H}_{DD}$), $\hat{H}_{EX}$ represents the exchange interaction, the $\hat{H}_{NZI}$ represents the nuclear Zeeman interaction, $\hat{H}_{ZFI}$ represents the zero field splitting, $\hat{H}_{NQI}$ represents the nuclear quadrupole interaction related to spins with nuclear spin quantum number $I > ½$ and $\hat{H}_{NNI}$ represents the spin-spin interaction between pairs of nuclear spins. The Hamiltonian operator can have three generic algebraic types of interaction terms that can be *linear* in the case of coupling to “external” vectors, such as magnetic field and orbital angular momentum, *bilinear* in the case of coupling between spins and *quadratic* in the case of the coupling between a spin and itself.

2.6.1 Zeeman interaction

The Zeeman term ($\widehat{H}_{EZI}$) in Equation 2.6.1 represents the interaction between the electron and the external magnetic field. It can be determined from equation Eq. 2.6.1.1:

$$\widehat{H}_{EZI} = \mu_B \hat{\vec{S}} \cdot g \cdot \vec{B} = (B_x\ B_y\ B_z) \begin{pmatrix} g_{xx} & g_{xy} & g_{xz} \\ g_{yx} & g_{yy} & g_{yz} \\ g_{zx} & g_{zy} & g_{zz} \end{pmatrix} \begin{pmatrix} \hat{S}_x \\ \hat{S}_y \\ \widehat{S}_z \end{pmatrix} \qquad \text{(Eq. 2.6.1.1)}$$

where $\vec{B}$ is the magnetic induction vector with B_x, B_y, B_z components, g is g-tensor that represents the anisotropy of this parameter in the three-dimensional space, $\hat{\vec{S}}$ represents the spin vector expressed in terms of Hermitian Cartesian component operators that are the Pauli matrices $\hat{S}_x$, $\hat{S}_y$ and $\hat{S}_z$ corresponding to spin operators in the three Cartesian projections of spin. The shift from the free-electron value (g_e = 2.002319) is due to three main contributions (37, 38):

$$g = g_e \mathbf{1}_3 + \Delta g^{RMC} + \Delta g^{GC} + \Delta g^{OZ/SOC} \qquad \text{(Eq. 2.6.1.2)}$$

where $\mathbf{1}_3$ is the 3x3 unit matrix, Δg^{RMC} and Δg^{GC} are first-order contributions that represent the relativistic mass (RMC, the observed mass of an object moving with respect to an observer) and gauge (GC) corrections, respectively. The first term can be expressed as:

$$\Delta g^{RMC} = -\frac{\alpha^2}{S} \sum_{\mu,\nu} P_{\mu,\nu}^{\alpha-\beta} \left\langle \varphi_\mu \middle| \hat{T} \middle| \varphi_\nu \right\rangle \qquad \text{(Eq. 2.6.1.3)}$$

where α is the fine structure constants, S the total spin of the ground state, $P_{\mu,\nu}^{\alpha-\beta}$ is the spin density matrix, φ_μ and φ_ν are the basis set and $\hat{T}$ is the kinetic energy operator. The second term can be expressed as:

$$\Delta g^{GC} = -\frac{1}{2S} \sum_{\mu,\nu} P_{\mu,\nu}^{\alpha-\beta} \left\langle \varphi_\mu \middle| \sum_A \xi(r_A) \left[\vec{r}_A\, \vec{r}_0 - \vec{r}_{A,r}\, \vec{r}_{0,s} \right] \hat{T} \middle| \varphi_\nu \right\rangle \qquad \text{(Eq. 2.6.1.4)}$$

where $\vec{r}_A$ is the position vector of the electron relative to the nucleus A, $\vec{r}_0$ is the position vector relative to the gauge origin, $\xi(r_A)$ depends on the effective charge of the nuclei and will be defined below. These two terms are usually small and have opposite signs so that their contributions tend to cancel out. The last term $\Delta g^{OZ/SOC}$ is a second-term contribution arising from the coupling of the orbital Zeeman (OZ) and the spin-orbit coupling (SOC) operators. The OZ contribution in the system Hamiltonian is given by:

$$\hat{H}^{OZ} = \frac{1}{2}\sum_i B_0 \cdot \hat{l}\,(i) \qquad \text{(Eq. 2.6.1.5)}$$

This term is referred to the gauge origin dependence and arises from the angular momentum of the i^{th} electron, $\hat{l}(i)$. The expression was derived using a gauge including atomic orbital (GIAO) approach. (39-41) Finally, the SOC term is a true two-electron operator, but it is often approximated by a one-electron operator involving adjusted effective nuclear charges. This approximation has been proven to work fairly well in the case of light atoms, providing results close to those obtained using more refined expressions for the SOC operator (42):

$$\hat{H}^{SOC} = \sum_{n,i} \xi\left(r_{i,n}\right)\, \hat{l}_n(i) \cdot \hat{S}\,(i) \qquad \text{(Eq. 2.6.1.6)}$$

where $\hat{l}_n$ is the angular momentum operator of the i^{th} electron, $\hat{l}(i)$ relative to the nucleus n and $\hat{s}(i)$ its spin-operator. The function $\xi\left(r_{i,n}\right)$ is defined as (43):

$$\xi\left(r_{i,n}\right) = \frac{\alpha^2}{2}\,\frac{Z_{eff}^{n}}{|r_i - R_n|^3} \qquad \text{(Eq. 2.6.1.7)}$$

where Z_{eff}^{n} is the effective nuclear charge of atom n at position R_n. The isotropic g-value is defined as:

$$g_{iso} = \frac{g_{xx} + g_{yy} + g_{zz}}{3} \qquad \text{(Eq. 2.6.1.8)}$$

The anisotropy of the g-tensor leads to orientation-dependent spectra for single crystals. The variation in g is classified as isotropic ($g_{xx} = g_{yy} = g_{zz}$), axial ($g_{xx} = g_{yy} \neq g_{zz}$) and rhombic ($g_{xx} \neq g_{yy} \neq g_{zz}$). There are also uniaxial, monoclinic and triclinic symmetries.

2.6.2 Hyperfine and dipolar interactions

The $\hat{H}_{HFI}$ term arises from the electron-nucleus interaction, in the case of a single electron, it is defined as:

$$\hat{H}_{HFI} = \hat{\vec{S}} \cdot A \cdot \hat{\vec{I}} = \begin{pmatrix} \hat{S}_x & \hat{S}_y & \hat{S}_z \end{pmatrix} \begin{pmatrix} A_{xx} & A_{xy} & A_{xz} \\ A_{yx} & A_{yy} & A_{yz} \\ A_{zx} & A_{zy} & A_{zz} \end{pmatrix} \begin{pmatrix} \hat{I}_x \\ \hat{I}_y \\ \hat{I}_z \end{pmatrix} \qquad \text{(Eq. 2.6.2.1)}$$

where, $\hat{\vec{S}}$ is the electron spin vector, A is the hyperfine tensor that represents the anisotropy of this parameter in the three-dimensional space and $\hat{\vec{I}}$ is the nuclear spin vector. Likewise with the *g*-tensor, the A-tensor can be classified as isotropic ($A_{xx} = A_{yy} = A_{zz}$), axial ($A_{xx} = A_{yy} \neq A_{zz}$) and rhombic ($A_{xx} \neq A_{yy} \neq A_{zz}$). In a molecule, a single electron is coupled with many nuclei and the Hamiltonian related to this term must be summed over all the contributions from the nuclei and is defined as:

$$\hat{H}_{HFI} = \sum_{k=1}^{m} \hat{\vec{S}} \cdot A_k \cdot \hat{\vec{I}}_k \qquad \text{(Eq. 2.6.2.2)}$$

where one electron is coupled to *m* different nuclei. The hyperfine interaction can be written as the sum of two terms related to the *Fermi contact interaction* ($\hat{H}_{FC}$) term and the electron-nuclear dipole-dipole coupling ($\hat{H}_{DD}$). The Fermi contact term can be derived from *1*st order perturbation theory and is expressed as:

$$\hat{H}_{FC} = A_{iso} \left(\hat{S}_x \hat{I}_x + \hat{S}_y \hat{I}_y + \hat{S}_z \hat{I}_z \right) \qquad \text{(Eq. 2.6.2.3)}$$

where $A_{iso} = \frac{8\pi}{3} g_e g_n \mu_B \mu_N \, |\psi_0 (0)|^2 = \frac{A_{xx} + A_{yy} + A_{zz}}{3}$ is the isotropic hyperfine coupling constant with electron spin density at the nucleus $|\psi_0 (0)|^2$. The Fermi contact interaction represents a direct interaction between the electron and the nucleus, it is isotropic due to the spherical symmetry of the *s* orbitals and does not change under rotations and is averaged to zero in liquid state systems. The electron-nuclear dipole-dipole coupling $\hat{H}_{DD}$ is a through space interaction, unlike the Fermi contact interaction that requires a non-zero electron density at the nucleus. It arises from electrons that have zero electron density at the nucleus and is anisotropic due to the non-spherical nature of the orbitals *p*, *d*, *f*, *etc*. hosting the unpaired electron, and can be derived as described in (44). This term depends on the orientation of the orbitals with respect to the applied magnetic field and does average to zero in non-viscous liquids, it can be expressed as:

$$\hat{H}_{DD} = \frac{\gamma_1\gamma_2\hbar^2\mu_0}{4\pi}\left(\frac{\hat{\vec{S}}\cdot\hat{\vec{I}}}{r_{SI}^3} - \frac{3\left(\hat{\vec{S}}\cdot\vec{r}_{SI}\right)\left(\vec{r}_{SI}\hat{\vec{I}}\right)}{r_{SI}^5}\right) \qquad \text{(Eq. 2.6.2.4)}$$

where $\gamma_1$and γ_2 are the gyromagnetic ratio of the nuctear and electron spins, respectiverly. Remembering the definition of magnetic moment in Eq. 2.5.2, the energy of the interaction between the two magnetic dipoles $\vec{\mu}_1$ and $\vec{\mu}_2$ is obtained replacing the operators, $\gamma_1\hbar\,\hat{\vec{S}}$ and $\gamma_2\hbar\,\hat{\vec{I}}$ with the vectors $\vec{\mu}_1$ and $\vec{\mu}_2$ in the Hamiltonian:

$$E = \frac{\mu_0}{4\pi}\left[\frac{\vec{\mu}_1\cdot\vec{\mu}_2}{r^3} - \frac{3(\vec{\mu}_1\cdot\vec{r})(\vec{r}\cdot\vec{\mu}_2)}{r^5}\right] \qquad \text{(Eq. 2.6.2.5)}$$

This energy depends on the relative orientation of the two magnetic dipoles with respect to $\vec{r}$. Considering Eq. 2.6.2.4 and the following expressions:

$$\hat{\vec{S}}\cdot\hat{\vec{I}} = \hat{\vec{S}}_x\hat{\vec{I}}_x + \hat{\vec{S}}_y\hat{\vec{I}}_y + \hat{\vec{S}}_z\hat{\vec{I}}_z \qquad \text{(Eq. 2.6.2.6)}$$

$$\hat{\vec{S}}\cdot\vec{r}_{SI} = \hat{\vec{S}}_x x + \hat{\vec{S}}_y\, y + \hat{\vec{S}}_z z \qquad \text{(Eq. 2.6.2.7)}$$

$$\hat{\vec{I}}\cdot\vec{r}_{SI} = \hat{\vec{I}}_x x + \hat{\vec{I}}_y\, y + \hat{\vec{I}}_z z \qquad \text{(Eq. 2.6.2.8)}$$

Substituting Eq. 2.6.2.6-8 in Eq. 2.6.2.4, it is obtained:

$$\begin{aligned}\hat{H}_{DD} = {} & \gamma_1\gamma_2\hbar^2 r^{-5}\Big[\hat{\vec{S}}_x\hat{\vec{I}}_x\,(r^2-3x^2) + \hat{\vec{S}}_y\hat{\vec{I}}_y\,(r^2-3y^2) + \hat{\vec{S}}_z\hat{\vec{I}}_z(r^2-3z^2) - \Big(\hat{\vec{S}}_x\hat{\vec{I}}_y + \\ & + \hat{\vec{S}}_y\hat{\vec{I}}_x\Big)3xy - \Big(\hat{\vec{S}}_y\hat{\vec{I}}_z + \hat{\vec{S}}_z\hat{\vec{I}}_y\Big)3yz - \Big(\hat{\vec{S}}_x\hat{\vec{I}}_z + \hat{\vec{S}}_z\hat{\vec{I}}_x\Big)3xz\Big]\frac{\mu_0}{4\pi}\end{aligned} \qquad \text{(Eq. 2.6.2.9)}$$

Expressing equation 2.6.2.9 in matrix form, it is obtained:

$$\hat{H}_{DD} = \omega_{DD}\left\{\begin{bmatrix}\hat{\vec{S}}_x & \hat{\vec{S}}_y & \hat{\vec{S}}_z\end{bmatrix}\begin{bmatrix}\frac{(r^2-3x^2)}{r^2} & \frac{-3xy}{r^2} & \frac{-3xz}{r^2}\\ \frac{-3xy}{r^2} & \frac{(r^2-3y^2)}{r^2} & \frac{-3yz}{r^2}\\ \frac{-3xz}{r^2} & \frac{-3yz}{r^2} & \frac{(r^2-3z^2)}{r^2}\end{bmatrix}\begin{bmatrix}\hat{\vec{I}}_x\\ \hat{\vec{I}}_y\\ \hat{\vec{I}}_z\end{bmatrix}\right\} \qquad \text{(Eq. 2.6.2.10)}$$

where $\omega_{DD} = \frac{\mu_0\,\gamma_1\gamma_2\,\hbar^2}{16\pi^3 r^3}$ is the dipolar coupling constant. Equation 2.6.2.10 can be abbreviated as:

$$\widehat{H}_{DD} = \omega_{DD}\left(\hat{\vec{S}} \cdot D \cdot \hat{\vec{I}}\right) \qquad \text{(Eq. 2.6.2.11)}$$

where D is the dipolar tensor. It is symmetric since $D_{\alpha\beta} = D_{\beta\alpha}$. If the inter-nuclear vector r lies along the z axis, the dipolar tensor becomes diagonal and the principal components of the dipolar tensor are axially symmetric or traceless, *i.e.* $D_{11} = D_{22}$ gives the following expression:

$$D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{bmatrix} \qquad \text{(Eq. 2.6.2.12)}$$

For this reason, in the case of EPR spectra of paramagnetic species having both isotropic and anisotropic hyperfine couplings and undergoing rapid tumbling in a low viscosity solution, the anisotropic term is averaged to zero and the observed hyperfine coupling corresponds to the isotropic part only. In spherical polar coordinates ($\frac{x}{r} = \sin\theta\cos\varphi$, $\frac{y}{r} = \sin\theta\sin\varphi$, $\frac{z}{r} = \cos\theta$), the dipolar Hamiltonian can be expressed in terms of the so-called dipolar alphabet:

$$\widehat{H}_{DD} = \omega_{DD}\left(\hat{A} + \widehat{B} + \hat{C} + \widehat{D} + \hat{E} + \widehat{F}\right) \qquad \text{(Eq. 2.6.2.13)}$$

where:

$$\hat{A} = \hat{\vec{S}}_z\hat{\vec{I}}_z\,(1\text{-}\ 3\,cos^2\,\theta) \qquad \text{(Eq. 2.6.2.14)}$$

$$\widehat{B} = -\tfrac{1}{4}(\,\hat{\vec{S}}_+\hat{\vec{I}}_- - \hat{\vec{S}}_-\hat{\vec{I}}_+)\,(1\text{-}\ 3\,cos^2\,\theta) \qquad \text{(Eq. 2.6.2.15)}$$

$$\hat{C} = -\tfrac{3}{2}(\,\hat{\vec{S}}_+\hat{\vec{I}}_z - \hat{\vec{S}}_z\hat{\vec{I}}_+)\,\sin\theta\,\cos\theta\,e^{-i\varphi} \qquad \text{(Eq. 2.6.2.16)}$$

$$\widehat{D} = -\tfrac{3}{2}(\,\hat{\vec{S}}_-\hat{\vec{I}}_z - \hat{\vec{S}}_z\hat{\vec{I}}_-)\,\sin\theta\,\cos\theta\,e^{i\varphi} \qquad \text{(Eq. 2.6.2.17)}$$

$$\hat{E} = -\tfrac{3}{4}\,\hat{\vec{S}}_+\hat{\vec{I}}_+\,sin^2\,\theta\,e^{-2i\varphi} \qquad \text{(Eq. 2.6.2.18)}$$

$$\widehat{F} = -\tfrac{3}{4}\,\hat{\vec{S}}_-\hat{\vec{I}}_-\,sin^2\,\theta\,e^{2i\varphi} \qquad \text{(Eq. 2.6.2.19)}$$

In the high-field approximation, where the electron Zeeman interaction is stronger than the dipole-dipole interaction (ω_{DD}), the non-secular terms $\hat{C}$ - $\widehat{F}$ can be neglected and only secular ($\hat{A}$) and pseudo-secular ($\widehat{B}$) terms remain. The corresponding Hamiltonian is:

$$\hat{H}_{DD}= \omega_{DD}\ (1-\ 3\,cos^2\theta)\left[\hat{\vec{S}}_z\hat{\vec{I}}_z\ -\ \tfrac{1}{4}(\ \hat{\vec{S}}_+\hat{\vec{I}}_- -\ \hat{\vec{S}}_-\hat{\vec{I}}_+)\right] \qquad \text{(Eq. 2.6.2.20)}$$

EPR distance measurements are in general performed at low-temperature using frozen solutions of spin-labelled proteins, this method is based on the measurement of the static dipole–dipole coupling between the spins of unpaired electrons localized on the NO groups of the nitroxide spin labels.

2.6.3 Dipolar interactions in the case of two electron spins

In the case of the two interacting electron spins $\hat{S}_1$ and $\hat{S}_2$, the static Hamiltonian can be written in terms of Larmor frequencies, ω_1 and ω_2:

$$\hat{H}_0\ (\hat{\vec{S}}_1\ ) = \omega_1\hat{\vec{S}}_1 \text{ and } \hat{H}_0\ (\hat{\vec{S}}_2) = \omega_2\hat{\vec{S}}_2 \qquad \text{(Eq. 2.6.3.1-2)}$$

Connecting equations 2.6.2.20 and 2.6.3.1-2, the following spin Hamiltonian is obtained:

$$\hat{H} = \omega_1\hat{\vec{S}}_1 + \omega_2\hat{\vec{S}}_2\ + \omega_{DD}\ (1-\ 3\ cos^2\theta)\left[\hat{\vec{S}}_{1,z}\hat{\vec{S}}_{2,z}\ -\ \tfrac{1}{4}(\ \hat{\vec{S}}_{1,+}\hat{\vec{S}}_{2,-} - \hat{\vec{S}}_{1,-}\hat{\vec{S}}_{2,+})\right]$$

(Eq. 2.6.3.3)

The coupling between these two spins is shown in Figure 2.6.3.1:

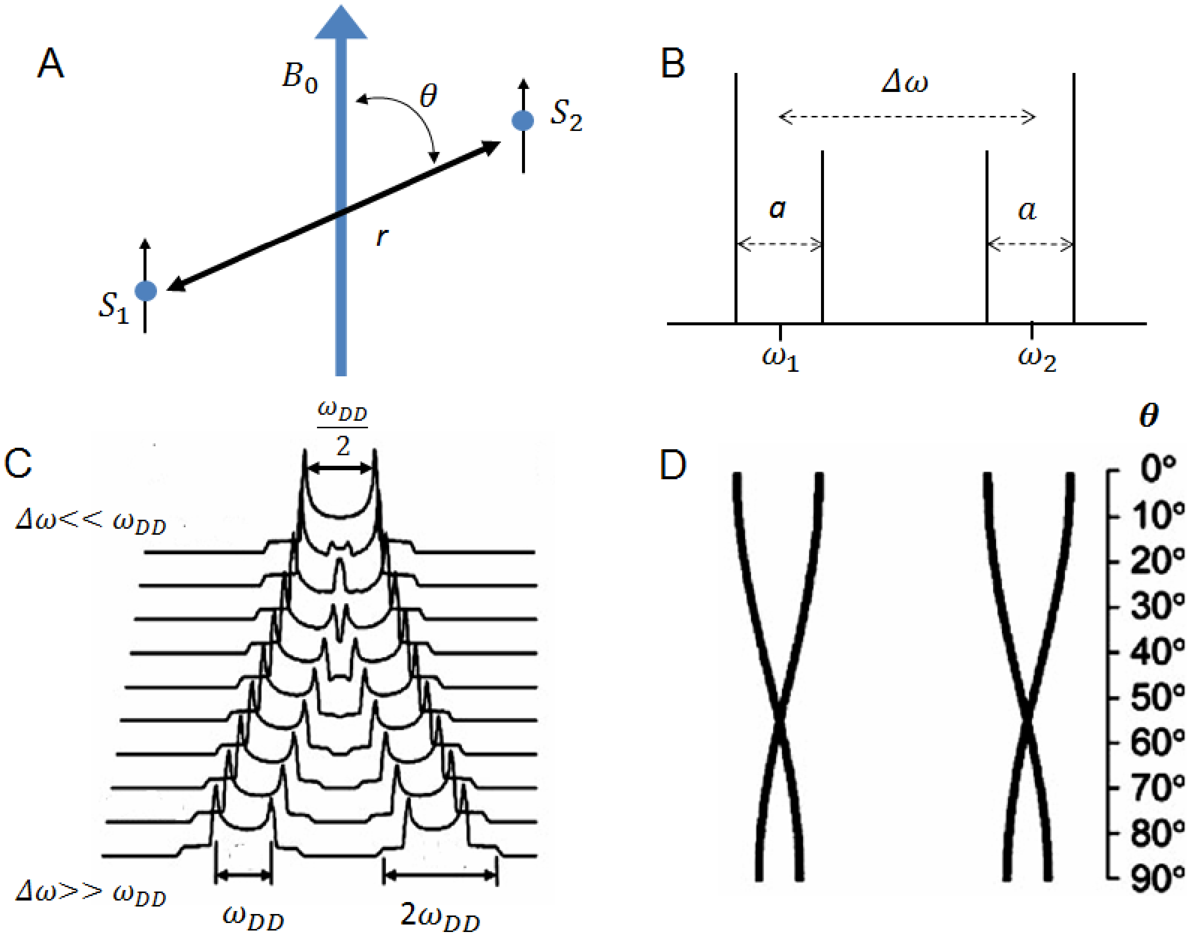


Figure 2.6.3.1: (A) A pair of coupled spins L and S connected by the vector r that is oriented along the z-axis in the molecular frame of reference: B_0 is the external magnetic field and θ is the angle between the direction of the magnetic field and r. (B) Dipolar spectrum of two electron spins resonating at ω_1 and ω_2 that is split by the dipolar coupling $\hat{A} = \omega_{DD}(1 - 3\cos^2\theta)$. (C) Pake pattern resulting from the dipolar coupling moving from weak ($\Delta\omega = |\omega_1 - \omega_2| \ll \omega_{DD}$) to strong coupling ($\Delta\omega = |\omega_1 - \omega_2| \gg \omega_{DD}$). (D) Roadmap showing the angular dependence of a (θ).[3]

In the case of weak coupling ($\Delta\omega = |\omega_1 - \omega_2| \ll \omega_{DD}$), the spectrum is a single Pake doublet with splitting equal to $\frac{\omega_{DD}}{2}$, this spectrum is given by only the secular terms of equation Eq. 2.6.2.13. Figure 2.6.3.1 (B C and D) shows the angular dependence of a on the angle θ, the dipolar spectrum is plotted as function of the angle θ giving all the values from -$2\omega_{DD}$ to +ω_{DD}. In the case of strong coupling ($\Delta\omega = |\omega_1 - \omega_2| \gg \omega_{DD}$), the pseudosecular term becomes important and cannot be omitted from equation 2.6.2.13. The dipolar spectrum is composed by two doublets separated by $\Delta\omega$ and the resonant frequency of each spin is split into a doublet separated by |a|. At intermediate cases ($\Delta\omega \simeq \omega_{DD}$), as $\Delta\omega$ decreases, at $\Delta\omega \simeq \omega_{DD}$, the two spectra start to merge into a single spectrum that is given by equation

---

[3] Figures C and D are adapted from 45. Bobart P. P. Freed J. H. Pulse Dipolar Electron Spin Resonance: Distance Measurements: Springer Berlin Heidelberg; 2013.

2.6.2.13 including both the secular and pseudosecular terms. In general, two electron spins in double labelled molecules resonate at two different frequencies due to the different orientations respect to the B_0 field that in turn provides different values of A- and g-tensors, and yields a single Pake doublet resulting from their common dipolar interaction. Therefore, the distance r is then immediately and accurately obtained from a measurement of ω_{DD}. In some cases, when ω_{DD} is sufficiently large, it can be determined directly from the static broadening (150 K) of the CW ESR (or FT ESR) (46) spectrum but this requires careful spectral simulations. In the CW EPR spectra, the dipolar couplings have a small effect since they are masked by the broadenings caused by the stronger interactions, such as the hyperfine interaction, especially at 9.4 GHz and 150 K. (45) Therefore, smaller couplings ω_{DD} require pulses ESR methods that are discussed below.

2.6.4 Electron-electron exchange interaction

The exchange contribution to the spin Hamiltonian arises from a significant overlap of the orbitals of two spins that lead to inter-electron Coulomb repulsion. The coupling of two atomic localized spins is described by the Heisenberg Hamiltonian[4]:

$$\hat{H}_{EX} = -2J\left[\hat{\vec{S}} \cdot \hat{\vec{I}}\right] \qquad \text{(Eq. 2.6.4.1)}$$

where $\hat{\vec{S}}$ and $\hat{\vec{I}}$ are two spins of ½ and J is the coupling constant. The total spin operator for two particle system is a linear combination of the individual spin operators:

$$\hat{\vec{L}}_{tot} = \hat{\vec{S}} + \hat{\vec{I}} \rightarrow \hat{\vec{L}}^2_{tot} = \hat{\vec{S}}^2 + \hat{\vec{I}}^2 + 2\,\hat{\vec{S}} \cdot \hat{\vec{I}} \qquad \text{(Eq. 2.6.4.2)}$$

where the product:

$$\hat{\vec{S}} \cdot \hat{\vec{I}} = \frac{1}{2}\left[\hat{\vec{L}}^2_{tot} - \hat{\vec{S}}^2 - \hat{\vec{I}}^2\right] \rightarrow$$

$$S \cdot I = \frac{1}{2}\left[\, L(L+1) - I(I+1) - S(S+1) \,\right] = \frac{1}{2}\left[L(L+1) - \frac{3}{2}\right] \qquad \text{(Eq. 2.6.4.3)}$$

---

[4] This Heisenberg Hamiltonian can be simplified to $\hat{H} = -\frac{1}{2}\left[\hat{\vec{S}} \cdot \hat{\vec{I}}\right]$, if it is assumed that the spin has only a z component.

The total angular momentum quantum number (L) can have two possible values equal to 0 and 1,[5] therefore, S·I can be equal to $\frac{1}{4}\,J$ and - $\frac{3}{4}\,J$. The $\hat{\vec{L}}_{tot}$ can have values equal to 0 and 2 since both $\hat{\vec{S}}^2$ and $\hat{\vec{I}}^2$ are equal to ¾. The eigenvalues of the Hamiltonian in equation 2.6.4.1 corresponds to two energy levels equal to $E_S = \frac{1}{4} J$ and $E_T = - \frac{3}{4} J$ corresponding to the singlet and triplet states with degeneracy $2L$+1 = 1 and 3, respectively. The value of J is given by $J = -\,(E_S - E_T)$ considering the wave-functions $\varphi_S(r)$ and $\varphi_I(r)$ *of the two spins* and the eigenstates obtained for L= 0, m_l =0 and L= 1, m_l =0, ±1. Four possible simultaneous eigenstates can be observed from the cross-product of the spin wave-function $\psi_{spin} = \chi_{S/T} = |\uparrow\rangle_S \otimes |\downarrow\rangle_I = |\uparrow\uparrow\rangle, |\downarrow\uparrow\rangle, |\uparrow\downarrow\rangle$ and $|\downarrow\downarrow\rangle$ along the z-axis corresponding to:

$$\hat{H}\,|\uparrow\uparrow\rangle = E_T\,|\uparrow\uparrow\rangle \quad \text{for } L = 1,\ m_l = +1$$

$$\hat{H}\,\frac{|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle}{\sqrt{2}} = E_T\,\frac{|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle}{\sqrt{2}} \quad \text{for } L = 1,\ m_l = 0$$

$$\hat{H}\,|\downarrow\downarrow\rangle = E_T\,|\downarrow\downarrow\rangle \quad \text{for } L = 1,\ m_l = -1$$

$$\hat{H}\,\frac{|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle}{\sqrt{2}} = E_S\,\frac{|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle}{\sqrt{2}} \quad \text{for } L = 0,\ m_l = 0 \qquad \text{(Eq. 2.6.4.4-7)}$$

Associated with the singlet state there is a symmetric wave-function ($\chi_S \rightarrow \frac{|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle}{\sqrt{2}}$), while the triplet states ($\chi_T \rightarrow T_+ = |\uparrow\uparrow\rangle, T_0 = \frac{1}{\sqrt{2}}\,(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle)$ and $T_- = |\downarrow\downarrow\rangle$) have to be anti-symmetric in its spatial indices[6]. The full wave-function ($\psi_{S/T}$) is given by the product of the space and spin wave-functions $\psi_{space}\,(x_1, x_2) \times \psi_{spin}\,(S, I)$, in the case of the singlet state it is equal to:

$$\psi_S = \frac{1}{\sqrt{2}}[\varphi_S(x_1)\varphi_I(x_2) + \varphi_I(x_1)\varphi_S(x_2)]\chi_S \qquad \text{(Eq. 2.6.4.8)}$$

In the case of triplet state it is equal to:

---

[5] The final total angular momentum quantum number (L) of a coupled system can take values of 0, a positive half integer or a whole integer. The allowed values of the total angular momentum differ from the maximum value (given by the sum of the single angular momentum quantum numbers) by one less fundamental unit of angular momentum.

[6] Considering that α = $|\uparrow\rangle$ and β =$|\downarrow\rangle$.

$$\psi_T = \frac{1}{\sqrt{2}}[\varphi_S(x_1)\varphi_I(x_2) - \varphi_I(x_1)\varphi_S(x_2)]\chi_T \qquad \text{(Eq. 2.6.4.9)}$$

The wave overall functions ψ_S and ψ_T must be anti-symmetric (because the electrons are Fermions). The Hartree-Fock energy of the singlet and triplet states is determined from the two-electron Coulomb integral[7]:

$$E_{S/T} = \int \psi_{S/T}^* \, \widehat{H} \, \psi_{S/T} \, dr \, ds \rightarrow$$

$$\int \psi_{spece}^* \, \psi_{spin}^* \, \widehat{H} \, \psi_{space}\psi_{spin} dr \, ds \rightarrow$$

$$\int \psi_{spin}^* \, \psi_{spin} \, ds \int \psi_{space}^* \, \widehat{H} \, \psi_{space} \, dr \qquad \text{(Eq. 2.6.4.10)}$$

Since the spin part is$\int \psi_{spin}^* \, \psi_{spin} \, ds$ = 1, never matter for the energy, therefore the energy is:

$$E_{S/T} = \int \psi_{S/T}^* \, \widehat{H} \, \psi_{S/T} \, dr_1 dr_2 \qquad \text{(Eq. 2.6.4.11)}$$

The Hartree-Fock energy of the singlet (E_S) and triplet (E_T) are different, two electrons have a minor electrostatic repulsion in the triplet state due to the antisymmetric symmetry of the spatial function as opposed to the singlet state. They can be defined as:

$$E_S = \frac{1}{2}\int [\varphi_S^*(x_1) \, \varphi_I^*(x_2) + \varphi_I^*(x_1)\varphi_S^*(x_2)] \, \frac{1}{r_{12}} \, [\varphi_S(x_1)\varphi_I(x_2) + \varphi_I(x_1)\varphi_S(x_2)] dr_1 dr_2 =$$
$$\int |\varphi_I(x_1)|^2 \, \frac{1}{r_{12}} \, |\varphi_S(x_2)|^2 \, dr_1 dr_2 + \int \varphi_S(x_1) \, \varphi_I^*(x_1) \, \frac{1}{r_{12}} \, \varphi_I(x_2)\varphi_S^*(x_2) \, dr_1 dr_2$$

(Eq. 2.6.4.12)

and

$$E_T = \frac{1}{2}\int [\varphi_S^*(x_1)\varphi_I^*(x_2) - \varphi_I^*(x_1)\varphi_S^*(x_2)] \, \frac{1}{r_{12}} \, [\varphi_S(x_1)\varphi_I(x_2) - \varphi_I(x_1)\varphi_S(x_2)] dr_1 dr_2 =$$

---

[7] The one-electron Coulomb integral has the general expression:
$\langle i \, |\widehat{H}| \, j \rangle = \int \chi_i^*(x_1) \, \widehat{H} \, (r_1) \, \chi_j \, (x_1) \, dx_1$
The two-electron Coulomb integral has the general expression:
$[i \, i \, | \, j \, j \,] = \int \chi_i^*(x_1) \, \chi_i \, (x_1) \, \frac{1}{r_{12}} \, \chi_j^* \, (x_2) \, \chi_j \, (x_2) \, dx_1 dx_2$
It is integrating over all the possible locations of the electrons. The term $\chi_i^*(x_1) \, \chi_i \, (x_1)$ indicates the probability of finding electron 1 in the orbital *i* that is located at position x_1, the term $\chi_j \, (x_2) \, \chi_j \, (x_2)$ indicates the probability of finding electron 2 in the orbital *j* that is located at position x_2. The term $\frac{1}{r_{12}}$ is the Coulomb repulsion between electron atx_1 and electron at x_2.
The two-electron exchange integral has the general expression:
$[i \, j \, | \, j \, i \,] = \int \chi_i^*(x_1) \, \chi_j \, (x_1) \, \frac{1}{r_{12}} \, \chi_j^*(x_2) \, \chi_i \, (x_2) \, dx_1 dx_2$

$$\int |\varphi_I(x_1)|^2 \frac{1}{r_{12}} |\varphi_S(x_2)|^2 \, dr_1 dr_2 \text{ - } \int \varphi_S(x_1)\, \varphi_I^*(x_1) \frac{1}{r_{12}} \varphi_I(x_2)\varphi_S^*(x_2) \, dr_1 dr_2$$

(Eq. 2.6.4.13)

Their difference is equal to:

$$E_S \text{ - } E_T = 2\int \varphi_S(x_1)\, \varphi_I^*(x_1) \frac{1}{r_{12}} \varphi_I(x_2)\varphi_S^*(x_2) \, dr_1 dr_2 \Rightarrow 2J \qquad \text{(Eq. 2.6.4.14)}$$

where the coupling constant $J = \int \varphi_S(x_1)\, \varphi_I^*(x_1) \frac{1}{r_{12}} \varphi_I(x_2)\varphi_S^*(x_2) \, dr_1 dr_2$ is the exchange integral that represents the difference in energy between singlet and triplet and represents the magnitude of the interaction, which is equal to *2J*. This shows that the eigenfunctions of the Hamiltonian in equation 2.6.4.1 are the singlet and triplet states and their energy gap obtained from the eigenvalues of the spin Hamiltonian is 2*J*, the effective Hamiltonian Hamiltonian to describe singlet and triplet statesis given by:

$$\hat{H} = \frac{1}{4}(E_S + 3E_T) - (E_S - E_T)\hat{\vec{S}} \cdot \hat{\vec{I}} \qquad \text{(Eq. 2.6.4.15)}$$

In equation 2.6.3.15, the term $E_S + 3E_T$ is constant and $E_S - E_T$ represents the spin-dependent term and is equal to 2*J*. Therefore, the Heisenberg interaction can be expressed using equation 2.6.4.16:

$$\hat{H}_{EX} = -2J\left[\hat{\vec{S}} \cdot \hat{\vec{I}}\right] = \left(\hat{I}_x\hat{S}_x + \hat{I}_y\hat{S}_y + \hat{I}_z\hat{S}_z\right) \qquad \text{(Eq. 2.6.4.16)}$$

Since the spin operator can be defined in terms of Pauli matrices, the Heisenberg interaction can be also expressed in terms of raising/lowering operators, in the case of spin-1/2, the product $\hat{\vec{S}} \cdot \hat{\vec{I}}$ (in Eq. 2.6.4.16) becomes equal to:

$$\hat{\vec{S}} \cdot \hat{\vec{I}} = \text{ - } \textstyle\sum_{i,j} J_{ij}\left[\frac{1}{2}\left(\hat{\vec{S}}_+\hat{\vec{I}}_- - \hat{\vec{S}}_-\hat{\vec{I}}_+\right) + \hat{\vec{S}}_z\hat{\vec{I}}_z\right] = \frac{1}{4}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 2 & 0 \\ 0 & 2 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad \text{(Eq. 2.6.4.17)}$$

Where *i* and *j* represent couples of spins. Therefore, it is possible to conclude that the two-electron exchange interaction effectively manifests itself as an isotropic spin-spin coupling. However, the exchange interaction can be more complicated including contributions that are not isotropic, such as super-exchange contributions originated by spin-orbit coupling observed in the magnetic compounds. The anisotropic contributions are mixed with other interactions, such as the inter-electron dipolar coupling and the zero-field splitting and cannot be separated during the analysis of experimental data. The *J* value becomes large enough to be measured only in the range of strong coupling, while it is ignored in the analysis of the Pake patterns in the case of weak dipolar coupling in which *J* is typically too small and can produce only a small inward or outward shift of the Pake doublet branches (Figure 2.6.4.1):

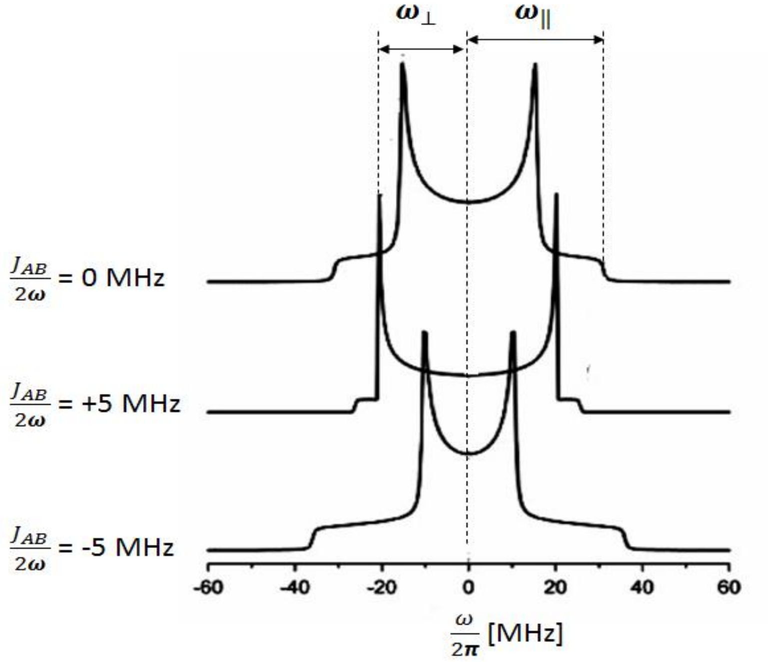


Figure 2.6.4.1: Pake patterns for two spins A and B separated by r_{AB} = 15 Å, $\frac{\omega_{DD}}{2\pi}$ =15.5 MHz and three different values of the exchange coupling constant $\frac{J_{AB}}{2\pi}$.[8]

The values J_{AB} can be determined from the from direct measurement of the peak to peak distances $\omega_{\perp}$and $\omega_{\parallel}$ and is equal to $J_{AB} = \frac{(2\omega_{\perp}+\omega_{\parallel})}{3}$, while the dipolar coupling constant is given by $\omega_{DD} = \frac{(\omega_{\parallel}-\omega_{\perp})}{3}$. In the case of coupling between two spins A and B, the dipolar coupling frequency ω_{AB} between them is equal to:

$$\omega_{AB}=\frac{\mu_0\mu_B^2}{4\pi\hbar}\frac{g_A g_B}{r_{AB}^3}(1-3\cos^2\theta)+J_{AB} \Rightarrow \omega_{DD}(1-3\cos^2\theta)+J_{AB} \qquad \text{(Eq. 2.6.4.18)}$$

---

[8] The diagram has been adapted from Fig. 3 in *Spin-labels and intrinsic paramagnetic centres in the Biosciences:* Orientation-selective DEER using rigid spin labels, cofactors, metals, and clusters. ed. C. R. Timmel and J. Harmer, Springer, 2012.

where μ_B is the Bohr magneton, g_A and g_B are the g-tensors of the spin A and B, respectively. When it is possible to measure the turning points $\omega_{\perp}$and $\omega_{\|}$ in the Pake spectrum in Figure 2.6.4.1, the ω_{DD} and J_{AB} can be separated. (47)

2.6.5 Nuclear Zeeman interaction

This term describes the interaction of a nuclear spin ($\hat{\vec{I}}$) with the external magnetic field ($\vec{B}$) and is described in equation 2.6.4.1:

$$\hat{H}_{NZI} = \gamma\, \hbar\, \vec{B}\, (1\text{-}\sigma)\, \hat{\vec{I}} \qquad \text{(Eq. 2.6.5.1)}$$

where γ is the gyromagnetic ratio constant, σ is the shielding constant that reflects a numeric correction proportional to the static applied external magnetic field considering that a local magnetic field is generated by the motion of the electrons in a molecule. In EPR, this interaction has a small effect on the spectrum but it is a source of interactions in NMR, in general it is isotropic and can be neglected due to the greater electron Zeeman interaction.

2.6.6 Zero-field splitting

The zero-field interaction reflects the strong dipole-dipole interactions between the electrons and is active also in absence of the external magnetic field. It is arises in systems with spin S > ½ such as transition metal ions with up to five unpaired d-electrons (*e.g.* high-spin Mn^{2+}) or spin triplets in organic molecules. The Hamiltonian describing the zero-field splitting in expressed in equation 2.6.5.1:

$$\hat{H}_{ZFI} = \hat{S}^T D\, \hat{S} = (\hat{S}_x\, \hat{S}_y\, \hat{S}_x) \begin{pmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{yx} & D_{yy} & D_{yz} \\ D_{zx} & D_{zy} & D_{zz} \end{pmatrix} \begin{pmatrix} \hat{S}_x \\ \hat{S}_y \\ \hat{S}_z \end{pmatrix} \Rightarrow$$

$$\hat{H}_{ZFI} = D_{xx}\hat{S}_{xx}^2 + D_{yy}\hat{S}_{yy}^2 + D_{zz}\hat{S}_{zz}^2 = D\left[\hat{S}_{zz}^2 - \frac{1}{3}\,\hat{S}\,(\hat{S}+1)\right] + E(\hat{S}_{xx} - \hat{S}_{yy}^2) \qquad \text{(Eq. 2.6.6.1)}$$

where $\hat{S}^T$ indicates the transpose of the spin vector $\hat{S}$, D is the traceless tensor. It is possible to write the Hamiltonian in terms of two constants $D = \frac{3}{2} D_{zz}$ and $E = \frac{D_{xx} - D_{yy}}{2}$. The zero-field interaction depends on the symmetry of the ligand field and the coupling of the electron spins. In the case of cubic symmetry $D = E = 0$ the zero-field splitting is not present.

2.6.7 Nuclear quadrupole interaction

Nuclei with $I \geq 1$ are characterized by a non-spherical charge distribution described by a nuclear electrical quadrupole moment eQ equal to:

$$eQ = \int \rho r^2 \,(3\cos^2\theta - 1) d\boldsymbol{\tau} \qquad \text{(Eq. 2.6.7.1)}$$

where e is the charge, ρ is the charge density in a volume element $d\boldsymbol{\tau}$ at a distance r from the center of the nucleus and making an angle θ to the nuclear spin quantisation axis. The sign of Q indicates the shape of the deformation, negative Q is due to the nucleus being flattened along the spin axis, while a positive Q gives an elongated nucleus. The nuclear electrical quadrupole moment does not interact directly with other spin vectors or the applied magnetic field, but through the electric field gradient (EFG) caused by the electrons and nuclei in close proximity to the nucleus. It is defined as:

$$\widehat{H}_{NQI} = -\frac{1}{6}\, eQ\, \nabla E \qquad \text{(Eq. 2.6.7.2)}$$

where $\nabla E = -\frac{\partial^2 V}{\partial x_i \partial x_j} = V_{ij}$ in which V is the electrostatic potential. If a suitable coordinate system is chosen the EFG can be represented by three principal axes V_{xx}, V_{yy} and V_{zz} with:

$$\eta = \frac{(V_{xx} - V_{yy})}{V_{zz}} \qquad \text{(Eq. 2.6.7.3)}$$

where $V_{zz} \geq V_{yy} \geq V_{xx}$ and $0 \leq \eta \leq 1$. The EFG can be specified by two parameters V_{zz} and η and the Hamiltonian can be written as:

$$\widehat{H}_{NQI} = \frac{e^2 Q}{4I(2I-1)}\left[3\hat{I}_z^2 - I(I+1) + \frac{\eta}{2}(\hat{I}_+^2 + \hat{I}_-^2)\right] \Rightarrow$$

$$\frac{e^2 Q}{4I(2I-1)}\,(\hat{I}_x\; \hat{I}_y\; \hat{I}_x) \begin{pmatrix} \eta - 1 & 0 & 0 \\ 0 & -\eta - 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} \hat{I}_x \\ I_y \\ \widehat{I_z} \end{pmatrix} = \frac{e^2 Q}{4I(2I-1)}\, \widehat{I^T} Q\, \hat{I} \qquad \text{(Eq. 2.6.7.4)}$$

where $I^2 = \hat{I}_{xx}^2 + \hat{I}_{yy}^2 + \hat{I}_{zz}^2$. In the EPR spectrum, the quadrupole interactions manifest themselves as shifts of the allowed resonance lines, these contributions are small and difficult to observe.

2.6.8 Nuclear spin-spin interaction

The dipole-dipole interaction between two nuclear spins $\hat{I}_i$ and $\hat{I}_k$ is expressed in equation 2.6.7.1:

$$\widehat{H}_{NNI} = \hat{I}_i \, d^{(\mathrm{i,k})} \, \hat{I}_k \qquad \text{(Eq. 2.6.8.1)}$$

where $d^{(i,k)}$ is the nuclear dipole-dipole tensor. In EPR spectroscopy this interaction is too small to be detected.

3. Methodologies

This chapter provides theoretical details about the experimental and computational methodologies employed to perform the EPR experiments and analyse the data, further details are reported in the material and methods sections of each paper.

3.1 CW and pulsed EPR methods

EPR experiments can be carried out using CW and pulsed modes. In a CW EPR experiment the spectrum is recorded by putting a sample into a microwave (MW) irradiation field with constant frequency ν and sweeping the external magnetic field, B_0 for a certain range until the resonance condition is fulfilled. In a pulsed EPR experiment, the spectrum is recorded by exciting a large frequency range simultaneously with a single high-power MW pulse of given frequency, ν at a constant B_0. Experiments can be performed using different microwave frequencies that are denoted as S band (3.5 GHz), X band (9.4 GHz), K band (20 GHz), Q band (34 GHz) and W band (94 GHz).

Both the techniques can be performed using samples at room temperature or in frozen state, however, pulsed EPR experiments are generally conducted using frozen solutions and typically based on the detection of the modulation of a spin echo. A Carr-Purcell echo (often mistakenly called “Hahn echo”) employes a $\pi/2$ pulse that brings the magnetization from the z axis to the x-y plane and causes the spins to presess at different rates, subsequently a second π pulse is applied that reverses the dephasing of the magnetization and refocuses the spins along the $-x$ axis, creating an echo (this will be discussed in more detail in paragraph 3.1.2). Spin-echo modulation is characterized by a temporal evolution that is governed by effects of spin relaxation, electron–electron spin dipolar and exchange couplings, and (pseudo-secular) electron-nuclear hyperfine and nuclear quadrupole couplings. There are two distinct mechanisms for spin-echo modulation; the first is the ESEEM mechanism that is applied to coupled heterospins characterized by anisotropic interactions and is therefore restricted to solids or high viscosity liquids; the second mechanism is applied to pairs of homonuclear spins with comparable Larmor frequency that are coupled through exchange or dipole-dipole interaction; the mechanism is at the basis of the four-pulse DEER employed in this work for the distance measurements within the Aurora-A kinase.

3.1.1 Four-pulse DEER

The DEER experiment was developed in Russia during the 1980s by Milov *et al.* (24, 25), subsequently a four-pulse version of the DEER was introduced by the group of Spiess (26) to alleviate issues related to the dead-time (time required to obtain the decay of the signal from high-power transmitter pulses to order to observe the echo signal) encountered in three-pulse DEER. More recently, a five-pulse sequence with suppression of nuclear spin diffusion has been published by the J. H. Freed group in order to suppress nuclear spin diffusion and extend the measurable range of distances. (48) The four-pulse DEER sequence employs detection microwave pulses at frequency (ω_A) that induces an electron spin echo (ESE) in the system that is followed by a pump pulse with frequency (ω_B).

In general, a two-pulse echo detected experiment is performed in order to set the pump and detection pulses before performing the DEER experiments. The echo-detected spectrum of a nitroxide spin label is the superposition of absorption spectra of *x*, *y*, *z* orientations (Figure 3.1.1.1).

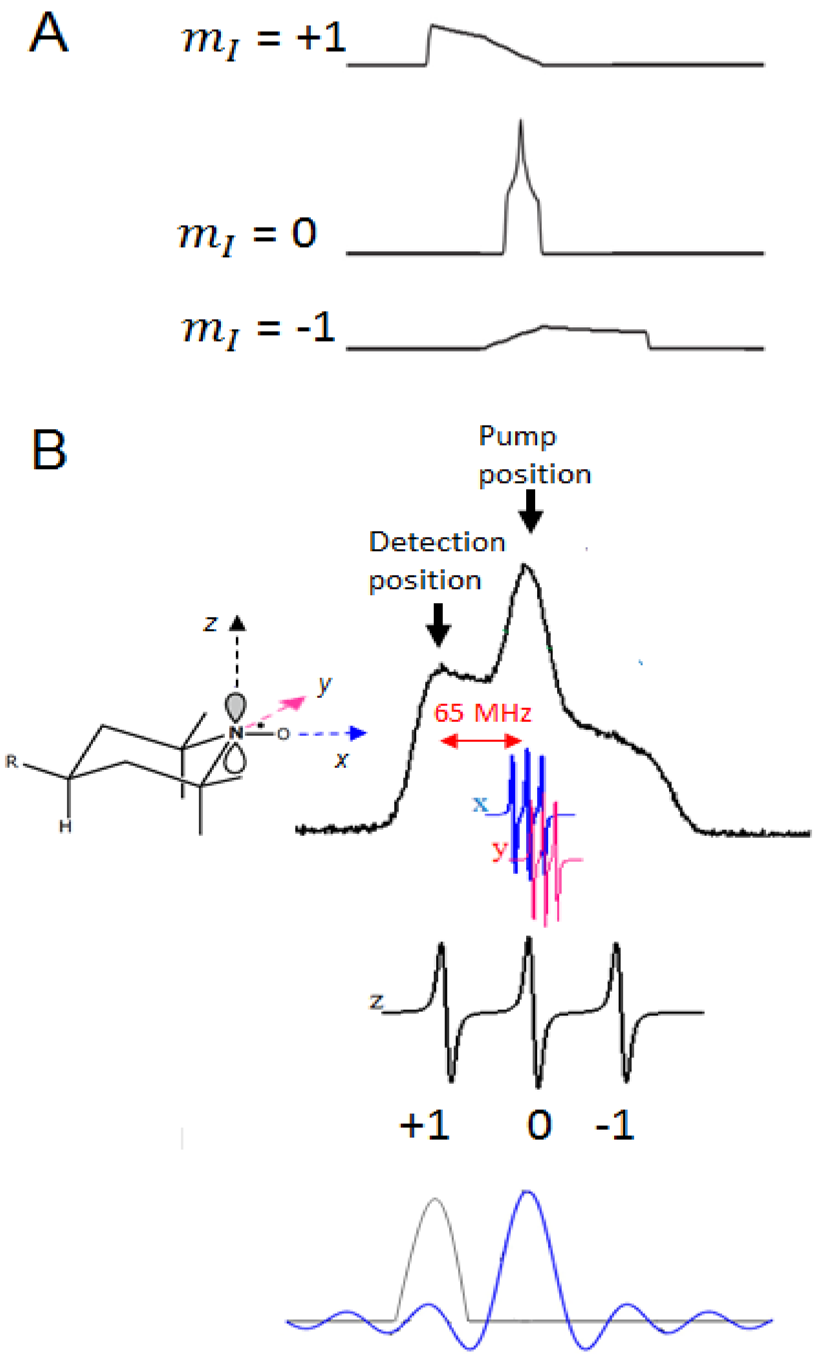


Figure 3.1.1.1: (A) Simulations of the components of the hyperfine-tensor of the nitroxide spectrum[9] (B) Echo detected field sweep spectrum of a nitroxide spin label measured at 9 GHz and 30 K. The spectral components *x* (pink lines), *y* (blue lines) and *z* (black lines) are shown below the spectrum. The arrows indicate the positions of the pump and detection (observe) pulses, their bandwidths are shown below the spectrum, the blue line is the bandwidth of the pump pulse and the grey line is the bandwidth of the detection pulse.

The unpaired electron has electron density delocalized between the oxygen and the nitrogen of the nitroxide group, the coupling between the unpaired electron and the nitrogen generates hyperfine splitting due to the m_I = 0, ±1 states. Frozen solution spectra show

---

[9] This picture was adapted from the tutorial *"Nitroxide Spectral Analysis"* by E. Bordignon.

significant overlap of the three components of the hyperfine. In order to set up a DEER experiment, the bandwidths of the detection and pump pulses (indicated in Figure 3.1.1.1) must not overlap. There are different criteria to choose the experimental parameters in order to obtain the best experimental results. One criteria is based on the optimization of the DEER modulation depth in which the pump pulse is set on the m_I = 0 component that coincides with the maximum of the field-swept nitroxide EPR spectrum (Figure 3.1.1.1) and allows the excitation of all orientations of the nitroxide with respect to the magnetic field. The detection/observer pulse is set on the m_I = +1 component that coincides with the low field edge of the field-swept nitroxide EPR spectrum. The low field edge, is a good candidate since there is a higher signal intensity as compared to placement on the high field edge and there is a difference of about 65 MHz from the pump frequency, which fulfils the non-overlap criterion. The usual experiment set-up of the DEER experiment places the pump pulse at the centre of the resonator, where the MW power, utilizing the greater bandwidth, the detection pulse on the left edge (Figure 3.1.1.2).

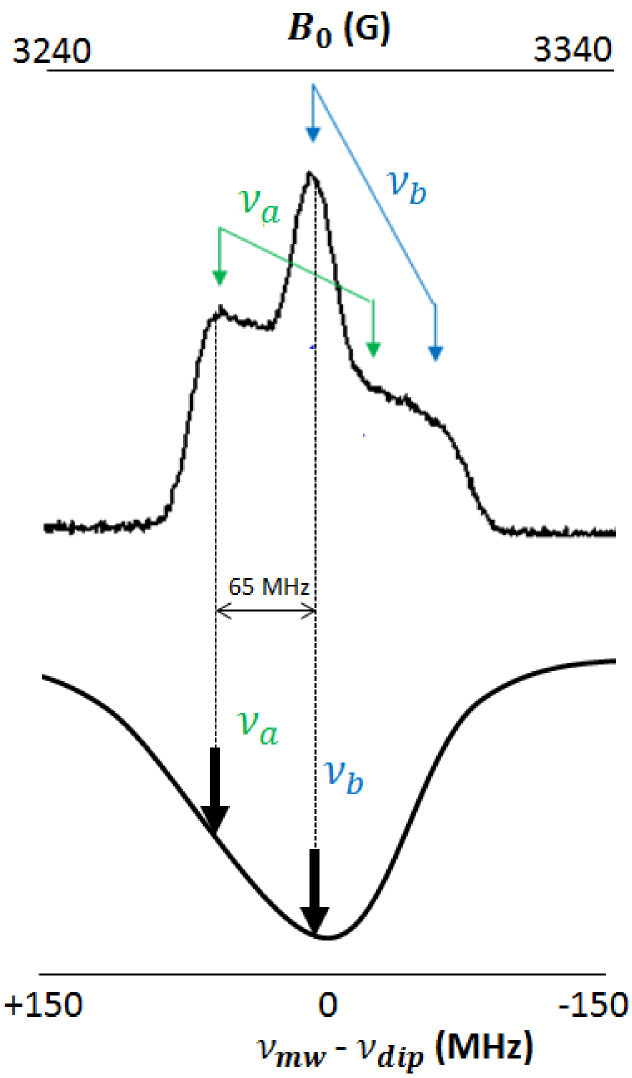


Figure 3.1.1.2: Echo-detected field-swept EPR spectrum of a nitroxide bi-radical and resonator dip profile, showing the optimum observer and pump pulse settings in the four-pulse DEER experiment with a frequency separation is equal to 65 MHz in order to avoid the overlap of the detection (ν_a) and pump frequencies (ν_b).

A nutation experiment is performed before the DEER experiment in order to set the length of the pump pulse. The pulse length obtained is normally shorter than the detection pulse length to maximise excitation with the available MW power.

3.1.2 Four-pulse DEER vector model

The four-pulse DEER experiment allows measurement of the electron-electron dipolar interaction ω_{AB} between two electron spins A and B. The vector diagram of the four-pulse showing the evolution of the magnetization in the DEER experiment is shown in Figure 3.1.2.1:

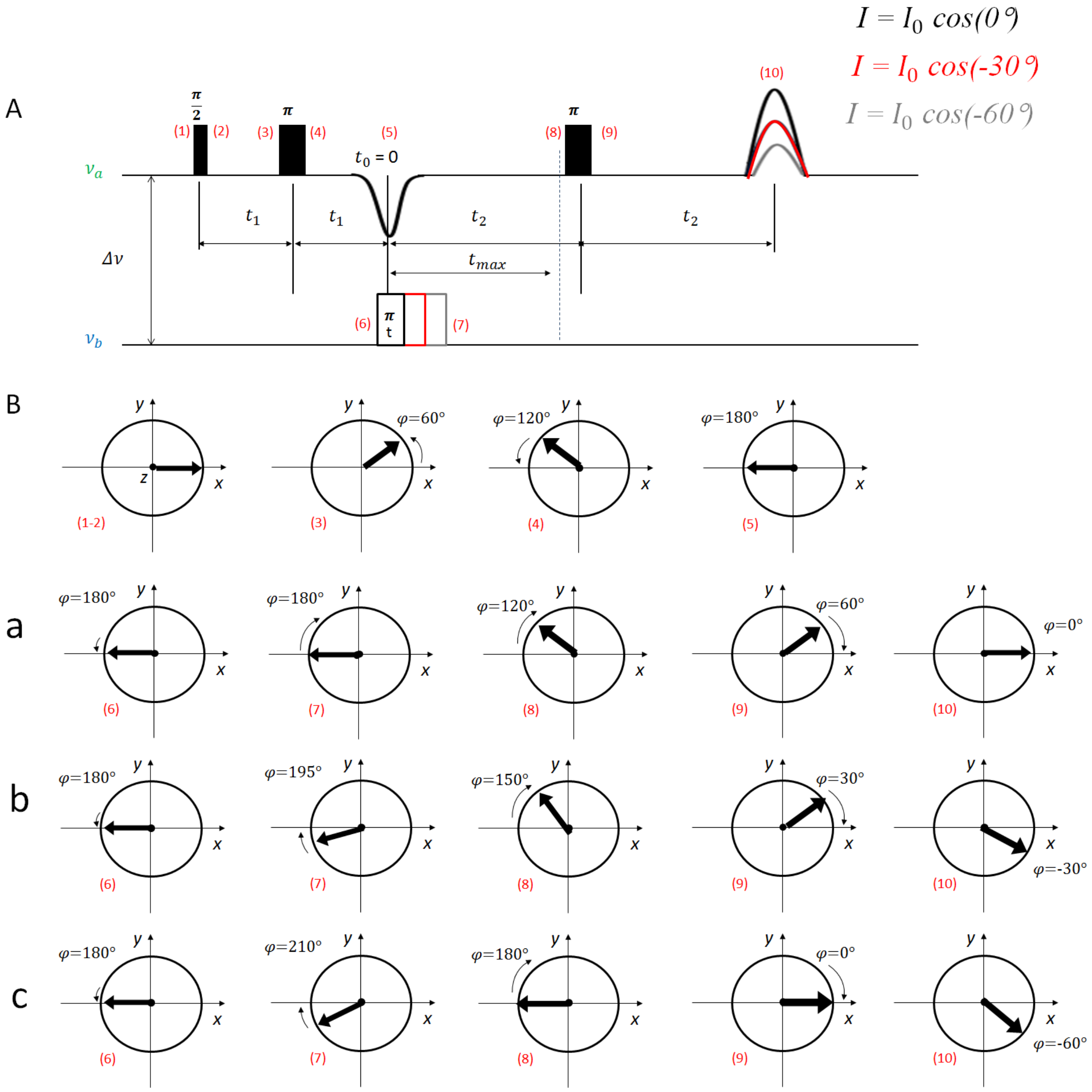


Figure 3.1.2.1: (A) Scheme of the 4-pulse DEER sequence, where ν_a represents the observe frequency and ν_b is the pump frequency, $\Delta\nu$ is the difference between them. t_1 represents the delay between the first $\frac{\pi}{2}$ pulse (1-2) and the second π pulse (3-4), and the second π pulse and t_0, while t_2 is the time between t_0 (5) and the second π pulse (8-9), and between the second π pulse (8-9) and the refocused echo (10). The inversion π pulse (6-7) at the pump frequency ω_b is applied at a variable time t with respect to the first echo and it is varied from t_0 to t_{max} (maximum dipolar evolution time) leading to the variation of the echo intensity. (B) Vector model of the DEER pulse sequence showing the evolution of the spins in the x-y plane after the pulses; a, b and c indicates the evolution of the spin after the **π** pump pulse applied at frequency ω_b at increasing values of t going from a to c.

In the first step of the experiment, spins A are excited with a $\frac{\pi}{2}$ microwave pulse with frequency ν_A along the *y*-axis which brings the spins into the *x*-*y* plane where they start precessing with angular rate of ω_A. Spins are de-phased by the presence of magnetic field inhomogeneities and by different resonance fields and spins start processing with slightly different angular rates. The dipolar field due to the spins B add or subtract from angular rate a quantity equal to $\pm\frac{1}{2}\omega_{AB}$ depending on the spin state of spins *B*. (49) The spins *A* continue to precess in the *x*-*y* plane, in order to extract the dipolar contribution from the total field experienced by A, a microwave pulse π with frequency ν_B is applied at t_0 = 0 to flip spins B and therefore reverses the dipolar contribution from + $\frac{1}{2}$ ω_{AB} to - $\frac{1}{2}$ ω_{AB}, and *vice versa* (Figure 3.1.2.2).

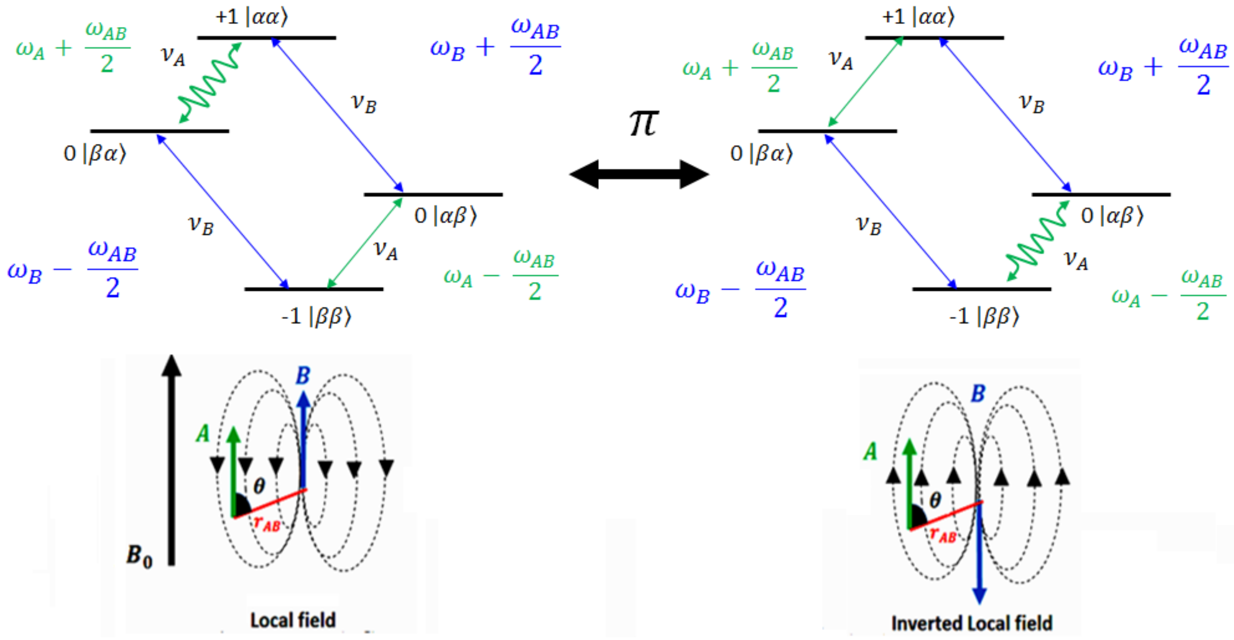


Figure 3.1.2.2: Energy level diagram, showing the π pump pulse inverting the state of spin B coupled to spin A (observer spin). The inversion of the local field at spin A exchanges coherence between the two transitions of spin A that differ in frequency by ω_{AB}.

The primary echo obtained has acquired a phase equal to 0° (Figure 3.1.2.1 B-a) with respect to the rotating frame at ω_A that corresponds to an echo with the maximum intensity. When the pumping pulse is moved from *t* = 0 to *t* = $\frac{1}{4t_2}$ and then *t* = $\frac{1}{2t_2}$, the phase lag is -30° (Figure 3.1.2.1 B-b) and -60° (Figure 3.1.2.1 B-c), respectively. The time t_0 is the time at which the dipolar evolution starts and co-insides with the time at which the the primary echo is generated and the dipolar phase at zero for all A spins. Shifting the starting point for dipolar evolution away from the second pulse by t_1 (as in three-pulse DEER) makes this pulse sequence dead-time-free with respect to dipolar evolution. The echo intensity is reduced to the cosine of the phase lag and the echo intensity oscillates with:

$$I(t) = I_0 \cos[\omega_{AB}(t_1 - t)] \qquad \text{(Eq. 3.1.2.1)}$$

where I_0 is the echo intensity in absence of the dipolar interactions, τ is the time between the $\frac{\pi}{2}$ and the first π pulse, t is the time between t_0 = 0 the and the ELDOR pulse (Figure 3.1.2.1). In the case of an homogeneous distributions of spins, such as a glass solution, equation 3.1.2.1 becomes:

$$I(t) = e^{-kCF_B|t_1 - t|} \qquad \text{(Eq. 3.1.2.2)}$$

where C is the concentration of spins A, F_B is the fraction of spins B excited by the pump pulse and $k = \frac{8\pi^2\mu_B^2 g_A g_B}{9\sqrt{3}\hbar}$. The transverse relaxation of spin A $\left(T_{2,A}\right)$ and couplings to other electron spins that are excited by the observer pulses lead to echo attenuation by a factor $e^{-2k\,(t_1 + t_2)}$, where $k = \frac{1}{T_{2,A}} + K_{ID}$ is the decay rate and $K_{ID} = C_A\ K_A$ is the instantaneous diffusion rate which is proportional to the concentration C_A of A spins. K_A is inversely proportional to the length of the observer π pulses and has an approximate value equal to 0.25 $\text{mM}^{-1}\ \mu\text{s}^{-1}$ for nitroxide spin labels at 9.6 GHz with an observer π pulse of length equal to 32 ns. (50)

3.1.3 Calculation of the DEER signal

The ESE signal decay due to the dipolar coupling can be written as:

$$V(t) = \langle \prod_{\substack{A=1 \\ A\neq B}}^{n} V_{AB}(2\tau) \rangle_{AB} \qquad \text{(Eq. 3.1.3.1)}$$

where V_{AB} = 1- λ $[1 - \cos(\omega_{AB}, t)]$ is the contribution from the A spin to the ESE decay due to the coupling with the B spins, $\langle \ldots \rangle_{AB}$ is the averaging over all the spatial spin distributions. Equation 3.1.3.1 assumes that the ESE is formed at $2t_1$, t is the t is the time between t_0 = 0 the pump pulse (as in Figure 3.1.2.1). The parameter λ corresponds to the fractions of spins A excited by the pump pulse. Considering that in general λ << 1, equation 3.1.3.1 should include also include a (*n*-1) term, hence it is obtained:

$$V(t) = \langle 1-(n-1)\lambda+\lambda\sum_{\substack{A=1\\A\neq B}}^{n}\cos(\omega_{AB},t)\rangle \Rightarrow 1\text{-}\,(n-1)\lambda+\lambda\,\langle\sum_{\substack{A=1\\A\neq B}}^{n}\cos(\omega_{AB},t)\rangle \qquad \text{(Eq. 3.1.3.2)}$$

where the third gives information about the distance distributions between coupled spins and goes to zero as *t* increases. The DEER signal is given by two contributions arising from intermolecular interactions between pairs of spins, and intramolecular interactions between the same kind of spins, the modulation of the echo arises the observed spins experiencing a local magnetic field occurring from dipolar interactions of nearby spins. Therefore, the DEER signal, *V(t)* is composed by the form factor *F(t)* that contains the DEER oscillations and the background function, *B(t)*:

$$V(t) = F(t)\,B(t) \qquad \text{(Eq. 3.1.3.3)}$$

where $B(t) = e^{-k\,t^{\frac{d}{3}}}$ is the background function in which *k* is the density of the coupled spins and *d* is the dimensionality of distribution and is equal to 3 for homogeneous frozen solutions. This function refers to the intermolecular contribution that has to be separated from the intramolecular contribution. In general, it is more advisable to measure the experimental background that is achieved performing DEER measurements of singly-labelled samples that are fitted with polynomial functions that are then subtracted from the double-labelled sample. The form factor has the following expression $F(t) = \frac{V(t)}{V(0)}$, where *V(0)* is the signal observed in absence of the pumping pulse, an ideal form factor is shown in Figure 3.1.3.1:

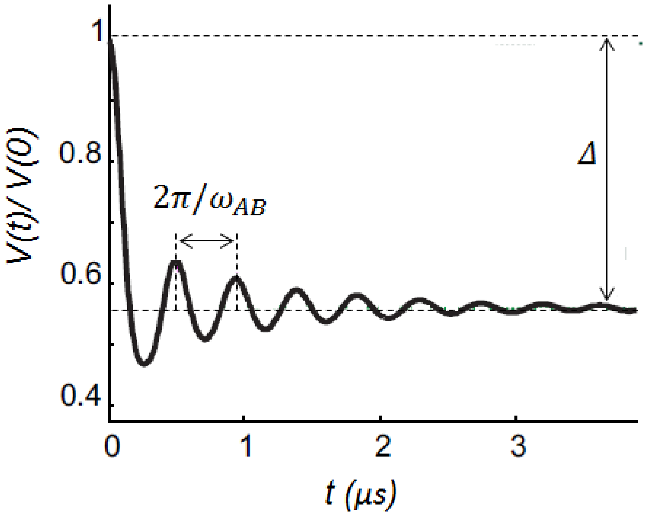


Figure 3.1.3.1: Plot of the $\frac{V(t)}{V(0)}$ in function of the time, with modulation depth, *Δ*.[10]

The modulation depth shown in Figure 3.1.3.1 is connected to the fraction of spins A excited by the pump pulse through the relation:

$$\Delta = 1 - (1-\lambda)^{n-1} \qquad \text{(Eq. 3.1.3.4)}$$

In the case of doubly-labelled proteins n is equal to 2, therefore $\Delta = \lambda$, indicating a correlation between modulation depth and spin-labelling efficiency (f) that is given by the ratio between the number of B spins and the number of possible B spin sites in the protein. This, can be defined such that the modulation depth is equal to λf. Measurements of Δ can provide information about the number of coupled spins (N) through:

$$N = 1 - \frac{\ln(1-\Delta)}{\ln(1-\lambda f)} \qquad \text{(Eq. 3.1.3.5)}$$

The spin labelling efficiency is connected to the sensitivity that in turn depends on the concentration of the sample, the optimal concentration in a typical 9 GHz DEER experiment with an observer π pulse length of 32 ns and pump pulse length of 12 ns is given by $\frac{1.38}{f t_{max}}$ and is equal to 50 μM considering a t_{max} of 25 μs and a complete labelling (f =1). The choice of t_{max} depends on the distance that needs to be measured and, experimentally is limited by $T_{2,A}$. Improvements in the sensitivity are obtained with the extension of $T_{2,A}$, since the signal

---

[10] Figure adapted from: Jeschke, G. DEER Distance Measurements on Proteins. *Annu. Rev. Phys. Chem.* 2012, *63*, 419.

intensity decays exponentially with $\frac{1}{T_{2,A}}$. Therefore, DEER experiments are typically performed in at low temperatures to increase $T_{2,A}$, further improvements in the sensitivity are given by using deuterated buffers and/or protein that further prolongate $T_{2,A}$ and allow distance measurements to 10 nm and beyond. Another important factor connected to the sensitivity is the frequency of measurement, a drastic improvement in sensitivity is obtained from 9 GHz to 34 GHz. Modern 34 GHz spectrometers employ a high-power setup and a probe head that allows oversized samples. Further improvements are obtained at 94 GHz but the SNR decreases due to the reduction of the sample volume and due to the less favourable ratio between excitation bandwidth and spectral width that leads to dramatic loss in spin inversion efficiency (51) but this can be corrected using a larger MW power. [43]

3.1.4 Analysis of the DEER traces

Distances between two spin labels are distributed into conformation distributions of the label and protein backbone, and this distribution can be characterized by its mean distance, width, and shape. Due to the dependence of inversion efficiency on the ratio between excitation bandwidth and on the dipole-dipole coupling, conformations with very short distances are underrepresented in the DEER signal. (52, 53) The reliability of the distance distribution depends very much on the maximum dipolar evolution time, in the case of $t_{max} = 2$ µs, the shape of the distance distribution is reliable up to a distance of about 3 nm, the mean distance $\langle r \rangle$ and the width σ_r are reliable up to a distance of 4 nm, whereas the mean distance, but not the width is reliable between 4 and 5 nm.

Therefore, the upper limit for obtaining an accurate mean distance can be approximated by $r_{max,\langle r \rangle} \sim 5 \sqrt[3]{t_{max} / (2\ \mu s)}$ nm, the width of the distance distribution is encoded in the decay rate of the dipolar oscillations in *F(t)*. Accordingly, several oscillations must be observed for an accurate determination of the width, which puts the upper limit for such determination to $r_{max,\sigma_r} \sim 4 \sqrt[3]{t_{max} / (2\ \mu s)}$ nm. The extraction of the distance distribution from a dipolar evolution function *V(t)* is an ill-posed problem meaning that small variations in the input data related to the noise can cause large variations in the output data, therefore, the SNR has to be as good as possible. (54, 55) An elegant way to determine the distance distribution from the DEER trace is to use Tikhonov regularization, which is considered to give the best compromise between artefact suppression and resolution. This method is implemented in DeerAnalysis software. (56) The approach relies on the computation of a simulated time-domain signal *S(t)* that represents the intramolecular dipolar modulation function and is given by:

$$S(t) = \int_0^\infty K(t,r)\, P(r)\; dr \qquad \text{(Eq. 3.1.4.1)}$$

where *K(t,r)* is a kernel function that describes the average of the dipolar interaction and considers an "ideal" isotropic sample, and *P(r)* is the distance distribution probability density function (PDF) ($\int \mathrm{d}r P(\mathrm{r}) = 1$) that describes the conformational heterogeneity of the ensemble of spin-labelled protein molecules in the sample. The kernel function has the analytical expression equal to:

$$K(t,r) = \int_0^1 \cos[(3x^2 - 1)\omega_{DD} t]dx \text{ with } \omega_{DD} = \frac{2\pi\ 52.04\ MHz\ nm^{-3}}{r^3} \qquad \text{(Eq. 3.1.4.2)}$$

This approach gives the best compromise between smoothness (artifact suppression) and resolution of the distance distribution and is quantified by a regularization parameter α. For a given α the optimal distance distribution *P(r)* is given by the minimization of the following function:

$$G_\alpha\ (\mathrm{P}) = \|S(t) - D(t)\|^2 + \alpha \left\| \frac{\partial^2}{\partial r^2}\ P(r) \right\|^2 \qquad \text{(Eq. 3.1.4.3)}$$

where the first term represents the mean square deviation between the simulated and experimental dipolar evolution function (characterized by the noise), while the second term is the regularization-parameter weighted square norm of the second derivative of *P(r)*, which is a measure for the smoothness of *P(r)*, the ‖... ‖ indicates the Euclidean norm. The parameter α is adimensional and in general set at 10. Values set to 0.001 introduce noise and artefacts, and values >>100 lead to stronger broadening of peaks in the distance distribution. The correct value of α is extracted from the L-curve that is a plot of $\log \eta(\alpha) = \|S(t) - D(t)\|_\alpha^2$ versus $\log \rho(\alpha) = \left\| \frac{\partial^2}{\partial r^2}\ P(r) \right\|_\alpha^2$, where $\rho(\alpha)$ quantifies the means square deviation and $\eta(\alpha)$ is the smoothness.

3.2 Simulations of the CW EPR spectra

In order to simulate the CW EPR spectra of the Aurora-A kinase, the stochastic Liouville equation (SLE) in the Fokker-Planck (FP) form has been employed using the FP equation (57, 58) that is a partial differential equation. The theory describes the time evolution of the probability density $P(\vec{x}, \hat{\rho}, t)$ of systems at different orientations $\vec{x}$ in different spin states $\hat{\rho}$. (59) This kind of theory is applied to systems outside the Redfield theory in which motions have long correlation times, τ_R >> 1 ns. It is derived from the Liouville-von Neumann equation that is a linear first-order partial differential equation that describes the dynamics of the density matrix $\hat{\rho}(t)$ as function of the time and has the following expression:

$$\frac{\partial \hat{\rho}(t)}{\partial t} = - i\hat{\hat{H}}\hat{\rho}(t) \;\Rightarrow\; \frac{\partial \hat{\rho}(t)}{\partial t} = - i\,(\hat{\hat{H}}_0 + \hat{\hat{H}}_1)\,\hat{\rho}(t) \qquad \text{(Eq. 3.2.1)}$$

where $\hat{\hat{H}}$ is the Hamiltonian superoperator that is given by $\hat{\hat{H}}_0$ and is the time-independent part of $\hat{\hat{H}}$, while $\hat{\hat{H}}_1$ is the time-dependent part of $\hat{\hat{H}}$ that is modulated by random molecular motion. The first part of this section concerns about the derivation of the FP equation or SLE equation, while the second part (after Eq. 3.2.14) describes the relevant equations that are implemented in the Spinach software (60) for simulations of the CW EPR spectra. The scripts used for simulations are shown in the Appendix I. The description below was developed following the lecture II, Module VI in the Spin dynamics portal (61).

Considering the joint probability distribution $P(\vec{x}, \hat{\rho}, t)$ of a physical system in a point $\vec{x}$ and in spin state $\hat{\rho}$, as function of the time, it is possible to find its dynamical equation assuming static spin dynamics and spatial probability density of molecules in a particular spin state evolving in accordance to some simple deterministic law:

$$\left.\frac{\partial P(\vec{x},\hat{\rho},t)}{\partial t}\right|_{\hat{\rho}} = \hat{D}\; P(\vec{x}, \hat{\rho}, t) \qquad \text{(Eq. 3.2.2)}$$

where ... $|_{\hat{\rho}}$ indicates static spin dynamics (spin variable static) and $\hat{D}$ is the diffusion operator describing the spatial dynamics of the system and is equal to:

$$\hat{D} = \frac{1}{6\,\tau_R}\,(\hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2) \qquad \text{(Eq. 3.2.3)}$$

where τ_R is the rotational correlation time, $\hat{L}_i^2$ represents the angular momentum (not spin) operators acting in the lab space to generate spatial rotations of the spin system from the starting Cartesian coordinates. Eq. 3.2.2 indicates that spin does not affect space dynamics. Considering a step forward in time Δt and fixing the space, it is possible to consider a step forward in time forward also for the density matrix $\hat{\rho} \longrightarrow e^{(-i\hat{\hat{H}}\Delta t)\hat{\rho}}$. The joint probability distribution becomes:

$$P(\vec{x}, \hat{\rho}, t + \Delta \mathrm{t})|_{\vec{x}} = P\left(\vec{x}, e^{(-i\hat{\hat{H}}\Delta t)\hat{\rho}}, t\right)\Big|_{\vec{x}} \qquad \text{(Eq. 3.2.4)}$$

where ... $|_{\vec{x}}$ indicates space static, the differential equation for the spin dynamics is derived taking the Taylor expansion of $P\left(\vec{x}, e^{(-i\hat{\hat{H}}\Delta t)\hat{\rho}}, t\right)$ respect to $\hat{\rho}$:

$$P\left(\vec{x}, e^{(-i\hat{\hat{H}}\Delta t)\hat{\rho}}, t\right) =$$

$$= P(\vec{x}, \hat{\rho}, t) + \left\langle \nabla_{\hat{\rho}}\, P(\vec{x}, \hat{\rho}, t) \middle| e^{(-i\hat{\hat{H}}\Delta t)}\hat{\rho} - \hat{\rho} \right\rangle + O\left(\left\| e^{(-i\hat{\hat{H}}\Delta t)}\hat{\rho} - \hat{\rho} \right\|^2\right)$$

(Eq. 3.2.5)

where $\nabla_{\hat{\rho}}$ is the gradient respect to $\hat{\rho}$, assuming a small time increment $\Delta t \longrightarrow 0$,
$e^{(-i\hat{\hat{H}}\Delta t)} \longrightarrow -i\hat{\hat{H}}$:

$$P(\vec{x}, \hat{\rho}, t) + \left\langle \nabla_{\hat{\rho}}\, P(\vec{x}, \hat{\rho}, t) \middle| -i\hat{\hat{H}}\hat{\rho} \right\rangle \Delta t + O\,(\Delta t^2) \qquad \text{(Eq. 3.2.6)}$$

This is the only term that survives when the Taylor expansion of the second term in Eq. 3.2.6 is taken respect to Δt. In the case of an infinitesimal displacement forward in time with static space coordinates, the following expression is obtained:

$$\frac{\partial P(\vec{x},\hat{\rho},t)}{\partial t}\bigg|_{\vec{x}} = \lim_{\Delta t \to 0} \frac{P(\vec{x},\hat{\rho},t+\Delta t) - P(\vec{x},\hat{\rho},t)}{\Delta t}$$

$$= \left\langle \nabla_{\hat{\rho}}\, P(\vec{x}, \hat{\rho}, t) \middle| -i\hat{\hat{H}}\Delta t \right\rangle$$

Taking $\nabla_{\hat{\rho}}$ out:

$$= \left\langle \nabla_{\hat{\rho}} \middle| P(\vec{x}, \hat{\rho}, t)(-i\hat{\hat{H}}\Delta t) \right\rangle - P(\vec{x}, \hat{\rho}, t) \left\langle \nabla_{\hat{\rho}} \middle| -i\hat{\hat{H}}\Delta t \right\rangle \qquad \text{(Eq. 3.2.7)}$$

The term $\left\langle \nabla_{\hat{\rho}} \middle| -i\hat{\hat{H}}\Delta t \right\rangle$ is zero[11], recombining the derived equations to recompose the full partial differential equation gives:

$$\frac{\partial P(\vec{x},\hat{\rho},t)}{\partial t} = \left\langle \nabla_{\hat{\rho}} \middle| P(\vec{x},\hat{\rho},t)(-i\hat{\hat{H}}\Delta t) \right\rangle + \hat{D}\ P(\vec{x},\hat{\rho},t) \qquad \text{(Eq. 3.2.8)}$$

In the case of an ensemble average density matrix that is an integral of the entire space weighted by the probability, gives:

$$\overline{\hat{\rho}(\vec{x},t)} = \int \hat{\rho}\, P(\vec{x},\hat{\rho},t) dV_p \qquad \text{(Eq. 3.2.9)}$$

where V_p is the volume of the spin space, the equation of the motion of the ensemble average density matrix become:

$$\frac{\partial \overline{\hat{\rho}(\vec{x},t)}}{\partial t} = \int \hat{\rho} \left(\frac{\partial P(\vec{x},\hat{\rho},t)}{\partial t} \right) dV_p = \int \langle \left\langle \nabla_{\hat{\rho}} \middle| P(\vec{x},\hat{\rho},t)(-i\hat{\hat{H}}\Delta t) \right\rangle + \ \hat{D}\ P(\vec{x},\hat{\rho},t) \rangle\, \hat{\rho}\, dV_p$$

(Eq. 3.2.10)

The first integral in equation 3.2.10 is solved using multi-dimensional integration by parts. The derivative of $\overline{\hat{\rho}(\vec{x},t)}$ respect to time will give the equation of the motion of the ensample averaging,[12] considering that the density matrix lives inside a restricted space like a sphere and is zero outside it, so the term $\oint f \langle \vec{g} | \overrightarrow{ds} \rangle$ is zero:

$$\int \left\langle \nabla_{\hat{\rho}} \middle| P(\vec{x},\hat{\rho},t) \left(-i\hat{\hat{H}}\Delta t \right) \right\rangle \hat{\rho}\, dV_p =$$

$$= -\mathrm{i}\hat{\hat{H}} \int \hat{\rho}\, P(\vec{x},\hat{\rho},t) dV_p + \hat{D} \int \hat{\rho}\, P(\vec{x},\hat{\rho},t) dV_p$$

$$= -i\, \hat{\hat{H}}\ \overline{\hat{\rho}(\vec{x},t)} + \hat{D}\overline{\hat{\rho}(\vec{x},t)} \qquad \text{(Eq. 3.2.11)}$$

Treating the density matrix as function of the spin and space; the spin operators will act on the spin part and the space operators will act on the space part. It is possible to obtain a joint density matrix that will take into account not only the distribution of the probability in the

---

[11] $\left\langle \nabla_{\hat{\rho}} \middle| -i\hat{\hat{H}}\Delta t \right\rangle$ = $-i\sum_k \frac{\partial}{\partial \rho_k} \left[\hat{\hat{H}}\ \hat{\rho} \right]_k$ expanding the scalar product, where *k* is linear index.
$= -i \sum_k \frac{\partial}{\partial \rho_k} \sum_m H_{km}\, \rho_m$ expanding the matrix vector product
$= -i\sum_k H_{kk} = 0$ since spin Hamiltonian is traceless the sum is zero.

[12] Integral over some volume: $\int f \langle \nabla | \vec{g} \rangle$ dV = $\oint f \langle \vec{g} | \overrightarrow{ds} \rangle - \int f \langle \vec{g} | \nabla f \rangle$ dV.

spin state within the ensemble but also the distribution of the probability in the space. The second integral is easy to solve, while it is equal to:

$$\int \hat{D}\, P(\vec{x},\hat{\rho},t)\,\hat{\rho}\, dV_p = \hat{D} \int P(\vec{x},\hat{\rho},t)\,\hat{\rho} dV_p = \hat{D}\,\overline{\hat{\rho}(\vec{x},t)}$$

(Eq. 3.2.12)

Putting the two parts together and removing bars:

$$\frac{\partial\hat{\rho}(\vec{x},t)}{\partial t} = -i\hat{\hat{H}}(\vec{x})\,\hat{\rho}(\vec{x},t) + \hat{D}\,\hat{\rho}(\vec{x},t)$$ (Eq. 3.2.13)

Equation 3.2.13 is the FP or SLE equation that can be solved numerically with all the terms represented as matrices. The density matrix $\hat{\rho}(\vec{x},t)$ and the spin Hamiltonian superoperator the $\hat{\hat{H}}(\vec{x})$ have such a representation by construction, while the lab space operators ($\hat{D}$) may be converted into matrices by choosing either a discrete grid of points or a basis set of continuous functions. There are two general ways to solve equation Eq. 3.2.13; in the case of rotational diffusion, it is possible to assume that a system will diffuse between discretized points on a sphere under $\hat{D}$ that is a matrix that will define density matrix at each point on the grid. So, there is a joint probability in spin and space and the total problem is a direct product of the space problem and spin problem. The second way to solve the equation, that describes the system as a continuous representation using basis sets in spin and space (eigenfunctions of the diffusion operator that are Wigner functions) instead of discretizing points on the sphere. Equation 3.2.12 can be also expressed in following form:

$$\frac{\partial\hat{\rho}(\vec{x},t)}{\partial t} = -i\hat{\hat{L}}(\vec{x},t)\,\hat{\rho}(\vec{x},t) + \hat{D}\,\hat{\rho}(\vec{x},t)$$ (Eq. 3.2.14)

where $\hat{\hat{L}}(\vec{x},t)$ is called the Liouvillian and it is responsible for spin dynamics and is equal to:

$$\hat{\hat{L}}(\vec{x},t) = \hat{\hat{H}}(\vec{x},t) + i\hat{\hat{R}} + i\hat{\hat{K}}$$ (Eq. 3.2.15)

where terms $\hat{\hat{R}}$ and $\hat{\hat{K}}$ are the relaxation and kinetic superoperators, respectively.The spin Hamiltonian superoperator $\hat{\hat{H}}(\vec{x}, t)$, in the case of a spin system composed by two spins S and L, the spin Hamiltonian is given by:

$$\hat{\hat{H}}(\alpha, \beta, \gamma) = \hat{\hat{H}}_{iso} + \sum_{k,m=-2}^{2} D_{k,m}^{(2)}(\alpha, \beta, \gamma)\, \hat{\hat{Q}}_{k,m} \qquad \text{(Eq. 3.2.16)}$$

where α, β and γ are the Euler angles, $D_{km}^{(2)}$ is the second rank Wigren matrix and $\hat{\hat{Q}}_{km}$ is the rotational basis operator that is a linear combinations of irreducible spherical tensors corresponding to the interactions within the spin system. A full expression of the spin Hamiltonian superoperator is given by (Eq. 3.2.17):

$$\hat{\hat{H}}(\vec{x}, t) = \hat{H}_{iso} +$$

$$+ \sum_{k,m=-2}^{2} D_{k,m}^{(2)} \left[\sum_{L} a_m^{(2)}(L) \hat{T}_k^{(2)}(L) + \sum_{L,S} a_m^{(2)}(L, S) \hat{T}_k^{(2)}(L, S) + \sum_{S} a_m^{(2)}(S, S) \hat{T}_k^{(2)}(S, S) \right]$$

(Eq. 3.2.17)

where $a_m^{(l)}$ and $T_m^{(l)}$are the coefficients of the irreducible spherical tensors and spherical tensor operators, respectively, in which *l* is the rank and *m* are the *2*(*l*+1) components of *l*. The term $\hat{H}_{iso}$ represents the isotropic part of the Hamiltonian that contains the zero-rank terms $a_0^{(0)}$ and $T_0^{(0)}$ related to the spin couplings, the other terms in the brackets correspond to linear, bilinear and quadratic couplings within the spin system and are represented by second-rank $a_m^{(2)}$ and $T_m^{(2)}$ terms. The term $\hat{\hat{Q}}_{km}$ in the brackets, no matter how large and complicated the spin system, has always just 25 rotational basis operators. Equation 3.2.14 can be solved using a block-tridiagonal system of differential equations with the standard numerical methods and has an analytical solution in terms of a continued fraction expansion. (62) This system is characterized by damping terms that grow quadratic with respect to that rank *l*, for spin system in liquid state the convergence is achieved around *l* =10. The dimension of the spatial basis grows cubically with *l* and exponentially with the number of spins in the system and matrices of astronomical dimensions are obtained that represents the primary practical problem for the calculation of the SLE equation, therefore, appropriate algorithms, such as the Lanczos algorithm (61) and appropriate work-stations are required to cope with these calculations. This formalism was used for the simulations of the EPR spectra of the Aurora-A kinase, further details are reported in Chapters 5 and 6.

3.3 Computational methods

3.3.1 Density functional theory (DFT)

The DFT is a computational method for quantum mechanical calculations in order to investigate structural, magnetic and electronic properties of periodic systems that include molecules, materials, etc. Unlike the Hartree-Fock theory that deals directly with many-body wavefunctions, DFT theory is based on functional (a function of another function) of the electron density that in turn is a function of space and time. The advantages of the DFT over the Hartree-Fock theory is related to an approximate treatment of the correlated motions of electrons and to the employment of the electron density that significantly speeds up calculations since it is a function of only three variables *x, y, z,* while the many-body electronic wavefunction is a function of 3N variables (the coordinates of all N atoms in the system). The theoretical calculations that include geometry optimizations and calculations of the EPR parameters in the framework of the DFT are based on Kohn-Sham theory that in turn is based on the two Hohenberg-Kohn theorems:

Theorem 1: The external potential, hence the total energy, is a unique functional of the electron density.
Theorem 2: The ground state energy can be obtained variationally: the density that minimises the total energy is the exact ground state density.

The Kohn-Sham (KS) equation is a specific solution of a one-particle Schrödinger wave equation. The resulting energy is composed by four contributions:

$$H_{KS}\Psi_i(r) = E_i\Psi_i(r) \qquad \text{(Eq. 3.3.1.1)}$$

where $H_{KS} = -\frac{1}{2}\nabla^2 + v_V(r) + v_J(r) + v_{XC}(r)$ is the Kohn-Sham Hamiltonian, $\Psi_i(r)$ with i = 1, ... ,n_α are the eigenfunctions corresponding to orthonormal orbitals, and $E_i = E^T + E^V + E^J + E^{XC}$ that are the corresponding eigenvalues representing to energy levels. The terms of the Hamiltonian represent potentials corresponding to the electronic energy composed by the kinetic energy (E^T), the electron-nuclear interaction energy (E^V), the Coulomb self-interaction of the electron density (E^J) and exchange-correlation energy that is part of the electron-electron repulsion (E^{XC}), that are all function of the electron density.

In the Kohn-Sham formulation of the DFT, the exchange-correlation functional includes terms accounting for the exchange and the electron correlation energies, the latter not being present in Hartree-Fock theory. Within the DFT theory, the E^{XC}can be calculated using to different methods, such as the local spin density approximation (LSDA) and the generalized gradient approximation (GGA), the former was developed to reproduce the exchange-correlation energy of a uniform electron gas and approximate the energy considering a constant local electron density, while the latter is based on the gradient-corrected functionals, which depend not only on the electron density, ρ but also on its gradient, $\nabla\rho$ leading to a large improvement to the LSDA results. There are also hybrid approximations for the exchange–hybrid correlation functional, such as the B3LYP functional that mix in a certain amount of the Hartree–Fock exchange energy (E_{HF}^X) into the exchange and correlation obtained from the GGA. The resulting energy is given by:

$$E_{hyb}^{XC} = \text{a } E_{HF}^X + (1-a)\, E_{GGA}^X + E_{GGA}^C \qquad \text{(Eq. 3.3.1.2)}$$

where a is a constant that can be fitted empirically or estimated theoretically. Equation 3.3.1.1 represents an eigenvalue problem and needs to be solved in a self-consistent way, following the scheme below:

1) Choose a set of wave functions, $\Psi_i(r)$.
2) Calculate the density, $\rho(r)$.
3) Calculate the total energy, E_{tot} using Eq. 3.3.1.3, below:

$$E_{tot}\,[\rho(r)] = \sum_i e_i = E_H\,[\rho(r)] + E_J[\rho(r)] + E_{XC}\,[\rho(r)] \qquad \text{(Eq. 3.3.1.3)}$$

The Kohn-Sham theory provides an approximate solution of the problem that is employed the practical DFT scheme. In order to perform DFT calculation, a representation of the wave functions is required. It is given the basis sets that are used to describe core and valence electrons in the KS orbitals of atoms and molecules and are used to perform calculations of structural, magnetic and electronic properties of periodic systems that include molecules, materials, etc. The KS orbitals can be expressed as a linear combination of basis functions to form the orbitals:

$$\Psi_i(r) = \sum_k c_k \, \varphi_k \, (r) \qquad \text{(Eq. 3.3.1.4)}$$

where c_k are coefficients and $\varphi_k(r)$ are basis functions that are composed of one or more primitive Gaussian functions or primitive Gaussian type orbitals (GTO) that have the following expression in Cartesian coordinates:

$$\varphi_k \, (x, y, z; l, m, n, \alpha, f) = N \, x^l y^m z^n e^{-\alpha f^2 r^2} \qquad \text{(Eq. 3.3.1.5)}$$

where N is the normalization function, x, y and z represent the Cartesian coordinates, α is a constant that determines the size of the orbital (the radial extent) of the Gaussian function, f is a scaling factor. Exponents l, m and n are not quantum numbers, $L = l+m+n$ is the angular momentum quantum number used to mark functions of *s*-type (L=0), *p*-type (L=1), *d*-type (L=2), etc. and are defined as:

$$L=0 \ \left\{ e^{-\alpha f^2 r^2} \right. \quad \text{to describe } s \text{ orbitals} \qquad \text{(Eq. 3.3.1.6)}$$

$$L=1 \begin{cases} xe^{-\alpha f^2 r^2} \\ ye^{-\alpha f^2 r^2} \\ ze^{-\alpha f^2 r^2} \end{cases} \quad \text{to describe } p \text{ orbitals} \qquad \text{(Eq. 3.3.1.7)}$$

$$L=2 \begin{cases} x^2e^{-\alpha f^2 r^2} \\ y^2e^{-\alpha f^2 r^2} \\ z^2e^{-\alpha f^2 r^2} \\ xye^{-\alpha f^2 r^2} \\ xze^{-\alpha f^2 r^2} \\ zye^{-\alpha f^2 r^2} \end{cases} \quad \text{to describe } d \text{ orbitals} \qquad \text{(Eq. 3.3.1.8)}$$

Additional polarization functions with a higher angular momentum number are introduced to allow atomic orbitals to change shape but not size, while the diffuse functions allow orbitals to occupy larger regions of space and are used to describe electrons that are far away from the nucleus, such as lone pairs, anions or system with negative charge or in their excited states. Basis sets (or Pople style basis sets) are indicated as:

*k-nlm++G*** or *k-nlm++G(idf, ipd)*

where *k* represents the number of primitives GTOs for the description of the core electrons, *n* for the for inner valence orbitals, *l* for the for medium valence orbitals and *m* for the outer valence orbitals, + means one set of *p* (*sp*) diffuse functions to be added to heavy atoms, ++ means one set of *p* (*sp*) diffuse functions added to heavy atoms and 1 set of *p* (*sp*) polarization functions added to H atom, * means 1 set of *d* polarization functions added to heavy atoms, ** means 1 set of *d* polarization functions added to heavy atoms and 1 set of *p* (*sp*) polarization functions added to H atom, *idf*, *jpd* means *i d* sets and 1 *f* set of polarization functions added to heavy atoms and *j p* sets and 1 *d* set of polarization functions added to H atom. Some basis sets contain only the *nl* numbers, these are called double-zeta basis sets, basis sets with all the three *nlm* numbers are called triple-zeta basis sets. In the case of the split-valence or valence-multiple basis sets, the core orbitals are represented by a single (contracted) basis function, while valence orbitals are split into arbitrarily many functions, for instance these are the 3-21G, 6-21G, 4-31G, 6-31G, and 6-311G basis sets. Another important kind of basis sets are the Slater-type orbital (STO) basis sets first introduced by Zener and Slater, which can provide the correct cusp behaviour in the vicinity of the nuclei, as well as the correct exponential decay at large distances considering$e^{-\zeta r}$, where ζ is the orbital exponent, instead of $e^{-\alpha f^2 r^2}$of the GTOs, but this evaluation requires multicenter integrals that are difficult and expensive.

In this work the DFT was employed for geometry optimizations, the calculation of torsional profiles of the MTSL spin label and the calculation of the EPR magnetic parameters (*g*- and *A*-tensors). Geometry optimizations and torsional profiles were performed using the B3LYP hybrid functional (63, 64) and the 6-31G(d) basis set (40). This level of theory is commonly used to provide reliable reproductions of the equilibrium geometries. The magnetic parameters were calculated using the B3LYP hybrid functional with the N07D basis sets in order to obtain accurate equilibrium and vibrationally averaged hyperfine coupling constants (hcc) for the ^{14}N nucleus. For accurate calculation of the EPR parameters, it is important to complete the core-valence *s* space with six *d* Gaussian functions that gives the 6-31G basis set (65-67). The EPRIII basis set employs only a set of five *d* Gaussian functions (68), while the N07D employs a set of six *d* functions and adds to the double-zeta description of valence orbitals a single sets of optimized core-valence *s* (on all atoms except H), diffuse *p* (on all atoms except H), polarization (on all atoms), and diffuse *d* (on O, F, Cl atoms) functions that improve its performance as compared with EPR II and EPR III. (66) The N07D basis set was optimized for second- and third-row atoms and has already been used with success for accurate calculation of the magnetic tensors (Δg_{ii} = ±0.0005, ΔA_{ii} = ±1 G) in gas phase and in solution of nitroxide radicals at a reasonable computational cost and can be downloaded from the DREAMSLAB website. (69)

3.3.2 Classical molecular dynamics (MD):

Classical molecular dynamics (MD) simulations are used for the characterization of the structure and dynamics of biological macromolecules, such as large proteins, cellular membranes, DNA etc. Using MD, it is possible to study thermodynamic properties and time-dependent phenomena, such as kinetic and a wide range of dynamical processes that occur on different time-scale. The current developments in MD allows simulation of systems having ~50,000 – 100,000 atoms, remarkable improvements in the performance of MD simulations are due to the use of high performance computing (HPC) and the development of the MD algorithms that include fine-tuning of energy calculations, parallelization, or the use of graphical processing units (GPUs).

In order to perform MD, the system has to be built and thereafter forces acting on each atom are calculated by deriving equations that compose the *force-field*, where potential energy is deduced from the molecular structure. Once the forces acting on individual atoms are obtained, MD are performed numerically integrating equations of the motion using the classical Newton's law in order to calculate accelerations and velocities and to update the atom positions for a given system:

$$F_i(r_1, r_2, \dots, r_N) = m_i a_i = m_i \frac{\partial v_i}{\partial t} = m_i \frac{\partial^2 r_i}{\partial t^2} \text{ with } i = 1, 2, \dots, N \quad \text{(Eq. 3.3.2.1)}$$

where r_i are the position vectors and F_i are the forces acting upon the N particles with mass (m_i), a_i and v_i represent acceleration and velocity, respectively. The forces F_i (r_N) are derived from a potential energy $V(\vec{R})$ that is a function of atomic coordinates ($\vec{R}$) and is given by the sum intra-molecular terms (bond, angle, torsion angle and out-of-plane torsions[13]) and non-bonded terms (van der Waals and Coulomb interactions). The AMBER force field (ff14SB) (70), within the AMBER software,(71) used in this work has well-defined bond

---

[13] The out-of-plane angle coordinates are defined by an angle between the bond vector and the plane containing the other two bond vectors

length, bond angle, partial charge, and van der Waals parameters for all the common amino acids and nucleotides, the potential energy has the following expression:

$$V(\vec{R}) = \sum_{bonds} K_b\ (r_i - r_0)^2 + \sum_{angles} K_\theta\ (\theta_i - \theta_0)^2 + \sum_{dihedrals} \frac{V_n}{2}[1 + \cos(n\varphi_i) - \gamma] + + \sum_{i,j} 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] + \sum_{ij} \frac{1}{4\pi\varepsilon_0} \frac{q_i q_j}{r_{ij}} \qquad \text{(Eq. 3.3.2.2)}$$

where r_i, θ_i and φ_i represent the values of the bond length, bond angle and torsion angle, while r_0 and θ_0 are the equilibrium bond length and bond angle, respectively. The force constants K_b, K_θ and $\frac{V_n}{2}$ depend on the strength of the bonds, *n* represents the periodicity and φ represent the shifts of the entire curve along the rotation angle axis. These terms can be obtained from *ab initio* quantum mechanical calculations using small model compounds and from comparisons to the geometry and vibrational spectra in the gas phase (IR and Raman spectroscopy). The last two sums represents the non-bonded interactions between couples of atoms *i* and *j*, the fourth term represents the pair potential, where σ_{ij} and ϵ_{ij} are the Lennard-Jones parameters and r_{ij} is the distance between couples of atoms. The last term represents the Coulombic interactions between partial charges q_i and q_j.

The current AMBER force field lacks parameters for the MTSL spin label, therefore in order to perform MD simulation, all the bond length, bond angle and torsion angle parameters for MTSL have to be generated. A general parameterization protocol for a non-standard residue, such as MTSL, for the extension of the AMBER force field was developed by Cotrnell *et al.* (72, 73) Therefore, the structure of the MTSL side chain was generated with GaussianView 5.0.9 and subsequently its geometry was optimized with the B3LYP hybrid functional using the 6-31G(d) basis set. In order to reproduce the electrostatic potential (ESP) and the hydrogen-bonding properties of the MTSL side chain the Hartree-Fock theory (74, 75) with the 6-31G(d) basis set. Finally, charges were produced by fitting the ESP using restrained electrostatic potential (RESP) following the procedure of Ciepllak *et al.* (14, 76) using the ANTECHAMBER module in the AMBER software, that generate the topology file with parameters of the MTSL side chain, the parameterisation of the MTSL determined in this work is reported in Appendix II. The MTSL was subsequently attached to the protein using the LEaP module AMBER, and the three steps, minimization, equilibration and production, of the MD simulation can be performed, details about these steps are given in Chapters 5 and 6. In the minimization step, the force field was assigned to the atoms in the system, minima and stable points on the potential energy surface are determined in order to begin dynamics. In the equilibration step, MD solves the equations of motion for a system of atoms, the solution for the equations of motion of a molecule represents the time evolution of the molecular motions that is the MD trajectory. The initial velocities are assigned using a random number generator using the constraint of the Maxwell-Boltzmann distribution.

In the production step, the MD data are collected, all the routines for the characterization of the structure and dynamics of the system, such as principal component and clustering (PCA) analysis and isotropic reorientational eigenmode (iRED) routine are performed on data from this step. All the scripts used in this work to perform the steps described above are reported in Appendix III.

3.3.3 Principal component analysis (PCA)

PCA is a technique that can be used to transform a series of coordinated observations into a set of orthogonal vectors called principal components (PCs) that are used to explain the variance of the data in the MD trajectory. The correlated internal motion of a system with N atoms can be described by a covariance matrix, C, with elements, c_{ij}:

$$c_{ij} = \langle (r_i - \langle r_i \rangle)(r_j - \langle r_j \rangle) \rangle \qquad \text{(Eq. 3.3.4.1)}$$

where $r_1, \ldots, r_{3N}$ are the mass-weighted Cartesian coordinates and $\langle \ldots \rangle$ represents the average over all sampled conformations. Diagonalization of this covariance matrix results in 3N eigenvectors (v_i) and eigenvalues (λ_i) which describe the modes of the collective motion and their respective amplitudes. The principal components, Vi:

$$V_i = v_i \;\; r \qquad \text{(Eq. 3.3.4.1)}$$

are projections of the data r = $(r_1, \ldots, r_{3N})T$ onto the eigenvectors, and can be used to describe the dynamics of the system. The eigenvectors are sorted such that their eigenvalues are in decreasing order, the eigenvector with the largest eigenvalue (the first PC) correspond to the highest proportion of variance within the data and to the predominant motion in the system. The second component is orthogonal to the first PC and corresponds to the second highest proportion of variance, and so on.

3.3.4 Clustering analysis

The cluster analysis is a technique for finding patterns within data and is based on clustering algorithms that group similar objects into subgroups, called clusters by minimizing intra-cluster and maximizing inter-cluster differences. (77) Clustering algorithms can be divided into partitional and hierarchical clustering approaches. (78) The partitional clustering approach is based on the K-means algorithm that places a predefined number of initial centroids (*K*) randomly and in the iterative process, *K* clusters are formed by assigning each data point to its closest centroid and a new centroid for each cluster is recomputed. These steps are repeated until the cluster membership is converged. The agglomerative hierarchical clustering is a collection of techniques that start with singleton clusters and proceeds to iteratively join the nearest cluster until all objects are grouped into one cluster. These techniques differ in the definition of cluster proximity and include the *single-linkage* algorithm that defines cluster proximity as the “closeness” between the nearest two objects that are in different clusters, the *complete-linkage* algorithm that considers the farthest two objects and *average-linkage* algorithm that uses the average pairwise proximities of all pairs of objects from different clusters.

3.3.5 Isotropic reorientational eigenmode (iRED) approach

For each MD snapshot, a $n(2L+1)$-dimensional vector $|Y(t)\rangle$ can be constructed from spherical harmonics $Y_{LM}(\Omega_j)$ evaluated at the n directions $\Omega_j(t) = (\theta_j(t),\ \varphi_j(t))$, j= 1, …, n of the inter-nuclear vectors (NH), where L is the rank of the spherical harmonics, M is the component of L. Therefore, $|Y(t)\rangle$ has the following expression:

$$|Y(t)\rangle = \left| \begin{matrix} Y_{L,-M}(\Omega_1), Y_{L,-M+1}(\Omega_1), \ldots, Y_{L,M-1}(\Omega_1),\ Y_{L,M}(\Omega_1), \ldots,\ Y_{L,-M}(\Omega_n),\ Y_{L,-M+1}(\Omega_n), \ldots, \\ Y_{L,M-1}(\Omega_n),\ Y_{L,M}(\Omega_n) \end{matrix} \right\rangle$$

(Eq. 3.3.3.1)

Subsequently, a *n(2L+1*) × *n(2L+1)* covariance matrix P is constructed:

$$P = \overline{|Y\rangle\langle Y|} - \overline{|Y\rangle} - \overline{\langle Y|} \qquad \text{(Eq. 3.3.3.2)}$$

where $\langle Y|$ is the complex-conjugate row vector of column vector $|Y\rangle$. The horizontal bar indicates an ensemble average over the N conformations or a time average over a trajectory, the term $\overline{|Y\rangle\langle Y|}$ represents the isotropic averaging that is obtained using the transformation properties of spherical harmonics under 3D rotation (demonstration omitted). This process generates a matrix M constructed by n×n elements:

$$M_{kl} = \overline{P_L(\cos(\Omega_k - \Omega_j))} \qquad \text{(Eq. 3.3.3.3)}$$

where $P_L(x)$ is Legendre polynomial of order L and $\Omega_k - \Omega_j$ denotes the angle between directions Ω_k and Ω_j. The reorientational eigenmodes and their amplitudes are the eigenvectors $|m\rangle$ (referring to N-H bonds in the protein backbone) and the eigenvalues $\lambda_m$of M are obtained in order to fulfil M $|m\rangle$= $\lambda_m|m\rangle$, respectively. The eigenmodes describe correlated modulations of different spin interactions and the eigenvalues are the amplitude fluctuations of the trajectory along each eigenmode. Therefore, in this approach, the eigenmodes and eigenvalues describing motions within the system are considered in terms of amplitudes and motional correlation effects experienced by the different spin interactions and no assumption about the existence of a molecular reference frame are taken into account. However, the matrix M contains no information on the motional correlation times associated to individual reorientational modes $|m\rangle$. This is obtained from correlation functions constructed projecting the vector $|Y(t)\rangle$ obtained from a snapshot at time t on the eigenvector $|q\rangle = |m\rangle \otimes |e_j\rangle$, where $|e_j\rangle$ are the $2L$+1-dimensional column vectors equal to $|e_1\rangle = (1,0,\dots,0)^T$, $|e_2\rangle = (0,1,\dots,0)^T$, …, $|e_{2L+1}\rangle = (0,0,\dots,1)^T$, in order to get:

$$a_q(t) = \langle q|Y(t)\rangle \qquad \text{(Eq. 3.3.3.4)}$$

and leads to a correlation function:

$$C_q(t) = C_{m,l}(t) = \sum_{l=-L}^{L} \langle a_{m,l} * (\tau + \mathrm{t}) \cdot a_{m,l} \rangle_\tau \qquad \text{(Eq. 3.3.3.5)}$$

where $q = (2L+1)m+l-L$ with $l=-L, \dots, L$ and $a_{m,l}$ are the time-dependent amplitudes obtained by projecting MD snapshots on the respective eigenmodes.The average extends over snapshots sampled at times $\tau = 0 \longrightarrow T\text{-}t$. Correlation functions for M are obtained by summing up the $2L + 1$ correlation functions, $C_{m,l}(t)$:

$$C_m(t) = \sum_{l=-L}^{L} C_{m,l}(t) \qquad \text{(Eq. 3.3.3.6)}$$

that belongs to the $2L$+1 degenerate modes q. If the $C_m(t)$ decays exponentially, the correlation time $\tau_m$associated with the mode $|m\rangle$ is determined by:

$$\tau_m \cong \frac{1}{[C_m(0)-C_m(t\to T)\,]} \int_0^T [C_m(0) - C_m(t \to T)\,]dt \qquad \text{(Eq. 3.3.3.7)}$$

where $C_m(t \to T)$ indicates the plateau value of $C_m(t)$:

$$C_m(t \to T) = \sum_{l=-L}^{L} C_{m,l}(t \longrightarrow T) \cong \sum_{l=-L}^{L} \left|\langle a_{m,l}(\tau)\rangle\right|^2 \qquad \text{(Eq. 3.3.3.8)}$$

for systems where internal and overall tumbling motions are separable, $C_{m,l}(t)$ and $C_m(t)$and their correlation times reflect both internal motions and overall motions.

Bibliography

1. Erlenbach N, Endeward B, Schops P, Gophane DB, Sigurdsson ST, Prisner TF. Flexibilities of isoindoline-derived spin labels for nucleic acids by orientation selective PELDOR. Physical chemistry chemical physics : PCCP. 2016;18(24):16196-201. Epub 2016/06/03.
2. Papaioannou D, Geibel S, Kunze MB, Kay CW, Waksman G. Structural and biophysical investigation of the interaction of a mutant Grb2 SH2 domain (W121G) with its cognate phosphopeptide. Protein science : a publication of the Protein Society. 2016;25(3):627-37. Epub 2015/12/10.
3. Zhang Z, Fleissner MR, Tipikin DS, Liang Z, Moscicki JK, Earle KA, et al. Multifrequency electron spin resonance study of the dynamics of spin labeled T4 lysozyme. The journal of physical chemistry B. 2010;114(16):5503-21. Epub 2010/04/07.
4. Groban ES, Narayanan A, Jacobson MP. Conformational changes in protein loops and helices induced by post-translational phosphorylation. PLoS computational biology. 2006;2(4):e32. Epub 2006/04/22.
5. Sezer D, Freed JH, Roux B. Parametrization, molecular dynamics simulation, and calculation of electron spin resonance spectra of a nitroxide spin label on a polyalanine alpha-helix. The journal of physical chemistry B. 2008;112(18):5755-67. Epub 2008/04/17.
6. Fajer MI, Li HZ, Yang W, Fajer PG. Mapping electron paramagnetic resonance spin label conformations by the simulated scaling method. Journal of the American Chemical Society. 2007;129(45):13840-6.
7. Budil DE, Sale KL, Khairy KA, Fajer PG. Calculating slow-motional electron paramagnetic resonance spectra from molecular dynamics using a diffusion operator approach. The journal of physical chemistry A. 2006;110(10):3703-13. Epub 2006/03/11.
8. Burgess SG, Concilio, M. G., Bayliss, R., Fielding, A. J. . Detection of ligand-induced conformational changes in the activation loop of Aurora-A kinase using PELDOR. ChemistryOpen, in press. 2016.
9. Glover DM, Leibowitz MH, Mclean DA, Parry H. Mutations in Aurora Prevent Centrosome Separation Leading to the Formation of Monopolar Spindles. Cell. 1995;81(1):95-105.
10. Fu JY, Bian ML, Jiang Q, Zhang CM. Roles of aurora kinases in mitosis and tumorigenesis. Mol Cancer Res. 2007;5(1):1-10.
11. Barr AR, Gergely F. Aurora-A: the maker and breaker of spindle poles. Journal of cell science. 2007;120(Pt 17):2987-96. Epub 2007/08/24.
12. Bischoff JR, Anderson L, Zhu YF, Mossie K, Ng L, Souza B, et al. A homologue of Drosophila aurora kinase is oncogenic and amplified in human colorectal cancers. Embo J. 1998;17(11):3052-65.
13. Barr AR, Gergely F. Aurora-A: the maker and breaker of spindle poles. Journal of cell science. 2007;120(17):2987-96.
14. Bayly CI, Cieplak P, Cornell WD, Kollman PA. A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges - the Resp Model. J Phys Chem-Us. 1993;97(40):10269-80.
15. Liu Y, Gray NS. Rational design of inhibitors that bind to inactive kinase conformations. Nat Chem Biol. 2006;2(7):358-64.
16. Dar AC, Shokat KM. The Evolution of Protein Kinase Inhibitors from Antagonists to Agonists of Cellular Signaling. Annual Review of Biochemistry, Vol 80. 2011;80:769-95.
17. Rowan FC, Richards M, Bibby RA, Thompson A, Bayliss R, Blagg J. Insights into Aurora-A kinase activation using unnatural amino acids incorporated by chemical modification. ACS Chem Biol. 2013;8(10):2184-91. Epub 2013/08/09.
18. Dodson C .A. Bayliss R. Activation of Aurora-A kinase by protein partner binding and phosphorylation are independent and synergistic. J Biol Chem 2011;287(2):1150-7.

19. Bouloc N, Large JM, Kosmopoulou M, Sun CB, Faisal A, Matteucci M, et al. Structure-based design of imidazo[1,2-a] pyrazine derivatives as selective inhibitors of Aurora-A kinase in cells. Bioorganic & Medicinal Chemistry Letters. 2010;20(20):5988-93.
20. Bavetsias V, Large JM, Sun CB, Bouloc N, Kosmopoulou M, Matteucci M, et al. Imidazo[4,5-b]pyridine Derivatives As Inhibitors of Aurora Kinases: Lead Optimization Studies toward the Identification of an Orally Bioavailable Preclinical Development Candidates. Journal of Medicinal Chemistry. 2010;53(14):5213-28.
21. Zavoisky YK. Ph.D. thesis, Kazan State University. 1944.
22. Di Valentin M, Bisol A, Agostini G, Fuhs M, Liddell PA, Moore AL, et al. Photochemistry of artificial photosynthetic reaction centers in liquid crystals probed by multifrequency EPR (9.5 and 95 GHz). Journal of the American Chemical Society. 2004;126(51):17074-86. Epub 2004/12/23.
23. Bender G, Stich TA, Yan L, Britt RD, Cramer SP, Ragsdale SW. Infrared and EPR spectroscopic characterization of a Ni(I) species formed by photolysis of a catalytically competent Ni(I)-CO intermediate in the acetyl-CoA synthase reaction. Biochemistry. 2010;49(35):7516-23. Epub 2010/07/31.
24. Milov AD, Salikhov KM, Shirov MD. Application of Eldor in Electron-Spin Echo for Paramagnetic Center Space Distribution in Solids. Fiz Tverd Tela+. 1981;23(4):975-82.
25. Milov AD, Ponomarev AB, Tsvetkov YD. Electron Electron Double-Resonance in Electron-Spin Echo - Model Biradical Systems and the Sensitized Photolysis of Decalin. Chem Phys Lett. 1984;110(1):67-72.
26. Pannier M, Veit S, Godt A, Jeschke G, Spiess HW. Dead-time free measurement of dipole-dipole interactions between electron spins. Journal of magnetic resonance. 2000;142(2):331-40. Epub 2000/01/29.
27. Ward R, Bowman A, Sozudogru E, El-Mkami H, Owen-Hughes T, Norman DG. EPR distance measurements in deuterated proteins. Journal of magnetic resonance. 2010;207(1):164-7. Epub 2010/09/02.
28. Misra SKB, P.P.; Freed, J. H. Calculation of Double-Quantum-Coherence Two-dimensional Spectra: Distance Measurements and Orientational Correlations. Appl Magn Reson 2009;36(2-4):237-58.
29. Hellmich UA, Lyubenova S, Kaltenborn E, Doshi R, van Veen HW, Prisner TF, et al. Probing the ATP hydrolysis cycle of the ABC multidrug transporter LmrA by pulsed EPR spectroscopy. Journal of the American Chemical Society. 2012;134(13):5857-62. Epub 2012/03/09.
30. Lovett JE, Bowen AM, Timmel CR, Jones MW, Dilworth JR, Caprotti D, et al. Structural information from orientationally selective DEER spectroscopy. Physical chemistry chemical physics : PCCP. 2009;11(31):6840-8. Epub 2009/07/30.
31. Fernandez A, Ferrando-Soria J, Pineda EM, Tuna F, Vitorica-Yrezabal IJ, Knappke C, et al. Making hybrid [n]-rotaxanes as supramolecular arrays of molecular electron spin qubits. Nature communications. 2016;7:10240. Epub 2016/01/09.
32. Spindler PE, Zhang Y, Endeward B, Gershernzon N, Skinner TE, Glaser SJ, et al. Shaped optimal control pulses for increased excitation bandwidth in EPR. Journal of magnetic resonance. 2012;218:49-58. Epub 2012/05/15.
33. Segawa TF, Doll A, Pribitzer S, Jeschke G. Copper ESEEM and HYSCORE through ultra-wideband chirp EPR spectroscopy. The Journal of chemical physics. 2015;143(4):044201. Epub 2015/08/04.
34. Doll A, Jeschke G. EPR-correlated dipolar spectroscopy by Q-band chirp SIFTER. Physical chemistry chemical physics : PCCP. 2016;18(33):23111-20. Epub 2016/08/06.
35. Srivastava M, Lindsay Anderson, C., Freed J. H. A New Wavelet Denoising Method for Selecting Decomposition Levels and Noise Thresholds. IEEE Access 2016;4: 3862 - 77.

36. Sidabras JW, Mett RR, Froncisz W, Camenisch TG, Anderson JR, Hyde JS. Multipurpose EPR loop-gap resonator and cylindrical TE011 cavity for aqueous samples at 94 GHz. Rev Sci Instrum. 2007;78(3).
37. Stone AJ. Gauge Invariance of the g Tensor. Proc R Soc London, Ser A. 1963;271:424.
38. McWeeny R. Methods of Molecular Quantum Mechanics. London AP, editor1992.
39. Neese F. Prediction of electron paramagnetic resonance g values using coupled perturbed Hartree-Fock and Kohn-Sham theory. Journal of Chemical Physics. 2001;115(24):11080-96.
40. Ditchfield R. Self-Consistent Perturbation-Theory of Diamagnetism .1. Gauge-Invariant Lcao Method for Nmr Chemical-Shifts. Mol Phys. 1974;27(4):789-807.
41. Cheeseman JR, Trucks GW, Frisch MJ. Calculation of nuclear magnetic resonance shielding tensors in large molecules: A comparison of MP2 and hybrid density functional methods for treating electron correlation. Abstr Pap Am Chem S. 1998;215:U218-U.
42. Malkina OL, Vaara J, Schimmelpfennig B, Munzarova M, Malkin VG, Kaupp M. Density functional calculations of electronic g-tensors using spin-orbit pseudopotentials and mean-field all-electron spin-orbit operators. Journal of the American Chemical Society. 2000;122(38):9206-18.
43. Koseki S, Schmidt MW, Gordon MS. Mcscf/6-31g(D,P) Calculations of One-Electron Spin-Orbit-Coupling Constants in Diatomic-Molecules. J Phys Chem-Us. 1992;96(26):10768-72.
44. Parker GW. Derivation of the electric dipole–dipole interaction as an electric hyperfine interaction. Am J Phys 1986;54:715.
45. Bobart PP, Freed JH. Pulse Dipolar Electron Spin Resonance: Distance Measurements: Springer Berlin Heidelberg; 2013.
46. Hustedt EJ, Smirnov AI, Laub CF, Cobb CE, Beth AH. Molecular distances from dipolar coupled spin-labels: the global analysis of multifrequency continuous wave electron paramagnetic resonance data. Biophysical journal. 1997;72(4):1861-77. Epub 1997/04/01.
47. Spin-labels and intrinsic paramagnetic centres in the Biosciences. Mingos DMP, editor: Springer; 2012.
48. Borbat PP, Georgieva ER, Freed JH. Improved Sensitivity for Long-Distance Measurements in Biomolecules: Five-Pulse Double Electron-Electron Resonance. J Phys Chem Lett. 2013;4(1):170-5.
49. Fajer PGB, Song L. ESR Spectroscopy in Membrane Biophysics: Springer US; 2007.
50. Jeschke G. DEER Distance Measurements on Proteins. Annual Review of Physical Chemistry, Vol 63. 2012;63:419-46.
51. Song LK, Larion M, Chamoun J, Bonora M, Fajer PG. Distance and dynamics determination by W-band DEER and W-band ST-EPR. Eur Biophys J Biophy. 2010;39(4):711-9.
52. Polyhach Y, Bordignon E, Jeschke G. Rotamer libraries of spin labelled cysteines for protein studies. Physical chemistry chemical physics : PCCP. 2011;13(6):2356-66. Epub 2010/12/01.
53. Milov AD, Naumov BD, Tsvetkov YD. The effect of microwave pulse duration on the distance distribution function between spin labels obtained by PELDOR data analysis. Applied magnetic resonance. 2004;26(4):587-99.
54. Jeschke G. DEER distance measurements on proteins. Annual review of physical chemistry. 2012;63:419-46. Epub 2012/03/13.
55. Höfer PH, R.; Maier,D. C.; Prisecaru, I.; Schmalbein, D. The super Q-FT accessory for pulsed EPR, ENDOR and ELDOR at 34 GHz. Bruker SpinReport. 2003;152:37.
56. Jeschke G, Chechik V, Ionita P, Godt A, Zimmermann H, Banham J, et al. DeerAnalysis2006 - a comprehensive software package for analyzing pulsed ELDOR data. Applied magnetic resonance. 2006;30(3-4):473-98.
57. Fokker AD. Die mittlere Energie rotierender elektrischer Dipole im Strahlungsfeld, Annalen der Physik. 1917:810-20.

58. Planck M. Über einen Satz der statistischen Dynamik und seine Erweiterung in der Quantentheorie, Sitzungsber. Kön. Preuss. Akad. Wiss. . 1917:324-41.
59. Edwards LJ, Savostyanov DV, Nevzorov AA, Concistre M, Pileio G, Kuprov I. Grid-free powder averages: on the applications of the Fokker-Planck equation to solid state NMR. Journal of magnetic resonance. 2013;235:121-9. Epub 2013/08/15.
60. Hogben HJ, Charnock GTP, Hore PJ, Kuprov I. Spinach – a software library for simulation of spin dynamics in large spin systems. J Magn Reson. 2011;208:179-94.
61. Moro GF, J. H. Calculation of ESR spectra and related Fokker–Planck forms by the use of the Lanczos algorithm J Chem Phys 1981;74:3757.
62. SpinDynamics portal, available from: http://spindynamics.org/
63. Lee C, Yang W, Parr R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. Phys Rev B. 1988(37):785–9.
64. Becke AD. A New Mixing of Hartree-Fock and Local Density-Functional Theories. Journal of Chemical Physics. 1993;98(2):1372-7.
65. Hermosilla L, De la Vega JMG, Sieiro C, Calle P. DFT Calculations of Isotropic Hyperfine Coupling Constants of Nitrogen Aromatic Radicals: The Challenge of Nitroxide Radicals. J Chem Theory Comput. 2011;7(1):169-79.
66. Barone V, Cimino, P. Validation of the B3LYP/N07D and PBE0/N07D Computational Models for the Calculation of Electronic g-Tensors. Chem Phys Lett. 2008;454:139-43.
67. Barone V, Bloino J, Biczysko M. Validation of the DFT/N07D computational model on the magnetic, vibrational and electronic properties of vinyl radical. Physical chemistry chemical physics : PCCP. 2010;12(5):1092-101. Epub 2010/01/23.
68. Barone V, Cimino P, Stendardo E. Development and validation of the B3LYP/N07D computational model for structural parameter and magnetic tensors of large free radicals. J Chem Theory Comput. 2008;4(5):751-64.
69. DREAMSLAB. Available from: http://dreamslab.sns.it/?pag=downloads.
70. Maier JA, Martinez C, Kasavajhala K, Wickstrom L, Hauser KE, Simmerling C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. J Chem Theory Comput. 2015;11(8):3696-713.
71. Case DA, Babin, V., Berryman, J.T., Betz, R.M., Cai, Q., Cerutti, D.S., Cheatham, III, T.E., Darden, T.A., Duke, R.E., Gohlke, H., Goetz, A.W., Gusarov, S., Homeyer, N., Janowski, P., Kaus, J., Kolossváry, I., Kovalenko, A., Lee, T.S., LeGrand, S., Luchko, T., Luo, R., Madej, B., Merz, K.M., Paesani, F., Roe, D.R., Roitberg, A., Sagui, C., Salomon-Ferrer, R., Seabra, G., Simmerling, C.L., Smith, W., Swails, J., Walker, R.C., Wang, J., Wolf, R.M., Wu X. and Kollman P.A. . 2014. p. AMBER 14, University of California, San Francisco. .
72. Aduri RP, B. T.; Saro, P.; Taniga,H.; Schlegel, H. B.; and SantaLucia, J. Jr. AMBER Force Field Parameters for the Naturally Occurring Modified Nucleosides in RNA. J Chem Theory Comput. 2007(3):1464-75.
73. Cornell WDC, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M., Jr.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P.A. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. J Am Chem Soc 1995;117:5179-97.
74. Hartree DR. The wave mechanics of an atom with a non-Coulomb central field Part I theory and methods. P Camb Philos Soc. 1928;24:89-110.
75. Fock V. Approximation method for the solution of the quantum mechanical multibody problems. Z Phys. 1930;61(1-2):126-48.
76. Cieplak PC, Cornell WD, Bayly C, Kollman, P. A. Application of the multimolecule and multiconformational RESP methodology to biopolymers: charge derivation for DNA, RNA, and proteins. J Comput Chem. 1995;16:1357-77.
77. Wolfram Research, Inc., Mathematica, Version 11.0, Champaign, IL 2016.

78. Tan PNS, M.; Kumar, V. Introduction to Data Mining, Addison-Wesley, chapter Cluster Analysis: Basic Concepts and Algorithms. . 2006:487–568.
79. Marchetto RS, S.; Nakaie, C. R. . A Novel Spin-Labeled Amino Acid Derivative for Use in Peptide Synthesis: (9-Fluorenylmethyloxycarbonyl-)2 ,2,6,6- tetramethylpiperidine-N-oxyl-4-amino-4-carboxylAicci d. J Am Chem SOC 1993;115:11042-3.
80. Stoller SS, G.; Baranova, T. Y.; Bennati, M.; Diederichsen, U. TOPP: A Novel Nitroxide-Labeled Amino Acid for EPR Distance Measurements Amino Acid for EPR Distance Measurements. Angew Chem Int Ed 2011;50:9743 –6.

4. Paper 1: Density functional theory studies of MTSL nitroxide side chain conformations attached to an activation loop

Maria Grazia Concilio,[1]* Alistair J. Fielding,[1]* Richard Bayliss,[2] and Selena G. Burgess[2]

[a]*The Photon Science Institute and School of Chemistry, EPSRC National EPR Facility and Service, University of Manchester, Manchester, M13 9PY, United Kingdom.*

[b]*Astbury Centre for Structural and Molecular Biology, Faculty of Biological Sciences, University of Leeds, Leeds LS2 9JT, United Kingdom.*

Abstract

A QM method rooted on DFT has been employed to determine conformations of the methane-thiosulfonate spin label (MTSL) attached to a fragment extracted from the activation loop of Aurora-A kinase. The features of the calculated energy surface revealed low energy barriers between isoenergetic minima, and the system could be described in a population of 76 rotamers that can be also considered for other systems since it was found that the χ_3, χ_4 and χ_5 do not depend on the previous two dihedral angles. Conformational states obtained were seen to be comparable to those obtained in the α-helix systems studied previously, indicating that the protein backbone does not affect the torsional profiles significantly and suggesting the possibility to use determined conformations for other protein systems for further modelling studies.

Corresponding authors:

*Maria Grazia Concilio
mariagrazia.concilio@postgrad.manchester.ac.uk

*Alistair J. Fielding
alistair.fielding@manchester.ac.uk

4.1 Introduction

The most widely employed spin label for studies of the structure and dynamics of biomolecules (1, 2), through electron paramagnetic resonance (EPR) spectroscopy, is the MTSL Figure 4.1.1. shows a model of MTSL attached to a fragment extracted from the activation loop of Aurora-A kinase, a serine/threonine protein kinase that regulates many cellular pathways and is overexpressed in a number of cancers (3, 4).

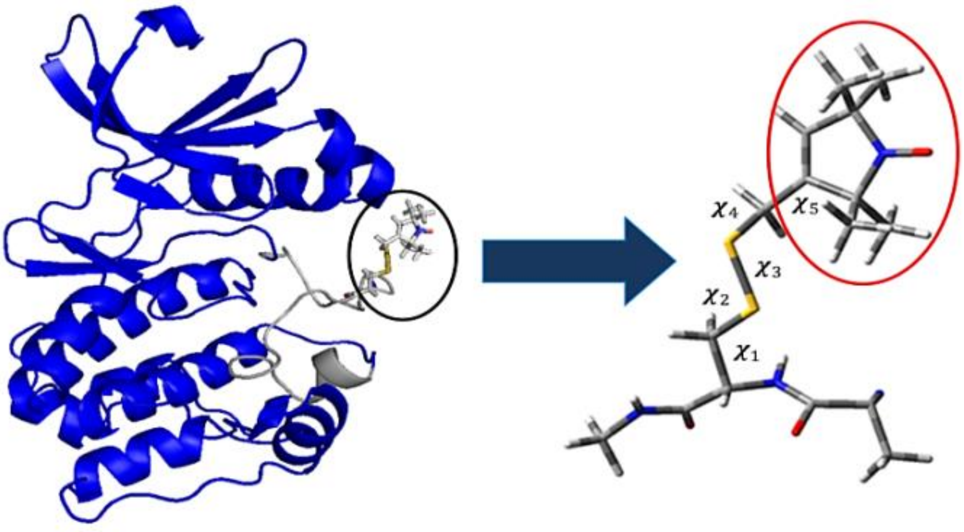


Figure 4.1.1: Structure of the Aurora-A kinase domain with the MTSL side chain (black circle) attached at position 288 within the activation loop which comprises residues 274–299 (grey). The unit peptide extracted for the DFT analysis is indicated by a blue arrow. All five dihedral angles of the MTSL side chain are enlarged and shown, with the pyrroline ring containing the nitroxide group highlighted in a red circle.

The MTSL is endowed of high flexibility and the knowledge of the potential surface is required in order to model its internal dynamics (5-7). In this work, a QM method based on DFT was employed to determine features of the energy surface of the five dihedral angles of MTSL. The DFT theory was employed since it induces the electron correlations effect, representing an appropriate method to reproduce short electrostatic interactions between sulphur atoms and backbone atoms (7). In the previous literature, similar approaches were adopted by Tombolato et al. (5) using the Hartree–Fock theory to study conformations of the MTSL in α-helix systems of the T4 Lysozyme protein (5-7) and in myoglobin (8), results obtained were used to complement subsequent MD studies (9, 10). Considering this system, the question arose whether the conformational states determined in the α-helix are the same in the activation loop of a different protein. A conformational analysis of the potential energy surfaces of the MTSL side chain attached to a fragment extracted from activation loop of Aurora-A kinase protein was performed with the purpose to characterize its geometrical parameters and describe the system in a limited number of rotamers. This work was carried out to establish a basis for more advanced modelling approaches involving MD simulations that require different initial starting conformations of the MTSL for appropriate statistical analyses.

4.2. Material and methods

4.2.1 QM calculations of the conformations of the MTSL

A short unit peptide was extracted from the X-ray crystal structure of the Aurora-A kinase domain (residues 122-403 C290A C393A; PDB 4CEG (11) with a resolution of 2.10 Å and R-value of 0.202) obtained after minimization and equilibration processes performed using the AMBER 14 package (12) in conjunction with ff14SB protein force field (13) in order to clean the structure and to remove bad contacts. Subsequently, the protein was solvated using the Extended Simple Point Charge (SPC/E) water model (9721 water molecules) in a truncated octahedral box with a buffer of 12 Å between the protein atoms and the edge of the box. Afterwards, a short energy minimization was performed in two steps using the Simulated Annealing with NMR-derived Energy Restraints (SANDER) module of AMBER. In the first stage, the water molecules and counter ions were relaxed with 200 cycles of minimization. In the second step, the entire system as a whole was relaxed with 1000 cycles of minimization. Subsequently, the system was heated at constant volume for 20 ps from 10 to 300 K with 10 kcal/mol weak restraints on the protein. This process was followed by two equilibration steps: the first was performed at constant pressure (1 atm) and temperature (300 K) for 200 ps with no restraints and the second was performed in a microcanonical (NVE) ensemble for 1 ns. Relaxed scans were performed subsequently using subsystems shown in Figure 4.2.1.1 where the MTSL side chain was gradually built by adding atoms to the unit peptide extracted from the crystal structure of the Aurora-A kinase.

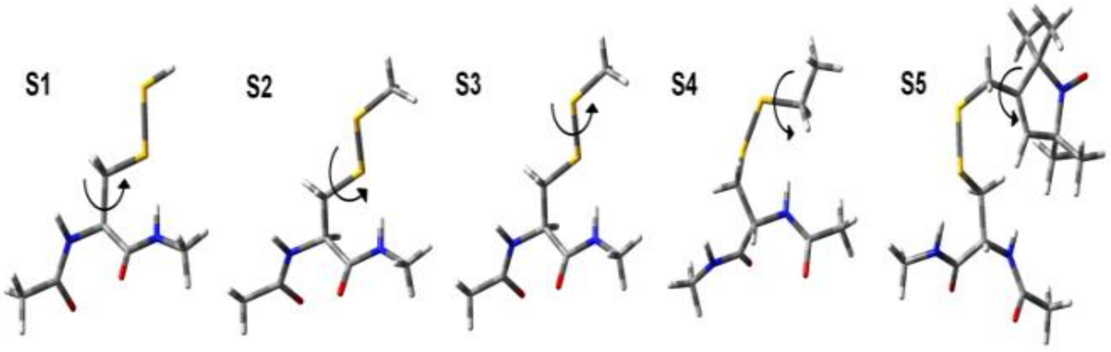


Figure 4.2.1.1: Subsystems considered for the calculations of the torsional energy profiles about the χ_1 (S1), χ_2 (S2), χ_3 (S3), χ_4 (S4) and χ_5 (S5) dihedral angle.

In order to provide reliable reproductions of the equilibrium geometries of the rotamers, the energy torsional profiles V_i (χ_i) were calculated by performing relaxed scans in thirty-seven steps of 10° with the B3LYP hybrid functional (14, 15) and the 6-31G(d) basis set (16) in gas phase around each dihedral angle (χ_1, χ_2, χ_3, χ_4 and χ_5) of the MTSL side chain. The same model system was used in previous work (5-7, 9, 10). The ab initio relaxed scans were performed in 37 steps of 10° using the *opt* = *mod redundant* keyword in the Gaussian 09 software (17) that fixes coordinates but optimizes or relaxes the other atoms. These scans helped to identify the minima of the torsional energy profiles of all five dihedral angles. The Gibbs free energies were determined with the B3LYP hybrid functional and the 6-31G(d) basis set using the *Freq* = *hindrot* keyword.

4.2.2 Experimental section

The expression construct for Aurora-A 122-403 T287A, T288C, C290A, C393A was produced in earlier work for recombinant expression of the protein in E. coli with a N-terminal TEV cleavable His6-tag. Expression and purification of the kinase was carried out as stated in Burgess & Bayliss (2015) with the following modifications. After TEV cleavage and affinity chromatography to remove the His6-tag and His6-tagged TEV, 5 mM DTT was added to Aurora-A containing fractions and left to incubate overnight at 4 °C to ensure C288 was in the reduced form to allow subsequent modification with MTSL. The reducing agent was removed by desalting the protein on a HiPrep 26/10 Desalting column as per the manufacturer's instructions (GE Healthcare) into 20 mM Tris pH 7.0, 0.2 M NaCl, 5 mM $MgCl_2$ & 10 % (v/v) glycerol (EPR buffer). 10-fold excess MTSL was added to the kinase and incubated overnight at 4 °C in order to spin label the kinase. Any remaining impurities, aggregated protein and excess MTSL were removed by size-exclusion chromatography using a HiLoad 16/600 Superdex 200 pg column as per the manufacturer's instructions (GE Healthcare) into EPR buffer. SDS-PAGE analysis was used to identify fractions of high purity MTSL-Aurora-A 122-403 T287A, T288C, C290A, C393A which were concentrated and flash-frozen for future experiments. 50 μM MTSL-Aurora-A 122-403 T287A, T288C, C290A, C393A was used for CW EPR studies. The spin labelling efficiency was equal to 86% as measured following a published procedure. (18) 93.778 GHz CW measurements were performed on a Bruker E560 spectrometer. The magnetic field was calibrated using a Mn2+ power standard (0.02% MgO) and the procedure described by O. Burghaus et al.(19) The rigid limits experimental spectrum was recorded at 150 K. Dual-scan measurements were made in order to avoid hysteresis effects and a modulation frequency of 100 KHz and low microwave power (0.004800 mW) were used to avoid distortion of the lineshape.

4.3. Results and Discussion

The MTSL side chain was gradually built in by adding atoms to the C_α atom of the CH_3–CO–NH–C_α–CO–NH–CH_3 fragment extracted from the minimized and equilibrated structure of the Aurora-A kinase domain in which the Cartesian coordinates of the peptide atoms were kept fixed in the configuration obtained after the previous scan. The relaxed scans were performed following the strategy shown in Figure 4.3.1. The energy minima of the χ_{i+1} dihedral angle were determined at the minima of the χ_i dihedral angle.

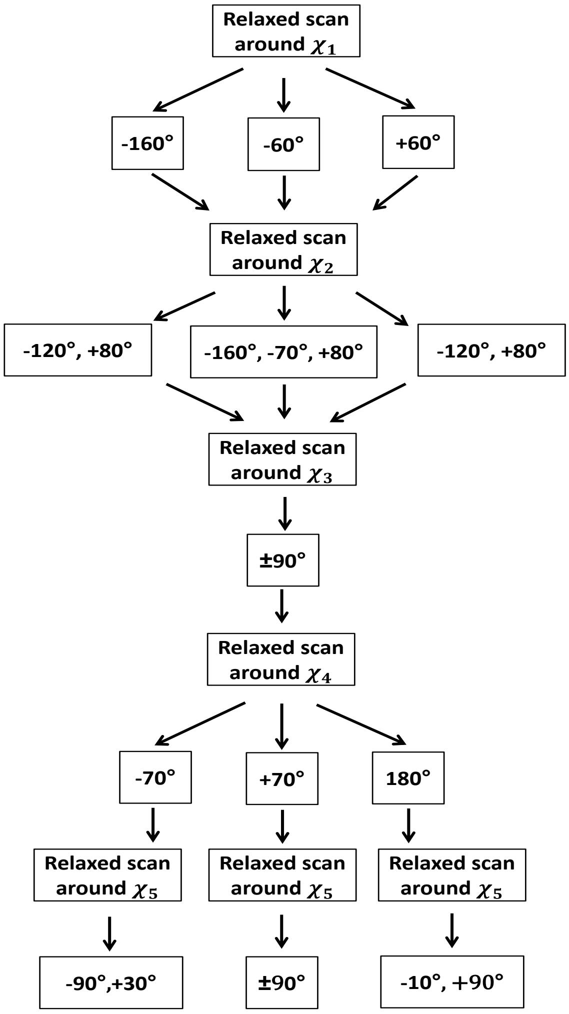


Figure 4.3.1: Strategy used to find minima of the torsional energy profile around χ_1, χ_2, χ_3, χ_4 and χ_5 dihedral angles from the relaxed scans. The first relaxed scan was performed around χ_1, and three minima at −160°, −60° and +60° were found. Subsequently, the energy torsional profile was calculated around χ_2 at the minima of and two minima, one broad between −160° and −120° and another at +80°, were found for χ_1 = −160° and +60°. Three minima (−160°, −70° and +80°) were found for χ_1= −60°. The relaxed scan around χ_3 was performed at the minima of χ_2, and two minima ±90° for all the possible combination of $\chi_1$and $\chi_2$were found. Three minima ±70° and 180° were found in the torsional profiles ofχ_4. The torsional profile ofχ_5 showed minima depending on the values of χ_5.

The first torsional energy profile of the dihedral angle, χ_1 ,was obtained by performing relaxed scans in the range from −180° to +180°, considering the rotation of the C_{α}–$C_{\beta}H_2$ group attached to the CH_3–CO–NH–C_{α}H–CO–NH–CH_3 fragment. The torsional energy profile showed three minima at the values of −160° and ±60° (Figure 4.3.2A).

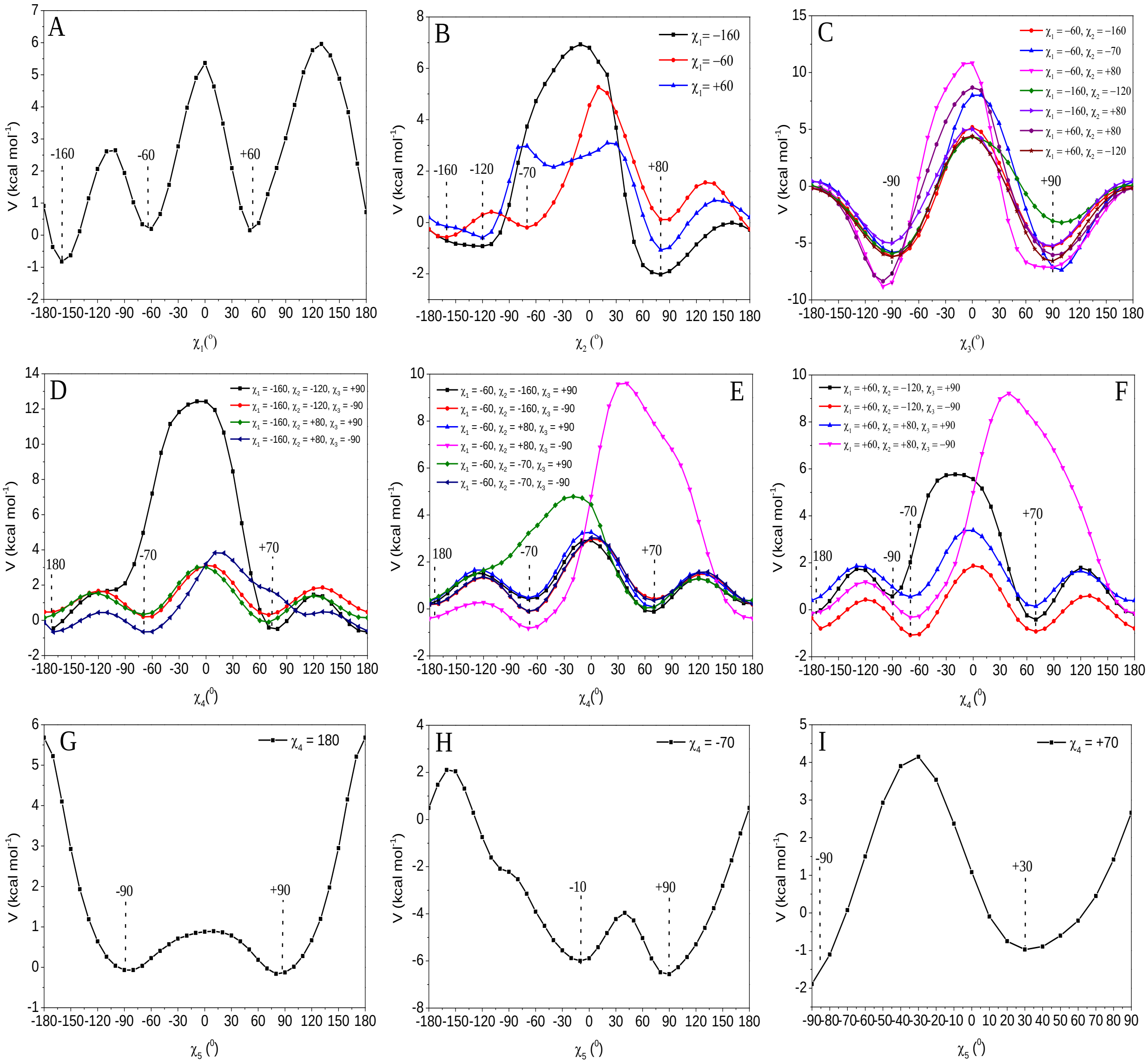


Figure 4.3.2: Energy torsional profiles about the five dihedral angles χ_1, χ_2, χ_3, χ_4 and χ_5 of the MTSL side chain. (A) Energy torsional profile about the dihedral angle, χ_1 (B) Energy torsional profiles about the dihedral angle, χ_2, at the minima of χ_3. (C) Energy torsional profiles about the dihedral angle, χ_3, at the minima of χ_2. (D-F) Energy torsional profile about the dihedral angle, χ_4 was calculated at the minima indicated in the insert. (G) Energy torsional profile about the χ_5 dihedral angle, χ_5 when χ_4 is equal to 180°. (H) Energy torsional profile about the dihedral angle when χ_4 is equal to −70°. (I) Energy torsional profile about the dihedral angle χ_5 when χ_4 is equal to +70.

The high energy barriers corresponded to eclipsed configurations, while the lower energy barriers correlated to configurations in which the S$_\gamma$ and S$_\delta$ atoms were interacting with the protein backbone. The first torsional profile (Figure 4.3.2 A) $V^{(1)}(\chi_1)$ showed three broad minima and two transitions between −160° ↔ −60° and −60° ↔ +60° that were separated by an energy barrier of ∼3 and ∼5 kcal/mol, respectively. These energy barriers are somewhat small, and transitions between them would be expected to occur frequently. Subsequently, the $V^{(2)}(\chi_2)$ torsional profiles were calculated considering the rotation around the $C_\beta H_2$–S_γ group attached to the C_α atom of the CH_3–NH–CO–C_αH–NH–CO–CH_3 fragment. The χ_1 dihedral angle was set at −160° and ±60° (minima in the previous scan). In the calculated $V^{(2)}(\chi_2)$ torsional profiles (Figure 4.3.2 B), energy barriers of ~1–2 kcal/mol between −120° ↔ +80° for χ_1 equal to −160° and +60°, and −160° ↔ −70° and −160° ↔ +80° for χ_1 equal to −60° were observed. These transitions would be expected to occur very frequently since they are separated by very small energy barriers, but are unlikely to do so considering the electrostatic interactions between the S_γ sulphur atom and the protein backbone. The torsional profiles about χ_3 were calculated at all these minima keeping the χ_1 and χ_2 dihedral angles fixed at the selected values shown in Figure 4.3.2C. For the calculation of the $V^{(3)}(\chi_3)$ torsional profiles, the rotation around the $C_\beta H_2 S_\gamma$–$S_\delta CH_2$ group attached to the C_α atom of the CH_3–NH–CO–C_αH–NH–CO–CH_3 fragment was considered. Similar energy torsional profiles and two minima at ± 90° separated by a higher energy barrier of ~14 kcal/mol were found for $V^{(3)}(\chi_3)$ for all seven possible combinations of the minima found for χ_1 and χ_2 (Figure 4.3.2 C). The energy minima corresponded to the structures stabilized by short electrostatic interactions between atoms of the MTSL side chain and the unit peptide model as shown in Figure 4.3.3:

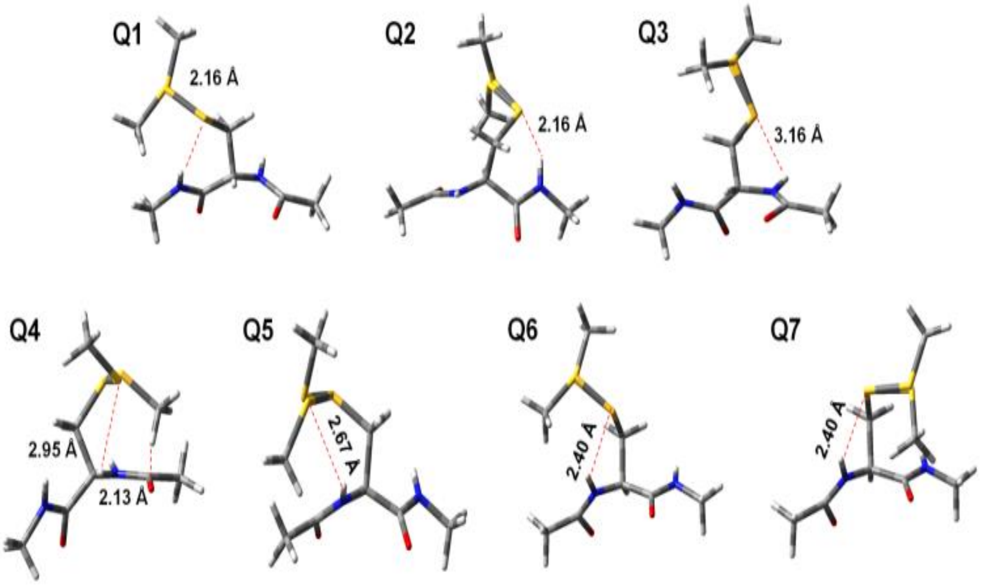


Figure 4.3.3: Interactions between atoms of the MTSL side chain and the unit peptide model for different values of χ_1, χ_2 and χ_3 (red dashed lines). Note that two conformations of χ_3 are shown in each panel. (Q1) Conformation at χ_1 = -160°, χ_2 = -120° and χ_3 = ±90°. (Q2) Conformation at χ_1 = -160°, χ_2 = +80° and χ_3 = ±90°. (Q3) Conformation at χ_1 = -60°, χ_2 = -160° and χ_3 = ±90°. (Q4) Conformation at χ_1 = -60°, χ_2 = -70° and χ_3 = ±90. (Q5) Conformation at χ_1 = -60°, χ_2 = +80° and χ_3 = ±90°. (Q6) Conformation at χ_1 = +60°, χ_2 = -120° and χ_3 = ±90°. (Q7) Conformation at χ_1 = +60°, χ_2 = +80° and χ_3 = ±90°.

The S_γ atom interacts with the NH group for χ_1 = −160° at = −120° (Figure 4.3.3 Q1), +80° (Figure 4.3.3 Q2) and for χ_1= −60° at χ_2= −160° (Figure 4.3.3 Q3). The conformation at χ_1= −60°, χ_2= −70° and χ_3 = + 90° was stabilized by two short electrostatic interactions between the S_δ atom of the MTSL chain and the $C_\alpha H$ group in the fragment, and the CH_2 group and the CO group (Figure 4.3.3 Q4). The S_δ atom interacts with the NH group for χ_1 = −60° and χ_2 = +80° (Figure 4.3.3 Q5). The conformation at χ_1= +60°, χ_2= −120° and +80° and χ_3= ±90° (Figure 4.3.3 Q6 and Q7) was stabilized by interactions between the S_γ atom of the MTSL and the NH group of the protein backbone. The remaining $V^{(4)}(\chi_4)$ and $V^{(5)}(\chi_5)$ torsional profiles (Figure 4.3.2 D-I) were calculated considering the rotations around the S_δ–CH_2 group (for χ_4) and S_δ–CH_2 pyrroline nitroxide ring (for χ_5) with the previous dihedral angles kept fixed at the selected values. Like for $V^{(3)}(\chi_3)$, the $V^{(4)}(\chi_4)$ torsional profiles (Fig. 4.3.2 D-F) were found to be independent from the values of the previous dihedral angles, probably due to the distance from the protein backbone. Similar torsional profiles for $V^{(4)}(\chi_4)$ and three minima (±70° and 180°) were observed for all possible combinations of χ_1, χ_2 and χ_3. Two low energy barriers of 1–2 kcal/mol were found between the minima at ±70° and 180°, and one high energy barrier between –70° and +70° was found. The $V^{(5)}(\chi_5)$ torsional profile was measured at ±70° and 180°, and three different profiles were observed (Figure 4.3.2 G-I). The $V^{(5)}(\chi_5)$ profile at χ_4 = +70° was calculated between –90° and +90° due to a clash between one of the methyl groups of the pyrroline nitroxide ring and the C_β carbon of the MTSL chain for angles over this range. The shape of the torsional profiles obtained in this work is similar to those seen in α-helices and values of the minima were found to be only slightly different (5-7, 9, 10), indicating the protein backbone does not significantly influence the torsional profiles. No relevant changes in the torsional profile of χ_1 and $\chi_2$were observed on extension of the atoms in the unit peptide (data not shown) but longer computational times. This was also observed for χ_3, χ_4 and χ_5. After calculation of the torsional energy profiles for all five dihedral angles in the unit peptide model, a population of 76 conformations was found at the minima of the torsional energy profiles. The torsional profiles showed low energy barriers from 1 to ~20 kcal/mol and isoenergetic minima in the potential energy surface, indicating that the rotamer population would be fully sampled at room temperature and conformational states are expected to be obtained in the same amount. Also, the Gibbs free energy of the different conformations were seen to be comparable and equal to −6,556,624, −6,556,624 and −6,556,623 kcal mol^{-1} for Q2, Q5 and Q7 (characterized by different χ_1), respectively. Similarly, it was observed for Q1 and Q2 (characterized by different χ_2) with Gibbs free energy equal to −6,556,624 and −6,556,622 kcal mol^{-1}, and for conformations with χ_3 equal to +90° and +90° that have a free energy equal to −6,556,240 and −6,556,230 kcal mol^{-1}. After having determined conformers of the MTSL side chain, we tested the effect of the geometric variation of the side chain on the magnetic parameters in order to exclude the contribution of the side chain to the EPR spectrum. Previous literature showed variations of A_{zz} and g_{xx} components upon geometrical variations (NO bond length and the CNOC out-of-plane dihedral angle) in the proxyl radical (20) and in aromatic radical rings (21). Hence, we tested the effect of the geometry of the MTSL side chain on the magnetic properties in order to exclude any contribution of the chain on the spectrum. Six conformations were selected at the minima of the torsional profiles (Fig. 4.3.2) with different values of the dihedral angles (Table 4.1 ESI in in electronic supporting information), and the magnetic parameters were computed at DFT level. Conformers showed comparable magnetic parameters, and minor changes were observed in the 94 GHz EPR spectra (Figure 4.1 ESI in electronic supporting information). This indicated that the addition of the side chain and variations of its geometry do not alter significantly the spin density and shape of the molecular orbital that remained well localized on the NO moiety, like observed in the case of the proxyl radical ring described in previous work (20).

4.4. Conclusions and future work

The conformational analysis of the MTSL side chain on a short fragment of the Aurora-A kinase activation loop revealed torsional profiles comparable to those observed in fragments of **α**-helix studied in previous work (5-7, 9, 10). This indicated that the backbone structure does not influence the torsional profiles significantly. A population of 76 conformers was found at the minima of the torsional profiles, and in addition, it was observed that the χ_3, χ_4 and χ_5 do not depend on the previous two dihedral angles, suggesting that the determined set of rotamers can be considered also for other systems. This analysis can be used to determine some starting conformations for MD simulations of the MTSL spin-labelled Aurora-A kinase or other systems. On the basis our own experience and previous work performed using MD of MTSL spin-labelled proteins (9, 10), it was observed that transitions of χ_4 and χ_5 are much faster than transitions of χ_1, χ_2 and χ_3, so the starting structures can be established fixing χ_1, χ_2 and χ_3. It was observed that conformers have comparable potential and Gibbs free energy allowing the fixing of values χ_1= ±60°, −160°, χ_2= +80° and χ_3= ±90° in order to perform more advanced modelling studies.

Acknowledgments

This work was supported by a studentship from Bruker Ltd. and a Cancer Research UK grant (C24461/A12772 to R.B). The authors would like to acknowledge the use of the EPSRC UK National Service for Computational Chemistry Software (NSCCS) and its staff (Dr. Alexandra Simperler and Dr. Helen Tsui for some technical advice) at Imperial College London in carrying this work. M. G. Concilio acknowledges Dr. A. Baldansuren, Dr. N. Burton and Dr. A. Simperler for useful discussions and feedback.

References

1. Hoff AJ. Advanced EPR: Applications in Biology and Biochemistry1989.
2.Drescher MJ, G. EPR spectroscopy: application in chemistry and biology. Springer, Berlin, ed2012.
3. Barr AR, Gergely F. Aurora-A: the maker and breaker of spindle poles. Journal of cell science. 2007;120(17):2987-96.
4. Bischoff JR, Anderson L, Zhu YF, Mossie K, Ng L, Souza B, et al. A homologue of Drosophila aurora kinase is oncogenic and amplified in human colorectal cancers. Embo J. 1998;17(11):3052-65.
5. Tombolato F, Ferrarini A, Freed JH. Modeling the effects of structure and dynamics of the nitroxide side chain on the ESR spectra of spin-labeled proteins. Journal of Physical Chemistry B. 2006;110(51):26260-71.
6. Tombolato F, Ferrarini A, Freed JH. Dynamics of the nitroxide side chain in spin-labeled proteins. Journal of Physical Chemistry B. 2006;110(51):26248-59.
7. Warshaviak DT, Serbulea L, Houk KN, Hubbell WL. Conformational Analysis of a Nitroxide Side Chain in an alpha-Helix with Density Functional Theory. Journal of Physical Chemistry B. 2011;115(2):397-405.
8. Kuprusevicius EW, G.; Oganesyan, V.S. Prediction of nitroxide spin label EPR spectra from MD trajectories: application to myoglobin. Faraday Discuss. 2011;148:283–98.
9. Sezer D, Freed JH, Roux B. Parametrization, molecular dynamics simulation, and calculation of electron spin resonance spectra of a nitroxide spin label on a polyalanine alpha-helix. The journal of physical chemistry B. 2008;112(18):5755-67. Epub 2008/04/17.
10. Sezer D, Freed JH, Roux B. Using Markov models to simulate electron spin resonance spectra from molecular dynamics trajectories. Journal of Physical Chemistry B. 2008;112(35):11014-27.
11. Burgess SG, Bayliss, R. . The Structure of C290A:C393A Aurora a Provides Structural Insights Into Kinase Regulation. Acta Crystallogr,SectF 2015;71(315).
12. Case DA, Babin, V., Berryman, J.T., Betz, R.M., Cai, Q., Cerutti, D.S., Cheatham, III, T.E., Darden, T.A., Duke, R.E., Gohlke, H., Goetz, A.W., Gusarov, S., Homeyer, N., Janowski, P., Kaus, J., Kolossváry, I., Kovalenko, A., Lee, T.S., LeGrand, S., Luchko, T., Luo, R., Madej, B., Merz, K.M., Paesani, F., Roe, D.R., Roitberg, A., Sagui, C., Salomon-Ferrer, R., Seabra, G., Simmerling, C.L., Smith, W., Swails, J., Walker, R.C., Wang, J., Wolf, R.M., Wu X. and Kollman P.A. . AMBER 14, University of California, San Francisco. 2014.

13. Maier JA, Martinez C, Kasavajhala K, Wickstrom L, Hauser KE, Simmerling C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. J Chem Theory Comput. 2015;11(8):3696-713.

14. Lee C, Yang, W., Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. Phys Rev B. 1988(37):785–9.

15. Becke AD. A New Mixing of Hartree-Fock and Local Density-Functional Theories. Journal of Chemical Physics. 1993;98(2):1372-7.

16. Ditchfield R. Self-Consistent Perturbation-Theory of Diamagnetism .1. Gauge-Invariant Lcao Method for Nmr Chemical-Shifts. Mol Phys. 1974;27(4):789-807.

17. Frisch MJ, Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Scalmani, G., Barone, V., Mennucci, B., Petersson, G. A., Nakatsuji, H., Caricato, M., Li, X., Hratchian, H. P., Izmaylov, A. F., Bloino, J., Zheng, G., Sonnenberg, J. L., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Vreven, T., Montgomery, J. A., Peralta, Jr., J. E., Ogliaro, F., Bearpark, M., Heyd, J. J., Brothers, E., Kudin, K. N., Staroverov, V. N., Kobayashi, R., Normand, J., Raghavachari, K., Rendell, A., Burant, J. C., Iyengar, S. S., Tomasi J., Cossi, M., Rega, N., Millam, J. M., Klene, M., Knox, J. E., Cross, J. B., Bakken, V., Adamo, C., Jaramillo, J., Gomperts, R., Stratmann, R. E., Yazyev, O., Austin, A. J., Cammi, R., Pomelli, C., Ochterski, J. W., Martin, R. L., Morokuma, K., Zakrzewski, V. G., Voth, G. A., Salvador, P., Dannenberg, J. J., Dapprich, S., Daniels, A. D., Farkas, Ö., Foresman, J. B., Ortiz, J. V., Cioslowski, J. and Fox, D. J. Gaussian 09, Revision D.01, Gaussian, Inc., Wallingford CT. 2009.

18. Persson M, Harbridge JR, Hammarstrom P, Mitri R, Martensson LG, Carlsson U, et al. Comparison of electron paramagnetic resonance methods to determine distances between spin labels on human carbonic anhydrase II. Biophysical journal. 2001;80(6):2886-97.

19. Burghaus O, Rohrer M, Gotzinger T, Plato M, Mobius K. A Novel High-Field High-Frequency Epr and Endor Spectrometer Operating at 3 Mm Wavelength. Meas Sci Technol. 1992;3(8):765-74.

20. Pavone M, Cimino P, Crescenzi O, Sillanpaa A, Barone V. Interplay of intrinsic, environmental, and dynamic effects in tuning the EPR parameters of nitroxides: Further insights from an integrated computational approach. Journal of Physical Chemistry B. 2007;111(30):8928-39.

21. Cimino P, Pedone A, Stendardo E, Barone V. Interplay of stereo-electronic, environmental, and dynamical effects in determining the EPR parameters of aromatic spin-probes: INDCO as a test case. Physical Chemistry Chemical Physics. 2010;12(15):3741-6.

22. Cheeseman JR, Trucks GW, Keith TA, Frisch MJ. A comparison of models for calculating nuclear magnetic resonance shielding tensors. Journal of Chemical Physics. 1996;104(14):5497-509.

23. Barone V, Bloino J, Biczysko M. Validation of the DFT/N07D computational model on the magnetic, vibrational and electronic properties of vinyl radical. Physical chemistry chemical physics : PCCP. 2010;12(5):1092-101. Epub 2010/01/23.

24. Barone V, Cimino P, Stendardo E. Development and validation of the B3LYP/N07D computational model for structural parameter and magnetic tensors of large free radicals. J Chem Theory Comput. 2008;4(5):751-64.

25. Adamo C, Cossi M, Barone V. An accurate density functional method for the study of magnetic properties: the PBE0 model. J Mol Struc-Theochem. 1999;493:145-57.

26. Rega N, Cossi M, Barone V. Development and validation of reliable quantum mechanical approaches for the study of free radicals in solution. Journal of Chemical Physics. 1996;105(24):11060-7.

27. Hermosilla L, de la Vega JMG, Sieiro C, Calle P. DFT Calculations of Isotropic Hyperfine Coupling Constants of Nitrogen Aromatic Radicals: The Challenge of Nitroxide Radicals. J Chem Theory Comput. 2011;7(1):169-79.
28. Barone VaC, P. . Validation of the B3LYP/N07D and PBE0/N07D Computational Models for the Calculation of Electronic g-Tensors. Chem Phys Lett. 2008;454:139-43.
29. DREAMSLAB. Available from: http://dreamslab.sns.it/?pag=downloads.
30. Tomasi J, Mennucci B, Cammi R. Quantum mechanical continuum solvation models. Chemical reviews. 2005;105(8):2999-3093. Epub 2005/08/12.
31. Hogben HJ, Krzystyniak M, Charnock GT, Hore PJ, Kuprov I. Spinach--a software library for simulation of spin dynamics in large spin systems. Journal of magnetic resonance. 2011;208(2):179-94. Epub 2010/12/21.
32. Biternas AG, Charnock GTP, Kuprov I. A standard format and a graphical user interface for spin system specification. Journal of magnetic resonance. 2014;240:124-31.
33. Nesmelov YE, Thomas DD. Protein structural dynamics revealed by site-directed spin labeling and multifrequency EPR. Biophysical reviews. 2010;2(2):91-9. Epub 2011/06/21.

Electronic supporting information (ESI)

Density functional theory studies of MTSL nitroxide side chain conformations attached to an activation loop

The EPR parameters were calculated using the Gauge-Independent Atomic Orbital (GIAO) (22) method, the B3LYP hybrid functional and the polarized spilt-valence N07D basis sets (23, 24) that provide values of the *g*- and *A*-tensors in a better agreement with the experimental values, as compared with the EPRII (25) and EPR III (26) basis sets. EPR-II is a double-ζ basis set with a single set of polarization functions and an enhanced *s* part, EPR-III is a triple-ζ basis set including diffuse functions, double *d*-polarizations and a single set of *f*-polarization functions.

In order to obtain accurate hyperfine coupling constants (hcc) for the ^{14}N nucleus, it is important to complete the core-valence *s* space with six *d* Gaussian functions that gives the 6-31G basis set (23, 27, 28). The EPRIII basis set employes only a set of five *d* Gaussian functions (24), while the N07D employs a set of six *d* functions and adds to the double-zeta description of valence orbitals a single sets of optimized core-valence *s* (on all atoms except H), diffuse *p* (on all atoms except H), polarization (on all atoms), and diffuse *d* (on O, F, Cl atoms) functions that improve ist performance as compared with EPR II and EPR III. (28) The N07D basis set was optimized for second- and third-row atoms and has already been used with success for accurate calculation of the magnetic tensors (Δg_{ii} = ±0.0005, ΔA_{ii} = ±1 G) in gas phase and in solution of nitroxide radicals at a reasonable computational cost and can be downloaded from the DREAMSLAB website. (29) The Polarizable Continuum Model (PCM) was used to describe solvation in water since the experimental EPR spectrum was measured in water. (30) All the EPR spectra were simulated using the open source Spinach software library (31), the Euler angles between frames were determined using Spinach GUI. (32) The static powder simulation of the W-band spectrum at 150 K was performed using the powder context with the Lebedev spherical grid rank 131. All the simulations were performed using one component and no line broadening parameters were included. Figure 4.1 ESI shows the comparison between experimental and simulated 94 GHz EPR spectra calculated using parameters magnetic parameters shown in Table 4.1 ESI.

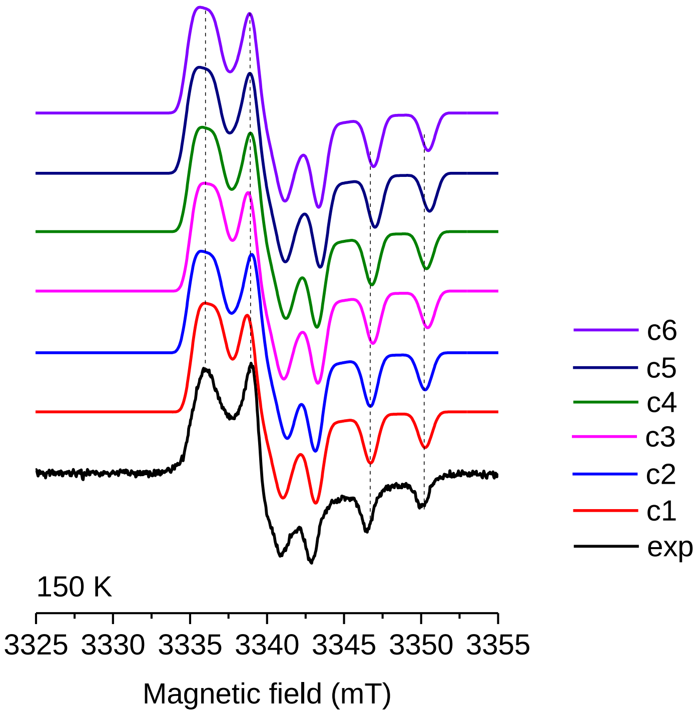


Figure 4.1 ESI: Comparison between experimental and theoretical 94 GHz EPR spectra calculated using conformations C1–C6 described in Table C.1. The experimental spectrum was measured at 150 K.

Table 4.1 ESI: Torsional angles of conformations (C1 to C6) of MSTL and corresponding calculated g-tensor. All the simulations were performed considering a molecular reference frame that has the origin in the N atom of the nitroxide with the z axis parallel to the N-p_z orbital and the x axis parallel to the NO bond, the g- and A-tensors collinear with it. α, β, γ are the Euler angles between the molecular frame and another frame that has origin in the C_α carbon with the z axis perpendicular to the plane N-C_α-CO and the x axis parallel to the C_α-C_b bond. Comparable values of the hyperfine A-tensor were obtained for the 6 conformations that are equal to A_{xx} = 8 Gauss, A_{yy} = 7 Gauss and A_{zz} = 34 Gauss.

	$\chi_1(°)$	$\chi_2(°)$	$\chi_3(°)$	$\chi_4(°)$	$\chi_5(°)$	α	β	γ	g_{xx}	g_{yy}	g_{zz}
C1	-160	180	-90	180	-90	0	0	0	85	62	20
C2	-160	180	-90	180	+90	-60	0	0	86	61	20
C3	-160	180	-90	-70	+90	+20	0	0	85	61	19
C4	-160	180	+90	180	-90	+50	+150	+120	86	61	20
C5	-160	+80	-90	180	-90	+35	+80	-30	86	60	19
C6	-60	180	-90	180	-90	-30	+260	+10	86	62	19

Simulated EPR spectra were seen to reproduce well the experimental powder 94 GHz EPR spectrum of the MTSL spin-labelled Aurora-A kinase, the frequency of measurement and experimental conditions were chosen since they provide high resolution of x, y and z principal components of the g- and hyperfine A-tensors. (33) This indicated that the system can be described as an isotropic distribution of conformations with different geometry, and the orientation of A- and g-tensors with respect to specific reference frame was considered lost.

5. Paper 2: Characterization of the structural dynamics of the activation loop of Aurora-A kinase

Maria Grazia Concilio,[1]* Alistair J. Fielding,[1]* Richard Bayliss,[2] and Selena G. Burgess[2]

[a]*The Photon Science Institute and School of Chemistry, EPSRC National EPR Facility and Service, University of Manchester, Manchester, M13 9PY, United Kingdom.*

[b]*Astbury Centre for Structural and Molecular Biology, Faculty of Biological Sciences, University of Leeds, Leeds LS2 9JT, United Kingdom.*

Abstract

Kinase structural studies are mostly based on protein crystallography, which is limited by the requirement to trap molecules within a crystal lattice. Therefore, characterisation of the conformational dynamics of the kinase activation loop in solution is important to enhance our understanding of molecular processes related to diseases and to support the discovery of small molecule kinase inhibitors. Herein, molecular dynamics (MD) simulations were used to exhaustively sample the conformational space of activation loop of Aurora-A kinase and of methane-thiosulfonate spin label (MTSL) attached to the loop. MD was used to determine structural fluctuations, order parameters and rotational correlation times of the motion of the activation loop and of the MTSL. The theoretical data obtained were used as input for the calculation of room temperature 9.4 GHz continuous-wave (CW) electron paramagnetic resonance (EPR) spectra of the MTSL labelled Aurora-A kinase in solution. Comparison between simulated and experimental spectra revealed that the motion of the protein and spin label occurred on comparable timescales. This work is a starting point for deeper experimental and theoretical studies of the rotational and translational diffusion properties of Aurora-A kinase related to its biological activity.

Corresponding authors:

*Maria Grazia Concilio
mariagrazia.concilio@postgrad.manchester.ac.uk

*Alistair J. Fielding
alistair.fielding@manchester.ac.uk

5.1 Introduction

Site-directed spin labelling (SDSL) combined with EPR spectroscopy is a very powerful technique used widely for studying the structural properties and dynamical processes of biological systems. (1, 2) This has provided valuable information on many proteins such as T4 Lysozyme, (3) lipoxygenase L-1, (4) α-synuclein, (5) bacteriorhodopsin,(6) SNARE (7) and NavMs. (8) Studies on singly labelled proteins can reveal a wealth of information on the tumbling and diffusion properties of the target by analysis of the CW EPR lineshape. (3, 9) A spin label employed to study structural and dynamic properties of biomolecules by EPR spectroscopy is MTSL (Figure 5.1.1) that has been widely characterized in previous work (10, 11) and often the label of choice (12) in different biological systems.

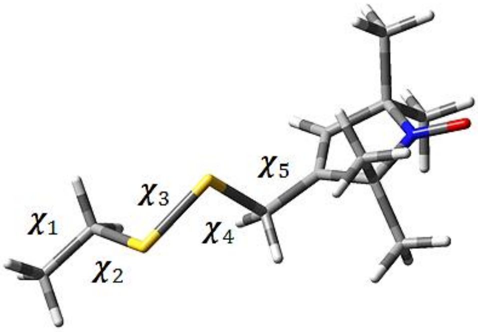


Figure 5.1.1: Structure of the MTSL side chain with five dihedral angles indicated. The nitrogen and the oxygen atoms of the nitroxide group (NO) are represented in blue and red, respectively.

In order to provide an interpretation of the CW EPR spectrum of MTSL spin labelled proteins several different computational approaches to determine the conformational space and dynamics of spin label and subsequently simulate the EPR spectrum have been suggested by several groups. Some approaches are based on Brownian dynamics (BD), introduced by *Robinson et al.* (13) and *Steinhoff et al.* (14) and MD methods introduced by *Budil et al.* (15) and *Oganesyan* (16-18) to generate stochastic dynamical trajectories of the spin labels and/or to derive diffusion parameters such as the rotational diffusion tensor, diffusion tilt angles and expansion coefficients of the orienting potentials. The parameters are then incorporated in the SLE equation for the calculation of the EPR spectrum. Other methods for the calculation of the EPR spectra include the stochastic Markov models (19-21) and temperature scaling or replica exchange methods, (15, 22) including simulated tempering (ST) and parallel tempering (PT) that have been developed to improve sampling and kinetic information. There are also integrated computational approaches (ICA) that link a quantum mechanical (QM) method rooted on the density functional theory (DFT) to the stochastic Liouville SLE equation in the Fokker-Planck form. A similar approach was used to calculate the CW EPR spectra of free radicals in their environments (23-25) and to study the structure and dynamics of the MTSL side chain in protein systems, such as T4 lysozyme. (26, 27) The models produced were in good agreement with experimental data with a reasonable computational cost (28-30) and consistent with data obtained by MD simulations. In these approaches QM methods were used to determine EPR parameters (*g*- and *A*-tensors), while coarse grained methods were used to produce a hydrodynamic model of the diffusion tensor (31) that are the input of the SLE equation for the calculation of the CW EPR spectra. (32) There are also computationally inexpensive approaches that can be used to predict the distribution of the conformations of the spin labels and distances between spin-labelled sites, such as Multiscale Modelling of Macromolecular System (MMM), (33) PRONOX (34) and MtsslWizard. (35) There are also MATLAB (36) based EPR simulation software, such as *EasySpin* (37) and *Spinach* (38) that generate spectra in good agreement with those produced by software developed by J. H. Freed's group implementing the microscopic-order-macroscopic-disorder (MOMD) (39) and slowly relaxing local structure (SRLS) (40) approaches based on the stochastic Liouville equation (SLE) developed by Kubo (41) in the 1963 and then, adapted for EPR simulations by Freed and co-workers for the calculation of slow motion and rigid limit spectra. (42-46)

In this work, we employed MD simulations to characterise the conformational dynamics of Aurora-A kinase's activation loop in solution. A clustering analysis (47) and the principal component (PC) analysis (48) were performed to determine structure populations and dynamics of the system from MD. Structural fluctuations were characterized through the root mean square fluctuations (RMSFs) and order parameters (S^2) calculated using the isotropic reorientational eigenmode (iRED) approach (49) that provided rotational correlation times, τ_R, related to the overall tumbling and reorientational motions of proteins. This approach can be applied only to N-H bonds within the protein backbone and is the only method to calculate order parameters values within the AMBER software and was used in this work to determine the mobility of residue 288 within the activation loop of the kinase, where the MTSL spin label was attached to perform. The τ_R values of MTSL were determined through the fit of the auto-correlation function related to the motion of vectors representing its bonds, as performed in previous work. (17, 50) MD were used to determine input parameters and τ_R for the simulation of the room temperature 9.4 GHz CW EPR spectra of spin labelled Aurora-A kinase.

The purpose of this work is not to propose new methodologies, but to combine existing theoretical (using MD simulations and QM methods) and experimental approaches (CW EPR spectroscopy) in order to characterize the conformational dynamics of the activation loop of Aurora-A kinase, which has never been studied in solution before using EPR spectroscopy. Our approach does not represent a simplification, since we found that the simulated 9.4 GHz and 94 GHz EPR spectra and information content obtained from our work are exactly the same as those obtained from the works cited above (16, 39, 51, 52).

The Aurora-A kinase is a serine/threonine protein kinase that regulates mitotic entry, centrosome maturation and bipolar spindle assembly and is overexpressed in a number of cancers including breast, colorectal, ovarian, and glioma (53, 54). Kinase activity is tightly regulated by conformational changes in a conserved region of the protein known as the activation loop upon phosphorylation of threonine residues at positions 287 and 288 and binding of the activator protein Targeting Protein for Xklp2 (Figure 5.1.2). (55, 56)

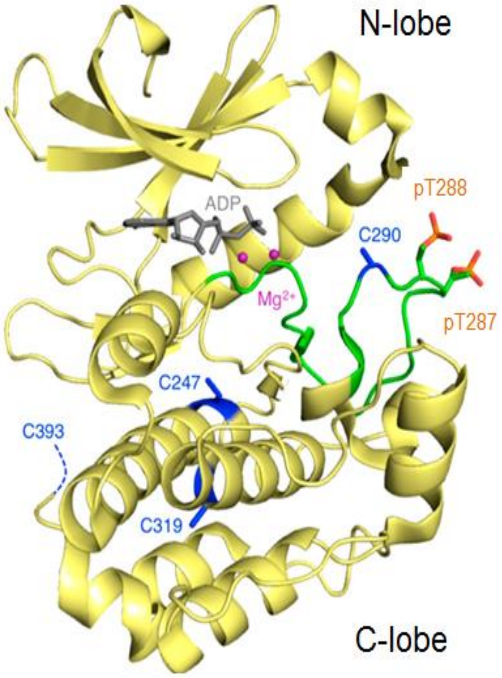


Figure 5.1.2: Cartoon representation of the kinase domain of Aurora-A (PDB 1OL7 (55)). The activation loop is coloured green (residues 274-299) with the phosphorylated threonine residues at position 287 and 288 shown as sticks. Cysteine residues are shown as blue sticks apart from Cys393, which is missing from the crystal structure. The adenosine diphosphate (ADP) is coloured grey with co-ordinated magnesium ions shown as magenta spheres.

The conformation of the activation loop may also be influenced by the binding of inhibitors to the active site of Aurora-A and, in the case of the potent and selective inhibitor MLN8054, the position of the activation loop's main chain is moved by up to 19 Å. (57) Our understanding of kinase structure is mostly based on protein crystallography, which is limited by the requirement of the protein to form crystals and often can not resolve mobile structural regions. (58, 59) Therefore, studies of kinase activation, through characterisation of activation loop conformations in solution are important to enhance our understanding of conformational processes related to diseases and to support the discovery of small molecule kinase inhibitors. In this work, MTSL was introduced into the activation loop of Aurora-A kinase at residue 288. To allow this modification, the wild-type threonine residue was mutated to cysteine, while the threonine at position 287 was mutated to alanine in order to observe conformational changes in a more dynamic activation loop. The solvent exposed cysteine residues at position 290 and 393 were mutated to alanine to prevent labelling with MTSL.

Exhaustively sampling the conformational space of the activation loop and MTSL using free MD simulations is a challenging task due to their weakly structured nature related to inherent flexibility. We also identified several limitations of theoretical approaches used in this work that can be a suggestion for future work on this topic.

5.2 Material and Methods

5.2.1 Force field parameterization of the MTSL side chain and MD simulation details

The latest AMBER force field (60) (ff14SB), recommended for the study of protein dynamics was extended in order to perform MD simulations with AMBER 2014. (61) The structure of the MTSL side chain was first optimized with the B3LYP hybrid functional (62, 63) using the 6-31G(d) basis set (64). This level of theory was used as it provided accurate experimental geometries of the MTSL spin label (19) in previous work. Subsequently, in order to reproduce the electrostatic potential and hydrogen-bonding properties of the MTSL side chain, the atom-centred point charges were calculated with Hartree-Fock theory (65, 66) and the 6-31G(d) basis set, these DFT calculations were performed using Gaussian 09 revision d.01. (67) The electrostatic potential was calculated using the restrained electrostatic potential (RESP) procedure. (68) The atom charges determined to extend the AMBER force field to the MTSL side chain were comparable with those reported in literature and are shown in Table 5.1 in electronic supporting information (ESI). (19, 30, 69) In order to parametrize the NO group, the bond angles, angle bending, regular torsions and non-bond interactions parameters were taken from Stendardo *et al* (30). MD simulations were carried out on MTSL spin labelled Aurora-A kinase, using the X-ray crystal structure of the Aurora-A kinase domain (residues 122-403 C290A C393A; PDB 4CEG (57) with a resolution of 2.10 Å and R-value of 0.202). The PDB file was edited in the website WHATIF (70) to mutate the native Threonine-288 to cysteine. Alanine substitutions were made at residues T287, C290 and C393 in order to produce a structural model similar to the experimentally studied protein. The missing crystallographic hydrogens and MTSL were added using the LEaP module AMBER. The protein was solvated using the Extended Simple Point Charge (SPC/E) water model (11896 water molecules) in a truncated octahedral box with a buffer of 12 Å between the protein atoms and the edge of the box. This water model was used as it predicts viscosities and diffusion constants closer to the experimentally observed data and reproduces crystallographic water positions more accurately than the commonly used Transferable Intermolecular Potential 3 Point (TIP3P) water model. (2, 71, 72) The initial configuration of the MTSL was determined fixing χ_1, χ_2 and χ_3 equal to -60^0, +80^0 and +90^0, respectively. These values were determined at the minima of the torsional profiles calculated in a previous work (73). The MD simulation was performed in three main steps: minimization, equilibration and production. Firstly, a short energy minimization was performed in two steps in order to clean the structure and to remove bad contacts using the Simulated Annealing with NMR-derived Energy Restraints (SANDER) module of AMBER. Three chloride ions were used to neutralize the net charge. In the first stage, the water molecules and counter ions were relaxed with 200 cycles of minimization. In the second step, the entire system as a whole was relaxed with 1000 cycles of minimization. Subsequently, the system was heated at constant volume for 20 ps from 10 K to 300 K with 10 kcal/mol weak restraints on the protein. This process was followed by two equilibration steps, the first was performed at constant pressure (1 atm) and temperature (300 K) for 200 ps with no restraints and the second was performed in a microcanonical (NVE) ensemble for 1 ns. The production step was performed in the NVE ensemble in five steps up to 470 ns collecting time windows equal to 270 ns, 320 ns, 370 ns, 420 ns and 470 ns in order to obtain a good statistical sampling quality. The NVE ensemble was used since the interest of this work is on the dynamics of a system that would be perturbed by the use of a thermostat. This system is large enough (with 14000 atoms in total) that microcanonical and canonical ensemble are almost the same, the Δt between the initial and the final frame in the production trajectory was equal to 10 K. All the calculations were performed using 1 fs integration time step and coordinates were collected every 2 ps, a total of 235000 frames were collected. The dynamical long-range electrostatic interactions were treated using a particle mesh Ewald (PME) algorithm with default parameters and a 10 Å cut-off Lennard-Jones. A clustering analysis was performed to identify different conformational states sampled during the MD simulation by grouping molecular structures into subsets based on their conformational similarity. The clustering analysis was performed using a version of CPPTRAJ in the AMBER software clustering 5 PC projections by RMSD using the average-linkage algorithm with a sieve of 250 frames. All the MD simulations were performed using one NVidia GTX 980 GPU and with AMBER software optimized to run entirely on a CUDA enabled NVIDIA GPU using a mixed-precision SPDP (single precision, double precision) model that is comparable with the double precision model on a central processing unit (CPU).(74)

5.2.2 Calculation of EPR parameters (g- and A-tensors) and EPR spectra

The EPR parameters were calculated using the Gauge-Independent Atomic Orbital (GIAO) (75) method, the B3LYP hybrid functional and the latest N07D basis set that has been used with success for accurate calculation of the magnetic tensors (Δg_{ii} = ±0.0005, ΔA_{ii} = ±1 G) in gas phase and in solution of nitroxide radicals at a reasonable computational cost (28, 29) and can be downloaded from the DREAMSLAB website (76). The Polarizable Continuum Model (PCM) was used to describe solvation in water since the experimental EPR spectrum was measured in water (77-79).

All the EPR spectra were simulated using the open source *Spinach* software library, version 1.7.2996.(38) The room temperature 9.4 GHz CW EPR spectrum was simulated using the *gridfree* context that uses Fokker-Planck formalism (80, 81) that generates spatial dynamics using lab space rotation generators for the three-dimensional rotational diffusion $D_R = \frac{1}{6\tau_R}\left(\hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2\right)$, where τ_R is the rotational correlation time and $\hat{L}_i$ represents the angular momentum (not spin) operators acting in the lab space to generate spatial rotations of the spin system from the starting Cartesian coordinates obtained from selected MD frames. The 9.4 GHz and 94 GHz EPR spectra, measured at 150 K, were simulated using Lebedev spherical powder grids with rank 131.

5.2.3 Experimental section

Spin-labelled Aurora-A 122-403 T287A, T288C, C290A, C393A was produced as stated in earlier work (82). 50 μM MTSL-Aurora-A 122-403 T287A, T288C, C290A, C393A was used for CW EPR studies. 9 GHz measurements were performed using a Bruker Micro EMX spectrometer at 298 K, the modulation frequency was set at 100 KHz and the microwave power at 2.0 mW. The spin labelling efficiency was equal to 86% as measured following a published procedure. (83) The 94 GHz CW measurements were performed on a Bruker E560 spectrometer. The magnetic field was calibrated using a Mn^{2+} power standard (0.02% MgO) and the procedure described by O. Burghaus *et al*. (84) Dual-scan measurements were made in order to avoid hysteresis effects.

5.3 Results and Discussion

5.3.1 Characterization of the dynamics of the MTSL spin label

During the MD simulation the MTSL fully probed the space around the point of attachment and explored all the regions surrounding the loop. Figure 5.3.1.1 shows the plot of the transitions of the five dihedral angles over the time and a comparison between dihedral transitions and oscillations in the root mean square deviation (RMSD) of the activation loop and the protein.

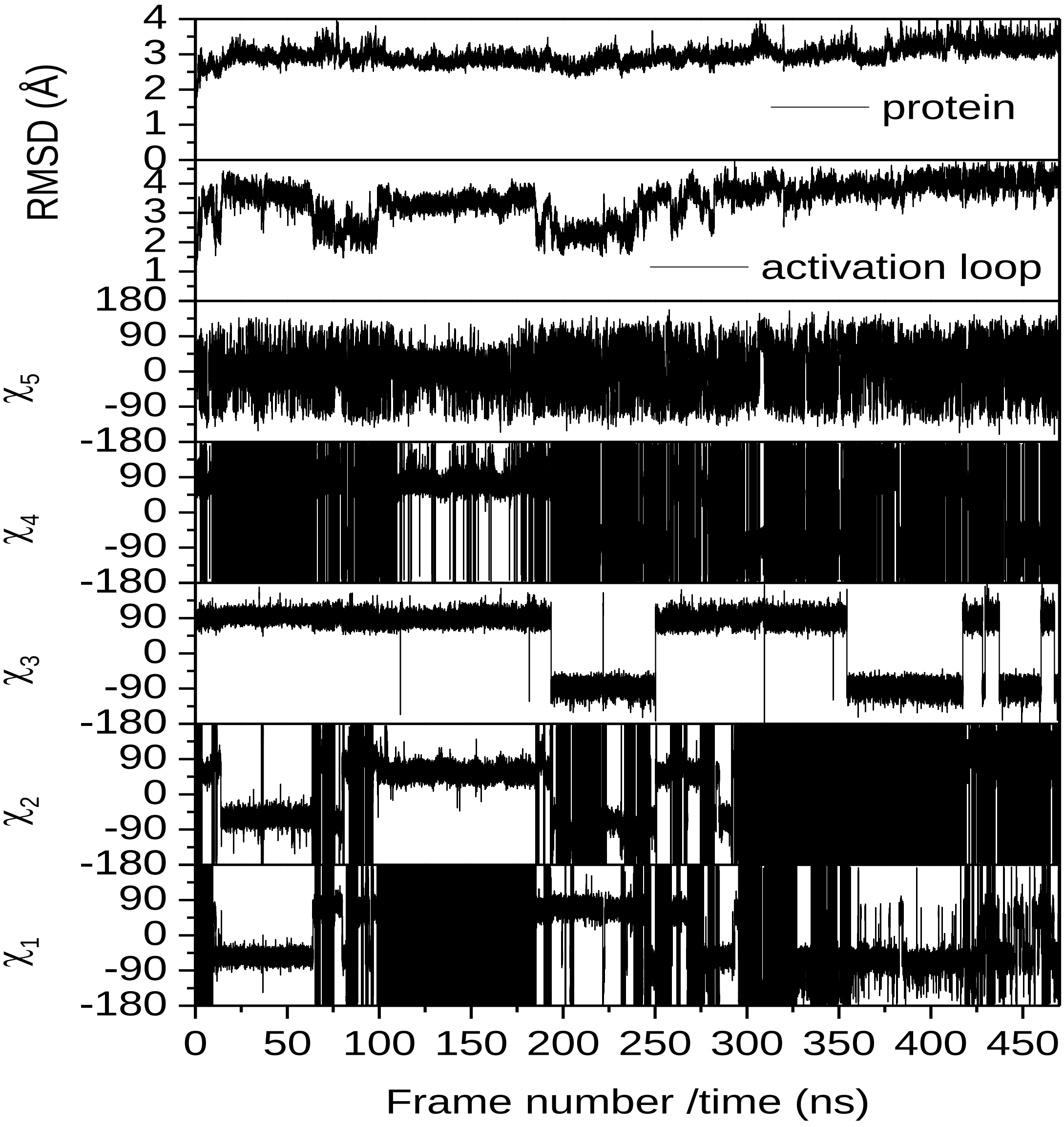


Figure 5.3.1.1: Plots of the transitions of the χ_1, χ_2, χ_3, χ_4 and χ_5 dihedral angles and of RMSD of the protein and of the activation loop and over 470 ns. The RMSDs were calculated considering the Cα, C, N, O atoms and the reference frame was the first frame after the equilibration step.

It was observed that the RMSDs of the activation loop and the protein reasonably converged at ~3-4 Å, indicating not a large variation from the initial X-ray crystal structure. The RMSD of the protein remained constant ~3 Å, while the RMSD of the activation loop was characterized by several fluctuations that corresponded to simultaneous transitions of χ_1 and χ_2 between 10 ns and 64 ns, and between 98 ns and 185 ns. After 250 ns, the RMSD of the activation loop became stable and less transitions of χ_1 and χ_2 were observed, this indicated a coupling between transitions of the dihedral angles and oscillations of the protein backbone. The coupling between transitions of dihedral angles close to each was also observed, *e.g.* the simultaneous transitions of χ_1, χ_2 and χ_3 dihedral angles between 192 ns and 249 ns, and χ_4 and χ_5 between 108 ns and 170 ns.

A comparison between the torsional energy profiles of the five dihedral angles calculated in previous work (73) and those obtained from MD (Figure 5.1 ESI) revealed that the values of the minima and the energy barriers, determined by QM calculations, were well reproduced by MD simulations. Three minima were found in the torsional profile of χ_1 corresponding to -160°,-60° and +60° and the same three minima were found in the plots of the energy torsional profiles determined from MD from which the minima of χ_2 were confirmed and so on for χ_3 and χ_4. The torsional profile of χ_5 had two minima at 0° and +90° in according to the conformation with χ_4 equal to -70° determined from the DFT. The plot of the transitions of the five dihedral angles over time showed slow transitions of χ_1 and χ_2 due to high energy barriers of 4-6 kcal mol^{-1} separating the conformational minima in the torsional energy profiles (Figure 5.1 ESI).Transitions of χ_3 were very slow and less frequent due to higher energy barriers of 10-20 kcal mol^{-1}, while transitions of χ_4 and χ_5 were fast due to low energy barrier of 1-2 kcal mol^{-1}. The Gibbs energy surfaces (Figure 5.2 in ESI) obtained from MD revealed a small difference of ~4.2 kcal mol^{-1} between high and low energy regions indicating that the rotamer population is expected to be fully sampled at room temperature and each conformation has the same overall orientational probability. In order to determine the dynamics of the MTSL, the auto-correlation functions of vectors connecting two atoms of the spin label (the full list is shown in Figure 5.3 and Table 5.2 in ESI) were calculated within the AMBER software using spherical harmonics with rank equal to 2, as described in (17, 85), corresponding to $\langle D^2_{00}\,(\theta(0))D^2_{00}(\theta(t))\rangle$ with D^2_{00} = (3cos$^2\theta(t)$-1)/2 and $\theta(t)$ representing the variation in the time *t* of the direction of the vector. The auto-correlation functions showed a multi-exponential behaviour, indicating different time scales of the motional contributions and were fitted using a bi-exponential function from which two correlation times, τ_1 and τ_2 were extracted. The fit to the auto-correlation function of the vector Cα-C3 was not feasible due to the trend of its curve and unreasonable results were obtained for the vector C3–S1. Values of τ_1 obtained for the vectors connecting atoms of the MTSL side chain were seen to be very different from each other and ranging between 10 ns and 57 ns. In the case of τ_2, quite comparable τ_2 values were obtained that decreased from 3.7 ns for the vector C3–S1 (close to the point of attachment of the spin label) to 2.7 ns for the vector C4-C5 (close to the pyrroline ring), in good agreement with the plots of the transitions of the dihedral angles that showed very slow motion about χ_3 and fast motions about χ_4 and χ_5. The fits to the auto-correlation functions of the vector related to the N-O group revealed values of τ_1 and τ_2 smaller than those observed for the other vectors, indicating that the motion of the N-O in the pyrroline ring is uncoupled from the motion of the spin chain. The two different correlation times obtained from the auto-correlation function can represent slow and fast rotameric transitions of the MTSL spin label. MD revealed that both configurations of the MTSL were fully exposed to the solvent (Figure 5.3.1.2 A) and interacting with residues within the N- and C-lobes of the protein (Figure 5.3.1.2 B and C).

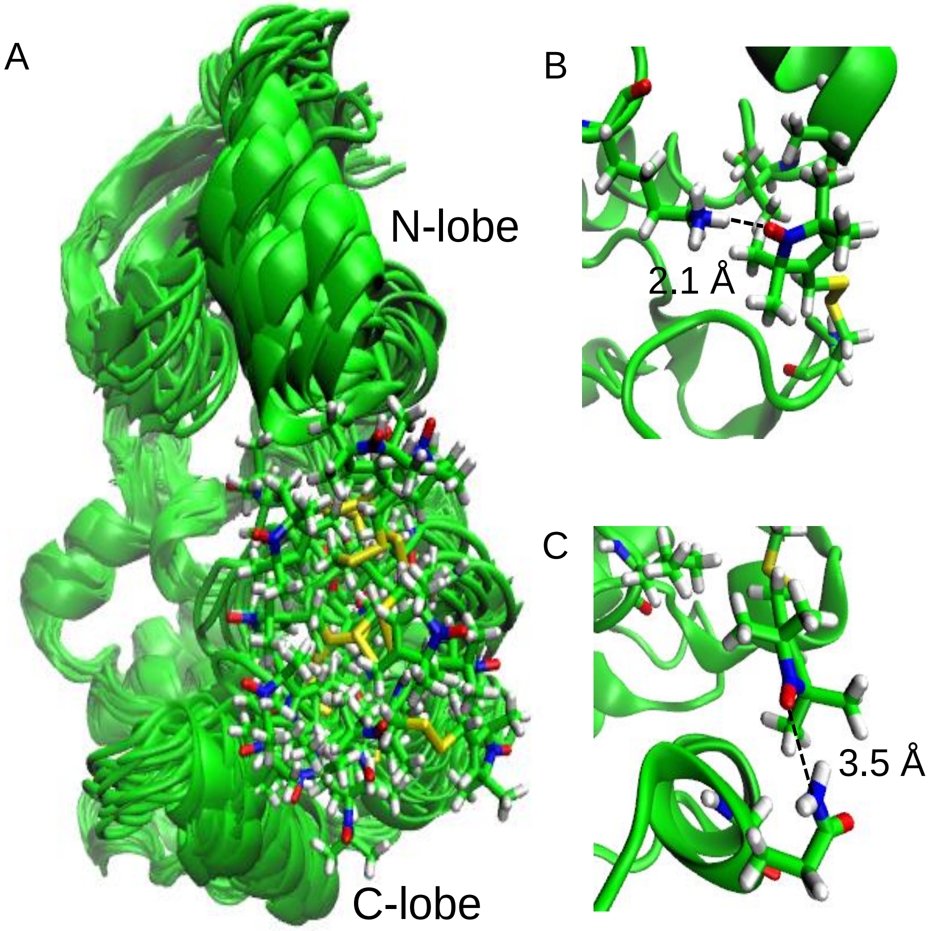


Figure 5.3.1.2: (A) Distribution in the space of the MTSL spin label obtained from the MD simulation, (B) Configuration of MTSL interacting with the lysine 46 N-lobe of the protein. (C) Configuration of MTSL interacting with the glutamine 209 in the C-lobe of the protein.

An isotropic and uniform distribution of conformations that remained constant in space and time was observed, similar results were obtained from the MMM rotamer library (Figure 5.4 in ESI). Faster transitions can be associated to configurations of the spin label fully exposed to the solvent, while slower transitions can be explained considering interactions between atoms of the spin label and the residues of the protein, however, a large exposure to the solvent was observed in all the possible configurations giving evidence of no restricted motions.

5.3.2 Characterization of the conformational states and the activation loop of dynamics of Aurora-A

The clustering and PC analyses were performed to identify the conformational states and to determine the five most dominant modes of motions of the activation loop. The clustering analysis was performed on five different regions of the full MD trajectory, specifically between 0 ns and 270 ns, between 270 ns and 320 ns, between 320 ns and 370 ns and between 420 ns and 470 ns. Figure 5.3.2.1 A shows the representative structures obtained with lowest RMSD to the centroid in the ten most populated clusters determined in the five regions.

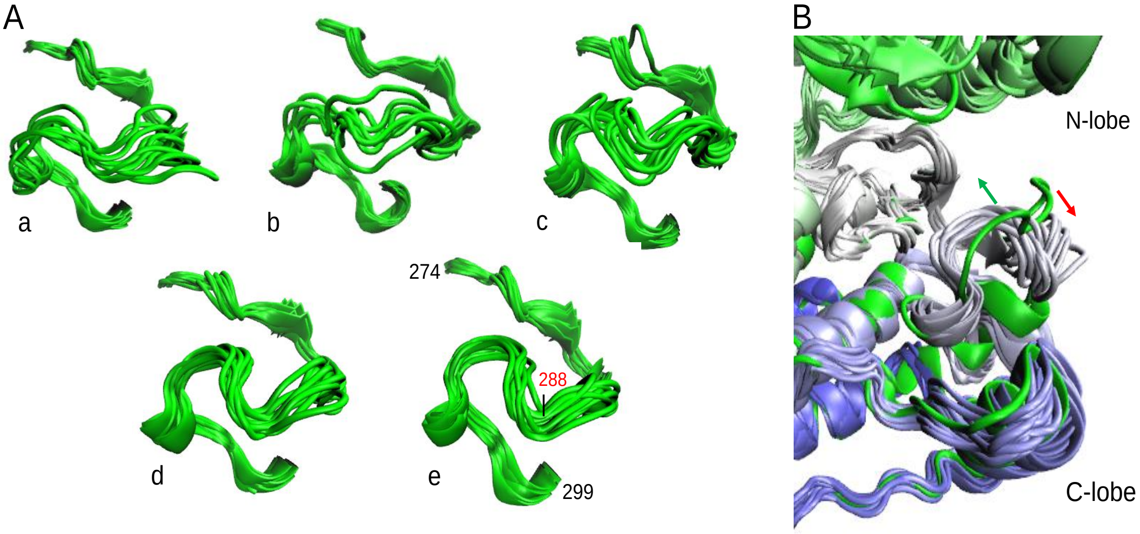


Figure 5.3.2.1: Comparison between conformational states of the Aurora-A kinase activation loop (from residues 274 to 299) obtained from the MD simulation. (A) Representative structures obtained from the clustering analysis at different time windows, (a) 270 ns, (b) 320 ns, (c) 370 ns, (d) 420 ns and (e) 470 ns. The position of residue 288 is indicated in red. (B) Comparison between the energy minimized X-ray structure of Aurora-A (green structure) and the 10 representative structures obtained after 470 ns (white structures). The green and the red arrows indicate the differences between the initial X-ray crystal structure and those obtained in the final region of the MD simulation.

The activation loop was seen to adopt different conformational states in the time windows at 270 ns, 320 ns and 370 ns, while no significant structural changes were observed in the time windows at 420 ns and 470 ns, indicating less dynamics on longer timescales. In the time window between 420 ns and 470 ns, high values of the pseudo F-Statistic (pFS) and the Davies-Bouldin Index (DBI) were obtained and equal to 1731 and 1.55, respectively, indicating a good quality of cluster analysis. Figure 5.3.2.1 B shows a comparison between the ten calculated conformations of the activation loop observed in the last MD time window (420 ns - 470 ns) and the energy minimized X-ray structure. The average RMSD was equal to 3.1 Å, indicating no significant changes from the starting structure. However, it was observed that residues from 289 to 295 moved inward toward the N-lobe of the protein, while residues from 288 to 282 moved toward regions more exposed to the solvent. Figure 5.3.2.2 shows five molecular structures obtained from the PC analysis performed in the last MD time window and the predominant modes of motion (indicated by the arrows) observed within the activation loop and the plot of the root mean squared fluctuations (RMSF) for each active mode.

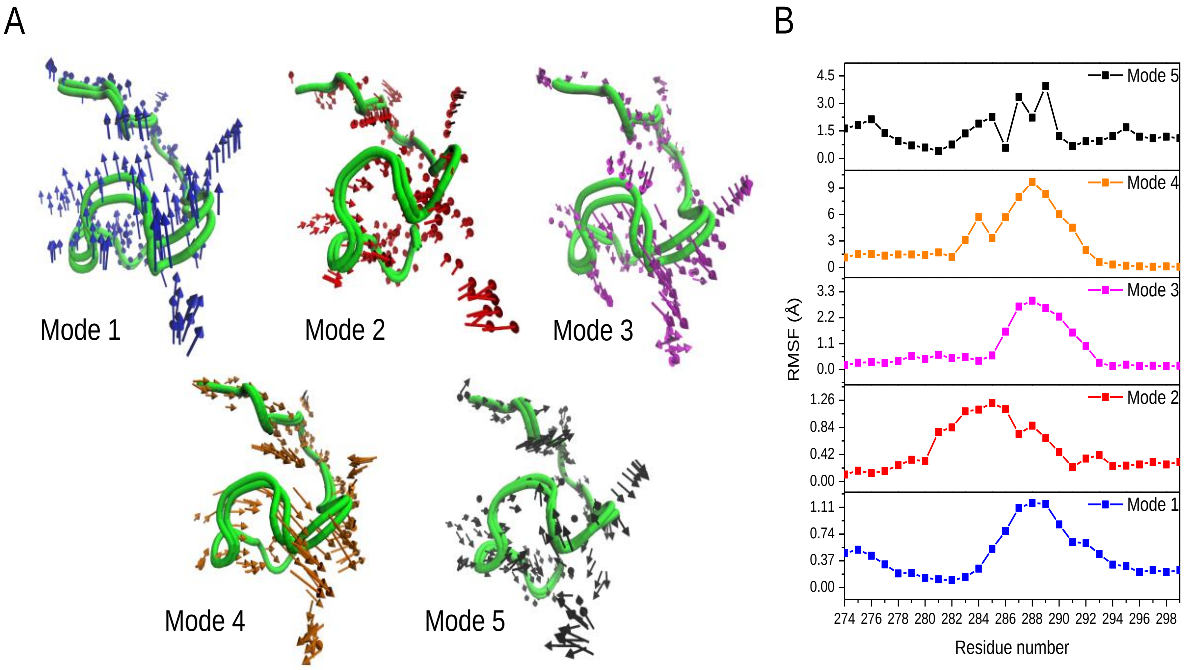


Figure 5.3.2.2: Molecular structures of the activation loop of Aurora-A kinase obtained from the projected frames of the first, second, third, fourth and fifth PC mode. The arrows represent the direction of the motions of the activation loop and the length represents the magnitude of the oscillations. (B) Corresponding plots of the RMSFs for each active mode.

The five PCs corresponded to global protein bending motions going from the smaller fluctuations ~1.1 Å observed in Mode 1 to the wider fluctuations ~ 4.5 Å and 9.0 Å observed in Mode 4 and 5, respectively. The plots of the RMFSs (Figure 5.3.2.2 B) showed large fluctuations, indicated by sharp peaks, in the region between residues 284 and 294 within the activation loop, while much smaller fluctuations were observed for the other residues of the activation loop. Sharper peaks were observed at identical positions indicating concentred fluctuations within the modes. In order to measure how well the motions converged to each other, the histograms of the PC projections were determined and significant overlap was obtained (Figure 5.5 in ESI). Higher fluctuations were observed for modes 3, 4 and 5 indicating higher flexibility when compared with modes 1 and 2, mode 4 was the most flexible. In Figure 5.3.2.2 A, it is possible to observe that the residues of the activation loop fluctuated in the same direction for modes 1, 2, 3 and 4, while fluctuations in different directions were observed for mode 5.

In order to determine timescales of the motions within the activation loop, the iRED analysis (49) was performed along the full MD trajectory in order to determine eigenmodes and amplitudes corresponding to tumbling and reorientational motions of activation loop. This approach incorporates motional correlations between reorientations of backbone N-H bonds in the form of a covariance matrix analysis of internuclear vector orientations that were represented by spherical harmonics of rank 2. In the iRED analysis, the correlation times τ_m of the loop motions, along the reorientational eigenmodes, can be estimated from the time-correlation function $C_{m,l}=\langle a_{m,l} * (\tau + t) a_{m,l}(\tau)\rangle_\tau$ calculated for each mode, where $a_{m,l}$ are the time-dependent amplitudes obtained by projecting MD snapshots on the eigenmodes calculated from corresponding to iRED vectors that were set equal to 26, corresponding to the number of residues of the activation loop (the number of defined iRED vectors have to match the number of eigenmodes calculated). The corresponding τ_m values of these correlation functions can be determined by exponential fitting using the equation $\tau_m \cong \frac{1}{[C_m(0) - C_m(t \to T)\,]} \int_0^T [C_m(0) - C_m(t \to T)\,]dt$, where $C_m(t \to T)$ indicates the plateau value of $C_m(t)$. The iRED approach was used also to calculate the order parameters S^2 of each residue of the activation loop in order to determine the local rigidity. These values corresponded to the amplitude of motion within a cone formed by bond vectors (N-H bonds within the activation loop) such that with increasing rigidity, a smaller cone angle would be produced. Site-specific values of S^2 can range between 0 for a completely disordered bond-vector, to 1 for a completely rigid bond-vector. (86) Figure 5.3.2.3 shows the $C_m(t)$, the corresponding τ_m calculated for each of the 26 modes and the order parameter of each residue of the activation loop calculated with $S^2 = \langle P_2(cos(\Omega_i - -\Omega_j))\rangle_{ij}$, where P_2 is the Legendre polynomial of rank 2 and $\Omega_i - \Omega_j$ is the orientational change of the internuclear vector when going from snapshot *i* to snapshot *j*.

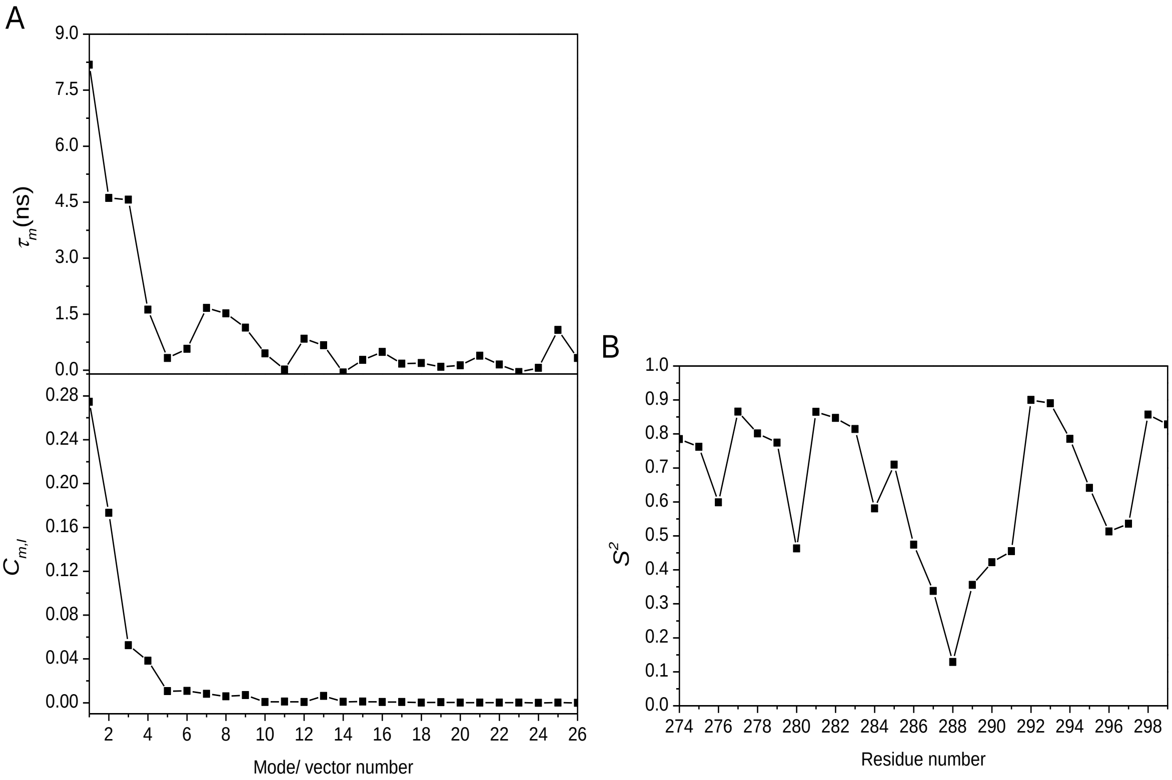


Figure 5.3.2.3: Results obtained from the iRED analysis performed along the full MD trajectory. (A) Time-correlation function, $C_{m,l}$ and corresponding correlation times, τ_m for each mode/ iRED vector. (B) Order parameters, S^2 for each residue of the activation loop.

The modes decay were in good approximation, monoexponentially, and corresponde to individual reorientational modes of motion and rotational tumbling. The extracted correlation times varied between 0.1 ns and 8 ns, indicating a wide distribution of motions within the activation loop and structural fluctuations dominated by slower time-scale events. The calculated S^2 for the residues of the activation loop varied between 0.75 and 0.13. The less flexible regions were observed between residues 274 - 283 and 292 - 299 where S^2 varied between ~0.85 and ~0.50. This is consistent with the plots of the RMSFs (Figure 5.3.2.2) that revealed higher flexibility between residues 284 and 292, where the most flexible residue is 288 (S^2 = 0.13) that corresponds to the point of attachment of the MTSL spin label and due to its length and flexibility increased the mobility of the residue.

5.3.3 Description of the 9.4 GHz CW EPR spectra of Aurora-A kinase

The room temperature 9.4 GHz CW EPR spectrum of Aurora-A kinase showed a lineshape very common in EPR that has been described in previous work as a superposition of two components using two different rotational correlation times, one arising from configurations of the spin label in restricted states and the other arising from configuration of the spin label highly exposed to the solvent. (16, 39, 51, 52) This spectrum reflects a slow tumbling regime motion even at room temperature justifying the inutility of the higher frequencies for the study of this particular construct. In this work, the principal values of *g*- and *A*-tensors were determined from the simulation of the 9.4 GHz EPR spectrum (Figure 5.3.3.1 A), measured at 150 K to freeze out all motion and were equal to g_{xx}=2.0074, g_{yy}=2.0068, g_{zz}=2.0028; A_{xx}= 4 G, A_{yy}= 6 G and A_{zz}= 33 G. These values are in good agreement (Δg ~ ±0.0005 and ΔA ~ ±2 G) with those obtained from 94 GHz measurements (Figure 5.5 in ESI) and DFT calculations performed using an energy optimised structure of the MTSL as shown in Figure 5.3.3.1. The 9.4 GHz EPR spectrum of Aurora-A kinase (Figure 5.3.3.1 B) was fitted using two components in a 1:1 ratio. The two components were obtained with τ_R= 3 ns and τ_R= 11 ns, indicating that 50 % of the spin label exhibited slow dynamics and the remaining 50% exhibited fast dynamics. A low amount of free spin label equal to 14 % was added to the spectrum in Figure 5.3.3.1 B to reproduce the sharp component at 335 mT.

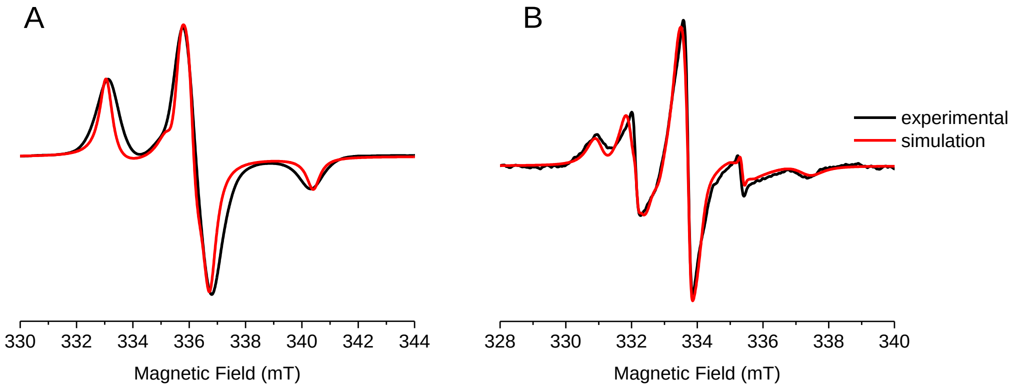


Figure 5.3.3.1: Comparison between experimental and simulated 9 GHz EPR spectra of spin labelled Aurora-A kinase measured at 150 K (A) and at 298 K (B), respectively.

The 1:1 ratio can be explained looking at the plot of the transitions in Figure 5.3.1.1, where the transitions of χ_1, χ_2 and χ_4 that were slow in the range 10-60 ns, 100-234 ns and 350-219 ns, that represent the 54% of the full MD trajectory.

In this work the best fits were obtained using S^2 equal to zero that is very close to the S^2 = 0.13 determined by iRED analysis (Figure 5.3.2.3 B), indicating a low restriction of the

N-H bond and high flexibility of the residue 288. This was also observed in the plot of the RMSFs in Figure 5.3.2.2 B. However, remembering the definition of order parameter that corresponds to the amplitude of motion within a cone formed by a bond vector, an appropriate description of the restriction of the motion of the spin system would require the calculation of the order parameter for the N-O bond, the iRED approach within the AMBER software can be applied only to N-H bonds and therefore only qualitative descriptions of the EPR spectra can be provided. However, an attempt to apply this routine to the MTSL has been carried out in this work by selecting each bond of the spin label, but unreasonable high (~0.75, not suitable to fit the room temperature EPR spectrum) order parameters were obtained, in contraposition to observation in Figure 5.3.1.2 A that showed that the spin label is largely exposed to the solvent. A similar result was also obtained in the case of the order parameters calculated from the angle obtained from the cross product between C_1-C_2 and the N-O vectors along the full MD trajectory. Therefore, these data are not reported in this work.
Figure 5.3.1.2 shows that the spin label is largely exposed to the solvent with no evidence of restriction. The fast motion observed in the EPR spectrum can arise from configurations of the spin label freely rotating in the solvent (Figure 5.3.1.2 A), while the slow motions can arise from configurations in which the motion is slowed down by interactions with the protein (Figure 5.3.1.2 B and C). Values of τ_1 equal to 15 ns and τ_2 equal to 1 ns determined from the auto-correlation function of the N-O vector were very close to those required to simulate the EPR spectrum, indicating that the EPR spectrum can arise from the single motion of the NO group. Also values of τ_2 between 2 ns and 3 ns (obtained for the dihedral angles χ_4 and χ_5), determined for the vectors of the side chain, were in good agreement with that those required to simulate the fast component in the EPR spectrum, while values of τ_1 were much higher than those required. This indicated that the interpretation of the exponential decay of the auto-correlation function might not be valid to describe rotational diffusion in the case of restricted motion and other approaches have to be developed. However, a firm conclusion cannot be given due to the lack of literature of a similar analysis performed on all the vectors of the side chain and further investigations are required in future work.

Values of the correlation times of the reorientational motion of the activation loop, determined from the iRED analysis (Figure 5.3.2.3), were equal to 8.1 ns, 4.5 ns, 1.5 ns and 0.1 ns, considering an average of 3.5 ns that is very close to correlation time of the motion of the MTSL spin label, we can conclude that the motion of the protein and spin label occur on the same timescale. The 11 ns needed to simulate the slow component of the EPR spectrum can arise from a motion of the spin label slowed down by interactions with residues of the protein. However, a quantitative description of the effect of the protein environment on the spin label would require a description of the potential energy surface of the motions of the spin label in the protein environment that is a complicated function of space and time. Order parameters have to be determined both at the point of attachment of the tether and on its top in order to determine the degree of restriction of the N-O bond by the protein environment. This would require the development of a new approach that would extend the iRED routine to MTSL side chain in order to provide input parameters for the simulations of the EPR spectra; this can be a suggestion for future work.

5.4 Conclusions

In this work we characterized the structural dynamics of the activation loop of the Aurora-A kinase, the clustering analysis revealed different conformational states between 0.1 ns and 370 ns, after 370 ns conformational states converged to comparable structures indicating less dynamics on longer time scales. A comparison between conformations obtained from MD and the initial X-ray crystal structure revealed small differences, residues from 289 to 295 moved inward toward the N-lobe of the protein, while residues from 288 to 282 moved toward regions more exposed to the solvent, providing new insight useful for the development of new kinase inhibitors. The PC analysis revealed the amplitude of the fluctuations of the residues of the activation loop, the most flexible region included residues between 284 and 292 where RMSFs oscillated between 1.1 Å and 9.0 Å. This result was confirmed by the iRED analysis that provided lower values of the ordering parameters in this region as compared with the other regions of the activation loop.

We also demonstrated that long MD simulations can provide exhaustive sampling of the MTSL rotamers, since multiple transitions of all the dihedral angles were observed, including those of the χ_3 dihedral angle that are not straightforward to obtain. The analysis of the auto-correlation function was performed using different vectors within the MTSL spin label and revealed the presence of slower and faster motions corresponding to configurations of the spin label fully exposed to the solvent and other interacting with other residues of the protein. This result was consistent with the room temperature 9.4 GHz CW EPR spectrum that was simulated summing two components with two different correlation times equal to 3 ns and 11 ns in good agreement with those determined from the auto-correlation function of the N-O bond, indicating the EPR spectrum strongly depends on the motion of the N-O group in the

MTSL side chain, as well as on the motion about the dihedral angle χ_4 and χ_5. Values of τ_2 extracted from the auto-correlation function related to vectors of the side chain representing χ_4 and χ_5 oscillated between 2 ns and 3 ns, while the values of τ_1 oscillating between 20 ns and 57 ns were seen to be too high to simulate the EPR spectra. This indicated that the auto-correlation function might be not valid to describe rotational diffusion in the case of restricted motion and further investigations are required. The iRED revealed correlation times that ranged between 0.1 ns and 8 ns, indicating a wide distribution of motions within the activation loop and structural fluctuations dominated by events on a slower time-scale that was comparable to that related to the motion of the MTSL. In this work we provided an overview of the structure and dynamics of the Aurora-A kinase activation loop and a next step for the testing of inhibitors and binding partners using the 9.4 GHz CW EPR spectrum of Aurora-A kinase.

Acknowledgments

This work was supported by a studentship from Bruker Ltd. to M.G.C. and by a Cancer Research UK grant (C24461/A12772) to R.B. The authors would like to acknowledge the use of the EPSRC UK National Service for Computational Chemistry Software (NSCCS) and its staff (Dr. Alexandra Simperler and Dr. Helen Tsui for some technical advice) at Imperial College London in carrying out this work and the EPSRC National EPR Facility at the University of Manchester. This work is a result of many discussions within the EPR and computational communities. M. G. C. acknowledges Dr. A. Baldansuren and Prof. D. Collison for useful discussions about the EPR lineshape analysis; Dr. A. Simperler for useful feedback.

References

1. Hoff AJ. Advanced EPR: Applications in Biology and Biochemistry1989.

2. Jeschke G. Conformational dynamics and distribution of nitroxide spin labels. Prog Nucl Mag Res Sp. 2013;72:42-60.

3. Zhang Z, Fleissner MR, Tipikin DS, Liang Z, Moscicki JK, Earle KA, et al. Multifrequency electron spin resonance study of the dynamics of spin labeled T4 lysozyme. The journal of physical chemistry B. 2010;114(16):5503-21. Epub 2010/04/07.

4. Gaffney BJ, Bradshaw MD, Frausto SD, Wu F, Freed JH, Borbat P. Locating a lipid at the portal to the lipoxygenase active site. Biophysical journal. 2012;103(10):2134-44. Epub 2012/12/04.

5. Drew SC, Leong SL, Pham CL, Tew DJ, Masters CL, Miles LA, et al. Cu2+ binding modes of recombinant alpha-synuclein--insights from EPR spectroscopy. Journal of the American Chemical Society. 2008;130(24):7766-73. Epub 2008/05/23.

6. Steinhoff HJ. Inter- and intra-molecular distances determined by EPR spectroscopy and site-directed spin labeling reveal protein-protein and protein-oligonucleotide interaction. Biological chemistry. 2004;385(10):913-20. Epub 2004/11/24.

7. Fasshauer D. Structural insights into the SNARE mechanism. Biochimica et biophysica acta. 2003;1641(2-3):87-97. Epub 2003/08/14.

8. Bagneris C, Decaen PG, Hall BA, Naylor CE, Clapham DE, Kay CW, et al. Role of the C-terminal domain in the structure and function of tetrameric sodium channels. Nature communications. 2013;4:2465. Epub 2013/09/21.

9. Sun Y, Zhang Z, Grigoryants VM, Myers WK, Liu F, Earle KA, et al. The internal dynamics of mini c TAR DNA probed by electron paramagnetic resonance of nitroxide spin-labels at the lower stem, the loop, and the bulge. Biochemistry. 2012;51(43):8530-41. Epub 2012/09/27.

10. McHaourab HS, Lietzow MA, Hideg K, Hubbell WL. Motion of spin-labeled side chains in T4 lysozyme. Correlation with protein structure and dynamics. Biochemistry. 1996;35(24):7692-704. Epub 1996/06/18.

11. Berliner LJ, Grunwald J, Hankovszky HO, Hideg K. A novel reversible thiol-specific spin label: papain active site labeling and inhibition. Analytical biochemistry. 1982;119(2):450-5. Epub 1982/01/15.

12. Fielding AJ, Concilio MG, Heaven G, Hollas MA. New developments in spin labels for pulsed dipolar EPR. Molecules. 2014;19(10):16998-7025. Epub 2014/10/25.

13. Robinson BH, Slutsky LJ, Auteri FP. Direct Simulation of Continuous Wave Electron-Paramagnetic Resonance-Spectra from Brownian Dynamics Trajectories. Journal of Chemical Physics. 1992;96(4):2609-16.

14. Steinhoff HJ, Hubbell WL. Calculation of electron paramagnetic resonance spectra from Brownian dynamics trajectories: Application to nitroxide side chains in proteins. Biophysical journal. 1996;71(4):2201-12.

15. Budil DE, Sale KL, Khairy KA, Fajer PG. Calculating slow-motional electron paramagnetic resonance spectra from molecular dynamics using a diffusion operator approach. The journal of physical chemistry A. 2006;110(10):3703-13. Epub 2006/03/11.

16. White G.F, Ottignon L, Georgiou T, Kleanthous C, Moore GR, Thomson AJ, et al. Analysis of nitroxide spin label motion in a protein-protein complex using multiple frequency EPR spectroscopy. Journal of magnetic resonance. 2007;185(2):191-203. Epub 2007/01/16.

17. Oganesyan V. S. A general approach for prediction of motional EPR spectra from Molecular Dynamics (MD) simulations: application to spin labelled protein. Physical chemistry chemical physics : PCCP. 2011;13(10):4724-37. Epub 2011/02/01.

18. Oganesyan VS, Chami F, White GF, Thomson AJ. A combined EPR and MD simulation study of a nitroxyl spin label with restricted internal mobility sensitive to protein dynamics. Journal of magnetic resonance. 2017; 274: 24–35. Epub 2017/03/16.

19. Sezer D, Freed JH, Roux B. Parametrization, molecular dynamics simulation, and calculation of electron spin resonance spectra of a nitroxide spin label on a polyalanine alpha-helix. The journal of physical chemistry B. 2008;112(18):5755-67. Epub 2008/04/17.

20. Sezer D, Freed JH, Roux B. Multifrequency electron spin resonance spectra of a spin-labeled protein calculated from molecular dynamics simulations. Journal of the American Chemical Society. 2009;131(7):2597-605. Epub 2009/02/05.

21. Sezer D, Freed JH, Roux B. Using Markov models to simulate electron spin resonance spectra from molecular dynamics trajectories. Journal of Physical Chemistry B. 2008;112(35):11014-27.

22. Fajer MI, Li HZ, Yang W, Fajer PG. Mapping electron paramagnetic resonance spin label conformations by the simulated scaling method. Journal of the American Chemical Society. 2007;129(45):13840-6.

23. Barone V, Polimeno A. Toward an integrated computational approach to CW-ESR spectra of free radicals. Physical chemistry chemical physics : PCCP. 2006;8(40):4609-29. Epub 2006/10/19.

24. Hermosilla L, Sieiro C, Calle P, Zerbetto M, Polimeno A. Modeling of cw-EPR spectra of propagating radicals in methacrylic polymerization at different temperatures. The journal of physical chemistry B. 2008;112(36):11202-8. Epub 2008/08/19.

25. Zerbetto M, Polimeno A, Barone V. Simulation of electron spin resonance spectroscopy in diverse environments: An integrated approach. Comput Phys Commun. 2009;180(12):2680-97.

26. Tombolato F, Ferrarini A, Freed JH. Dynamics of the nitroxide side chain in spin-labeled proteins. Journal of Physical Chemistry B. 2006;110(51):26248-59.

27. Tombolato F, Ferrarini A, Freed JH. Modeling the effects of structure and dynamics of the nitroxide side chain on the ESR spectra of spin-labeled proteins. Journal of Physical Chemistry B. 2006;110(51):26260-71.

28. Barone V, Cimino P, Stendardo E. Development and validation of the B3LYP/N07D computational model for structural parameter and magnetic tensors of large free radicals. J Chem Theory Comput. 2008;4(5):751-64.

29. Barone VaC, P. . Validation of the B3LYP/N07D and PBE0/N07D Computational Models for the Calculation of Electronic g-Tensors. Chem Phys Lett. 2008;454:139-43.

30. Stendardo E, Pedone A, Cimino P, Menziani MC, Crescenzi O, Barone V. Extension of the AMBER force-field for the study of large nitroxides in condensed phases: an ab initio parameterization. Physical Chemistry Chemical Physics. 2010;12(37):11697-709.

31. Barone V, Zerbetto M, Polimeno A. Hydrodynamic Modeling of Diffusion Tensor Properties of Flexible Molecules. J Comput Chem. 2009;30(1):2-13.

32. Polimeno A, Barone, V. and Freed, J.H. Stochastic Methods for Magnetic Resonance Spectroscopies In Computational Spectrocopy2012.

33. Polyhach Y, Bordignon E, Jeschke G. Rotamer libraries of spin labelled cysteines for protein studies. Physical chemistry chemical physics : PCCP. 2011;13(6):2356-66. Epub 2010/12/01.

34. Hatmal MM, Li Y, Hegde BG, Hegde PB, Jao CC, Langen R, et al. Computer modeling of nitroxide spin labels on proteins. Biopolymers. 2012;97(1):35-44. Epub 2011/07/28.

35. Hagelueken G, Ward R, Naismith JH, Schiemann O. MtsslWizard: In Silico Spin-Labeling and Generation of Distance Distributions in PyMOL. Applied magnetic resonance. 2012;42(3):377-91. Epub 2012/03/27.

36. Penalva-Arana DC, Forshay K, Johnson PTJ, Strickler JR, Dodson SI. Chytrid infection reduces thoracic beat and heart rate of Daphnia pulicaria. Hydrobiologia. 2011;668(1):147-54.

37. Stoll S, Schweiger A. EasySpin, a comprehensive software package for spectral simulation and analysis in EPR. Journal of magnetic resonance. 2006;178(1):42-55.

38. Hogben HJ, Krzystyniak M, Charnock GT, Hore PJ, Kuprov I. Spinach--a software library for simulation of spin dynamics in large spin systems. Journal of magnetic resonance. 2011;208(2):179-94. Epub 2010/12/21.

39. Budil DE, Lee S, Saxena S, Freed JH. Nonlinear-least-squares analysis of slow-motion EPR spectra in one and two dimensions using a modified Levenberg-Marquardt algorithm. J Magn Reson Ser A. 1996;120(2):155-89.

40. Polimeno A, Freed JH. Slow Motional Esr in Complex Fluids - the Slowly Relaxing Local-Structure Model of Solvent Cage Effects. J Phys Chem-Us. 1995;99(27):10995-1006.

41. Kubo R. Stochastic Liouville Equations. J Math Phys. 1963;4(2):174-&.

42. Freed JH, Bruno GV, Polnaszek CF. Electron Spin Resonance Line Shapes and Saturation in Slow Motional Region. J Phys Chem-Us. 1971;75(22):3385-+.

43. Schwartz LJ, Stillman AE, Freed JH. Analysis of Electron-Spin Echoes by Spectral Representation of the Stochastic Liouville Equation. Journal of Chemical Physics. 1982;77(11):5410-25.

44. Nevzorov AA, Freed JH. Spin relaxation by dipolar coupling: From motional narrowing to the rigid limit. Journal of Chemical Physics. 2000;112(3):1413-24.

45. Mason RP, Freed JH. Estimating Microsecond Rotational Correlation Times from Lifetime Broadening of Nitroxide Electron-Spin Resonance-Spectra near Rigid Limit. J Phys Chem-Us. 1974;78(13):1321-3.

46. Schneider DJ, Freed, J.H. . Calculating Slow Motional Magnetic Resonance Spectra: A User's Guide, In Spin Labeling: Theory and Applications, Vol. III. Biological Magnetic Resonance, Plenum: New York ed1989.

47. Shao J, Tanner SW, Thompson N, Cheatham TE. Clustering Molecular Dynamics Trajectories: 1. Characterizing the Performance of Different Clustering Algorithms. J Chem Theory Comput. 2007;3(6):2312-34. Epub 2007/11/01.

48. Galindo-Murillo R, Roe DR, Cheatham TE. On the absence of intrahelical DNA dynamics on the mu s to ms timescale. Nature communications. 2014;5.

49. Prompers JJ, Bruschweiler R. General framework for studying the dynamics of folded and nonfolded proteins by NMR relaxation spectroscopy and MD simulation. Journal of the American Chemical Society. 2002;124(16):4522-34.

50. Kemmerer S, Voss JC, Faller R. Molecular dynamics simulation of dipalmitoylphosphatidylcholine modified with a MTSL nitroxide spin label in a lipid membrane. Biochimica et biophysica acta. 2013;1828(11):2770-7. Epub 2013/08/21.

51. Fanucci GE, Cafiso DS. Recent advances and applications of site-directed spin labeling. Current opinion in structural biology. 2006;16(5):644-53. Epub 2006/09/05.

52. White GF, Field S, Marritt S, Oganesyan VS, Gennis RB, Yap LL, et al. An EPR spin label study of the quinol oxidase, E. coli cytochrome bo3: a search for redox induced conformational changes. Biochemistry. 2007;46(9):2355-63. Epub 2007/02/10.

53. Barr AR, Gergely F. Aurora-A: the maker and breaker of spindle poles. Journal of cell science. 2007;120(Pt 17):2987-96. Epub 2007/08/24.

54. Bischoff JR, Anderson L, Zhu YF, Mossie K, Ng L, Souza B, et al. A homologue of Drosophila aurora kinase is oncogenic and amplified in human colorectal cancers. Embo J. 1998;17(11):3052-65.

55. Rowan FC, Richards M, Bibby RA, Thompson A, Bayliss R, Blagg J. Insights into Aurora-A kinase activation using unnatural amino acids incorporated by chemical modification. ACS chemical biology. 2013;8(10):2184-91. Epub 2013/08/09.

56. Bayliss R, Sardon T, Vernos I, Conti E. Structural basis of Aurora-A activation by TPX2 at the mitotic spindle. Molecular cell. 2003;12(4):851-62. Epub 2003/10/29.

57. Burgess SG, Bayliss, R. . The Structure of C290A:C393A Aurora a Provides Structural Insights Into Kinase Regulation. Acta Crystallogr,SectF 2015;71(315).

58. Endicott JA, Noble ME, Johnson LN. The structural basis for control of eukaryotic protein kinases. Annual review of biochemistry. 2012;81:587-613. Epub 2012/04/10.

59. Huse M, Kuriyan J. The conformational plasticity of protein kinases. Cell. 2002;109(3):275-82.

60. Maier JA, Martinez C, Kasavajhala K, Wickstrom L, Hauser KE, Simmerling C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. J Chem Theory Comput. 2015;11(8):3696-713.

61. Case DA, Babin, V., Berryman, J.T., Betz, R.M., Cai, Q., Cerutti, D.S., Cheatham, III, T.E., Darden, T.A., Duke, R.E., Gohlke, H., Goetz, A.W., Gusarov, S., Homeyer, N., Janowski, P., Kaus, J., Kolossváry, I., Kovalenko, A., Lee, T.S., LeGrand, S., Luchko, T., Luo, R., Madej, B., Merz, K.M., Paesani, F., Roe, D.R., Roitberg, A., Sagui, C., Salomon-Ferrer, R., Seabra, G., Simmerling, C.L., Smith, W., Swails, J., Walker, R.C., Wang, J., Wolf, R.M., Wu X. and Kollman P.A. . AMBER 14, University of California, San Francisco. 2014.

62. Lee CT, Yang WT, Parr RG. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron-Density. Phys Rev B. 1988;37(2):785-9.

63. Becke AD. Density-functional thermochemistry. III. The role of exact exchange. The Journal of chemical physics. 1993(5648–5652).

64. Ditchfield R, Hehre WJ, Pople JA. Self-Consistent Molecular-Orbital Methods .9. Extended Gaussian-Type Basis for Molecular-Orbital Studies of Organic Molecules. Journal of Chemical Physics. 1971;54(2):724.

65. Hartree DR. The wave mechanics of an atom with a non-Coulomb central field Part I theory and methods. P Camb Philos Soc. 1928;24:89-110.

66. Fock V. Approximation method for the solution of the quantum mechanical multibody problems. Z Phys. 1930;61(1-2):126-48.

67. Frisch MJ, Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Scalmani, G., Barone, V., Mennucci, B., Petersson, G. A., Nakatsuji, H., Caricato, M., Li, X., Hratchian, H. P., Izmaylov, A. F., Bloino, J., Zheng, G., Sonnenberg, J. L., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Vreven, T., Montgomery, J. A., Peralta, Jr., J. E., Ogliaro, F., Bearpark, M., Heyd, J. J., Brothers, E., Kudin, K. N., Staroverov, V. N., Kobayashi, R., Normand, J., Raghavachari, K., Rendell, A., Burant, J. C., Iyengar, S. S., Tomasi J., Cossi, M., Rega, N., Millam, J. M., Klene, M., Knox, J. E., Cross, J. B., Bakken, V., Adamo, C., Jaramillo, J., Gomperts, R., Stratmann, R. E., Yazyev, O., Austin, A. J., Cammi, R., Pomelli, C., Ochterski, J. W., Martin, R. L., Morokuma, K., Zakrzewski, V. G., Voth, G. A., Salvador, P., Dannenberg, J. J., Dapprich, S., Daniels, A. D., Farkas, Ö., Foresman, J. B., Ortiz, J. V., Cioslowski, J. and Fox, D. J. Gaussian 09, Revision D.01, Gaussian, Inc., Wallingford CT. 2009.

68. Bayly CI, Cieplak P, Cornell WD, Kollman PA. A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges - the Resp Model. J Phys Chem-Us. 1993;97(40):10269-80.

69. Cornell WD, Piort, C. , Bayly,C. I., Gould, I. R., Merz, K. M. Jr., Ferguson, D. C., Spellmeyer, D.C., Fox , T., Caldwell, J. W. and Kollman, P. A. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules J Am Chem Soc. 1996;118 (9):2309–.

70. Vriend G. What If - a Molecular Modeling and Drug Design Program. J Mol Graphics. 1990;8(1):52-&.

71. Hakansson P, Westlund PO, Lindahl E, Edholm O. A direct simulation of EPR slow-motion spectra of spin labelled phospholipids in liquid crystalline bilayers based on a molecular dynamics simulation of the lipid dynamics. Physical Chemistry Chemical Physics. 2001;3(23):5311-9.

72. Becker O. M. MADJ, Roux B., Watanabe M. . Computational Biochemistry and Biophysics2001.

73. Concilio MG, Fielding AJ, Bayliss R, Burgess SG. Density functional theory studies of MTSL nitroxide side chain conformations attached to an activation loop. Theoretical chemistry accounts. 2016;135(4):1-6.

74. Salomon-Ferrer R, Gotz AW, Poole D, Le Grand S, Walker RC. Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald. J Chem Theory Comput. 2013;9(9):3878-88. Epub 2013/09/10.

75. Cheeseman JR, Trucks GW, Keith TA, Frisch MJ. A comparison of models for calculating nuclear magnetic resonance shielding tensors. Journal of Chemical Physics. 1996;104(14):5497-509.

76. DREAMSLAB. Available from: http://dreamslab.sns.it/?pag=downloads.

77. Tomasi J, Mennucci B, Cammi R. Quantum mechanical continuum solvation models. Chemical reviews. 2005;105(8):2999-3093. Epub 2005/08/12.

78. Barone V, Bencini A, Cossi M, Di Matteo A, Mattesini M, Totti F. Assessment of a combined QM/MM approach for the study of large nitroxide systems in vacuo and in condensed phases. Journal of the American Chemical Society. 1998;120(28):7069-78.

79. Hermosilla L, de la Vega JMG, Sieiro C, Calle P. DFT Calculations of Isotropic Hyperfine Coupling Constants of Nitrogen Aromatic Radicals: The Challenge of Nitroxide Radicals. J Chem Theory Comput. 2011;7(1):169-79.

80. Kuprov I. Fokker-Planck formalism in magnetic resonance simulations. Journal of magnetic resonance. 2016;270:124-35. Epub 2016/07/30.

81. Edwards LJ, Savostyanov DV, Nevzorov AA, Concistre M, Pileio G, Kuprov I. Grid-free powder averages: on the applications of the Fokker-Planck equation to solid state NMR. Journal of magnetic resonance. 2013;235:121-9. Epub 2013/08/15.

82. Concilio MG, Fielding AJ, Bayliss R, Burgess SG. Density functional theory studies of MTSL nitroxide side chain conformations attached to an activation loop. Theoretical chemistry accounts. 2016;135:97. Epub 2016/04/14.

83. Persson M, Harbridge JR, Hammarstrom P, Mitri R, Martensson LG, Carlsson U, et al. Comparison of electron paramagnetic resonance methods to determine distances between spin labels on human carbonic anhydrase II. Biophysical journal. 2001;80(6):2886-97.

84. Burghaus O, Rohrer M, Gotzinger T, Plato M, Mobius K. A Novel High-Field High-Frequency Epr and Endor Spectrometer Operating at 3 Mm Wavelength. Meas Sci Technol. 1992;3(8):765-74.

85. Wong V, Case DA. Evaluating rotational diffusion from protein MD simulations. The journal of physical chemistry B. 2008;112(19):6013-24. Epub 2007/12/07.

86. Kleckner IR, Foster MP. An introduction to NMR-based approaches for measuring protein dynamics. Bba-Proteins Proteom. 2011;1814(8):942-68.

Electronic supporting information (ESI)

Characterization of the structural dynamics of the activation loop of Aurora-A kinase

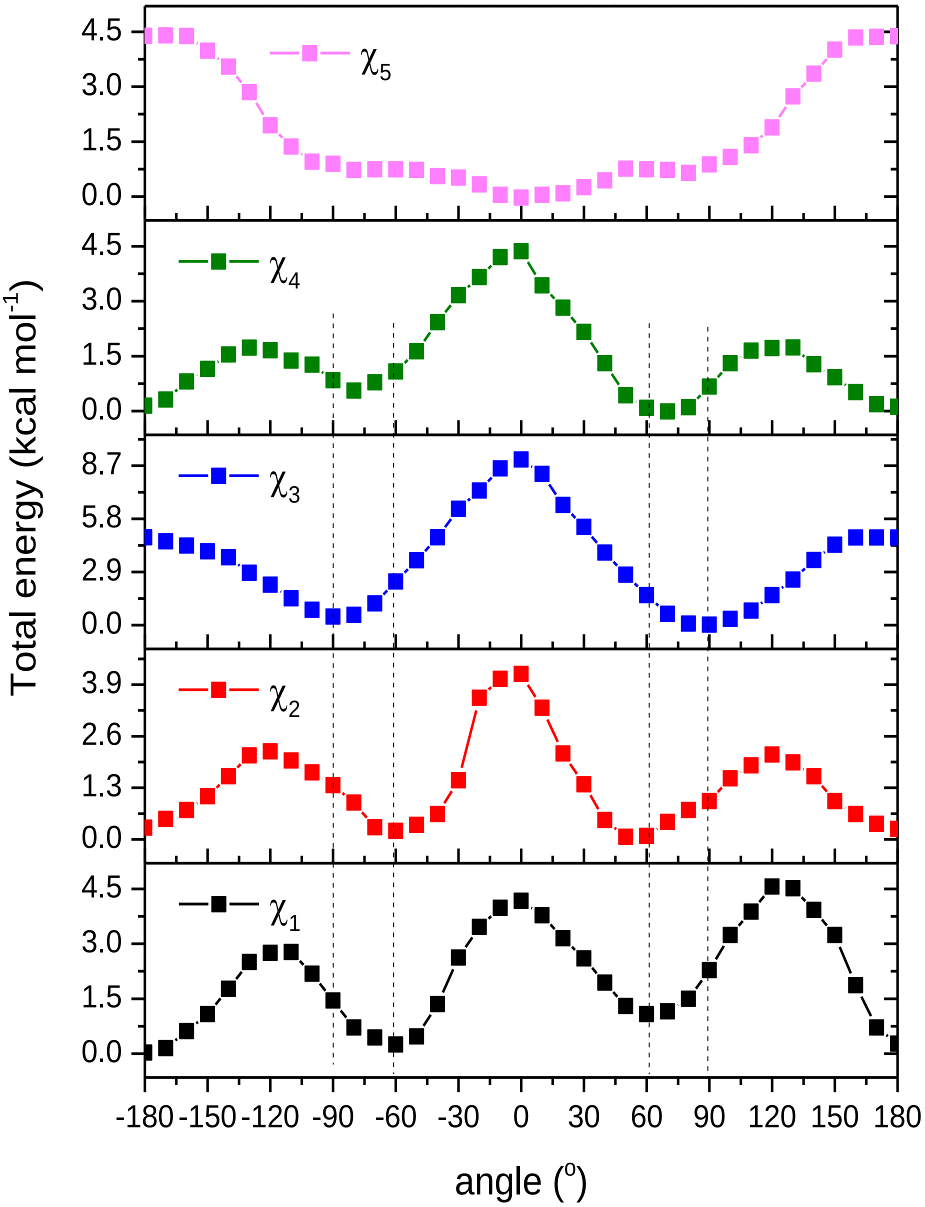


Figure 5.1 ESI: Energy torsional profiles related to the five dihedral angles χ_1, χ_2, χ_3, χ_4 and χ_5 of the MTSL side chain obtained from the MD simulation. Vertical dotted bars indicate the relevant minima.

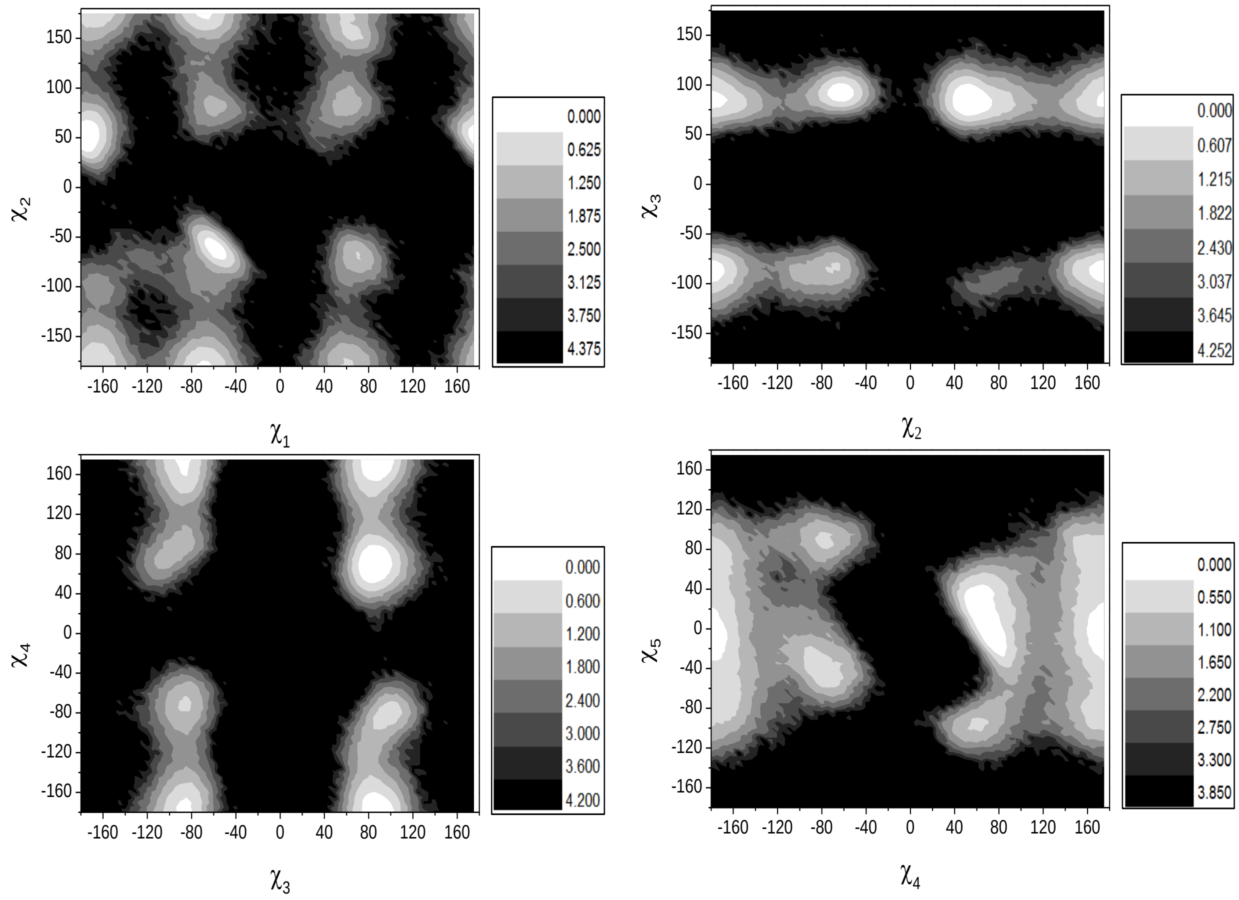


Figure 5.2 ESI: Plots of the Gibbs energy of the five χ_i-χ_{i+1} dihedral angles of the MTSL estimated from the bin populations using $G_i = -k_B T \ln (N_i/N_{max})$, where k_B is the Boltzmann's constant, T is the temperature, N_i is the population of the bin i and N_{max} is the population of the most populated bin.

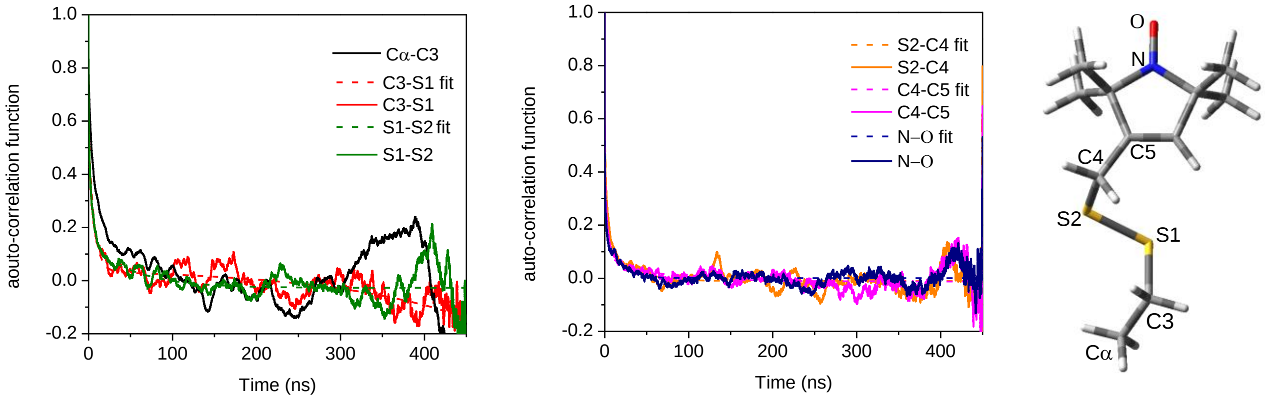


Figure 5.3 ESI: Auto-correlation functions of vectors corresponding to bonds between couple of atoms of the MTSL spin label (insert on the right). The fit (dashed lines) was performed using a bi-exponential function ($y = y_0 + A_1 e^{-x/\tau_1} + A_2 e^{-x/\tau_2}$) from which two correlation times were extracted and reported in Table 2.

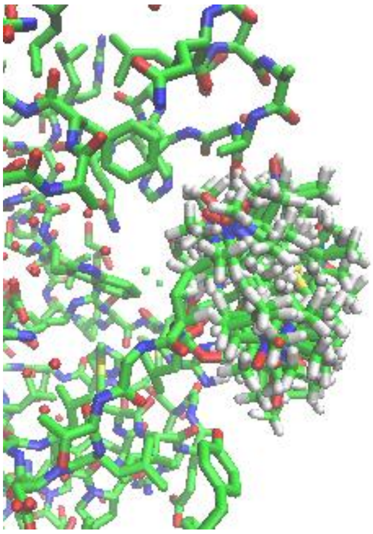


Figure 5.4 ESI: Rotamer distribution produced by the MMM software.

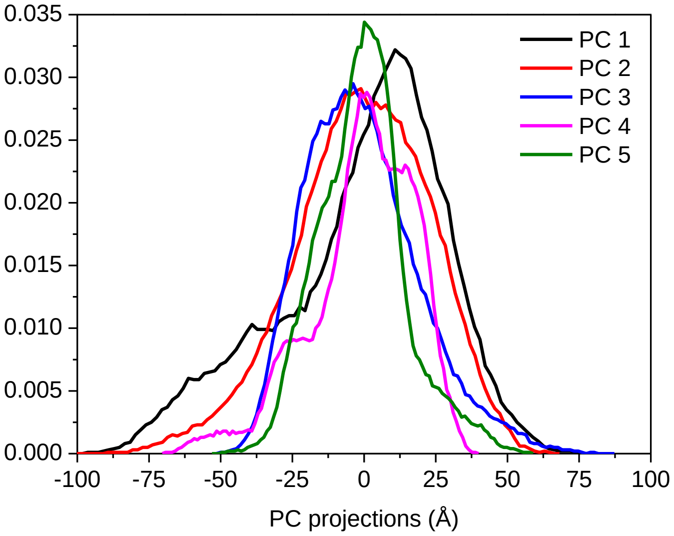


Figure 5.5 ESI: Histograms of the projections for the five most predominant modes of motions from the MD simulation. All protein residues were considered.

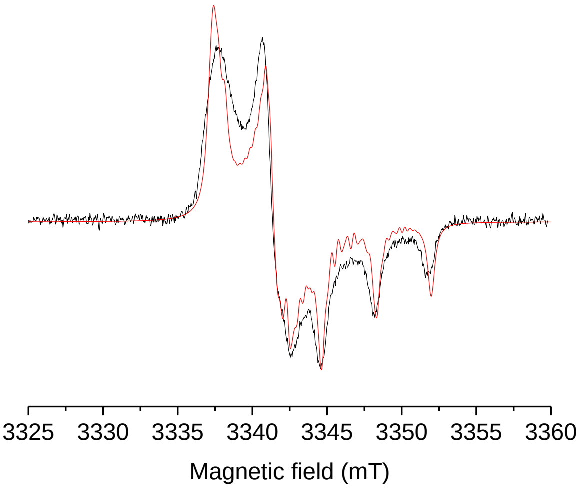
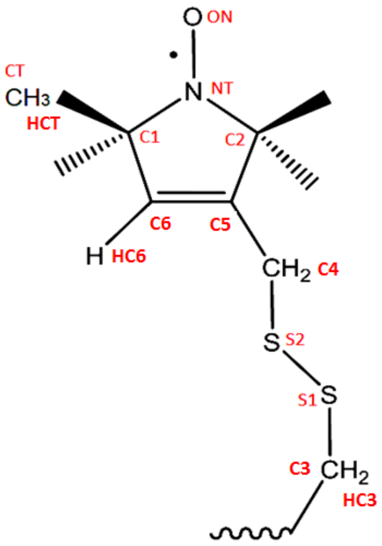


Figure 5.6 ESI: Comparison between experimental and simulated 94 GHz EPR spectra of spin labelled Aurora-A measured at 150 K.

Table 5.1 ESI: Partial charges determined to extend the AMBER force field to the MTSL side chain.



Atom type	Partial charge
C3	-0.178
C3H	0.154
S1	-0.226
S2	-0.124
C4	0.154
C5	-0.063
C6	-0.275
C6H	0.188
C1	0.355
C2	0.362
CT	-0.237
HCT	0.0655
NT	0.137
ON	-0.435

Table 5.2 ESI: Values of τ_1 and τ_2 determined from the bi-exponential fitting of the auto-correlation functions showed in Figure 5.3 ESI for each vector representing bonds of the MTSL (structure shown in Figure 5.3 ESI). It was not possible to perform the fit of the auto-correlation function of Cα-C3 due to its trend and values are not reported.

Vector	**τ_1(ns)**	**τ_2(ns)**
Cα-C3	/	/
C3–S1	/	6.5
S1-S2	53	3.7
S2-C4	27	2.6
C4-C5	58	2.7
N-O	15	1.0

6. Paper 3: Characterization of conformational states and dynamics of phosphorylated activation loops of the Aurora-A kinase

Maria Grazia Concilio,[1*] Alistair J. Fielding,[1*] Richard Bayliss,[2] and Selena G. Burgess[2]

Abstract

The catalytic activity of most protein kinases is regulated through phosphorylation of a conserved region, known as the activation loop. Several studies using X-ray crystallography and mass spectrometry have been performed previously to characterize activity and regulation of Aurora-A kinase and a correlation between conformation and phosphorylation states has been found. However, a limited number of studies have been performed in solution. This work reports a combined study employing both MD simulations and CW EPR spectroscopy to characterize conformational states and dynamics of phosphorylated activation loops of Aurora-A kinase in solution. MD revealed that conformational states were in good agreement with the published X-ray crystal structures and a comparison with the un-phosphorylated activation loop showed that the latter is more flexible due to the lack of interactions between phosphorylated threonine within the activation loop and surrounding residues. This result was consistent with experimental CW EPR data that revealed reduction of mobility of the MTSL-spin labelled phosphorylated activation loop. However, MD revealed that in both the cases the most flexible region was that localized between residues 282 and 294, which include residues 287 and 288 that are crucial for the catalytic activity of Aurora-A kinase.

Corresponding authors:

*Maria Grazia Concilio
mariagrazia.concilio@postgrad.manchester.ac.uk

*Alistair J. Fielding
alistair.fielding@manchester.ac.uk

6.1 Introduction

Aurora-A is a mitotic serine/threonine protein kinase that controls many cellular pathways and its activity is tightly regulated by changes of conformation upon phosphorylation of two regulatory phosphorylation sites on its activation loop, T287 and T288 within the catalytic domain of the protein. [1-3] The transfer of phosphate from ATP to residues T287 and T288, catalysed by Aurora-A itself, is required for activation of the kinase. In turn, Aurora-A activity is crucial for correct mitotic progression. Abnormal Aurora-A activity is implicated in cancer.

The activation loop of Aurora-A kinase adopts specific conformational states due to the presence of phospho-threonine residues at positions 287 and 288 that interact with nearby residues leading to specific arrangements that can prevent substrate binding and a consequent reduction in the catalytic efficiency of Aurora-A. [2] Therefore, characterisation of the dynamics and conformational states of phosphorylated kinase activation loops in solution is important to enhance the understanding of molecular processes related to diseases and to support the discovery of small molecule kinase inhibitors. So far, the conformational states of Aurora-A kinase have been mainly determined using X-ray crystallography [1, 3, 4] and no information about structure and dynamics of the protein are available in solution. In this work, MD simulations were performed to determine the predominant conformational states and modes of motion of the phosphorylated activation loop, the X-ray crystal structures with PDB ID 4BN1 [2] and 1OL5 [5] were used for the simulations. In order to further address the dynamics of the system, room temperatures CW EPR measurements were performed by introducing a common spin label, the methane-thiosulfonate spin label (MTSL), [6] into the activation loop of Aurora-A kinase at position 284. Opportune comparisons were made with the un-phosphorylated activation loop studied in previous work [7, 8] in which the spin label was introduced at position 288 and the phosphorylated threonine at 287 was mutated to alanine (Figure 6.1.1 A). Figure 6.1.1 shows the differences between the phosphorylated and un-phosphorylated structures investigated.

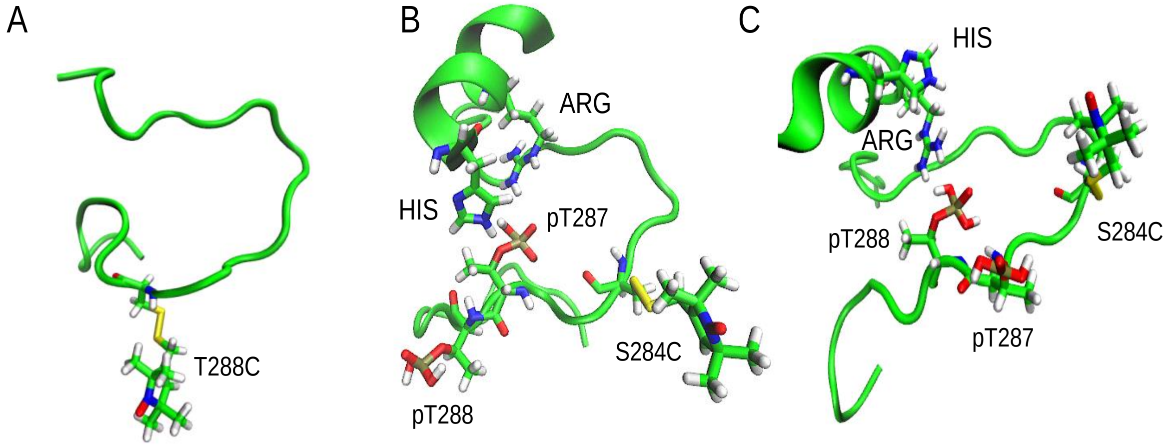


Figure 6.1.1: Energy minimized MTSL spin-labelled activation loops (residues 274 to 299) of Aurora-A using different X-ray crystal structures. (A) Structure 4CEG showing the un-phosphorylated activation loop with the MTSL attached at position 288, (B) structure 4BN1 showing a phosphorylated activation loop with the MTSL attached at 284, (C) structure 1OL5 showing a phosphorylated activation loop with the MTSL attached at 284. The labels ARG and HIS labels indicate arginine and histidine residues, respectively.

Figure 6.1.1 B and C shows MTSL attached at position 284 and the two phosphorylated threonine residues at positions 287 and 288, in structure 4BN1 the phosphate on residue 287 is rotated inward towards the activation loop and interacts with histidine 176 and arginine 180, while the phosphate on residue 288 is exposed to the solvent. In contrast, structure 1OL5, shows the phosphate group on residue 287 pointing outwards while, the phosphate on residue 288 points inwards. Structural differences in the starting X-ray structure might corresponds to differences in structure and dynamics in solution, the lack of good NMR data encouraged the study of this system through MD simulations and EPR spectroscopy. The primary use of MD is to elucidate the dynamics of the MTSL spin-labelled Aurora-A, but also to determine configurations the MTSL spin labels for the interpretation of the experimental EPR data.

6.2 Material and methods

Spin-labelled Aurora-A 122-403 S284C, C290A, C393A and Aurora-A 122-403 T287A, T288C, C290A, C393A were produced and the CW EPR measurements recorded as stated in earlier work [8-10]. The latest AMBER force field (ff14SB) [11] was extended in order to carry out MD simulations on MTSL spin labelled Aurora-A kinase using the AMBER 14 software [12], following the procedure described in previous work [8]. Phosphorylation of Aurora-A 122-403 S284C, C290A, C393A on Thr288 was confirmed by Western blotting using an anti-phosphoT288 Aurora-A antibody. Details of the minimization and equilibration steps can be found in the reference [8]. The production step was performed in the NVE ensemble in three stages up to 370 ns, collecting time windows equal to 270 ns, 320 ns and 370 ns. The ΔT between the initial and the final frame was equal to 10 K. A clustering analysis was performed on 5 principal components (PC) in each time window using the K-means algorithm [13, 14]. All EPR spectra were simulated using the open source *Spinach* software library [15], version 1.7. The CW EPR spectra were simulated using the *gridfree* context employing the Fokker-Planck formalism [16, 17].

6.3 Results and discussion

6.3.1 Characterization of the conformational states of the activation loop of Aurora-A kinase

The predominant conformational states adopted by the phosphorylated activation loop of Aurora-A kinase were determined by performing a clustering analysis in different regions of the full MD trajectory, specifically between 0 ns - 270 ns, 270 ns - 320 ns, and 320 ns – 370 ns. Figure 6.3.1.1 shows the representative structures obtained in each region of the MD simulation with the lowest RMSD to the centroid in the ten most populated clusters. Conformational states obtained for the phosphorylated activation loops (Figure 6.3.1.1 A and B) were then compared with conformations obtained from the un-phosphorylated activation loop studied in previous work [8] (Figure 6.3.1.1 C).

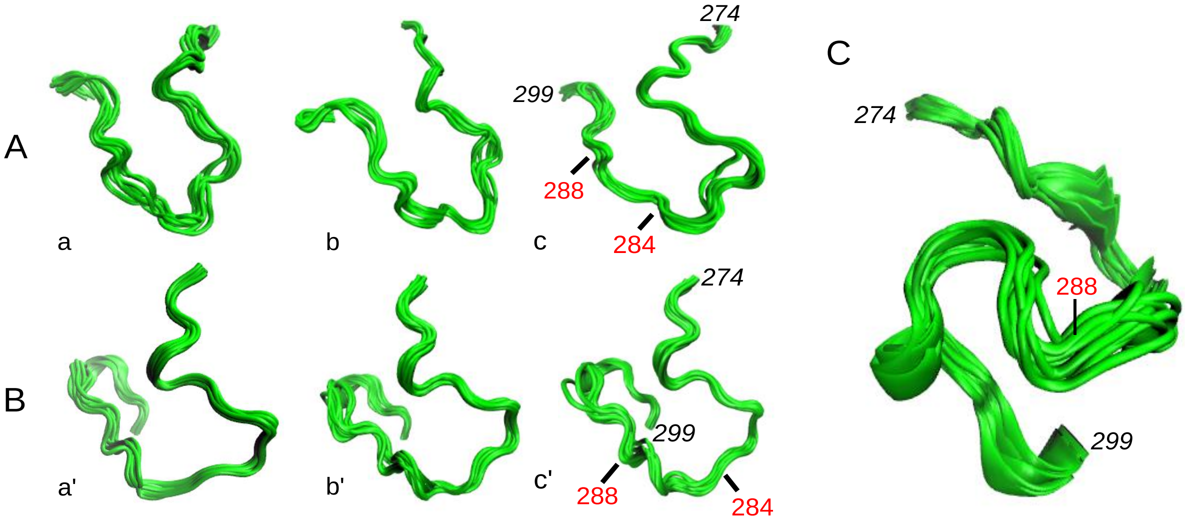


Figure 6.3.1.1: Comparison between conformational states of the Aurora-A kinase activation loop (from residues 274 to 299) at different time windows, (a-a’) 270 ns, (b-b’) 320 ns and (c-c’) 370 ns. (A) Representative structures obtained using the X-ray structure, 1OL5. (B) Representative minima structures obtained using the X-ray structure,

4BN1. The position of residues 284 and 288 is indicated in red. (C) Representative structures obtained using the X-ray structure, 4CEG.

The clustering analysis revealed that both phosphorylated activation loops (using PDB ID 1OL5 and PDB ID 4BN1 for simulation) adopted similar conformations in the three time windows as shown in Figure 6.3.1.1. This indicated less conformational movement in comparison with the un-phosphorylated activation loop which showed a wider distribution of conformational states between 0 ns and 420 ns [8] and similar conformational states after 420 ns (Figure 6.3.1.1 C). In this work, data over 420 ns were not reported since results were comparable to those observed in the time windows between 0 and 370 ns. Phosphorylated and un-phosphorylated activation loops adopted different conformations arising from specific inter-and intra-molecular interactions between residues within the activation loop and other residues of the protein (Figure 6.3.1.2).

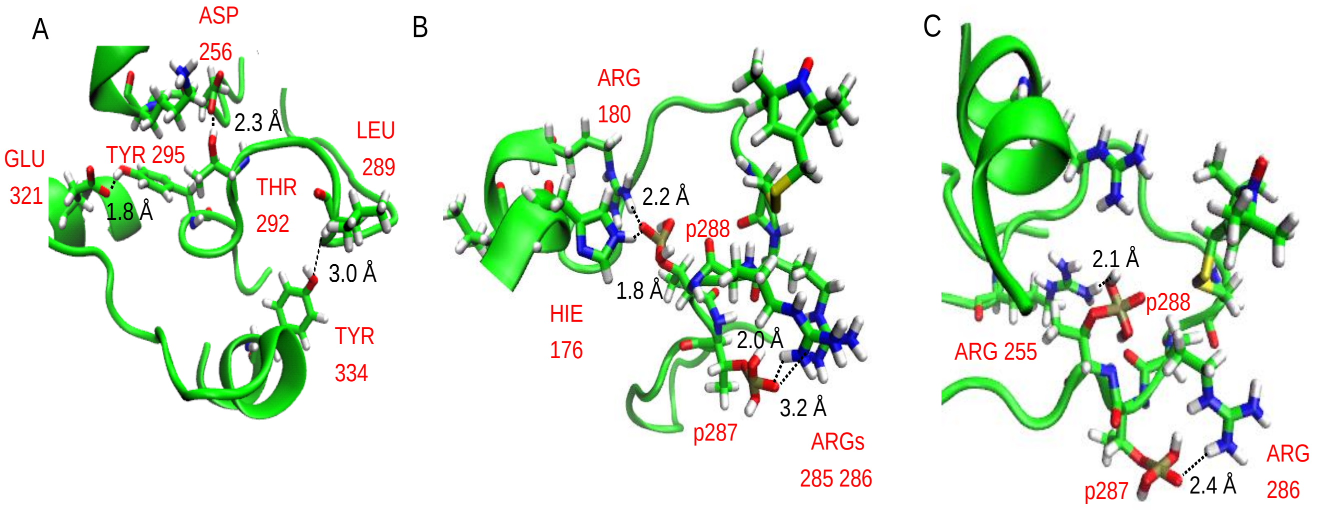


Figure 6.3.1.2: Interactions between the activation loop and other residues of the protein obtained from the clustering analysis between 320 and 370 ns using the X-ray crystal structures, (A) PDB ID 4CEG, (B) PDB ID 4BN1 and (C) PDB ID 1OL5 for simulations. The MTSL is attached at position 284 in Figures B and C.

Phosphorylated activation loops adopted a similar shape which is different to that observed in 4CEG (studied in previous work) due to the presence of inter- and intra-molecular interactions between the phosphorylated threonine residues at positions 287 and 288 and histidine and arginine residues of the protein. Figure 6.3.1.2 A shows the interactions between leucine 289 with the tyrosine 334, the threonine 292 with the ascorbic acid 256 and tyrosine 295 with the glutamate 321. In Figure 6.3.1.2 B depicting the activation loop of 4BN1, it is possible to observe interactions between the phosphate group at position 287 and both histidine 176 and arginine 180 in the C-lobe of Aurora-A, and between the phosphate at 288 and both arginine residues at 286 and 285 within the activation loop in the X-ray crystal structure 4BN1. Similarly, interactions between the phosphate at 287 and the arginine at position 286, and between the phosphate at 288 and the arginine at position 255 were observed in 1OL5 (Figure 6.3.1.2 C). Figures 6.3.1.2 B and C also show configurations of the MTSL spin label attached at residue 284 fully exposed to the solvent. These interactions were comparable with those observed in the X-ray crystal structure 4BN1 (Figure 6.3.1.3) that showed the phosphorylated threonine at 287 oriented inward towards the activation loop and the phosphorylated threonine at 288 oriented outward towards the solvent, and *vice versa* in the crystal structure 1OL5. Figure 6.3.1.3 shows a comparison between the conformational states of the activation loop of Aurora-A kinase determined in the last time window (320 ns - 370 ns) of the MD simulation and the initial X-ray crystal structures 4BN1 and 1OL5 obtained after the minimization step.

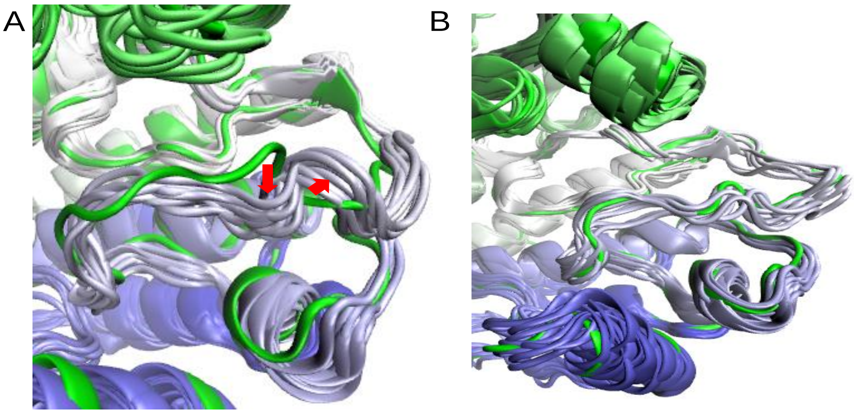


Figure 6.3.1.3: Comparison between the energy minimized X-ray structure of Aurora-A (green) and the 10 representative structures obtained from the MD simulation (white). (A) Structures obtained from the X-ray crystal structure 4NB1. The red arrows indicate the differences found between the initial X-ray crystal structure and the configurations obtained from the MD simulation. (B) Structures obtained from the X-ray crystal structure 1OL5.

MD revealed an outward movement of residues 284 to 287 toward the solvent for structure 4BN1 (as indicated by red arrows), while no significant changes were observed for structure 1OL5. The RMSD between the initial X-ray crystal structure and the configurations obtained from the MD simulation were equal to 1.6 Å (4BN1) and 1.2 Å (1OL5), respectively. The convergence over the two full trajectories is shown in Figure 6.1 in ESI. The activation loop in the structure 1OL5 appeared to be more elongated and exposed to the solvent for residues 280 to 286, as compared with the 4BN1 structure and the un-phosphorylated 4CEG structure investigated in previous work [8].

6.3.2 Characterization of the main mode of Aurora-A kinase activation loop motion

A PC analysis was used to describe the flexibility of the phosphorylated activation loop. Figure 6.3.2.1 shows molecular structures and corresponding plots of the root-mean square fluctuations (RMSFs) obtained from each PC in the final region of the MD simulation (320 ns to 370 ns) in order to determine the predominant modes of motions of the activation loop in the structures 4BN1 and 1OL5.

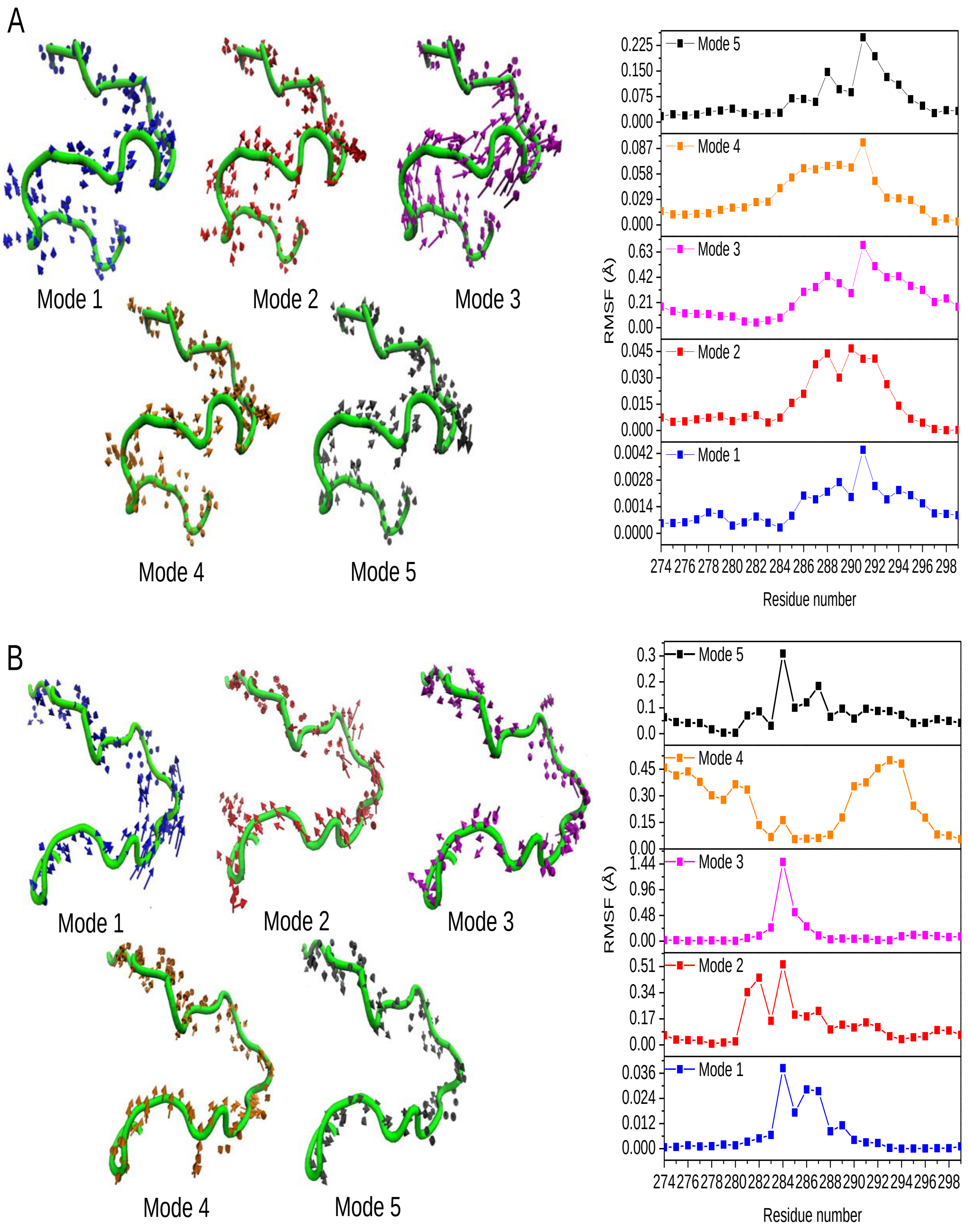


Figure 6.3.2.1: Molecular structures of the activation loop of Aurora-A kinase obtained from the projected frames of the first, second, third, fourth and fifth PC mode, respectively, and corresponding plots of the RMSFs for each active mode, going from the smaller fluctuations (mode 1) to the wider fluctuations (mode 5). Arrows represent the direction of the motions of the activation loop and the length represents the magnitude of the oscillations. (A) Representative structures obtained using the 4BN1 crystal structure. (B) Representative structures obtained using the 1OL5 crystal structure.

The five PCs corresponded to global protein bending motions in both 1OL5 and 4BN1. One difference between the two phosphorylated activation loops was the fluctuation of all residues of 4BN1 in the same direction, while residues fluctuated in different directions in 1OL5 (see direction of arrows in Figure 6.3.2.1). Fluctuations observed in the structure 4BN1 were about one order of magnitude smaller than those observed in 1OL5 in the modes from 1 to 4 (Figure 6.3.2.1). This can be explained considering that in structure 4BN1 more electrostatic interactions between the phosphate and the amino groups were observed than those observed in structure 1OL5 (Figure 6.3.1.2 B and C). Both the phosphorylated activation loops showed fluctuations smaller than those observed in the un-phosphorylated activation loop (structure 4CEG) studied in previous work [8], probably due to stronger electrostatic interactions between phosphorylated threonine residues at positions 287 and 288 and the arginine residues within the activation loop. In the plots of the RMSFs for 4BN1, larger fluctuations (indicated by sharp peaks) were observed in the region between residues 284 and 298. The most flexible residue corresponded to the glycine at 291, while much smaller fluctuations were observed for the other residues of the activation loop. In the plots of the RMSFs for 1OL5, smaller distributions of larger fluctuations were observed and the shape depended on the modes of motion. Similar distributions were observed in modes 1, 2 and 5 between residues 280 and 290, with the largest fluctuation centred at residue 284 that is more mobile due to the high flexibility of the spin label. A similar effect was observed in previous work for the structure 4CEG [8]. In mode 3, the most flexible residue was 284, while in mode 4 larger fluctuations were observed in the regions between residues 274 and 282, and 289 and 296, and smaller fluctuations were observed at residue 284. These motions arose from the presence of bending motions that pulled residue 284 outward from the activation loop and stretched the other residues in mode 3, the opposite was observed in mode 4. In order to determine timescales of the motions within the activation loop, the iRED approach [18] was employed to calculate correlation times (τ_m) of the loop motions and the order parameters (S^2) of each residue along the full MD trajectory. Through this approach, implemented in the AMBER software, the values of τ_m, along the reorientational eigenmodes, can be estimated from the time-correlation functions $C_{m,l} = \langle a_{m,l} * (\tau + t) a_{m,l}(\tau) \rangle_\tau$ calculated for each mode, where $a_{m,l}$ are the time-dependent amplitudes obtained by projecting MD snapshots on the eigenmodes calculated from corresponding iRED vectors that were set equal to 26, corresponding to the number of N-H groups belonging to residues of the activation loop (the number of defined iRED vectors have to match the number of eigenmodes calculated). The corresponding τ_m values of these correlation functions can be determined by exponential fitting using the equation $\tau_m \cong \frac{1}{[C_m(0)-C_m(t\to T)]} \int_0^T [C_m(0) - C_m(t \to T)]\, dt,$ where $C_m(t \to T)$ indicates the plateau value of $C_m(t)$. All the data obtained from the iRED analysis are shown in Figures 6.2 and 6.3 in ESI, values of S^2 observed for the phosphorylated activation loops were, for the most of the residues, higher than those observed for the un-phosphorylated activation loop, confirming higher rigidity. In the case of the structure 4BN1, the most flexible region (with the smallest S^2) was that located between residues 284-292, that corresponded to the region in which the RMSFs were seen to be wider (Figure 6.3.2.1. A), and with highest intensity centered at 288 and 291. These residues had S^2 equal to 0.40 and 0.30, respectively. In the case of the structure 1OL5, the most flexible residues corresponded to 284 and 289, with S^2 equal to 0.11 and 0.37, respectively, that occur in the region between 283 and 289 in which wider fluctuations were observed.

The extracted correlation times (Figure 6.3 in ESI) for the phosphorylated structures (1OL5 and 4BN1) varied between 0.1 ns and 8 ns, indicating a wide distribution of motions and structural fluctuations dominated by slower time-scale events, similarly observed in structure 4CEG. However, the mono-exponential decay went to zero slower in the phosphorylated structures, indicating slower motions.

6.3.3 Description of the 9.4 GHz EPR spectrum of MTSL spin-labelled Aurora-A kinase at residue 284

The room temperature CW EPR spectra of MTSL spin-labelled Aurora-A kinase at residues 284 and 288 were compared in order to determine differences in dynamics between the phosphorylated and un-phosphorylated activation loops, respectively. The spectra of spin-labelled Aurora-A 122-403 S284C, C290A, C393A (Figure 6.3.3.1 A) showed an increase in the resonances at 331 mT and 337 mT and a line broadening of 0.1 mT, indicating a lower mobility of the spin label. This was reflected in the simulated EPR spectra (Figure 6.3.3.1 B) as the rotational correlation time was increased from 11 ns to 17 ns. The lineshape resulted from a convolution of motions arising from configurations of the spin label whose motion is slowed down by interactions with residues of the protein and the other arising from configurations where the spin label is fully exposed to the solvent and completely free to rotate. Both the spectra (Figure 6.3.3.1 C and D) were simulated with values of g- and A-tensors equal to $g_{xx} = 2.0071,\ g_{yy} = 2.0069,\ g_{zz} = 2.0028;\ A_{xx} =\ 8$ G, $A_{yy} =\ 7.5$ G and $A_{zz} =\ 37$ G.

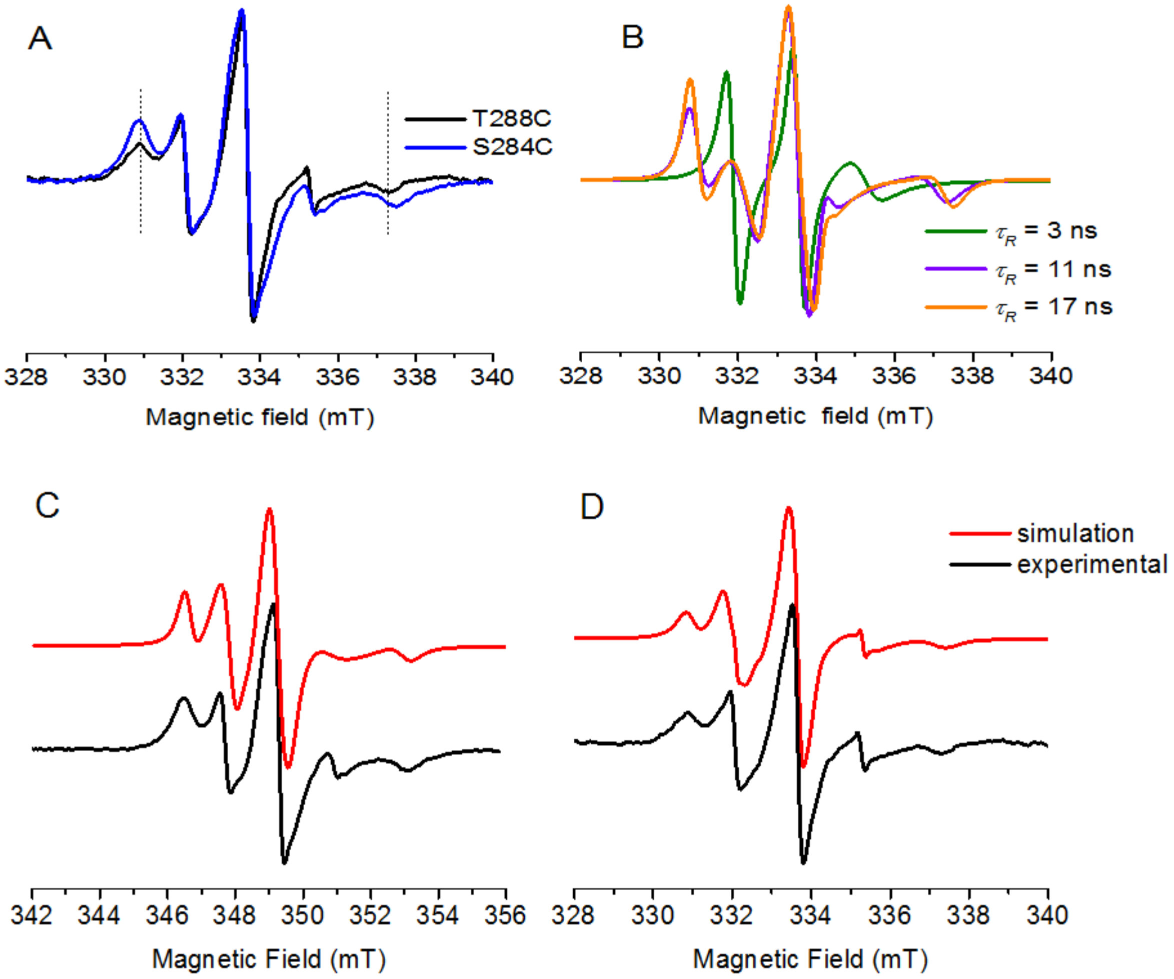


Figure 6.3.3.1: Comparison between experimental and simulated 9.4 GHz EPR spectra of MTSL spin-labelled Aurora-A at residues 288 and 284. (A) Experimental EPR spectra of MTSL spin-labelled Aurora-A at residues 288 and 284, measured at 298 K. (B) Comparison simulated EPR spectra with different rotational correlation times, τ_R and constant *g*- and *A*-tensors. (C) Comparison between simulated and experimental 9 GHz EPR spectra for Aurora-A spin-labelled at 284. (D) Comparison between simulated and experimental 9 GHz EPR spectra for Aurora-A spin-labelled at 288.

The EPR spectra of Aurora-A kinase spin-labelled at 284 and 288 showed a lineshape that is quite common in EPR and is described in previous work as a superposition of two components [8, 19-22]. The EPR spectra of MTSL spin-labelled Aurora-A kinase at residue 284 (Figure 6.3.3.2 C) was obtained by summing a 1:1 ratio of two spectra with τ_R= 3 ns and τ_R= 17 ns, respectively, while the EPR spectrum of Aurora-A kinase spin-labelled at residue 288 (Figure 6.3.3.1 D) was obtained by summing a 1:1 ratio two of spectra with τ_R= 3 ns and τ_R= 11 ns, respectively. The analysis of spectra suggested that in both the spectra, 50 % of the spin label exhibited slow dynamics, while the remaining 50 % exhibited fast dynamics.

In order to provide an interpretation of the experimental EPR data from the MD data, the analysis of the auto-correlation function, corresponding to $\langle D_{00}^{2}\,(\theta(0))D_{00}^{2}(\theta(t))\rangle$ with D_{00}^{2} = (3cos$^{2}\theta(t)$-1)/2 and $\theta(t)$ representing the variation in the time t of the direction of the vector, was performed. In the previous work [8], it was observed that the fit to the auto-correlation function of the vector related to the N-O group of the MTSL provided values of τ_1 and τ_2 closer to the correlation times required to simulate the room temperature EPR spectrum of the un-phosphorylated Aurora-A. In this work, fits to the auto-correlation functions related to the phosphorylated structures provided values of τ_1 = 8.1 ns and τ_2= 0.49 ns for 4BN1, and values of τ_1 = 4.2 ns and τ_2= 0.24 ns for 1OL5, respectively, (Figure 6.4 ESI). These values were found to be smaller than those used to fit the EPR spectrum of the phosphorylated Aurora-A. In the case of the structure 4BN1, the plot of the auto-correlation function was very noisy and the values extracted might not be accurate. In contrast, in the case of the structure 1OL5, τ_1 = 4.2 ns was quite close to the correlation time representing the fast motion in the system. Considering previous results [8], discussed in Chapter 5, and those obtained from this work, it is possible to conclude that MD can partially provide information for the interpretation of the EPR spectrum of the Aurora-A kinase. The 1:1 ratio can be explained looking at the plot of the transitions in Figure 6.5 ESI, for the structure 4BN1 (likewise 1OL5, data not shown for sake of clarity), where the transitions of χ_1, χ_2 and χ_4 were faster in the range 190 ns 335 ns, and represent the ~40% of the full MD trajectory. As stated above, the MD simulation was used to determine order parameters related to the flexibility of the protein backbone and determine configurations of MTSL spin label in the space. In both the 1OL5 and 4BN1 crystal structures, MD revealed configurations of the spin label exposed to the solvent (Figure 6.3.3.2 A) and interacting with residues of the protein (Figure 6.3.3.2 B and C).

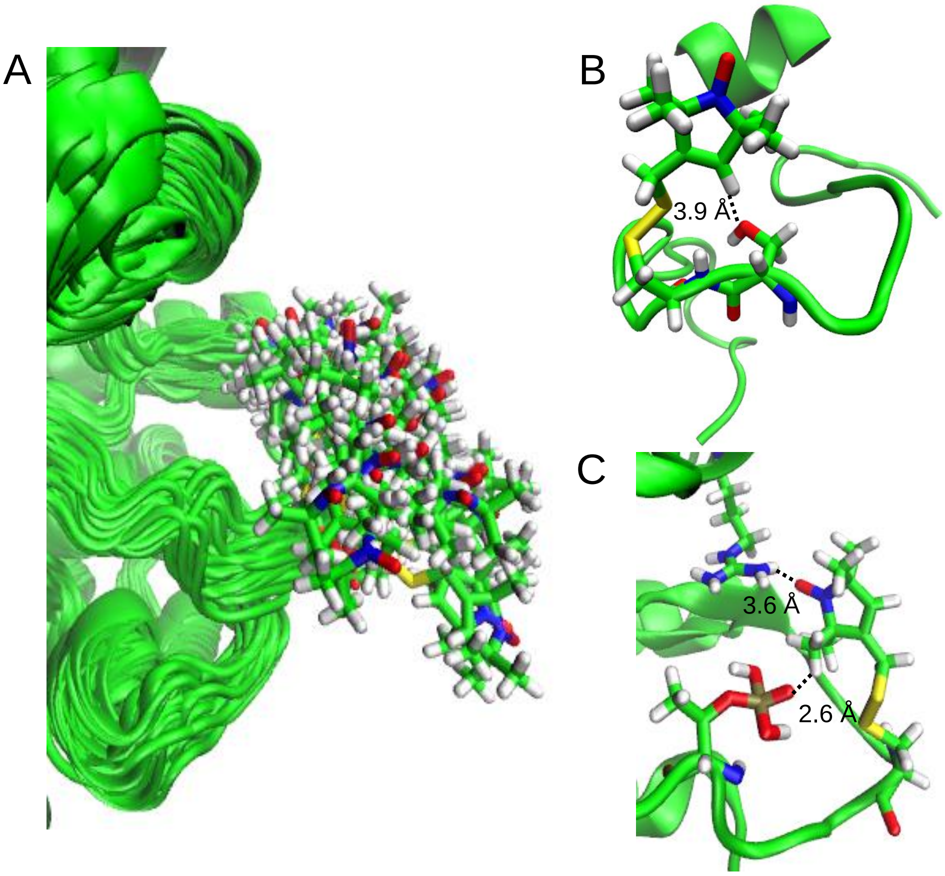


Figure 6.3.3.2: (A) Distribution in the space of the MTSL spin label obtained from the MD simulation, (B) Configuration of MTSL interacting with serine 284 in the crystal structure 1OL5. (C) Configuration of MTSL interacting with phosphorylated threonine at position 288 and the arginine at position 220 in the crystal structure 1OL5. Similar configurations were observed for the 4BN1 crystal structure.

The experimental data were interpreted considering the fast motions arising from configurations of the spin label exposed to the solvent, while slow motions arising from the configurations of the spin label in which the motion is slowed down by interactions within the protein domain. Although the iRED approach can be applied only to N-H bonds within the protein backbone, it is also possible to conclude that results obtained from EPR were consistent with those obtained from MD that revealed the phosphorylated activation loop is less flexible than un-phosphorylated activation loop. This confirmed previous results from X-ray crystallography that showed the rigidity of many phosphorylated activation loops in kinase proteins. [2]

6.4 Conclusions and future work

In this work, the main differences of conformation and dynamics between the phosphorylated and un-phosphorylated activation loops of Aurora-A kinase revealed by MD simulations were reported. Phosphorylated activation loops adopted similar conformations due to the presence of interactions between phosphorylated threonine residues at positions 287 and 288 and the other residues within and surrounding the activation loop. These conformations were seen to be different from those observed in the un-phosphorylated activation loop studied in previous work and involved phosphorylated threonine at 287 and 288 and histidine and arginine residues that arranged the activation loop in more constrained conformations reducing its mobility. Conformational states revealed from MD were consistent with those observed in the crystal structure 1OL5, while an outward movement of residues 284 to 287 toward the water exposed regions was observed in the crystal structure 4BN1. The amplitude of the RMSFs observed in both the phosphorylated structures 4BN1 and 1OL5 were of about one order of magnitude smaller than those observed in the un-phosphorylated activation loop. Values of the order parameters S^2 in the phosphorylated structures oscillated between 0.10 - 0.30 and 0.70 in the region between residues 282 and 292, while were seen to be equal to 0.80 - 0.90 for the other residues. Phosphorylated activation loops were seen to be less flexible than un-phosphorylated activation loops, but in both the cases the most flexible was that between residues 282 - 290 that includes residues 287 and 288 that are crucial for Aurora-A kinase activity. Experimental CW EPR data revealed an increase of the rotational correlation time when the MTSL spin label is at 284 and residues 287 and 288 are phosphorylated, in accordance with MD data that revealed higher rigidity of the phosphorylated activation loops. This work can be used to support the research for the development of kinase inhibitors and binding partners, which may alter the conformation and dynamics of the activation loop.

Acknowledgments

This work was supported by a studentship from Bruker Ltd. to M.G.C. and by a Cancer Research UK grant (C24461/A12772) to R.B. The authors would like to acknowledge the use of the EPSRC UK National Service for Computational Chemistry Software (NSCCS) and its staff (Dr. Alexandra Simperler and Dr. Helen Tsui for some technical advice) at Imperial College London in carrying out this work and the EPSRC National EPR Facility at the University of Manchester. This work is a result of many discussions within the EPR and computational communities. M. G. C. acknowledges

Dr. A. Baldansuren and Prof. D. Collison for useful discussions about the EPR lineshape analysis; Dr. A. Simperler for useful feedback.

References

1. Endicott JA, Noble ME, Johnson LN. The structural basis for control of eukaryotic protein kinases. Annual review of biochemistry. 2012;81:587-613. Epub 2012/04/10.
2. Bayliss R, Fry A, Haq T, Yeoh S. On the molecular mechanisms of mitotic kinase activation. Open biology. 2012;2(11):120136. Epub 2012/12/12.
3. Nolen B, Taylor S, Ghosh G. Regulation of protein kinases; controlling activity through activation segment conformation. Molecular cell. 2004;15(5):661-75. Epub 2004/09/08.
4. Wang XJ, Chu HY, Lv MJ, Zhang ZH, Qiu SW, Liu HY, et al. Structure of the intact ATM/Tel1 kinase. Nature communications. 2016;7.
5. Case DA, Babin, V., Berryman, J.T., Betz, R.M., Cai, Q., Cerutti, D.S., Cheatham, III, T.E., Darden, T.A., Duke, R.E., Gohlke, H., Goetz, A.W., Gusarov, S., Homeyer, N., Janowski, P., Kaus, J., Kolossváry, I., Kovalenko, A., Lee, T.S., LeGrand, S., Luchko, T., Luo, R., Madej, B., Merz, K.M., Paesani, F., Roe, D.R., Roitberg, A., Sagui, C., Salomon-Ferrer, R., Seabra, G., Simmerling, C.L., Smith, W., Swails, J., Walker, R.C., Wang, J., Wolf, R.M., Wu X. and Kollman P.A. . AMBER 14, University of California, San Francisco. 2014.
6. Zhao Z, Wu H, Wang L, Liu Y, Knapp S, Liu QS, et al. Exploration of Type II Binding Mode: A Privileged Approach for Kinase Inhibitor Focused Drug Discovery? ACS chemical biology. 2014;9(6):1230-41.
7. Dodson CA, Kosmopoulou M, Richards MW, Atrash B, Bavetsias V, Blagg J, et al. Crystal structure of an Aurora-A mutant that mimics Aurora-B bound to MLN8054: insights into selectivity and drug design (vol 427, pg 19, 2009). Biochem J. 2010;427:551-.
8. Solanki S, Innocenti P, Mas-Droux C, Boxall K, Barillari C, van Montfort RLM, et al. Benzimidazole Inhibitors Induce a DFG-Out Conformation of Never in Mitosis Gene A-Related Kinase 2 (Nek2) without Binding to the Back Pocket and Reveal a Nonlinear Structure-Activity Relationship. J Med Chem. 2011;54(6):1626-39.
9. Huang ZF, Tan L, Wang HY, Liu Y, Blais S, Deng JJ, et al. DFG-out Mode of Inhibition by an Irreversible Type-1 Inhibitor Capable of Overcoming Gate-Keeper Mutations in FGF Receptors. ACS chemical biology. 2015;10(1):299-309.
10. Shults MD, Imperiali B. Versatile fluorescence probes of protein kinase activity. Journal of the American Chemical Society. 2003;125(47):14248-9.
11. Simard JR, Getlik M, Grutter C, Pawar V, Wulfert S, Rabiller M, et al. Development of a Fluorescent-Tagged Kinase Assay System for the Detection and Characterization of Allosteric Kinase Inhibitors. Journal of the American Chemical Society. 2009;131(37):13286-96.
12. Zondlo SC, Gao F, Zondlo NJ. Design of an Encodable Tyrosine Kinase-Inducible Domain: Detection of Tyrosine Kinase Activity by Terbium Luminescence. Journal of the American Chemical Society. 2010;132(16):5619-+.
13. Carmena M, Earnshaw WC. The cellular geography of aurora kinases. Nat Rev Mol Cell Bio. 2003;4(11):842-54.
14. Barr AR, Gergely F. Aurora-A: the maker and breaker of spindle poles. Journal of cell science. 2007;120(17):2987-96.
15. Fernandez A, Ferrando-Soria J, Pineda EM, Tuna F, Vitorica-Yrezabal IJ, Knappke C, et al. Making hybrid [n]-rotaxanes as supramolecular arrays of molecular electron spin qubits. Nature communications. 2016;7:10240. Epub 2016/01/09.

16. Asteriti IA, Rensen WM, Lindon C, Lavia P, Guarguaglini G. The Aurora-A/TPX2 complex: A novel oncogenic holoenzyme? Bba-Rev Cancer. 2010;1806(2):230-9.
17. Bayliss R, Sardon T, Vernos I, Conti E. Structural basis of Aurora-A activation by TPX2 at the mitotic spindle. Molecular cell. 2003;12(4):851-62. Epub 2003/10/29.
18. Milov AD, Ponomarev AB, Tsvetkov YD. Electron Electron Double-Resonance in Electron-Spin Echo - Model Biradical Systems and the Sensitized Photolysis of Decalin. Chem Phys Lett. 1984;110(1):67-72.
19. Milov AD, Salikhov KM, Shirov MD. Application of Eldor in Electron-Spin Echo for Paramagnetic Center Space Distribution in Solids. Fiz Tverd Tela+. 1981;23(4):975-82.
20. Jeschke G. DEER Distance Measurements on Proteins. Annual Review of Physical Chemistry, Vol 63. 2012;63:419-46.
21. Borbat PP, Freed JH. Measuring distances by pulsed dipolar ESR spectroscopy: Spin-labeled histidine kinases. Method Enzymol. 2007;423:52-+.
22. Bhatnagar J, Sircar R, Borbat PP, Freed JH, Crane BR. Self-Association of the Histidine Kinase CheA as Studied by Pulsed Dipolar ESR Spectroscopy. Biophysical journal. 2012;102(9):2192-201.
23. Fu Z, Aronoff-Spencer E, Backer JM, Gerfen GJ. The structure of the inter-SH2 domain of class IA phosphoinositide 3-kinase determined by site-directed spin labeling EPR and homology modeling. P Natl Acad Sci USA. 2003;100(6):3275-80.
24. Rowan FC, Richards M, Bibby RA, Thompson A, Bayliss R, Blagg J. Insights into Aurora-A kinase activation using unnatural amino acids incorporated by chemical modification. ACS chemical biology. 2013;8(10):2184-91. Epub 2013/08/09.
25. Concilio MG, Fielding AJ, Bayliss R, Burgess SG. Density functional theory studies of MTSL nitroxide side chain conformations attached to an activation loop. Theoretical chemistry accounts. 2016;135(4):1-6.
26. G. R. Eaton SSE, D. P. Barr, R. T. Weber. ed.; Springer. ed2010.
27. Pannier M, Veit S, Godt A, Jeschke G, Spiess HW. Dead-time free measurement of dipole-dipole interactions between electron spins. Journal of magnetic resonance. 2000;142(2):331-40. Epub 2000/01/29.
28. Jeschke G, Chechik V, Ionita P, Godt A, Zimmermann H, Banham J, et al. DeerAnalysis2006 - a comprehensive software package for analyzing pulsed ELDOR data. Applied magnetic resonance. 2006;30(3-4):473-98.
29. Polyhach Y, Bordignon E, Jeschke G. Rotamer libraries of spin labelled cysteines for protein studies. Physical chemistry chemical physics : PCCP. 2011;13(6):2356-66. Epub 2010/12/01.
30. Concilio MG, Fielding AJ, Bayliss R, Burgess SG. Density functional theory studies of MTSL nitroxide side chain conformations attached to an activation loop. Theoretical chemistry accounts. 2016;135:97. Epub 2016/04/14.
31. De Groot JEH, J. V. Anzola, A. Motamedi, M. Yoon, Y. L. Wong, D. Jenkins, H. J. Lee, M. B. Martinez, R. L. Davis, T. C. Gahman, A. Desai, A. K. Shiau. Frontiers in oncology. 2015;5.
32. Manchado E, Guillamot M, Malumbres M. Killing cells by targeting mitosis. Cell Death Differ. 2012;19(3):369-77.
33. Bouloc N, Large JM, Kosmopoulou M, Sun CB, Faisal A, Matteucci M, et al. Structure-based design of imidazo[1,2-a] pyrazine derivatives as selective inhibitors of Aurora-A kinase in cells. Bioorg Med Chem Lett. 2010;20(20):5988-93.
34. Bavetsias V, Large JM, Sun CB, Bouloc N, Kosmopoulou M, Matteucci M, et al. Imidazo[4,5-b]pyridine Derivatives As Inhibitors of Aurora Kinases: Lead Optimization Studies toward the Identification of an Orally Bioavailable Preclinical Development Candidates. Journal of Medicinal Chemistry. 2010;53(14):5213-28.
35. Manfredi MG, Ecsedy JA, Meetze KA, Balani SK, Burenkova O, Chen W, et al. Antitumor activity of MLN8054, an orally active small-molecule inhibitor of Aurora A kinase. P Natl Acad Sci USA. 2007;104(10):4106-11.

36. Sloane DA, Trikic MZ, Chu MLH, Lamers MBAC, Mason CS, Mueller I, et al. Drug-Resistant Aurora A Mutants for Cellular Target Validation of the Small Molecule Kinase Inhibitors MLN8054 and MLN8237. ACS chemical biology. 2010;5(6):563-76.
37. Brockmann M, Poon E, Berry T, Carstensen A, Deubzer HE, Rycak L, et al. Small Molecule Inhibitors of Aurora-A Induce Proteasomal Degradation of N-Myc in Childhood Neuroblastoma. Cancer Cell. 2013;24(1):75-89.

Electronic supporting information (ESI)

Characterization of conformational states and dynamics of phosphorylated activation loops of the Aurora-A kinase

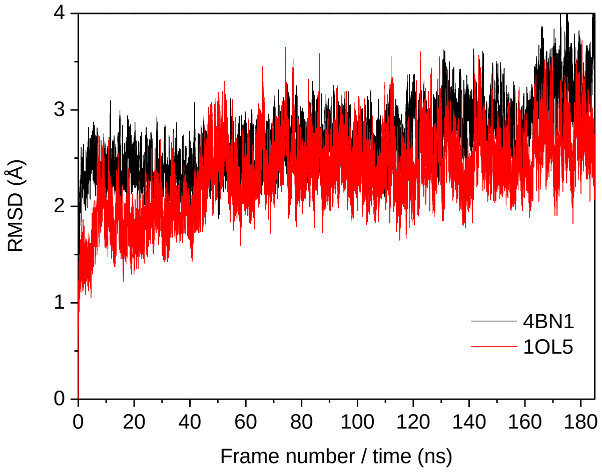


Figure 6.1 ESI: Plots of the root mean square deviation (RMSD) of the protein backbone considering the Cα, C, N, O atoms, the reference frame was the first frame after the equilibration step. The RMSD was seen reasonably converged to ~2.4 Å for 1OL5 and to ~2.8 Å for 4BN1.

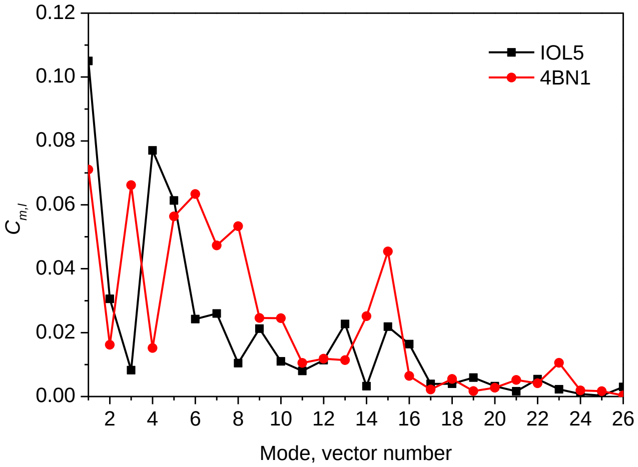


Figure 6.2 ESI: Time-correlation function, $C_{m,l}$ and corresponding correlation times for each mode/ iRED vector for structures 1OL5 and 4BN1.

The modes decay mono-exponentially to zero in both the structures 1OL5 and 4BN1, and corresponded to individual reorientational modes of motion and rotational tumbling that are shown in Figure 6.3 ESI.

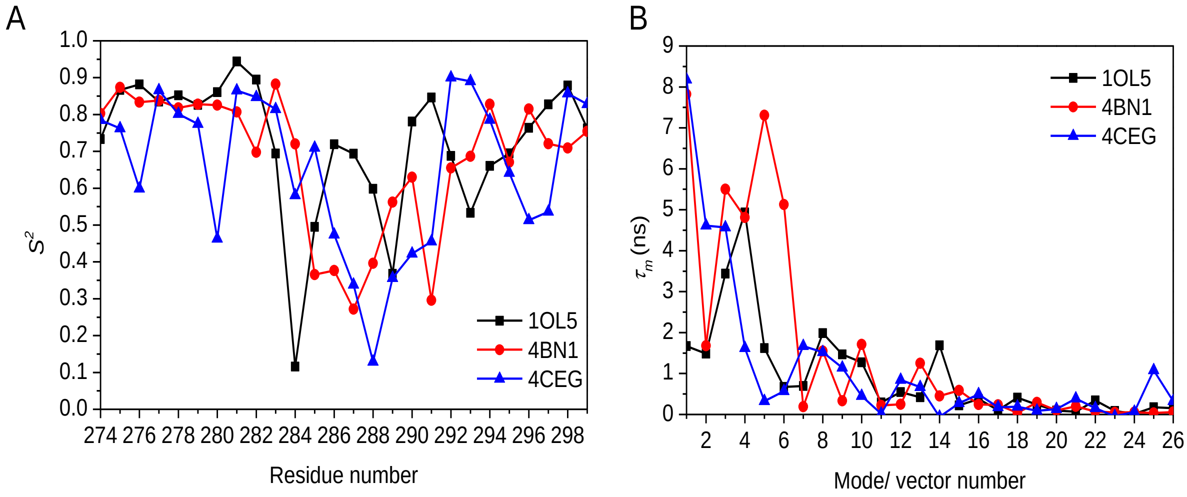


Figure 6.3 ESI: Results obtained from the iRED analysis performed along the full MD trajectory. (A) Comparison between rotational correlation times, τ_m for each mode/ iRED vector obtained for structures 1OL5, 4BN1 (phosphorylated) and 4CEG (un-phosphorylated). (B) Order parameters, S^2 for each residue of the activation loop for structures 1OL5, 4BN1 (phosphorylated) and 4CEG (un-phosphorylated).

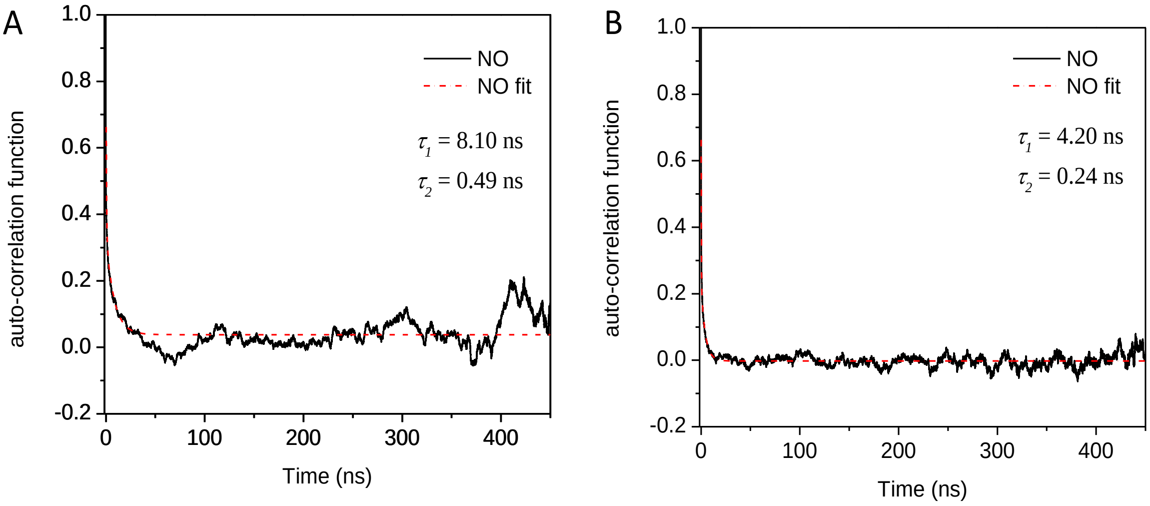


Figure 6.4 ESI: Auto-correlation function of vector corresponding to the NO bond of the MTSL spin label in the X-ray structure 4BN1 (A) and X-ray structure 1OL5 (B). The fit (dashed lines) was performed using a bi-exponential function ($y = y_0 + A_1 e^{-x/\tau_1} + A_2 e^{-x/\tau_2}$) from which two correlation times were extracted, they are reported in the insert.

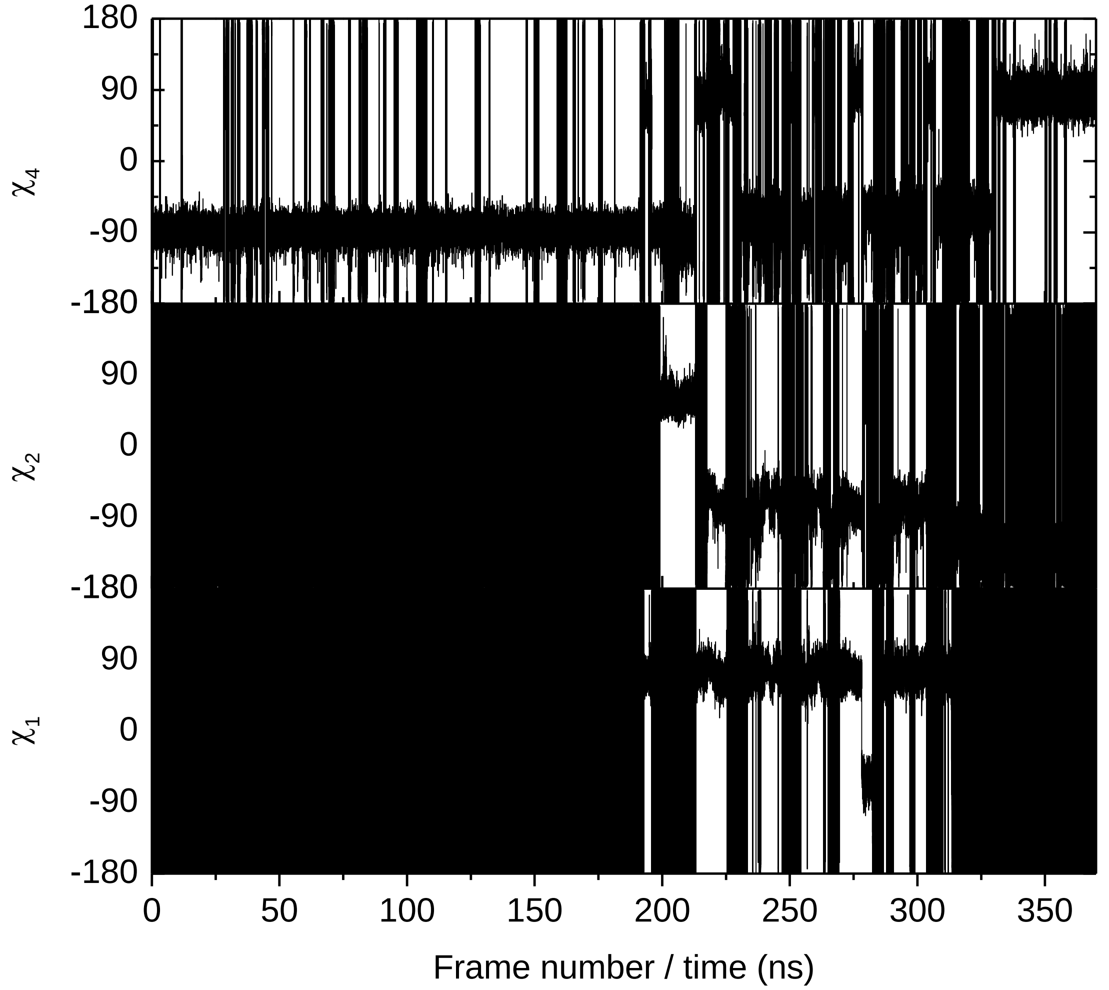


Figure 6.5 ESI: Plots of the transitions of the χ_1, χ_2 and χ_4 dihedral angles of the MTSL, from the X-ray crystal structure 4BN1.

7. Paper 4: Detection of Ligand-induced Conformational Changes in the Activation Loop of Aurora-A Kinase by PELDOR Spectroscopy

Selena G. Burgess,[a] Maria Grazia Concilio,[b] Richard Bayliss,*[a] and Alistair J. Fielding*[b]

[a]Astbury Centre for Structural and Molecular Biology, Faculty of Biological Sciences, University of Leeds, Leeds LS2 9JT, United Kingdom.
[b]The Photon Science Institute and School of Chemistry, EPSRC National EPR Facility and Service, University of Manchester, Manchester, M13 9PY, United Kingdom.

Abstract

Conformational movement of the kinase activation loop is thought to be crucial for regulation of its activity, but in many cases the position of the activation loop in solution is unknown. Protein kinases are an important class of therapeutic target and kinase inhibitors are classified by their effect on the activation loop. In this work, PELDOR spectroscopy and site-directed spin labelling were employed to monitor conformational changes through the insertion of MTSL on the dynamic activation loop and a stable site on the outer surface of the enzyme. The action of different ligands such as TPX2 and inhibitors could be discriminated as well as their ability to lock the activation loop in a fixed conformation. This study provides evidence for structural adaptations that could be used for drug design and a methodological approach that has potential to characterize inhibitors in development,

Corresponding authors:

*Prof. R. Bayliss

R.W.Bayliss@leeds.ac.uk

*Alistair J. Fielding
alistair.fielding@manchester.ac.uk

7.1 Introduction

Protein kinases are signalling enzymes that are regulated by molecular-switch mechanisms. (1, 2) Different protein kinases have very similar structures when active, but can adopt a variety of divergent conformations when inactive. One common mechanism of kinase activation is phosphorylation of the activation loop, a dynamic region that spans a consensus sequence from Asp–Phe–Gly (DFG) to Ala–Pro–Glu (APE). (3) Active kinases adopt a DFG-in conformation needed for the correct placement of a catalytic Asp residue and for the central region of the activation loop to form part of the binding site for protein and peptide substrates. Despite intensive studies of kinases by X-ray crystallography, (4, 5) there are few studies of conformational dynamics of the activation loops in solution. Protein kinases are frequently dysregulated in human disease and are a common target for the development of new therapeutics. Most protein kinase inhibitors (6) compete for the ATP binding site and are further classified as either Type I, which bind to the kinase in its active state, or Type II, which bind to an inactive DFG-out conformation of the kinase and occupy an additional hydrophobic pocket within the active site. However, validation of the effect of inhibitors requires determination of the kinase-inhibitor structure by X-ray crystallography, and some compounds that would be classified as Type I based on their chemical structures induce DFG-flipped conformations of the activation loop that resemble Type II (7-9) inhibitors. Methods for analysing the activation loop conformation in solution would be invaluable in the development and characterization of kinase inhibitors, enabling classification in the absence of crystal structures. (10-12) Aurora kinases constitute a family of serine–threonine protein kinases whose localization and activities are precisely choreographed as a cell progresses through mitosis.(13-15) They play a major role in cell cycle progression and map to a chromosome region that is frequently amplified in tumors. (16) Aurora-A is activated by phosphorylation on Thr-288 and by the TPX2. (17) Crystal structures suggest that this involves a lever-arm-like movement of the Aurora-A activation loop from a relatively mobile conformation to a conformation that is stabilized by being hooked onto a short helical region in TPX2 (Figures 7.1.1 A and B).

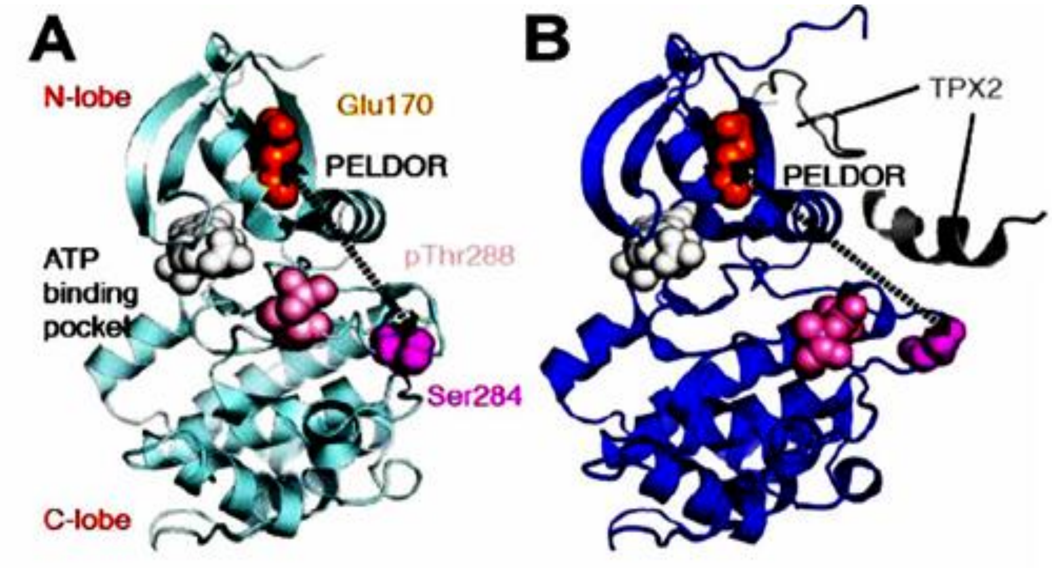


Figure 7.1.1: (A) Cartoon of Aurora-A kinase (PDB:1OL7) showing the MTSL labeling sites (Glu170, Thr288 and Ser284). (B) Cartoon of Aurora-A with TPX bound (PDB:1OL5).

However, this model has yet to be probed in solution. The application of SDSL and PELDOR spectroscopy (18, 19) was investigated to measure distances between pairs of spin labels attached to Aurora-A in the presence or absence of ligands and the TPX2 protein. PELDOR separates dipole–dipole coupling between spins, which is inversely proportional to the cube of their distance. It can measure distances between spin labels on the nanometer scale (1.5–10 nm) (20) Dipolar spectroscopy has been successfully employed to study kinases. (21-23) However, in the best of the knowledge of the author, this approach has not been applied to a kinase activation loop.

7.2 Material and Methods

The expression constructs for Aurora-A 122-403 E170C, S284C, C290A, C393A and Aurora-A 122-403 E170C, T287A, T288C, C290A, C393A were produced by site-directed mutagenesis of the expression vector, pET30TEV Aurora-A 122-403 C290A, C393A produced in earlier work.(24) Expression and purification of the kinase was carried out as previously described. (25) The resulting protein was subject to buffer exchange into 20 mM Tris pH 7.0, 0.2 M NaCl, 5 mM $MgCl_2$ & 10% (v/v) glycerol made with deuterium oxide using a 5 ml Desalting HP column as per the manufacturer’s instructions (GE Healthcare). 70-100 μM MTSL-Aurora-A variant containing 30% glycerol/d6-glycerol was used for the PELDOR studies. The spin labelling efficiency was equal to 85-90% as measured following a previous published procedure.

(26) The S284C variant was shown to be phosphorylated on Thr288 by Western Blot analysis using a phosphospecific antibody (data not shown). PELDOR measurements were performed at 50 K on a Bruker Elexsys 580 spectrometer operating at either 9.4 or 34 GHz. The four-pulse DEER sequence (27) $\pi/2\nu_{obs}–\tau_1–\pi\nu_{obs}–t–\pi\nu_{pump}–(\tau_1+\tau_2–t)–\pi\nu_{obs}–\tau_2–$echo was applied. Phase-cycling was applied. Measurements at 9.4 GHz used: $\pi/2\ \nu_{obs}$ pulse length of 16 ns, $\pi\nu_{obs}$ pulse length of 32 ns and $\pi\nu_{pump}$ pulse length of 32 ns. Pump pulses were applied at the maximum of the field sweep spectrum with the observe pulses 65 MHz lower. τ_1 was varied by incrementing the first $\pi\nu_{obs}$ pulse position over eight steps of 8 ns and 56 ns for averaging of the proton and deuterium nuclear modulation, respectively. Measurements at 34 GHz used: $\pi/2\nu_{obs}$ pulse length of 20 ns, $\pi\nu_{obs}$ pulse length of 40 ns and $\pi\nu_{pump}$ pulse length of 50 ns. The separation between the pump and observe pulse was 55 MHz. The software DEERAnalysis2013 (28) was used to remove the background due to intermolecular interactions using an exponential background decay and to calculate the interspin distance distribution by Tikhonov regularization. Distributions were calculated using a rotamer library using MMM software.(29)

7.3 Results and discussion

The structure of Aurora-A kinase to identify suitable pairs of sites for site-specific labeling was analyzed. Glu170 on the αB-helix of the kinase N-lobe was identified as a stable position that does not respond to inhibitors, and Ser284 and Thr288 were chosen on the activation loop were the MTSL was added. Using SDSL and PELDOR spectroscopy, distances were measured with and without the presence of a ligand, ADP and TPX2 (Figure 7.3.1).

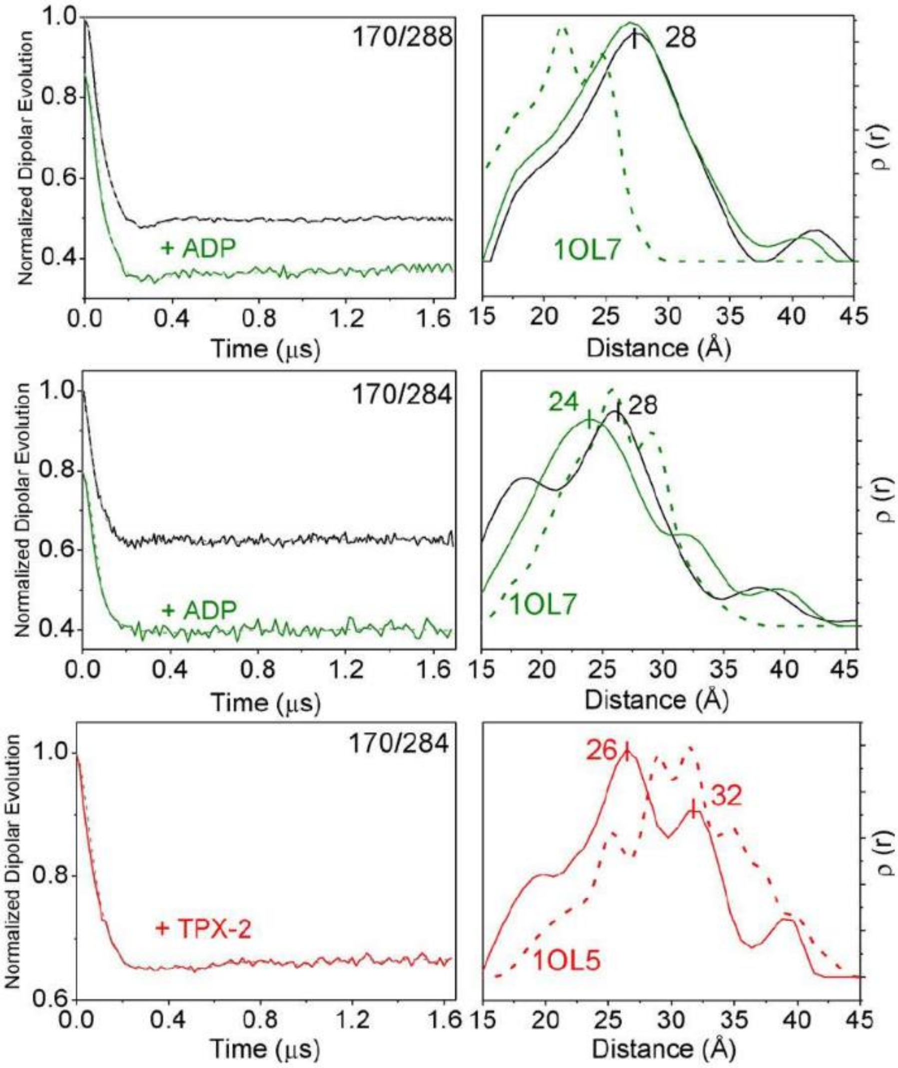
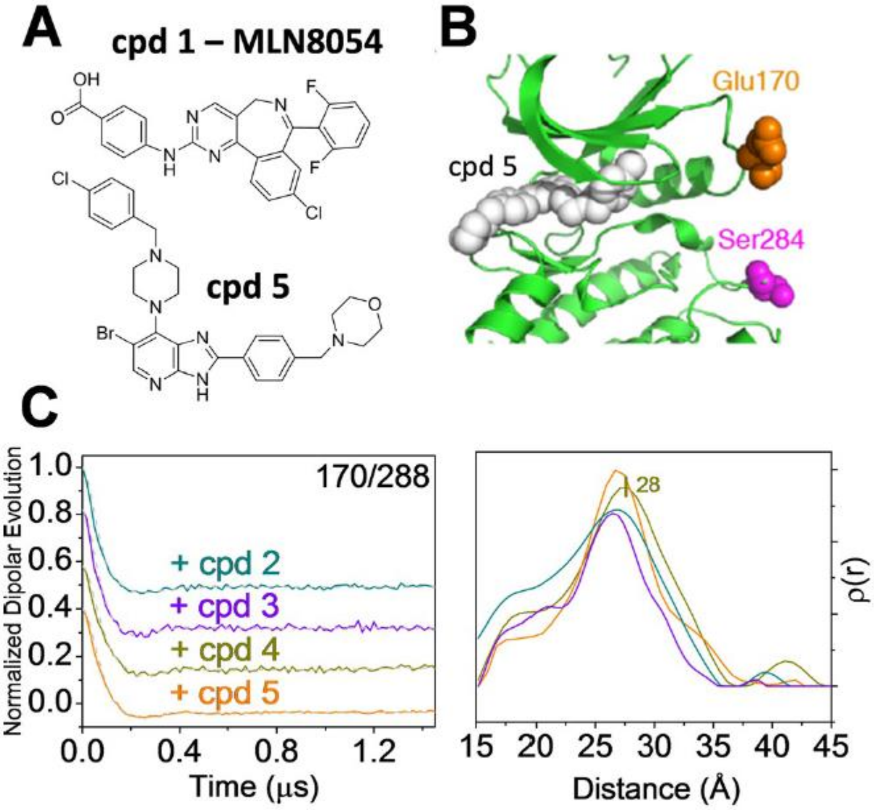


Figure 7.3.1: Background-corrected PELDOR traces at 9.4 GHz for MTSL-labeled Aurora-A kinase variants 284 and 288 with and without (black lines) a four-fold excess of ADP and TPX2 (left column). Form factor fits are given as a dashed line. Corresponding distance distributions (right columns) derived using Tikhonov regularization (α = 100). Rotamer library-derived distributions are given as dashed lines. All PELDOR traces before background correction are given in Figure 7.1 ESI in electronic supporting information.

The PELDOR traces for MTSL-labeled Aurora-A variants E170C/T284C and E170C/T288C showed a shallow oscillation, indicative of a wide distribution of conformational states. The extracted distance distribution showed a dominant mean distance at 28 Å. Based on the acquisition window, the data are of sufficient quality to establish from the distribution a reliable mean and width with approximate shape. The wide distribution suggests spatial disorder of the Aurora-A kinase activation loop, consistent with crystal structures, as well as multiple internal rotamers of the MTSL. (30) Comparison was made to the distribution created from a rotamer library (29) using PDB:1OL7. Although, a similar distribution width was found, the mean distance was offset in the case of the E170C/T288C variant, suggesting a difference in the conformation of the activation loop in solution. CW EPR spectra showed no significant dipolar broadening of the EPR line width, indicating that distances <15 Å are not significantly populated (Figure 7.2 ESI in electronic supporting information). Little difference in the distance distributions was observed with and without the ADP (present in the crystal structures) for the E170C/T288C variant; however, the E170C/T284C variant showed a shift of the mean distance to 24 Å. The effect of TPX2 on the activation loop of Aurora-A phosphorylated on T288 was investigated with a PELDOR experiment on MTSL-labelled E170C/T284C Aurora-A, this resulted in a pronounced change in the distance distribution compared to Aurora-A without ADP, consistent with a different conformational state of the activation loop (Figure 7.3.1). There was a sharp peak in the distance distribution at 34 Å, which is in reasonable agreement with the conformational change observed in the crystal structure. Small-molecule inhibitors of Aurora kinases, such as compounds 1 and 5 (Figure 7.3.2 A), have been identified as potential cancer therapeutics, (31) and clinical trials are ongoing. (32) To validate the site-directed-labeling EPR approach, we tested imidazo [1,2-a] pyrazine-based (compounds 2 and 3) (33) and imidazo[4,5-b] pyridine-based inhibitors (compounds 4 and 5), (34) which dock into the ATP binding site but do not affect the activation loop, as observed by X-ray crystallography (Figures 7.3.2 B and 7.3 B ESI in electronic supporting information). (33, 34) The results showed little change in the mean and width of the distribution (Figure 7.3.2 C), indicating minimal perturbation of the activation loop, a distinct difference compared to TPX2.


A
cpd 1 – MLN8054
OH
O
N
H
N
N
N
F
F
Cl
Cl
cpd 5
N
N
Br
N
N
N
H
N
O
B
Glu170
cpd 5
Ser284
C
Normalized Dipolar Evolution
170/288
+ cpd 2
+ cpd 3
+ cpd 4
+ cpd 5
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.4
0.8
1.2
Time (μs)
28
ρ(r)
15
20
25
30
35
40
45
Distance (Å)

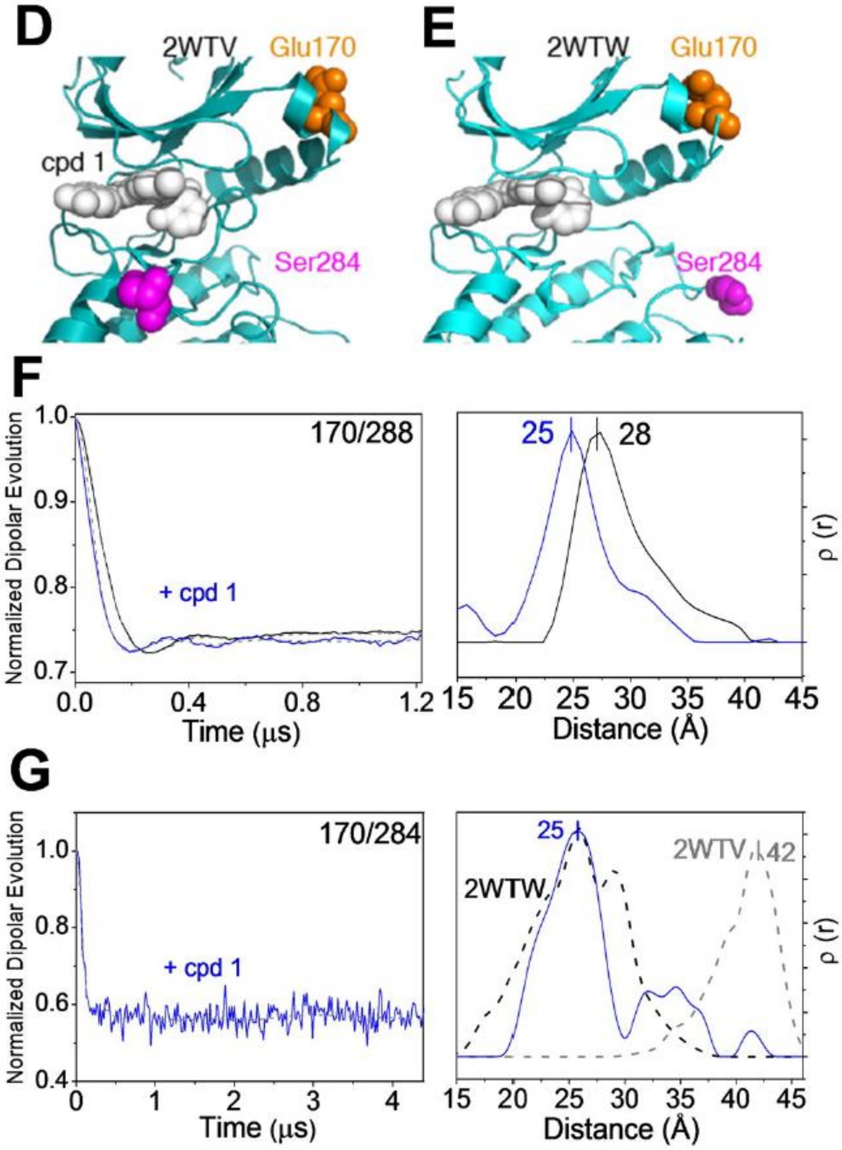


Figure 7.3.2: (A) Chemical structures of compounds 1 and 5. (B) Cartoon of the Aurora-A binding site for compound 5 in PDB:2X6G. (C) Background-corrected PELDOR traces at 9.4 GHz for MTSL-labeled Aurora-A E170C/T288C with a four-fold excess of inhibitors (left column). Form factor fits are given as a dashed line. Distance distributions derived using Tikhonov regularization (α=100) (right column). (D) Cartoon of the Aurora-A binding site of compound 1 in PDB: 2WTV. (E) Cartoon of the Aurora-A binding site of compound 1 in PDB: 2WTW. (F) The same as (C) but with cpd 1, but at 34 GHz. (G) The same as (C) for Aurora-A E170C/S284C. Rotamer library-derived distributions are given as dashed lines. All PELDOR traces before background correction are given in Figure 7.4 ESI in electronic supporting information.

A crystal structure of compound 1 (35) bound to Aurora-A (PDB: 2WTV) showed that the drug induces an ordered and closed, inactive, conformation of the Aurora-A activation loop (Figure 7.3.2 D). (7) The conformational change in the activation loop moves the Cα of Val279 by approximately 19 Å compared with its position in the ADP-bound structure. However, other crystal structures (Figure 7.3.2 E) showed an alternative conformation of

Aurora-A bound to compound 1 (PDB codes: 2WTW and 2X81). (7, 36) The specific conformation induced by MLN8237, an Aurora-A inhibitor closely related to compound 1, might explain its activity in neuroblastoma cells. (37) The conformation of the kinase in solution in the presence of the compound is an open and important question. PELDOR experiments on MTSL-labelled E170C/T288C Aurora-A in the presence of compound 1 revealed a clear change in the oscillation, reflected by a change in the major peak of the distance distribution (Figure 7.5 ESI in electronic supporting information). The distance distribution appears to narrow with the presence of compound 1 (absent with compounds 2–5), suggesting a greater order of the conformation of the activation loop. Further experiments at 34 GHz (Figure 7.3.2 F) gave a well-defined oscillation and confirmed the result. The distribution is narrower and shifts to a shorter dominant distance of 25 Å, as compared to Aurora-A alone. Key measurements were repeated to evaluate changes in MTSL-labelled Aurora-A E170C/S284C on ligand binding (Figure 7.3.2 G). Upon addition of compound 1, a change in the mean distance and a narrowing of the distribution were observed, which is absent in the case of compound 5 (Figure 7.6 ESI in electronic supporting information). Interestingly, the rotamer library produced with PDB: 2WTV (Figure 7.3.2 G, grey dash) showed a significant difference with a mean distance of 42 Å, suggesting the conformation of the activation loop is more akin to the DFG-in state adopted in PDB: 2WTW, which differs from that induced by compounds 2–5, because the N and C lobes of the kinase are twisted. There is perhaps only a minor contribution from the DFG-up state in PDB: 2WTV. Subsequent addition of TPX2 caused a broad distance distribution profile, which is distinct from that of TPX2 and compound 1 alone, indicating a dynamic conformation that results from the opposing tendencies of TPX2 and compound 1 (Figure 7.7 ESI in electronic supporting information). Recently, the structures of a number of mammalian kinases bound to ligands and substrates have been solved using X-ray crystallography. Although structures for open and closed forms are available for certain kinases, it is becoming clear that they occupy several conformational states, and that our knowledge of these states in solution is poor. In particular, the dynamics of the activation loop and how it responds to ligands is poorly characterized. PELDOR is complementary to other biophysical methods used to study kinase activation loops and holds many advantages, such as the smaller size of their labels compared to Förster resonance energy transfer (FRET) and no issues with limitations of molecular weight or difficulties with chemical exchange in NMR. However, one weakness of methods such as PELDOR and FRET is that the measurements are of the synthetic modifications of the protein, not the main chains and side chains of the amino acids. Indeed, the full extent of the movement of the activation loop will be understood through ongoing work by using molecular dynamics to implement the distance constraints provided by the PELDOR measurements to model the full conformational change of the activation loop. Inarguably, a combination of approaches will be required to solve this problem.

7.4 Conclusions and future work

The studies here provide key information on adaptations to inhibitor binding. These results confirm that PELDOR is a suitable approach to discriminate between different classes of ligand that do, or do not, influence the conformation of the kinase. Overall, the results suggest a correlation between the trends found for inhibitors and TPX2 with PELDOR spectroscopy and the observed conformational changes in the activation loop as judged by X-ray crystallography. Although the differences in the mean of the PELDOR distributions for the inhibitors are small, they show a clear trend and are able to discriminate whether or not inhibitor binding to the ATP site influences the activation loop. This work suggests that PELDOR spectroscopy could be a potential method to classify inhibitors by their effects on the conformation of the kinase activation loop as well as the study of native substrates in solution. Moreover, the method has sufficient sensitivity to detect the influence of TPX2 on the activation loop of Aurora-A in the absence of an inhibitor, and shows that this is altered in the presence of compound 1. The ultimate aim of this work is to exploit the information gleaned from these studies to generate plausible models of the structure and dynamics of human kinases. This may enable the development of inhibitors that exploit specific conformational states that are not accessible for study by other methods.

Acknowledgements

This work was supported by funding from Bruker (A.J.F. and M.G.C.) and a Cancer Research UK grant (C24461A12772 to R.B.). The authors would like to acknowledge the use of the EPSRC National EPR Service and Facility.

References

1. Endicott JA, Noble ME, Johnson LN. The structural basis for control of eukaryotic protein kinases. Annual review of biochemistry. 2012;81:587-613. Epub 2012/04/10.

2. Bayliss R, Fry A, Haq T, Yeoh S. On the molecular mechanisms of mitotic kinase activation. Open biology. 2012;2(11):120136. Epub 2012/12/12.

3. Nolen B, Taylor S, Ghosh G. Regulation of protein kinases; controlling activity through activation segment conformation. Molecular cell. 2004;15(5):661-75. Epub 2004/09/08.

4. Wang XJ, Chu HY, Lv MJ, Zhang ZH, Qiu SW, Liu HY, et al. Structure of the intact ATM/Tel1 kinase. Nature communications. 2016;7.

5. Case DA, Babin, V., Berryman, J.T., Betz, R.M., Cai, Q., Cerutti, D.S., Cheatham, III, T.E., Darden, T.A., Duke, R.E., Gohlke, H., Goetz, A.W., Gusarov, S., Homeyer, N., Janowski, P., Kaus, J., Kolossváry, I., Kovalenko, A., Lee, T.S., LeGrand, S., Luchko, T., Luo, R., Madej, B., Merz, K.M., Paesani, F., Roe, D.R., Roitberg, A., Sagui, C., Salomon-Ferrer, R., Seabra, G., Simmerling, C.L., Smith, W., Swails, J., Walker, R.C., Wang, J., Wolf, R.M., Wu X. and Kollman P.A. . AMBER 14, University of California, San Francisco. 2014.

6. Zhao Z, Wu H, Wang L, Liu Y, Knapp S, Liu QS, et al. Exploration of Type II Binding Mode: A Privileged Approach for Kinase Inhibitor Focused Drug Discovery? ACS chemical biology. 2014;9(6):1230-41.

7. Dodson CA, Kosmopoulou M, Richards MW, Atrash B, Bavetsias V, Blagg J, et al. Crystal structure of an Aurora-A mutant that mimics Aurora-B bound to MLN8054: insights into selectivity and drug design (vol 427, pg 19, 2009). Biochem J. 2010;427:551-.

8. Solanki S, Innocenti P, Mas-Droux C, Boxall K, Barillari C, van Montfort RLM, et al. Benzimidazole Inhibitors Induce a DFG-Out Conformation of Never in Mitosis Gene A-Related Kinase 2 (Nek2) without Binding to the Back Pocket and Reveal a Nonlinear Structure-Activity Relationship. J Med Chem. 2011;54(6):1626-39.

9. Huang ZF, Tan L, Wang HY, Liu Y, Blais S, Deng JJ, et al. DFG-out Mode of Inhibition by an Irreversible Type-1 Inhibitor Capable of Overcoming Gate-Keeper Mutations in FGF Receptors. ACS chemical biology. 2015;10(1):299-309.

10. Shults MD, Imperiali B. Versatile fluorescence probes of protein kinase activity. Journal of the American Chemical Society. 2003;125(47):14248-9.

11. Simard JR, Getlik M, Grutter C, Pawar V, Wulfert S, Rabiller M, et al. Development of a Fluorescent-Tagged Kinase Assay System for the Detection and Characterization of Allosteric Kinase Inhibitors. Journal of the American Chemical Society. 2009;131(37):13286-96.

12. Zondlo SC, Gao F, Zondlo NJ. Design of an Encodable Tyrosine Kinase-Inducible Domain: Detection of Tyrosine Kinase Activity by Terbium Luminescence. Journal of the American Chemical Society. 2010;132(16):5619-+.

13. Carmena M, Earnshaw WC. The cellular geography of aurora kinases. Nat Rev Mol Cell Bio. 2003;4(11):842-54.

14. Barr AR, Gergely F. Aurora-A: the maker and breaker of spindle poles. Journal of cell science. 2007;120(17):2987-96.

15. Fernandez A, Ferrando-Soria J, Pineda EM, Tuna F, Vitorica-Yrezabal IJ, Knappke C, et al. Making hybrid [n]-rotaxanes as supramolecular arrays of molecular electron spin qubits. Nature communications. 2016;7:10240. Epub 2016/01/09.

16. Asteriti IA, Rensen WM, Lindon C, Lavia P, Guarguaglini G. The Aurora-A/TPX2 complex: A novel oncogenic holoenzyme? Bba-Rev Cancer. 2010;1806(2):230-9.

17. Bayliss R, Sardon T, Vernos I, Conti E. Structural basis of Aurora-A activation by TPX2 at the mitotic spindle. Molecular cell. 2003;12(4):851-62. Epub 2003/10/29.

18. Milov AD, Ponomarev AB, Tsvetkov YD. Electron Electron Double-Resonance in Electron-Spin Echo - Model Biradical Systems and the Sensitized Photolysis of Decalin. Chem Phys Lett. 1984;110(1):67-72.

19. Milov AD, Salikhov KM, Shirov MD. Application of Eldor in Electron-Spin Echo for Paramagnetic Center Space Distribution in Solids. Fiz Tverd Tela+. 1981;23(4):975-82.

20. Jeschke G. DEER Distance Measurements on Proteins. Annual Review of Physical Chemistry, Vol 63. 2012;63:419-46.

21. Borbat PP, Freed JH. Measuring distances by pulsed dipolar ESR spectroscopy: Spin-labeled histidine kinases. Method Enzymol. 2007;423:52-+.

22. Bhatnagar J, Sircar R, Borbat PP, Freed JH, Crane BR. Self-Association of the Histidine Kinase CheA as Studied by Pulsed Dipolar ESR Spectroscopy. Biophysical journal. 2012;102(9):2192-201.

23. Fu Z, Aronoff-Spencer E, Backer JM, Gerfen GJ. The structure of the inter-SH2 domain of class IA phosphoinositide 3-kinase determined by site-directed spin labeling EPR and homology modeling. P Natl Acad Sci USA. 2003;100(6):3275-80.

24. Rowan FC, Richards M, Bibby RA, Thompson A, Bayliss R, Blagg J. Insights into Aurora-A kinase activation using unnatural amino acids incorporated by chemical modification. ACS chemical biology. 2013;8(10):2184-91. Epub 2013/08/09.

25. Concilio MG, Fielding AJ, Bayliss R, Burgess SG. Density functional theory studies of MTSL nitroxide side chain conformations attached to an activation loop. Theoretical chemistry accounts. 2016;135(4):1-6.

26. G. R. Eaton SSE, D. P. Barr, R. T. Weber. ed.; Springer. ed2010.

27. Pannier M, Veit S, Godt A, Jeschke G, Spiess HW. Dead-time free measurement of dipole-dipole interactions between electron spins. Journal of magnetic resonance. 2000;142(2):331-40. Epub 2000/01/29.

28. Jeschke G, Chechik V, Ionita P, Godt A, Zimmermann H, Banham J, et al. DeerAnalysis2006 - a comprehensive software package for analyzing pulsed ELDOR data. Applied magnetic resonance. 2006;30(3-4):473-98.

29. Polyhach Y, Bordignon E, Jeschke G. Rotamer libraries of spin labelled cysteines for protein studies. Physical chemistry chemical physics : PCCP. 2011;13(6):2356-66. Epub 2010/12/01.

30. Concilio MG, Fielding AJ, Bayliss R, Burgess SG. Density functional theory studies of MTSL nitroxide side chain conformations attached to an activation loop. Theoretical chemistry accounts. 2016;135:97. Epub 2016/04/14.

31. De Groot JEH, J. V. Anzola, A. Motamedi, M. Yoon, Y. L. Wong, D. Jenkins, H. J. Lee, M. B. Martinez, R. L. Davis, T. C. Gahman, A. Desai, A. K. Shiau. Frontiers in oncology. 2015;5.

32. Manchado E, Guillamot M, Malumbres M. Killing cells by targeting mitosis. Cell Death Differ. 2012;19(3):369-77.

33. Bouloc N, Large JM, Kosmopoulou M, Sun CB, Faisal A, Matteucci M, et al. Structure-based design of imidazo[1,2-a] pyrazine derivatives as selective inhibitors of Aurora-A kinase in cells. Bioorg Med Chem Lett. 2010;20(20):5988-93.

34. Bavetsias V, Large JM, Sun CB, Bouloc N, Kosmopoulou M, Matteucci M, et al. Imidazo[4,5-b]pyridine Derivatives As Inhibitors of Aurora Kinases: Lead Optimization Studies toward the Identification of an Orally Bioavailable Preclinical Development Candidates. Journal of Medicinal Chemistry. 2010;53(14):5213-28.

35. Manfredi MG, Ecsedy JA, Meetze KA, Balani SK, Burenkova O, Chen W, et al. Antitumor activity of MLN8054, an orally active small-molecule inhibitor of Aurora A kinase. P Natl Acad Sci USA. 2007;104(10):4106-11.

36. Sloane DA, Trikic MZ, Chu MLH, Lamers MBAC, Mason CS, Mueller I, et al. Drug-Resistant Aurora A Mutants for Cellular Target Validation of the Small Molecule Kinase Inhibitors MLN8054 and MLN8237. ACS chemical biology. 2010;5(6):563-76.

37. Brockmann M, Poon E, Berry T, Carstensen A, Deubzer HE, Rycak L, et al. Small Molecule Inhibitors of Aurora-A Induce Proteasomal Degradation of N-Myc in Childhood Neuroblastoma. Cancer Cell. 2013;24(1):75-89.

Electronic supporting information (ESI)

Detection of Ligand-induced Conformational Changes in the Activation Loop of Aurora-A Kinase by PELDOR Spectroscopy

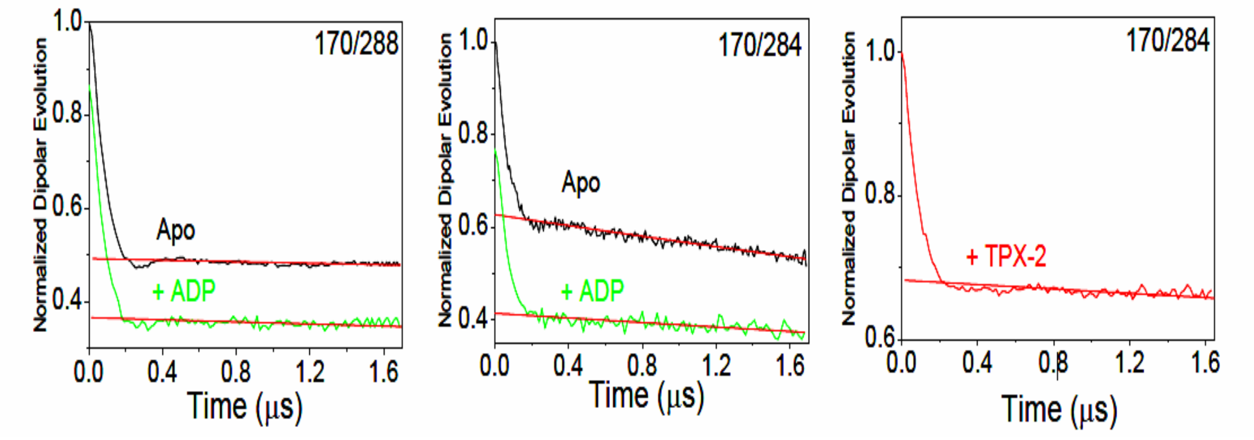


Figure 7.1 ESI: Raw PELDOR traces at 9 GHz for MTSL labelled Aurora-A kinase variants and with 4-fold ADP and TPX2. The background is shown as a red line.

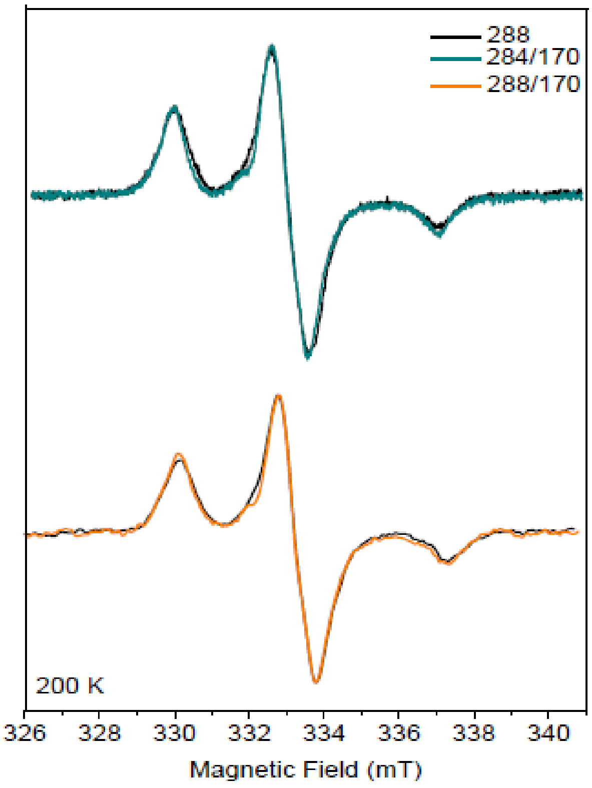


Figure 7.2 ESI: GHz continuous-wave EPR spectra of singly MTSL labelled Aurora-A kinase at T288C (black) and doubly MTSL labelled Aurora-A kinase at positions E170C/ S284C (cyan) and E170C/S288C (orange). The spectra were recorded at 200 K using 0.1 mT modulation amplitude. All samples contained 30 % glycerol.

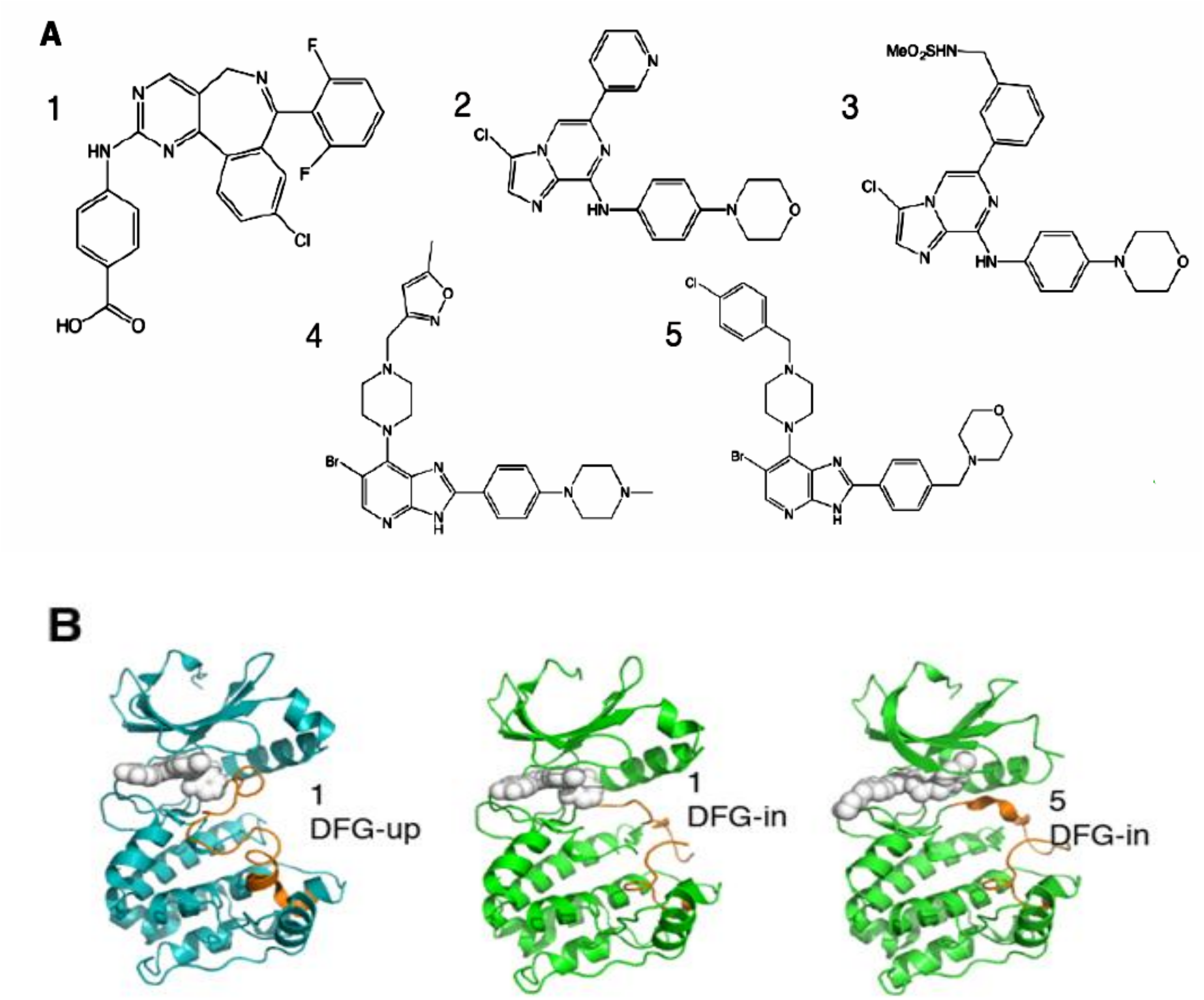


Figure 7.3 ESI: (A) Chemical structures of compounds used in this study. Compound 1 is MLN8054. (B) Crystal structures of Aurora-A/inhibitor complexes. In each case, the compound is shown as white spheres and the number relates to the compound number in panel A. From top to bottom, the PDB codes of the crystal structure shown are 2WTV, 2WTW and 2X6D.

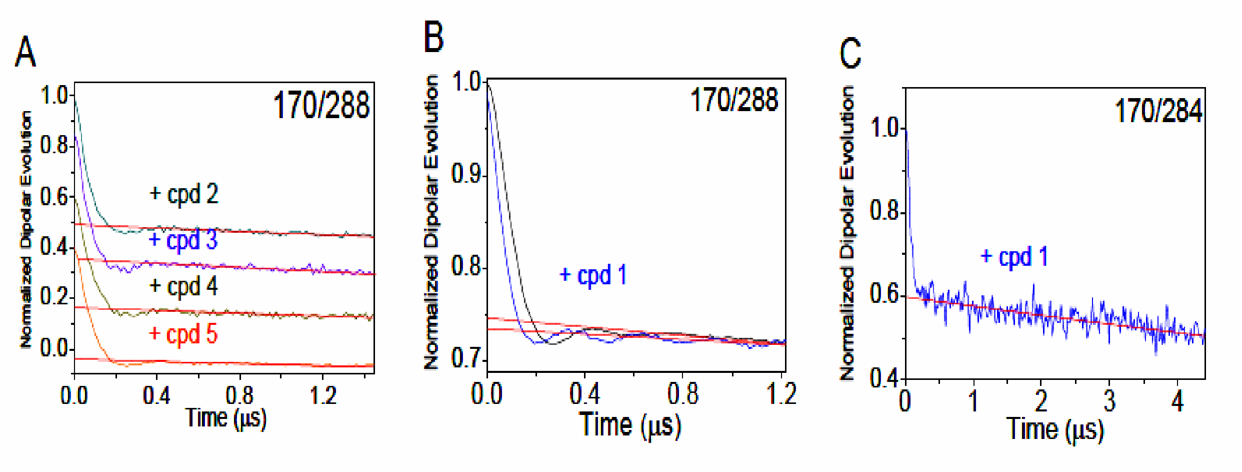


Figure 7.4 ESI: (A/C) Raw PELDOR traces at 9 GHz for MTSL labelled Aurora-A kinase variants with 4-fold excess of inhibitors. (B) Raw PELDOR trace at 34 GHz for MTSL labelled Aurora-A kinase variant with 4-fold excess of inhibitor. The background is shown as a red line.

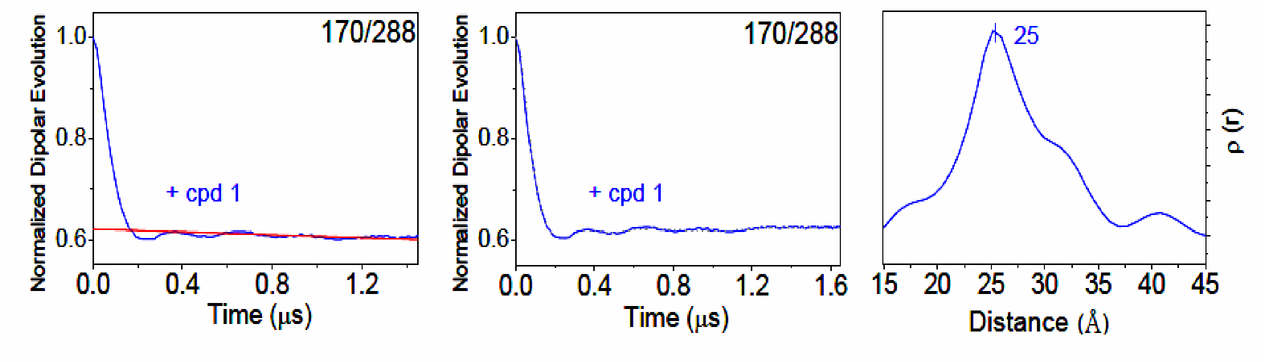


Figure 7.5 ESI: (left) Raw PELDOR trace at 9 GHz for MTSL labelled Aurora-A kinase variant 170/288 with 4-fold excess of cpd1. The background is shown as a red line. (middle) Background subtracted PELDOR trace. Form factor fits are given as a dashed line. (right) Distance distributions derived by Tikhonov regularization (α = 100).

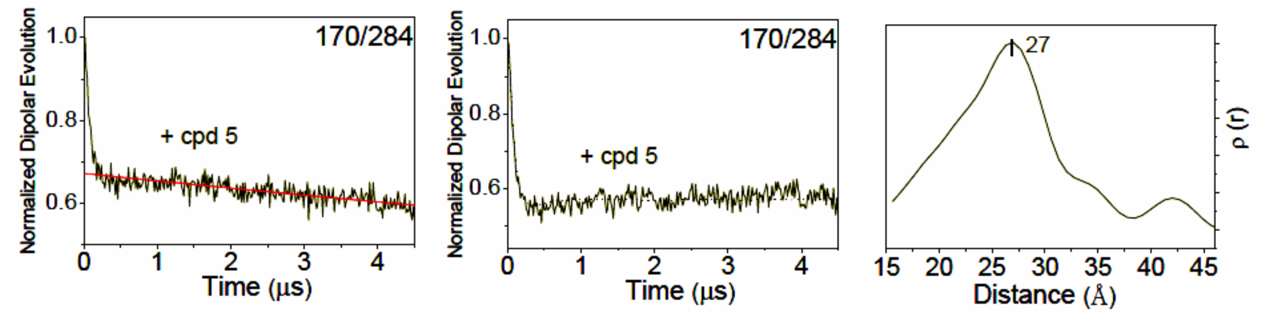


Figure 7.6 ESI: (left) Raw PELDOR trace at 9 GHz for MTSL labelled Aurora-A kinase variant 170/284 with 4-fold excess of cpd5. The background is shown as a red line. (middle) Background subtracted PELDOR trace. Form factor fits are given as a dashed line. (right) Distance distributions derived by Tikhonov regularization (α = 100).

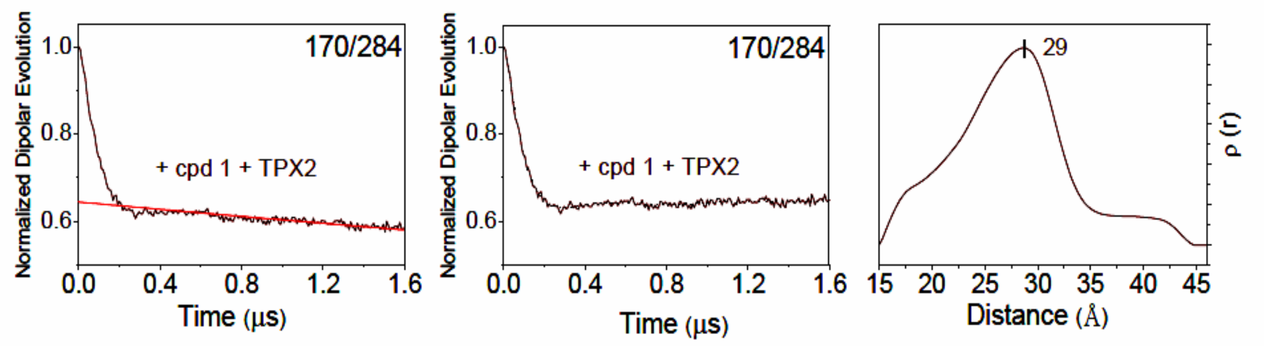


Figure 7.7 ESI: (left) Raw PELDOR trace at 9 GHz for MTSL labelled Aurora-A kinase variant 170/284 with 4-fold excess of cpd1 + TPX2. The background is shown as a red line. (middle) Background subtracted PELDOR trace. Form factor fits are given as a dashed line. (right) Distance distributions derived by Tikhonov regularization (α = 100).

8. Summary and suggestions for future work

In this project extensive computational and EPR studies were performed in order to characterize the conformational dynamics of the Aurora-A kinase protein's activation loop. The full work was summarized in four papers.

The first paper describes a detailed conformational study using QM calculations, based on DFT, employed for the characterization of the conformational states of the MTSL side chain. This work has been performed in order to determine the effect of the side chain on the EPR parameters (*g*- and *A*-tensors) required for the simulations of the CW EPR spectra and to determine some starting conformations for the MD studies on the full spin-labelled system. A population of 76 rotamers was found at the minima of the potential energy surfaces, simulated 94 GHz CW EPR spectra using conformers of the MTSL with different geometry were seen to be similar to each other and well-reproduced the experimental data, indicating that the side chain does not affect the spin density in the ring.

The second paper describes a computational study, based on classical MD simulations, performed in order to determine conformational states in which both the activation loop and spin label were engaged, as well as structural fluctuations, order parameters and rotational correlation times related to the motion of the full spin-labelled system. A comparison between conformations obtained from MD and the initial X-ray crystal structure revealed small differences, residues from 289 to 295 moved inward toward the N-lobe of the protein, while residues from 288 to 282 moved toward regions more exposed to the solvent. The analysis of the auto-correlation function, using different vectors within the MTSL spin label, revealed the presence of slower and faster motions corresponding to configurations of the spin label fully exposed to the solvent and other interacting with residues of the protein. The room temperature CW EPR spectra were simulated using two correlation times equal 3 ns and 11 ns that were very close to the values extracted from the auto-correlation function of the N-O nitroxide group, as well as to those related to the fast motions of the dihedral angles χ_4 and χ_5, indicating that the internal dynamics of the spin label arise predominantly from the motions around these bonds. Comparisons between simulated and experimental EPR spectra revealed that the motion of the protein and spin label occurred on a comparable timescale that ranged between 0.1 ns and 10 ns, indicating that the dynamics of the Aurora-A was dominated by events occurring on the nanosecond timescale.

The third paper describes a comparison between structure and dynamics in un-phosphorylated and phosphorylated activation loops, using MD simulations. It was found that there were small differences between the un-phosphorylated structures obtained from MD and the initial X-ray crystal structure, while no significant changes were found for the case of the phosphorylated structures. MD showed that the phosphorylated activation loop was characterized by less dynamics than the un-phosphorylated activation loop. In both cases the most flexible region was that localized between residues 282 and 294, which includes residues T287 and 288, which are crucial for the catalytic activity of Aurora-A kinase. The CW EPR spectra were simulated using rotational correlation times extracted from the analysis of the auto-correlation function analysis. In the case of the phosphorylated activation loop it was observed that correlation times related to the fast motions related of the N-O nitroxide group and χ_4 dihedral angle could simulate the EPR spectrum, confirming that the room temperature EPR spectra of Aurora-A kinase prevalently arise from the motion around these bonds.

The fourth paper describes experimental studies using PELDOR measurements in order to determine the average distances and distance distributions between spin-label pairs introduced in the C- and N-lobes of the protein. In this work two mutants were studied, the first was spin-labelled at positions T284 (activation loop) and E170 (helix), and the second was spin-labelled at positions T288 (activation loop) and E170 (helix). In both cases, broad distance distributions and multiple populations were identified, indicating the high flexibility of the structures. The effect of a class of inhibitors on the distance distributions was tested, variations were observed only upon addition of the inhibitor MLN8054, ATP and TPX2 in the first mutant, while no significant changes were observed in the case of the second mutant. The addition of the inhibitor MLN8054 and ATP decreased the values of the mean distance, while the addition of TPX2 increased it, confirming results obtained by X-ray crystallography that showed the inhibitor and ATP binding in the ATP binding pocket pulled closer the N- and the C-lobe, while the TPX2, pulled apart the N- and the C lobes.

In conclusion, MD simulations provided insights about structure and dynamics in solution of un-phosphorylated and phosphorylated activation loops there were consistent with those obtained by X-ray crystallography which showed that phosphorylated activation loops are more rigid that un-phosphorylated activation loops, while CW EPR studies gave insights into the overall dynamics of the spin-labelled system. This work revealed also the ability of the PELDOR experiment to classify inhibitors by their effects on the conformation of the kinase activation loop and provided substantial information about the structure and dynamics of the Aurora-A kinase protein and is the basis for further research for the development of classes of inhibitors.

Looking into the future, computationally, more sophisticated stochastic models, that combine directly EPR and MD simulation, could be developed in order to connect conformational states of the activation loop with the conformational states of the label that can allow quantitative interpretation of the EPR spectra from MD simulation. This can be supported by more research for the development of spin labels that are more rigid than MTSL, such as 4-amino-1-oxyl-2,2,6,6-tetramethyl-piperidine-4-carboxylic acid (TOAC), (79) and 4-(3,3,5,5-tetramethyl-2,6-dioxo-4-oxylpiperazin-1-yl)-l-phenylglycine (TOPP), (80) *etc.* for easier characterization of the conformational dynamics.

A further suggestion is to use QM/MM MD models to provide a description of the potential that the spin labels experience from both neighbouring side groups and protein backbone. It is well-known that this is a complicated function of space and time and the order parameters have to be determined both at the point of attachment of the tether and on its ring in order to determine the degree of restriction of the N-O bond by the protein environment. This would be a new approach specialized for the characterization of the dynamics of nitroxide spin labels that can provide input parameters for simulations and quantitative interpretation of the EPR spectra. After several discussions within the EPR and computational communities and in the best of knowledge of the author of the thesis, similar approaches do not yet exist, but research is ongoing.

9. Appendices

Appendix I

Spinach scripts

Script used in the Spinach software to perform simulations of the 94 GHz and 9 GHz EPR spectra of the singly MTSL spin labelled Aurora-A kinase at 150 K shown in Figure 4.1 ESI in Chapter 4, Figure 5.3.3.1 A in Chapter 5 and Figure 5.6 ESI.

```
% Isotopes
sys.isotopes={'14N', 'E'};

% Coordinates
inter.coordinates={[-2.734793 0.644417 -0.224154];
                   [ ]};
% Coupling Matrices
inter.coupling.matrix=cell(2);
inter.coupling.matrix{1, 2}=1e6*gauss2mhz([4e+000 -1.490e+000 -6.428e+000; -
1.490e+000 6e+000 5.524e+000; -6.428e+000 5.524e+000 3.3e+001]);

% Scalar Couplings
inter.coupling.scalar=cell(2);

% Zeeman Interactions
inter.zeeman.matrix=cell(1, 2);
inter.zeeman.matrix{2}=[2.0074 -0.0010219 0.0015196; -0.0010219 2.0068 -
0.0013239; 0.0015196 -0.0013239 2.0028];

% Magnet induction
sys.magnet=0.3499;

% Relazation theory
inter.relaxation='damp';
inter.damp_rate=5e7;

% Basis set
bas.formalism='sphten-liouv';
bas.approximation='none';

% Disable trajectory-level SSR algorithms
sys.disable={'trajlevel'};

% Spinach housekeeping
spin_system=create(sys,inter);
spin_system=basis(spin_system,bas);

% Sequence parameters
parameters.spins={'E'};
parameters.rho0=state(spin_system,'L+','E');
parameters.coil=state(spin_system,'L+','E');
parameters.decouple={};
parameters.offset=-0.15e8;
```

```
parameters.sweep=5.6e8;
parameters.npoints=1650;
parameters.zerofill=1650;
parameters.axis_units='mT';
parameters.grid='lebedev_ab_rank_131';
parameters.derivative=1;
parameters.invert_axis=0;

% Simulation
fid=powder(spin_system,@acquire,parameters,'esr');

% Fourier transform
spectrum=fftshift(fft(fid,parameters.zerofill));

% Plotting
plot_1d(spin_system,real(spectrum),parameters);
```

Scripts used in the Spinach software to perform simulations of the CW EPR spectra of the singly MTSL spin labelled Aurora-A kinase protein measured at 298 K shown in Figure 5.3.3.1 B and Figure 6.3.3.1 in Chapter 6 are shown below.

```
% Isotopes
sys.isotopes={'14N', 'E'};

% Coordinates
inter.coordinates={[-2.734793 0.644417 -0.224154];
                   [ ]};
% Coupling Matrices
inter.coupling.matrix=cell(2);
inter.coupling.matrix{1, 2}=1e6*gauss2mhz([4e+000 -1.490e+000 -6.428e+000; -
1.490e+000 6e+000 5.524e+000; -6.428e+000 5.524e+000 3.3e+001]);

% Scalar Couplings
inter.coupling.scalar=cell(2);

% Zeeman Interactions
inter.zeeman.matrix=cell(1, 2);
inter.zeeman.matrix{2}=[2.0074 -0.0010219 0.0015196; -0.0010219 2.0068 -
0.0013239; 0.0015196 -0.0013239 2.0028];

 % Magnet induction X-band
sys.magnet=0.3499;

% Basis set
bas.formalism='sphten-liouv';
bas.approximation='none';

% SLE housekeeping
spin_system=create(sys,inter);
spin_system=basis(spin_system,bas);

% SLE parameters
parameters.max_rank=10;
parameters.tau_c=17e-9;
```

```
parameters.rho0=state(spin_system,'L+','E');
parameters.coil=state(spin_system,'L+','E');
parameters.decouple={};
parameters.spins={'E'};
parameters.sweep=[-2.2e8 2e8];
parameters.npoints=1500;
parameters.zerofill=1500;
parameters.axis_units='mT';
parameters.invert_axis=0;

% SLE simulation
spectrum_sle=gridfree(spin_system,@slowpass,parameters,'esr');
spectrum_sle=fdvec(spectrum_sle,5,1);

% SLE plotting
plot_1d(spin_system,real(spectrum_sle),parameters);
```

Appendix II

Force field parameterization of the MTSL spin label

Information related to extension of the AMBER force field (ff14SB) to the structure of the MTSL side chain shown in Figure I.1.

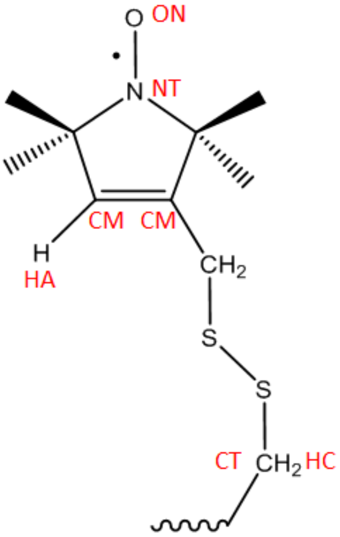


Figure I.1: Structure of the MTSL side chain used for the extension of the AMBER force field (ff14SB). Atom types used in the tables from 1 to 5 are highlighted in red. CT indicates any sp^3 carbons, CM indicates any sp^2 carbons, HC indicates any hydrogens attached to sp^3 carbons, HA indicates the hydrogen attached to the sp^2 carbon and S indicates the sulphur atoms. The new atom types related to the nitroxide group (NO) are ON in the case of the oxygen and NT in the case of the nitrogen.

The potential energy functions on which the AMBER force field used in this work is based on Equation 3.3.2.2. The first three summations represent stretch (Table I.1), bend (Table I.2) and torsion terms (Table I.3 and I.4). The final sum is referred to interactions for atoms (*i* and *j*) separated by more than three covalent bonds and represents van der Waals (Table I.5) and Coulomb (Table I.6) interactions.

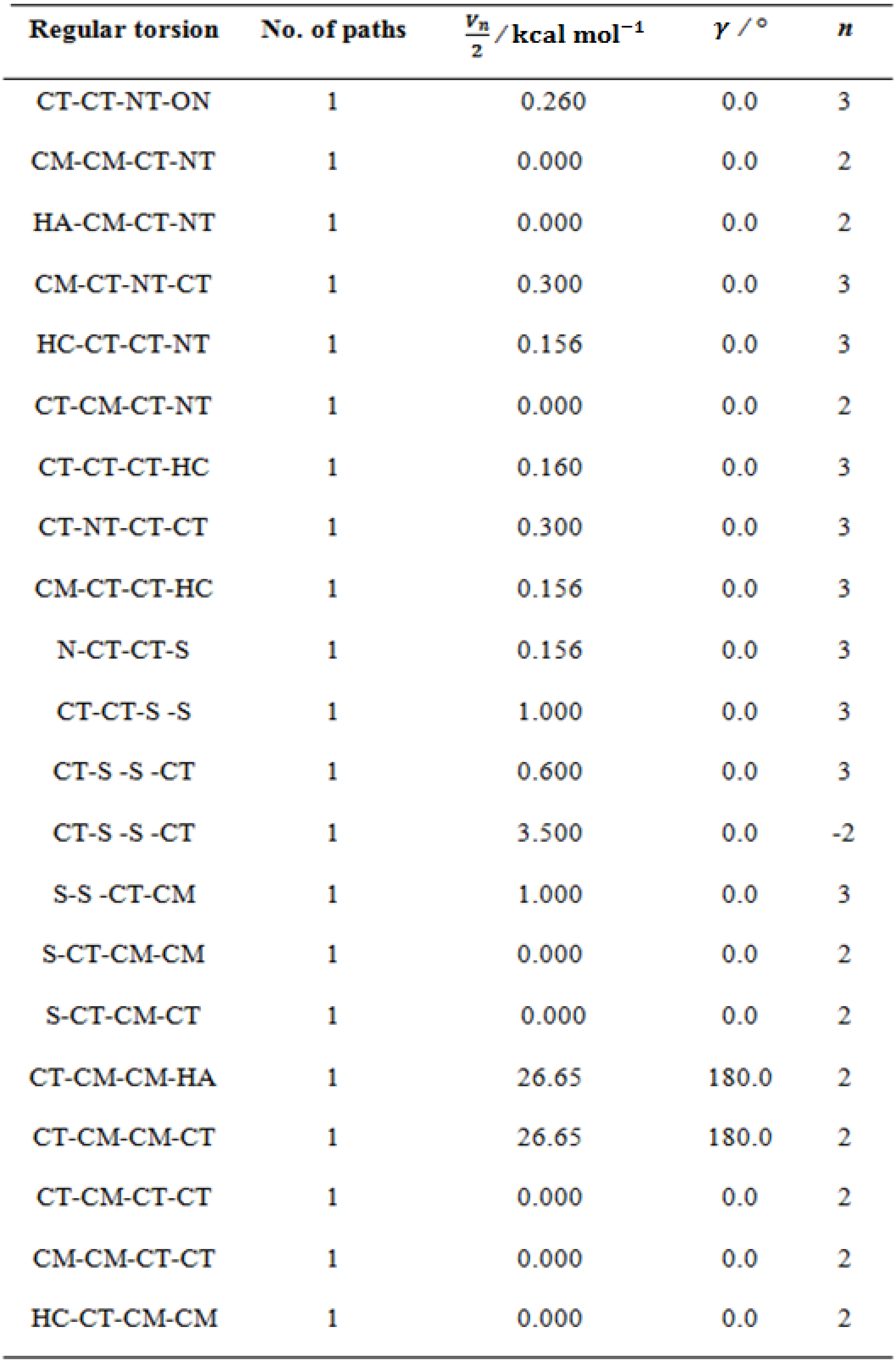
Table I.1: Bond starching values calculated for the MTSL side chain. k_r and r_0 represent the force constant and the equilibrium bond length, respectively.

Bond stretching	$\boldsymbol{k_r\,/\,kcal\,mol^{-1}}$ Å^{-2}	$\boldsymbol{r_0\,/\,\mathring{A}}$
NT-ON	360.00	1.26
CT-NT	370.00	1.47
CT-CT	310.10	1.54
CT-HC	337.30	1.09
CM-CM	447.30	1.43
CT-CM	318.30	1.51
CM-HA	367.00	1.087
S-S	166.00	2.05
CT-S	225.80	1.82

Table I.2: Angle bending values calculated for the MTSL side chain considered in this work. k_θ end θ_0 represent the force constant and the equilibrium bond angle, respectively.

Angle bending	$\boldsymbol{k_\theta / kcal\, mol^{-1}\, rad^{-2}}$	$\boldsymbol{\theta_0 / °}$
CT-NT-ON	82.00	117.50
CT-NT-CT	50.00	120.00
NT-CT-CT	60.18	110.38
CM-CT-NT	66.47	111.52
CT-CM-CM	64.33	123.42
CT-CM-CT	62.70	116.52
CM-CM-HA	50.30	119.70
HA-CM-CT	35.15	123.44
CM-CT-CT	63.53	111.44
HC-CT-CM	47.03	110.49
S-CT-CM	63.62	104.97
CT-S-S	68.34	103.10
CT-CT-S	50.10	112.69
CT-CT-CT	40.21	110.63
CT-CT-HC	47.37	110.05
HC-CT-HC	39.43	108.35

Table I.3: Dihedral angle values calculated for the MTSL side chain. $\frac{V_n}{2}$ represents the amplitude of the curve, n represents the periodicity and φ represents shifts the entire curve along the rotation angle axis.



Regular torsion	No. of paths	$\frac{V_n}{2}$/kcal mol^{-1}	γ /°	n
CT-CT-NT-ON	1	0.260	0.0	3
CM-CM-CT-NT	1	0.000	0.0	2
HA-CM-CT-NT	1	0.000	0.0	2
CM-CT-NT-CT	1	0.300	0.0	3
HC-CT-CT-NT	1	0.156	0.0	3
CT-CM-CT-NT	1	0.000	0.0	2
CT-CT-CT-HC	1	0.160	0.0	3
CT-NT-CT-CT	1	0.300	0.0	3
CM-CT-CT-HC	1	0.156	0.0	3
N-CT-CT-S	1	0.156	0.0	3
CT-CT-S -S	1	1.000	0.0	3
CT-S -S -CT	1	0.600	0.0	3
CT-S -S -CT	1	3.500	0.0	-2
S-S -CT-CM	1	1.000	0.0	3
S-CT-CM-CM	1	0.000	0.0	2
S-CT-CM-CT	1	0.000	0.0	2
CT-CM-CM-HA	1	26.65	180.0	2
CT-CM-CM-CT	1	26.65	180.0	2
CT-CM-CT-CT	1	0.000	0.0	2
CM-CM-CT-CT	1	0.000	0.0	2
HC-CT-CM-CM	1	0.000	0.0	2

Table I.4: Improper torsion values calculated for the MTSL side chain considered in this work. $\frac{V_n}{2}$ represents the amplitude of the curve, n represents the periodicity and φ represents shifts the entire curve along the rotation angle axis.

Improper torsion	$\frac{V_n}{2}$ **(kcal mol^{-1})**	φ / °	n
CM-CT-NT-CT	1.1	180.0	2
CM-CT-CM-CT	1.1	180.0	2
CM-CT-CM-HA	1.1	180.0	2

Table I.5: Non-bonded interaction values calculated for the MTSL side chain. The A and B parameters control the depth and position of the potential energy well for a given pair of non-bonded interacting atoms.

Non-bonded interactions	R^a/ Å	ε^b/ kcal mol^{-1}
ON	1.6750	0.2445
NT	1.9000	0.2600
CT	1.9080	0.1094
HC	1.4870	0.0157
S	2.0000	0.2500
CM	1.9080	0.0860
HA	1.4870	0.0157

$a\ R_{ij} = R_i + R_j$

$b\ \varepsilon_{ij} = (\varepsilon_i \varepsilon_j)^{1/2}$

Note that: $A_{ij} = \varepsilon_{ij}\,(R_{ij})^{12}$

$B_{ij} = 2\,\varepsilon_{ij}\,(R_{ij})^{6}$

Appendix III

AMBER scripts

1) The following script was written following AMBER manual and the online AMBER tutorials to build non-standard residue (the MTSL side chain) within the protein system

```
#To calculate charges:
p uhf/6-31g(d) geom=connectivity iop(6/33=2) pop=mk
# Load Leap and AMBER force field:
$AMBERHOME/bin/tleap -s -f $AMBERHOME/dat/leap/cmd/leaprc.ff14SB
$AMBERHOME/bin/xleap -s -f $AMBERHOME/dat/leap/cmd/leaprc.ff14SB
#Prepare .prepin files:
$AMBERHOME/exe/antechamber -i  MTSL.out  -fi gout  -o  mts.prepin  -fo prepi  -c resp  -s 2
-rn MTS -at amber  -nc -1
#To convert MTS.prepin files to MTS.frcmod files:
$AMBERHOME/exe/parmchk  -i  MTS.prepin  -f  prepi  -o  MTS.frcmod
#To load frcmod files in Leap:
loadamberparams  MTS.frcmod
#To load .prepin files:
loadamberprep MTS.prepin
#To make .lib files in Leap:
saveoff MTS  MTS.lib
#To make MTS.prmtop  and MTS.inpcrd files in Leap:
saveamberparm MTS MTS.prmtop   MTS.inpcrd
#To fix head and tail:
desc MTS
set  MTS head MTS.MTS.N3
set  MTS tail MTS.MTS.C6
#To save to library:
saveoff 4ceg_mtsl 4ceg_mtsl.lib
#To load lib. File:
loadoff 4ceg_mtsl.lib
#To save pdb files:
savepdb 4ceg_mtsl 4ceg_mtsl.pdb
```

2) The following scripts were used to perform the three steps of the MD simulation: minimization, equilibration and production.

Minimizations:

- Minimization with Cartesian restraints for the solute

```
&cntrl
 imin=1, maxcyc=200,
 ntpr=5,
 ntr=1,
 &end
 Group input for restrained atoms
 100.0
 RES 1 265
 END
END
```

- Minimization of the entire system

```
&cntrl
 imin=1, maxcyc=200,
 ntpr=5,
&end
```

Equilibrations:

- Constant volume equilibration

```
&cntrl
  imin   = 0,
  irest  = 0,
  ntx    = 1,
  ntb    = 1,
  cut    = 10.0,
  ntr    = 1,
  ntc    = 2,
  ntf    = 2,
  tempi  = 0.0,
  temp0  = 300.0,
  ntt    = 3,
  gamma_ln = 1.0,
  nstlim = 20000, dt = 0.001
  ntpr = 2000, ntwx = 2000, ntwr = 20000
 /
Keep protein fixed with weak restraints
1.0
RES 1 265
END
```

```
END
```

- Constant pressure and temperature equilibration with no restraints (200 ps):

```
&cntrl
imin = 0, irest = 1, ntx = 7,
  ntb = 2, pres0 = 1.0, ntp = 1,
  taup = 2.0,
  cut = 10.0, ntr = 0,
  ntc = 2, ntf = 2,
  tempi = 300.0, temp0 = 300.0,
  ntt = 3, gamma_ln = 1.0,
  nstlim = 200000, dt = 0.001,
  ntpr = 2000, ntwx = 2000, ntwr = 20000
 /
```

- Last equilibration in NVE ensemble (1 ns):

```
&cntrl
  imin = 0, irest = 1, ntx = 7,
  ntb = 1,
  ntt = 0,
  cut = 9.0,
  ntc = 2, ntf = 2,
  tol = 0.0000005,
  ntr = 0,
  iwrap = 1,
  nstlim = 1000000, dt = 0.001,
  ntpr = 2000, ntwx = 2000, ntwr = 20000
 /
```

Production:

- Production in NVE ensemble

```
&cntrl
  imin = 0, irest = 1, ntx = 7,
  ntb = 1,
  ntt = 0,
  cut = 9.0,
  ntc = 2, ntf = 2,
  tol = 0.0000005,
  ntr = 0,
  iwrap = 1,
  nstlim = 70000000, dt = 0.001,
  ntpr = 2000, ntwx = 2000, ntwr = 20000
```

/

3) To perform analysis with CPPTRAJ from AmberTools 14:

```
$AMBERHOME/bin/cpptraj
```

- To load topology and trajectory files, respectively:

```
parm 4ceg_mtsl.prmtop
trajin 4ceg_mtsl.mdcrd
```

- To calculate RMSD:

```
rms ToFirst :1-265&!@H= first out rmsd1.agr mass
run
xmgrace rmsd .dat
```

- To perform dihedral analysis

```
dihedral “name_dihedral_angle” :residue_number@atom_type :
residue_number@atom_type : residue_number@atom_type :
residue_number@atom_type out name_dihedral_angle.dat
```

- To perform the PCA

```
rms first :1-265&!@H= out 4ceg_rmsd.dat
run
average crdset cpu-gpu-average
run
createcrd cpu-gpu-trajectories
run
crdaction cpu-gpu-trajectories rms ref cpu-gpu-average :1-265&!@H=
run
crdaction cpu-gpu-trajectories matrix covar \
name cpu-gpu-covar :1-265&!@H=
run
analysis diagmatrix cpu-gpu-covar out cpu-gpu-evecs.dat \
vecs 5 name myEvecs3 \
nmwiz nmwizvecs 5 nmwizfile dna_loop_4nb1.nmd nmwizmask :1-265&!@H=
crdaction cpu-gpu-trajectories projection GPU modes myEvecs3 \
beg 1 end 5  :1-265&!@H=
run
hist GPU:1 bins 100 out cpu-gpu-hist.agr norm name GPU-1
hist GPU:2 bins 100 out cpu-gpu-hist.agr norm name GPU-2
hist GPU:3 bins 100 out cpu-gpu-hist.agr norm name GPU-3
hist GPU:4 bins 100 out cpu-gpu-hist.agr norm name GPU-4
hist GPU:5 bins 100 out cpu-gpu-hist.agr norm name GPU-5
```

- To perform the clustering analysis

```
rms first :1-265@CA
run
average Avg.rst
run
createcrd crd1
run
# Fit to average structure, save fit coordinates.
reference Avg.rst [avg]
run
crdaction crd1 rms ref [avg] :1-265@CA
# Calculate eigenvalues/eigenvectors.
crdaction crd1 matrix covar :1-265@CA name myCovar
run
analysis diagmatrix myCovar out evecs.dat vecs 5 name MyModes1
# Project fit coordinates onto eigenvectors.
crdaction crd1 projection P5 modes MyModes1 beg 1 end 5 :1-265@CA
# Perform clustering using projections.
runanalysis cluster C5data P5  :1-265 crdset crd1 \
```

In the case of the K-means algorithm:

```
kmeans clusters 10 \
out cnumvtime.dat repout rep repfmt pdb \
summary summary.dat info info.dat cpopvtime cpop.agr normframe \
loadpairdist pairdist CpptrajPairDist \
singlerepout singlerep.nc singlerepfmt netcdf
```

%average linkage algorithm

```
runanalysis cluster C3data P3  :1-265 crdset crd1 \
hieragglo epsilon 5 averagelinkage \
sieve 250 out cnumvtime.dat repout rep repfmt pdb \
summary summary.dat info info.dat cpopvtime cpop.agr normframe \
loadpairdist pairdist CpptrajPairDist \
singlerepout singlerep.nc singlerepfmt netcdf
#To determine lowest RMSD to centerid:
parm rep.c4.pdb
reference rep.c4.pdb parm rep.c4.pdb [c4]
trajin 4ceg_mtsl.mdcrd
rms c4 ref [c4] :1-265@CA out c4.arg
```

- To perform the iRED analysis

```
vector v1 : residue_number@atom_type ired : residue_number@atom_type
vector v2 : residue_number@atom_type ired : residue_number@atom_type
...
vector vn : residue_number@atom_type ired : residue_number@atom_type
matrix ired name matired order 2
diagmatrix matired vecs “number of the vectors (n)” out ired.vec name ired.vec
ired relax freq 500 order 2 tstep 1.0 tcorr “number of the vectors (n)” out output.out noefile noe modes ired.vec
```

Appendix IV

Further Ph.D. publications

1. Casadei, C. M.; Gumiero, A.; Metcalfe, C. L.; Murphy, E. J. Concilio, M. G.; Schradere, T. E.;Basran, J.; Fielding, A.J. ; Ostermann, A.; Blakeley, M. P.; Raven, E. L.; Moody, P.C. E. Neutron cryo-crystallography captures the protonation state of ferryl heme in a peroxidase. *Science.* **2014**, 345, 193-197. http://science.sciencemag.org/content/345/6193/193.long (*Number of citations: 36*).

2. Fielding, A. J.; Concilio, M. G.; Heaven, G.; Hollas, M. A. New Developments in Spin Labels for Pulsed Dipolar EPR. *Molecules.* **2014**, *19*, 16998-17025. http://www.mdpi.com/1420-3049/19/10/16998 (*Number of citations: 13*).