

Young Modellers Forum – 1st December 2006

Abstracts for Poster Presentations

Computational infrared spectroscopy studies of enzyme intermediates

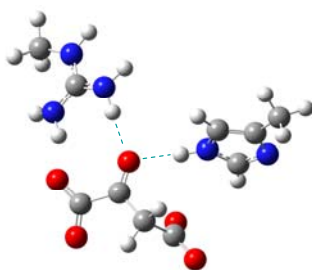
Michelle Szeto, Marc Van der Kamp, Senthil Kittusamy, Jeremy Harvey, Adrian Mulholland and Fred Manby

School of Chemistry, University of Bristol

Infra red spectroscopy has been used experimentally for many years to probe enzyme reaction intermediates.¹ Our interests lie in trying to understand a range of hydrolytic enzymes whose mechanism involves addition of a nucleophile to a carbonyl group thus yielding an oxyanion intermediate. The short lifetime of such intermediates means that they are difficult to probe experimentally. Studying the infra-red activity of the carbonyl precursor in the enzyme substrate complex can serve as a proxy to understanding the stabilisation of the oxyanion intermediate, as a consequence of its interactions with the protein environment. The carbonyl oxygen and the anionic oxygen atom are located in a site known as the oxyanion hole. In this site, the two anionic oxygen atoms accept multiple hydrogen bonds donated by the protein backbone and side chains. In the enzyme substrate complex, these hydrogen bonds polarise the C-O bond, thus shifting its vibrational frequency. IR spectroscopy has been used in numerous studies to understand the fine details of this interaction and its effects on reactivity.

We have implemented a QM/MM methodology to compute IR frequencies for large systems such as enzymes. This presentation will describe the application of this approach to probe the polarisation of the carbonyl bond of oxaloacetate in the active site of Citrate synthase (below). Our initial results are in good agreement with experimental studies.²

In addition we are studying the acyl-enzyme intermediate of a class C β lactamase with the antibiotic, Aztreonam, with the same method. Experimentally four possible conformations have been observed by characteristic IR absorption bands.³ Initial computational results for this system will also be presented.



OAA in the oxyanion hole of Citrate synthase

References

1. Wharton, C. W. (2000) *Nat. Prod. Rep.* 17, 447-453
 2. Kurz, L. C. (1987) *Biochemistry.* 26, 2623-2627
 3. Wilkinson, A. S., Bryant, P. K., Meroueh, S. O., Page, M. G. P., Mobashery, S., and Wharton, C. W. (2003) *Biochemistry* 42, 1950-1957
-

Interaction of DNA with Groove Binding Ligands

Mark Broome, *P.M.Rodger*

University of Warwick, Coventry, CV4 7AL

Synthetic molecules that target the major groove in a sequence-selective way are a major goal in molecular medicine. Recently a major step has been taken toward achieving this goal: a family of compounds has been developed from which a number have experimentally been seen to bind in the major groove of DNA. Experimental techniques have provided some information regarding the binding strength and preferred binding sites of the cylinder on DNA. From all the experimental data it is clear that the cylinder binds in the major groove and is able to induce dramatic conformational changes in the DNA; these are unprecedented effects with synthetic DNA binders. However, gaining molecular level information in such a macromolecular system is challenging. Molecular dynamics (MD) simulations can provide information at the molecular level that is complementary to experiment and therefore are an ideal way to get a better understanding of this system. In this work we present the results of various MD simulations designed to probe the DNA-cylinder system.

We have studied the effect of using CHARMM22 or CHARMM27 as the force-field for the simulations. Results showed that uncomplexed DNA simulated with CHARMM22 was less stable than the comparable strand of DNA simulated with CHARMM27.

Investigations into the effects of the cylinders charge and shape are also reported. Multi-nanosecond simulations were performed using two related synthetic cylinders, one with two Fe(II) metal centers and the other with two Cu(I) centers, and DNA.

Finally the role of DNA within the system was investigated by performing a series of simulations of the cylinders with d(ATATATATATAT)₂, d(CGCGCGCGCGCG)₂ and d(CGCGCATATACG). Simulations with this variety of DNA strands has only produced one system (C_{Cu}²⁺ with d(ATATATATATAT)) that appears to be major groove binding. Unfortunately due to the force field over stabilizing the B-Form of the DNA there is not a lot of DNA activity to report, compared to CHARMM22 simulations in which dramatic coiling was observed.

Computational studies of α -lactone reactivity

Mahmoud E. S. Soliman, J. Grant Buchanan, Giuseppe D. Ruggiero, Ian H. Williams

Department of Chemistry, University of Bath, Bath BA2 7AY, United Kingdom

α -Lactones are Cinderella species, whose existence is often overlooked. Known to polymerise or decompose readily, they generally do not survive warming in solution and are unstable at room temperature. Yet we have compelling experimental evidence for their spontaneous intermediacy in the rapid room-temperature aqueous bromination of disodium salts of derivatives of maleic and fumaric acids.

Density functional theoretical studies within the polarised continuum model for aqueous solvation show that the cyclic halonium ion formed from addition of Cl^+ or Br^+ to alkenes $^-\text{O}_2\text{C}(\text{R}')\text{C}=\text{C}(\text{R}'')\text{CO}_2^-$ with two carboxylate substituents are transition structures for interconversion of α -lactones. Internal rotation to a different conformer allows intramolecular nucleophilic attack by the second carboxylate group on the α -lactone leading to β -lactone products as found experimentally. The cyclic halonium ion from addition of Cl^+ or Br^+ to an alkene $\text{R}_2\text{C}=\text{CHCO}_2^-$ with one carboxylate substituent is also a transition structure for degenerate rearrangement of an α -lactone, but (depending on the nature of the β -substituents R) a β -lactone product may be obtained experimentally from electrophilic halogenation of α,β -unsaturated carboxylic acid salts. To account for this observation, we propose a dyotropic rearrangement of the α -lactone into the β -lactone, and predict that the barrier for this novel process in water is very low when $\text{R} = \text{CH}_3$. DFT/PCM methods agree qualitatively with a hybrid DFT/MM approach including explicit solvation by several hundred MM water molecules.

Depending on the nature of the α -substituents, hydrolysis may compete with intramolecular rearrangements, but DFT/PCM studies confirm that the site of attack is not $\text{C}=\text{O}$ but C_α . This may explain the range of products observed in aqueous bromination of disodium citraconate and mesaconate ($\text{R}' = \text{H}$, $\text{R}'' = \text{CH}_3$).

Study of the Conformational Dynamics of HIV-1 Protease using Reversible Digitally Filtered Molecular Dynamics

Sarah Williams, *Jonathan Essex*

School of Chemistry, University of Southampton, Highfield, Southampton, SO17 1BJ, UK

Reversible digitally filtered molecular dynamics (RDFMD) [1-3] is a method which enhances conformational dynamics through the application of digital filters to the internal velocities of a system to selectively enhance or suppress vibrational motion. This method has been applied to several protein systems, including the HIV enzymes.

The role of HIV-1 protease in the life-cycle of HIV involves the hydrolysis of viral polyproteins into functional proteins, which are essential for viral assembly and subsequent activity, making the enzyme an obvious target for drug development. However, mutations of this protein occur rapidly due to the high replication rate of the virus, heightened by selective pressure exerted by the current inhibitors which affects their long-term effectiveness.

The methods of conventional molecular dynamics and RDFMD have been applied to both the wild-type and mutant HIV-1 protease in the apo [4] and inhibitor-bound forms to evaluate the effects of mutations and inhibitor binding on the conformational dynamics of the enzyme. The RDFMD simulations of the apo wild-type enzyme reveal numerous flap opening events. The effect of the enzyme mutation and ligand binding on the conformations accessible to protein will be described. The RDFMD simulations are able to accelerate infrequent large-scale conformational changes, thereby revealing new conformations which were not present in simulations carried out using conventional molecular dynamics.

References

- [1] S. C. Phillips, M. T. Swain, A. P. Wiley, J. W. Essex and C. M. Edge, *J. Phys. Chem. B.*, **107**, 2098-2110 (2003)
- [2] S. C. Phillips, J. W. Essex and C. M. Edge, *J. Chem. Phys.*, **112**, 2586 (2000)
- [3] A. P. Wiley, M. T. Swain, S. C. Phillips and J. W. Essex, *J. Chem. Theory Comput.*, **1**, 24-35 (2005)
- [4] A. P. Wiley, S. L. Williams and J. W. Essex, to be submitted

Conformational Analysis and Molecular Dynamics at the Receptor Level of the Immunodominant Myelin Basic Protein Epitope 87-99 Implicated in Multiple Sclerosis, and its Antagonists Linear Altered Peptide Ligands

Efthimia Mantzourani, Simona Golič Grdadolnik, Andrea Brancale, John Matsoukas, Thomas Mavromoustakos, *Theodore Tselios, James Platts*

University of Patras, Department of Chemistry, 265 00, Rion, Patras, Greece;
Department of Chemistry, Cardiff University, P.O. Box 912, CF10 3TB, Cardiff, Wales, United Kingdom

Multiple Sclerosis (MS) is a myelin-specific CD4+ T-cell mediated autoimmune disease. Antigen-specific therapies that suppress autoreactive T cells or switch the immune response are possible therapeutic approaches. Myelin Basic Protein (MBP) is considered to be a candidate autoantigen, supported by evidence depicting that MBP₇₂₋₈₅ has the potential of inducing Experimental Autoimmune Encephalomyelitis, an animal model of MS, when injected in rats. Altered Peptide Ligands (APLs) of MBP have been designed to interfere with the formation of the trimolecular complex MHC–APL–T-cell Receptor (TCR), and hence to inhibit the disease. In particular, the linear APLs [Arg⁹¹, Ala⁹⁶]

MBP₈₇₋₉₉ and [Ala^{91,96}] MBP₈₇₋₉₉ demonstrate antagonistic activity. We performed conformational analysis of the APLs to study the structural causes of this antagonistic activity. A combination of 2D NMR and molecular dynamics was applied. From this, putative bioactive conformations were proposed, *i.e.* conformations that enable binding with the MHC, but prevent binding to the TCR and therefore T-cell activation. These conformations were then used as starting points for binding with a crystal structure of MHC isolated from a diseased patient. The APL-MHC complexes were subjected to molecular dynamics in aqueous solvent, and the same procedure was applied to the MPB₈₅₋₉₈ – MHC complex (PDB code 1ymm), allowing direct comparison of agonist (disease-inducing) and antagonist peptides (disease-preventing). The network of hydrogen bonds between peptide backbone and active site, hydrophobic interactions between peptide side chains and the binding site, and the solvent accessible surface area of bound peptides were studied, and a structural motif for antagonistic activity sought. For the APLs it was found that once bound to the MHC, significant changes occur in the orientation of the amino acids that serve as TCR anchors. Several possible mechanisms of action can be proposed from these findings. The TCR may not be able to bind at all to the APL-MHC complexes. If it binds, it must be in the conventional way, inducing an immune response and protection against disease. Alternatively, regulatory T-cells might be activated, maintaining peripheral tolerance by suppression of autoreactive lymphocytes that avoid negative selection. These findings could lead to the rational design of potent peptidomimetic compounds, active against MS.

Reference

Mantzourani, E.D.; Tselios, T.V.; Golič Grdadolnik, S.; Brancale, A.; Platts, J.A.; Matsoukas, J.M.; Mavromoustakos, T.M. A putative bioactive conformation for the altered peptide ligand of myelin basic protein and inhibitor of experimental autoimmune encephalomyelitis [Arg⁹¹, Ala⁹⁶] MBP₈₇₋₉₉ *J. Mol. Graph. Mod.* (2005) 25/1, 17-29

Sequence-Structure-Function Relationships of Viper Venom Serine Proteases

Sakthivel Vaiyapuri^a, Jon Gibbins^a, Rob Harrison^b and Gail Hutchinson^a

^a*School of Biological Sciences, University of Reading, Reading, United Kingdom*

^b*School of Tropical Medicine, University of Liverpool, Liverpool, United Kingdom*

Snake venom serine proteases (SVSPs) act on a variety of components of the blood coagulation cascade such as kininogen, fibrinogen and plasminogen. They are widely distributed in the family of viperidae and a few have been identified in hydrophidae and colubridae. This study mainly focuses on Viper Venom Serine Proteases (VVSPs). Despite their high sequence similarity VVSPs are quite narrow in their substrate specificity. This contrasts with thrombin, a multifunctional enzyme mainly involved in the blood coagulation cascade. However the relationship between the specificity of VVSPs and their amino acid sequences is poorly understood. Moreover many of the VVSPs whose sequences are known remain functionally uncharacterized. This project

aims to understand better the relationships between sequence, structure and function for VVSPs. This would help to predict the functions of uncharacterized VVSPs from sequence and also help in attempts to develop novel drug molecules for the treatment of viper bites. All known complete VVSPs amino acid sequences, together with information about their function were collected from the NCBI database. In total 130 VVSPs were collected, of which only 33 have been functionally characterized. A multiple sequence alignment of the sequences was generated using ClustalW. This was used to generate phylogenetic trees showing the functional and evolutionary relationships between the sequences. The sequence alignment was used to identify key catalytic residues and other residues involved in the specific function of individual proteins. Structural information about these was determined using known crystal structures of TSVPA, DAASPs and ACC-C, and homology models of VVSPs have been made to compare with other determined structures and their functions. Post translational modification sites in VVSPs were predicted as they play a key role in substrate recognition and specificity. A web-based database to make sequence and functional information about VVSPs generally available is being developed. In order to increase the information available about sequence and structure of VVSPs, the venom of *Bitis gabonica* has been obtained and we are currently aiming to purify, functionally characterize and crystallize the serine proteases present in this venom.

VSTx1, a Modifier of Kv Channel Gating, Localizes to the Interfacial Region of Lipid Bilayers

Chze Ling Wee, Daniele Bemporad, Zara A. Sands, Alessandro Grottesi, and *Mark S. P. Sansom*

Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, U. K.

VSTx1 is a tarantula venom toxin which binds to the archaebacterial voltage-gated potassium channel KvAP. VSTx1 is thought to access the voltage sensor domain of the channel via the lipid bilayer phase. In order to understand its mode of action and implications for the mechanism of channel activation, it is important to characterize the interactions of VSTx1 with lipid bilayers. Atomistic and coarse-grained (CG) molecular dynamics (MD) simulations have been used to explore VSTx1 localization and interactions with zwitterionic (POPC) and with anionic (POPE/POPG) lipid bilayers. In particular, a series of atomistic MD simulations have been used to explore the net drift of VSTx1 relative to the center of a bilayer, starting from different locations of the toxin. The preferred location of the toxin is at the membrane/water interface. Although there are differences between POPC and POPE/POPG bilayers, in both cases the toxin forms favorable interactions at the interface, maximizing H-bonding to lipid headgroups and to water molecules while retaining interactions with the hydrophobic core of the bilayer. Furthermore, an extended CG simulation of 200 ns duration reveals dynamic partitioning of VSTx1 into the interface of a POPC bilayer. The preferential location of VSTx1 at the interface is discussed in the context of Kv channel gating models and provides support

for a mode of action in which the toxin interacts with the Kv voltage sensor “paddle” formed by the S3 and S4 helices.

References:

Bemporad, D., C. L. Wee, Z. Sands, A. Grottesi, and M. S. P. Sansom. 2006. VSTx1, a Modifier of Kv Channel Gating, Localizes to the Interfacial Region of Lipid Bilayers. *Biochemistry*. 45:11844-11855.

Contemporary QSAR Classifiers Compared¹

Craig L. Bruce,[†] James L. Melville,[†] Stephen D. Pickett,[‡] *Jonathan D. Hirst*[†]

[†] *School of Chemistry, University of Nottingham, University Park, Nottingham NG7 2RD, UK*

[‡] *GlaxoSmithKline, Gunnels Wood Road, Stevenage, Hertfordshire SG1 2NY, UK*

The use of modern data mining techniques can be directly implemented for QSAR studies. Such techniques include decision tree, support vector machine and random forest. We have compared a selection of these classifiers using eight literature datasets, with two descriptor sets, one whole molecule and one fragment-based. Using the WEKA data mining workbench we have parameter tuned the support vector machine and random forest. The effect of applying resampling techniques, such as bagging and boosting, on various classifiers is also explored.

The results obtained from the classifiers improve when resampling techniques are applied, as well as making them more robust. Parameter tuning also has a positive effect, especially on the random forest, improving it from the most variable classifier to the most robust, 2.8% to 1.3%. A parameter tuned random forest produces the best predictions on 4 datasets, using the whole molecule descriptors, and on the remaining 4 datasets is within 2.1% of the best classifier.

To establish which classifier performs best we conduct multiple comparison statistical tests, namely the Friedman and Nemenyi test.² This is a two-stage, nonparametric statistical approach.

Our results show that the random forests have comparable performance to the current benchmark, support vector machine, it therefore should not be overlooked as a QSAR classifier. Support vector machine have limited interpretability, whereas the random forest with its bio-inspired tree structure allows further interpretation.

References:

1. Bruce, C. L.; Melville, J. L.; Pickett, S. D.; Hirst, J. D. Contemporary QSAR Classifiers Compared. *J. Chem. Inf. Mod.* Submitted.
2. Demšar, J., Statistical Comparisons of Classifiers over Multiple Data Sets. *J. Mach. Learn. Res.* **2006**, 7, 1-30

In silico simulation of the self association of amyloidogenic peptides.

Oliver Clarke, Geraint Thomas, Malcolm McLean and *Martin Parker*.

Institute of Molecular and Cellular Biology, University of Leeds.

Over 20 diseases have been linked with amyloid fibrils, including Alzheimer's, Parkinson's and prion diseases. The amyloidogenic proteins do not share native structures or sequence identity, though the fibrils themselves share a distinct morphology; a cross- β spine. A detailed molecular mechanism of amyloid fibril formation has not been established, partly due to the difficulty in studying fibril formation *in vivo/vitro*. Therefore, they are an excellent subject for study *in silico*. Using a simplified off-lattice protein model, RAFT (1), and a novel Monte Carlo sampling algorithm developed in our lab (2), Boltzmann-weighted conformations have been calculated for the self-associations of 15 peptides of between 4 and 9 residues in length that have been shown experimentally to form amyloid fibrils. Three of each peptide were placed in a box with periodic boundaries. Boltzmann-weighted conformations calculated for each three peptide system are predominantly β -sheet conformations. This result shows that the simple model used is capable of emulating the forces that drive the self association of amyloid peptides.

References

1. Gibbs, N., Clarke, A., and Sessions, R. Ab Initio Structure Prediction Using Physiochemical Potentials and a Simplified Off-lattice Model, *PROTEINS: Structure, Function, and Genetics* (2001) **43**, 186-202.
2. Thomas, G., Sessions, R., and Parker, M. Density guided importance sampling: application to a reduced model of protein folding, *Bioinformatics* (2005) **21**, 2839-2843.

The Importance of Water Molecules and Charge Assistance for the Inhibition of TGT

Ritschel, T.¹, Stengl, B.¹, Hörtner, S.², Diederich, F.², Klebe, G.¹

¹*Institute of Pharmaceutical Chemistry, Philipps-University Marburg, Marbacher Weg 6, 35032 Marburg/Lahn, Germany*

²*Department of Chemistry and Applied Biosciences, Swiss Federal Institute of Technology Zurich, ETH Zurich, 8093 Zurich, Switzerland*

The enzyme tRNA-Guanine Transglycosylase (TGT) has been identified as a valuable target for structure-based drug design. Several series of structurally diverse ligands have been designed, synthesized and characterized by crystal structure analysis and enzyme kinetics. TGT catalyzes the exchange of guanine in the wobble position of the anticodon loop of tRNA^{His, Asn, Asp, Tyr} by the modified base preQ₁. The modified tRNAs take influence on the translation efficacy of pathogenicity factors of *Shigella flexneri*, the causative agent of shigellosis. Therefore efficient inhibition of TGT could lead to new selective antibiotics. Virtual screening and structure-based drug design have suggested various ligand scaffolds to compete with the guanine of the cognate tRNA. Among these the most promising scaffold is the recently discovered *lin*-benzoguanine (6-aminoimidazol[4,5-g]quinazolin-8(7H)-one) [3]. Substitution of *lin*-benzoguanine by aliphatic and aromatic side chains to address a small hydrophobic subpocket did not result in a significantly improved affinity. To penetrate into the hydrophobic pocket, the ligand side chains have to perturb and partially replace a cluster of water molecules, adjacent to catalytically relevant residues. We believe that these waters are involved in a charge compensation in the binding pocket and conclude that their presence is a crucial structural element of TGT. Any attempts to replace these waters appear to be detrimental to binding affinity.

This observation provides a new concept for the next drug design cycle. Instead of replacing the water cluster we want our ligands to actively participate and support the water cluster. We performed a ReliBase search to compare the locations of the water molecules in the catalytic site of different TGT crystal structures to derive a consensus picture of this water. The established synthesis of *lin*-benzoguanines provides two possible attachment points for substituents to subsequently replace one or more members of the water cluster by functional groups imitating their interactions. To facilitate synthesis only commercial available substituents were selected and attached to the scaffold. Docking of the assembled model compounds was performed with GOLD and the water cluster was partially replaced by the polar side chains which involve in the hydrogenbond network of the water molecules [1]. A rescoring of the docking solutions was performed with DrugScore^{CSD} and faced with that of known inhibitors. In order to assess putative charges of protonation states of all ligands upon binding we performed Poisson-Boltzmann calculations based on PEOE_PB charges recently developed in our group [2]. A potential change of protonation will produce charges on our ligands. Accordingly, normal hydrogen bonds will turn into enhanced charge-assisted ones. Both aspects, the unperturbed water cluster and the putative charge assistance seem to be a crucial factor for ligand binding affinity.

References:

- [1] Czodrowski, P., Dramburg, I., Sotriffer, Ch., Klebe, G. Development, validation, and application of adapted PEOE charges to estimate pK_a values of functional groups in protein-ligand complexes, *Proteins: Structure, Function, and Bioinformatics*, in press
- [2] Verdonk, M. L., Chessari, G., Cole, J. C., Hartshorn, M. J., Murray, C. W., Nissink, J. W., Taylor, R. D. Taylor, R. Modeling Water Molecules in Protein-Ligand Docking Using GOLD, *J Med Chem* (2005) **48**, 6504-6515
-

GARLig: A fully automated tool for reagent enumeration and subset selection of de novo skeletons via genetic algorithms

Patrick Pfeffer, Gerhard Klebe

Department of Pharmaceutical Chemistry, Philipps-University, Marbacher Weg 6, 35032 Marburg, Germany

In combinatorial chemistry, molecules are assembled by linking suitable reagents taken from a large fragment space of starting materials. Often the number of molecules possible to enumerate greatly exceeds the amount feasible to handle for in depth *in silico* analysis and even more for synthetic realization. Genetic algorithms mimic Darwinian evolution and provide a powerful tool to make such exhaustive searches computationally tractable.

GARLig (**G**enetic **A**lgorithm using **R**eagents to compose **L**igands) has been designed in order to enumerate large fragment spaces of starting materials and to perform subset selections which satisfy target-specific 3D-scoring criteria (fitness). The algorithm uses fragment subsets from the Zinc database [1] and a knowledge-based scaffold library which considers a set of synthesis rules to assemble larger molecules from starting material fragments. Ligand optimization is performed by means of crossover and mutation operators to avoid large product space docking.

GARLig generates promising subsets for chemical synthesis and shows that this tool may find novel architectures for de novo design projects.

References

[1] Irwin and Shoichet, *J. Chem. Inf. Model.* 2005;45(1):177-82

[2] Douguet D, Thoreau E, Grassy G, *J. Comput. Aided Mol. Des.*, Vol. 14, No. 5. (July 2000), pp. 449-466

Aromaticity of the metal dithiolene ring during the nitrate reduction pathway

Ganga Periyasamy*, Ian H Hillier and Neil A Burton

School of Chemistry, The University of Manchester.

Metal complexes containing dithiolene ligands have attracted much attention due to their interesting properties-notably their versatile redox behaviour, novel geometries and their presence as catalytic centres in Mo and W enzymes. Nitrate reductase is one of the enzymes which contain two molybdopterin units and catalyses the reduction of nitrate to nitrite ($\text{NO}_3^- + 2\text{e}^- + 2\text{H}^+ \rightarrow \text{NO}_2^- + \text{H}_2\text{O}$). Previous studies have suggested that there may be two different pathways for nitrate reduction involving either $\text{M}^{\text{V}}-\text{NO}_3$ (path1) or $\text{M}^{\text{IV}}-\text{NO}_3$ (path2) intermediates. In both cases, the metal oxidation state changes from four to six, which highly affects the electronic structure of the dithiolene ligand. The stability of

structures along these reaction pathways is largely determined by the degree of electron delocalisation (or aromaticity) in the metal-dithiolene cofactor. Computationally, a useful measure of aromaticity is through the effect of diatropic ring currents induced in a molecule by an external magnetic field. In this poster, we shall analyse the change the aromaticity of the metal dithiolene ring during the nitrate reduction reaction pathway using these methods.
