

POSTER PRESENTATIONS

Can Clefts Aid the Prediction of Protein-Protein Interfaces?

Nick Burgoyne and *Richard Jackson*

Institute of Molecular and Cellular Biology, Faculty of Biological Science, University of Leeds, Leeds, LS2 9JT

Protein-protein interactions are now seen as key drug targets, playing as important a pharmaceutical role as protein ligand interactions. The protein surface used to interact with other proteins is much harder to predict than protein interfaces that bind ligands. This is mainly due to the vast differences seen in the physical properties of interacting proteins. Existing protein-interface prediction methods attempt to score protein-interface 'likeness' using the knowledge of existing protein-complexes. It has been suggested that there are a finite number of interactions between protein folds, of which only a fraction are known [1]. Current protein-interface methods attempt to differentiate interacting from non-interacting patches based on the physical and chemical properties across the whole patch. However experimental evidence has shown that the stability of many protein complexes is mainly determined by only a fraction of their interface residues. These important residues form two groups, most form clefts, while the others bind within the clefts of the interacting protein. Prediction of these clefts is relevant to protein docking in that they undergo little deformation on association with the partner protein, and may indicate suitable areas to target when inhibiting the interaction.

Current work aims to assess the possibility of predicting protein-protein interfaces by identifying clefts across the protein surface and then ranking them according to different criteria. Clefts that are likely to be involved in an interaction are first identified using the ligand pocket detection program Q-SiteFinder [2]. They are then re-ranked according to properties known to be useful in protein stability and the prediction of protein interfaces (hydrophobicity, conservation, estimated desolvation energy of the cleft surfaces and the electrostatic field within the clefts). Results show there is some correlation when interface clefts are ranked by the desolvation of the site for all proteins. Other properties seem to correlate differently depending whether the complex represents an enzyme/inhibitor, antibody/antigen or other type of interaction. This work is in development and indications are that by clustering favourably ranked sites the results will improve.

References:

1. Aloy P and Russell R (2004) Ten Thousand Interactions for the Molecular Biologist, *Nature Biotechnology*, **22**, 1317-1321
2. Laurie ATR and Jackson RM (2005) Q-SiteFinder: an Energy Based Method for the Prediction of Protein-Ligand Binding Sites, *Bioinformatics*, **21**, 1908-1916

A Graph Theoretic Survey of RNA Base Triple Interactions

Mohd Firdaus Raih¹, Anne-Marie Harrison, *Peter Willett*², *Peter J. Artymiuk*¹

¹ *Krebs Institute for Biomolecular Research, Department of Molecular Biology and Biotechnology, University of Sheffield, Western Bank, Sheffield, S10 2TN and*

² *Department of Information Studies, University of Sheffield, Western Bank, Sheffield S10 2TN.*

RNA function is dependent on its folding into closely packed helices. Hydrogen bond interactions between RNA base residues play a major part in stabilizing these sometimes complex folds. Base triple interactions, particularly those with at least two hydrogen bonds to each base residue, are expected to be highly stable and therefore may have an important role in forming functional RNA tertiary structures. Our program, NASSAM, is a pattern matching algorithm using graph theory, that allows fast and efficient searching of databases of nucleic acid structures based on a simplified vectorial representation of the nucleic acid bases (Harrison *et al.* 2003). This method allows such a search to be carried out rapidly, efficiently and provides output to easily locate known or novel base triple interactions in RNA structure by using input search patterns consisting of either known or theoretical pattern formations. A set of filters were then used to further screen NASSAM output. The NASSAM searches yielded a number of novel base triple interactions. The process developed here will be a valuable method for studying models and mechanisms of RNA structural interactions which are formed from smaller subsets of base interactions such as base triples, quartets, quintets and the A-minor motifs.

References:

Harrison, A-M, South, D.R., Willett, P., and Artymiuk, P.J. (2003) Representation, searching and discovery of patterns of bases in complex RNA structures. *Journal of Computer-Aided Molecular Design*, **17**, 537-549.

Scoring Functions and Enrichment: A Case Study on Hsp 90

Chrysi Konstantinou Kirtay¹, John. B. O. Mitchell¹ and James. A. Lumley²

¹*Unilever Centre for Molecular Science Informatics, Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, UK*

²*Arrow Therapeutics Ltd, Britannia House, 7 Trinity Street, London SE1 1DA, UK*

The need for fast and accurate scoring functions has been driven by the increased use of virtual screening (*in silico*) twinned with high-throughput screening as a method to rapidly identify potential candidates in the early stages of drug development. We have been examining [1] the ability of the most common docking/scoring functions (**GOLD v2.2 GOLDscore**, **CHEMscore**, **DOCKscore**, **PMFscore**, **BLEEPscore** and **Consensus Scoring**) [2] to discriminate correctly and efficiently between active and non-active compounds among a library of ~3,600 diverse decoy compounds in a virtual screening against heat shock protein 90 (Hsp 90). We test the common assumption that the best-ranked pose selected during docking is a suitable pose for further scoring studies and show how consideration of multiple poses clearly

improves enrichment. Among the five scoring functions, **GOLDscore** (as implemented in **SYBYL 7.0**) appears to show the best trend in enrichment (85% actives within the first 40% of the screening library) based on the best-ranked **GOLD v2.2** pose. However, by considering the best score independently for each scoring function, with the exception of **PMFscore** which was consistently the poorest performer, we find that the enrichment is further improved (90% actives by screening 20% of the library). Overall we demonstrate the validity of virtual screening as a method for identifying new leads and we believe that the outcome of this study provides a useful insight into the setting up of an effective docking/scoring protocol, resulting in enrichment of 'hit' lists in the drug discovery process.

References:

[1] Konstantinou Kirtay, C., Mitchell, J.B.O. and Lumley, J.A., Knowledge Based Potentials: the Reverse Boltzmann Methodology, Virtual Screening and Molecular Weight Dependence, *QSAR & Combinatorial Science* (2005) **24**, 527-536.

[2] Marsden, P.M., Puvanendrapillai, D., Mitchell, J.B.O. and Glen, R.C., Predicting Protein-Ligand Binding Affinities: A Low Scoring Game? *Org. Biomol. Chem.* (2004) **2**, 3267-3273 and references therein.

QSAR Analysis of NK3 Activity

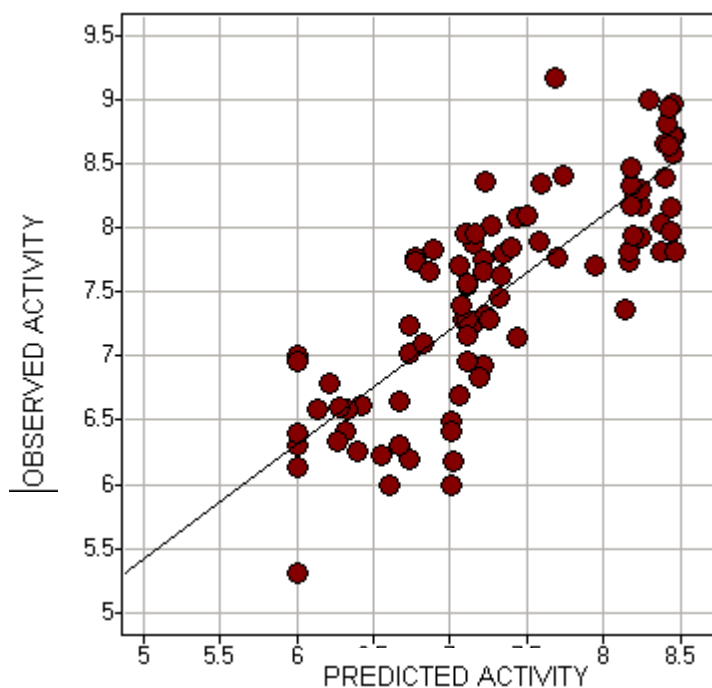
Natalie Akoma-Mordi, Christophe Buyck, Trevor Howe, Benoit De Boeck

Johnson & Johnson Pharmaceutical Research & Development, A division Janssen pharmaceutica, Turnhoutseweg 30, B-2340 Beerse, Belgium.

An extensive NK3 dataset containing 413 JNJ compounds was assembled. In an attempt to understand features for NK3 activity a retrospective QSAR study was undertaken.

The dataset and concomitant descriptors (1D & 2D) used for model construction were selected by design of experiment ¹ to ensure orthogonality and complete chemical representation. JADES ², an in-house neural network/genetic algorithm, determined the active or redundant role of each descriptor in the model building process. Based on the complexity and interpretability of the descriptors employed, two types of models were constructed; a chemically interpretable model and a generalized model using chemically 'uninterpretable' descriptors. Model validation techniques employed include Y-scrambling, cross-validation and test set predictions ³. Resulting models are able to explain some aspects of the NK3 pharmacophore.

Figure 1: Graph of a models prediction for the test set vs. the observed activity values.



References

- [1] Umetrics® http://www.umetrics.com/methodtech_doe.asp
 [2] Yasri, A., Hartsough, D., J. Chem Inf. Comput. Sci., 2001, **41**, 1218-1227.
 [3] Kubinyi, H., Quant. Struct.-Act. Relat., 2002, **21**, 348-356.

The Application of RAPPER to and it's Implications for the Limitations in Comparative Modelling

Nicholas Furnham, Paul I.W. de Bakker, Mark A. DePristo, David F. Burke, Tom L. Blundell

Department of Biochemistry, Cambridge University, Sanger Building, 80 Tennis Court Road, Cambridge, CB2 1GA

Limitations in experimental methods in determining the three dimensional structure of proteins has led to the development of a number of computational approaches to predict protein structure. One such class of methods is comparative modelling which uses information derived from homologous structures. Although comparative modelling is a relatively mature field, challenges still exist in producing reliable high quality models. Here we apply RAPPER, a novel restraint based conformational space search engine, to the problem and use it to address the limitations in comparative modelling.

RAPPER uses the principle of satisfaction of spatial restraints¹ derived from a structurally superimposed set of homologous structures. From this superimposition a number of different types of spatial restraints can be described for each residue to be built. For example main chain (principally C α atoms) can be defined as an ellipsoid generated from the union of the set of spherical restraints centred on the equivalent atom position from each of the templates. The restraint network is then solved by iteratively building residues from N to C termini using a branch-and-bound algorithm. Due to the way these restraints are described it is possible to generate an 'at-best'

scenario using non-native restraints (informed mode), analogous to the $\text{C}\alpha$ -trace problem² of regenerating native structure from just the native $\text{C}\alpha$ atoms.

A modelling data set was derived from the HOMSTRAD database of ten families to represent the four main SCOP classes of fold space. Each family has at least five constituent members, one that could be defined as the target, while the others have a range of percentage sequence identity to it with structures resolved to at least 2Å resolution. Models were generated using a number of combinations of different templates to the target. Using a range of indices it could be seen that models built on single templates were as good as the equivalent models generated by MODELLER³, the clear best in-class program for comparative modelling.

RAPPER's use of the available templates for models built using all the templates in comparison with the informed mode showed that if a more intelligent use of the templates could be used RAPPER could improve its performance. Further more the informed mode out performed MODELLER indicating there is a more general room for improvement in comparative modelling. In order to address this two preliminary filters to the restraint network based on differential geometry, pattern recognition and χ angle conservation was implemented. Models generated using this filtered network were much closer to the 'at-best' scenario. As we approach the theoretical limit of knowledge based methods attentions needs to focus on the final refinement to the native structure.

References:

1. Sali, A. & Blundell, T.L. Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol* **234**, 779-815 (1993).
2. DePristo, M.A., De Bakker, P.I., Shetty, R.P. & Blundell, T.L. Discrete restraint-based protein modeling and the $\text{C}\alpha$ -trace problem. *Protein Sci* **12**, 2032-46 (2003).
3. Fiser, A.S. & Sali, A. MODELLER: Generation and refinement of homology-based protein structure models. in *Macromolecular Crystallography, Pt D*, Vol. 374 461-+ (2003).

QM/MM Study of the Intramolecular Proton Transfer in Glycine

Phaedra A. Williams, Gus D. Ruggiero, Ian H. Williams

Department of Chemistry, University of Bath, Bath, UK, BA2 7AY

Many interesting chemical and biochemical processes in solution or in enzyme active sites involve the separation of opposite charges (or their neutralisation in the reverse reaction). Any methodology for realistic computational modelling of reactions in condensed media must be capable of treating the energetics of these charge-separation processes reliably. But how do we know how well any particular method performs? It is necessary to compare calculated and experimental energies for some standard reaction.

An attractive candidate for a standard reaction is the intramolecular proton transfer from the carboxyl to the amino group of neutral glycine $\text{HO}_2\text{CCH}_2\text{NH}_2$ to give the zwitterionic form ${}^-\text{O}_2\text{CCH}_2\text{NH}_3^+$. The Gibbs free energy for this process in aqueous solution is obtained simply from the difference in the $\text{p}K_a$ values of the two functional

groups, which has been reliably determined by experiment. It is well known that neutral glycine is more stable than the zwitterionic form in the gas phase, but that aqueous solvation reverses this preference. Owing to its small size, glycine has been the subject of numerous computational studies by a variety of methods.¹ Many studies report the relative energies of the several different conformers, but few draw attention to the computed values of the energy difference between the neutral and zwitterionic forms; perhaps this is because generally the results are very bad! A recent paper described an *ab initio* (Car-Parinello) molecular dynamics study, but this involved only 52 water molecules in the unit cell with periodic boundary conditions.²

We have used a hybrid QM/MM approach with the semiempirical AM1 method describing the glycine solute in a cubic box containing over 1000 TIP3P water molecules, to which periodic boundary conditions are applied. We have employed *NVT*-ensemble molecular dynamics with umbrella sampling to compute the potential of mean force along a reaction coordinate for intramolecular proton transfer, using the DYNAMO library. This approach yields the AM1/TIP3P differential free energy of solvation which, when added to the intrinsic gas-phase energy difference between the neutral and zwitterions forms computed at the MP2/6-311++G** level, provides a result in good agreement with experiment.

References:

1. I. Tuñón and E. Silla, *J. Phys. Chem. A* (1998) **102**, 8673-8678.
2. K. Leung and S.B. Rompe, *J. Chem. Phys.* (2005) **122**, 184506.

Regioselectivity of Soluble Epoxide Hydrolase: A QM/MM study

Simon Hoyle, *Dr A J Mulholland*

University of Bristol, School of Chemistry, Bristol, BS8 1TS

Soluble Epoxide Hydrolase (sEH) catalyses the addition of water to epoxide compounds yielding the corresponding diol compound¹. sEH has two main functions in biochemical systems. The first role is detoxification as many epoxide compounds are toxic²; its second is in chemical mediation³. Epoxide hydrolases can also be used to produce enantiopure epoxides for use in synthesis⁴. This Quantum Mechanics/Molecular Mechanics (QM/MM) study investigates the reasons behind the regioselectivity in sEH. With this model we can also predict the regioselectivity of the catalysis. Predictions are made using the geometric molecular similarity of the transition state to the Michaelis Complex (MC). This method is also compared to predictions using Near Attack Conformations.

Molecular Dynamics (MD) with Umbrella Sampling was used in conjunction with the Weighted Histogram Analysis Method (WHAM) to model the first step in the sEH reaction: the nucleophilic attack on the epoxide ring. Three enzyme/substrate complexes were considered; sEH with *trans*-Stilbene Oxide, and sEH with two conformations of *trans*-Diphenylpropene Oxide. Molecular similarity investigations required unrestrained simulations of the three sEH/substrate complexes.

It was found, from the MD/Umbrella Sampling simulations, that sEH does have a noticeable regioselectivity in two of the sEH/substrate complexes. This

regioselectivity was attributed to the stabilising contribution of the tyrosine 465 residue to the epoxide oxygen. Predictions of the regioselectivity, using the unrestrained simulations, confirmed the regioselectivity of sEH with the substrates used in the MD/Umbrella Sampling simulations.

References:

- [1] Armstrong, R. N. & Cassidy, C. S. New structural and chemical insight into the mechanism of the epoxide hydrolases. *Drug Metabolism Reviews*. (2000) **32**, 327-338.
 - [2] Wixtrom, R. N. & Hammock, B. D. *Biochemical Pharmacology and toxicology* (eds. Zakim, D. & Vessey, D. A.) (Wiley, New York, 1985).
 - [3] Moghaddam, M. F. et al. Bioactivation of leukotoxins to their toxic diols by epoxide hydrolase. *Nature Medicine*. (2001) **3** 562-566.
 - [4] Archelas, A. & Furstoss, R. Synthetic applications of epoxide hydrolases. *Current Opinion in Chemical Biology*. (2001) **5** 772-119.
-

Molecular Dynamics Simulations of LDAO detergent micelles and of Mystic

Emi Psachoulia, P.J. Bond, Mark S.P. Sansom

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

Mistic is a small bacterial integral membrane protein from *Bacillus subtilis*. It is an unusual protein as it folds and inserts into the lipid bilayer independently of the targeting machinery that in general cells use to insert proteins into membranes. Its structure was determined by NMR spectroscopy [ref. 1]. It was solubilized in lauryl dimethylamine oxide (LDAO) forming a protein-detergent complex (micelle) containing ~50 molecules of LDAO. In this project, the conformational stability and the behaviour of Mystic in a detergent micelle environment was studied via molecular dynamics simulations. Moreover, LDAO molecules were modelled and used to study their ability to produce a stable micelle. For all simulations, the GROMACS package (www.gromacs.org) was used.

Firstly, two simulation systems were set up with LDAO detergents only: one with 50 LDAO molecules randomly placed in a simulation box (self-assembled) and one with a 50-LDAO preformed micelle [ref. 2,3], both solvated in water. The simulations ran for 10ns. In the self-assembled simulation, the LDAO detergents formed a micelle and the analysis showed that both micelles were equal in size and shape and that both simulations reached equilibrium in the end of the 20ns.

Then, 3 more simulations were set up: one with Mystic and a 50-LDAO preformed micelle, one with Mystic and 50 LDAO detergents randomly placed in a simulation box and one with Mystic and 78 (~the aggregation number, ref. 4) LDAO detergents randomly placed in a simulation box. All simulations ran for 20ns-just finished. Preliminary analysis showed that the protein is very mobile. All simulations were extended for 10ns. The simulations will be analysed in more detailed.

Moreover, some more simulations will be set up, for example Mystic in different bilayer environments, where it can be observed how Mystic inserts in the bilayer.

1. Roosild, T.P.; Greenwald, J.; Vega, M.; Riek, R.; Choe, S., NMR structure of Mystic, a membrane-integrating protein for membrane protein expression, *SCIENCE*, 2005, 301, 1317-1321
2. Bond, P.J.; Sansom, M.S.P., Membrane protein dynamics versus environment: simulations of OmpA in a micelle and in bilayer, *J. Mol.Biol.*, 2003, 329, 1035-1053
3. Bond, P.J.; Cuthbertson, J.M.; Deol, S.S., Sansom, M.S.P., MD simulations of spontaneous membrane protein/detergent micelle formation, *JACS*, 2004, 126, 15948-15949
4. Herrmann, K.W.J., *Colloid Interface Sci.* 22, 352 (1996)

QSPR investigation into solvation properties of platinum complexes

Steven P. Oldfield, Maria M. Reif, *James A. Platts*

School of Chemistry, Cardiff University, Park Place, Cardiff, CF10 3AT, UK

Cisplatin is the archetypal metal-based DNA drug with more than \$500M annual worldwide sales. However, it has serious problems in terms of its toxicity, limited spectrum of activity and resistance; prompting the intense search for better compounds. One of the major causes of cisplatin resistance is reduced cellular uptake, and drugs with increased lipophilicity have been shown to cross cell membranes more effectively¹.

Here we present mathematical models (QSPR relations) to predict partition coefficients, logP(octanol/water) values, of cisplatin analogues. The molecular electrostatic potential (MEP), $V(r)$ created around isolated molecules has been found by Politzer² to describe the ability to form non-covalent interactions. We therefore calculated a range of Politzer-type surface descriptors for optimised structures at the B3LYP/6-31+G(d,p) (cep-121G for Pt) level, giving a model with R^2 of 0.89 (RMSE 0.34).

Ionisation energy is closely connected to both the chemical potential and the polarisability³ and is intimately related to common concepts like hardness, softness and electronegativity which may also affect solubility. We therefore calculated descriptors based on the surface mapped with ionization potential, giving a model with R^2 of 0.92 (RMSE 0.34).

In conclusion, it is known that the initial phases of biological recognition are mediated by non-covalent interactions and are dominated by electrostatic forces. This current study shows the importance of ESP measures (which encode long-range coulombic forces) in modelling solvation. A description of polarisability is also needed, as this accounts for shorter-range dispersion-type forces. Improvement over the quantum mechanical polarisability is achieved using IP surface-type descriptors. This work should have great significance to the pharmaceutical industry allowing cheap in silico screening of potential unsynthesised platinum compounds using calculated molecular properties alone.

References:

1. Conradi R.A., Burton P.S., Borchardt R.T.; '*Lipophilicity in Drug Action and Toxicology*'; ed. Pliska V., Testa B., Vanderwaterbeemd H.; (1996); Weinheim; p233
 2. Politzer P., Truhlar D.G.; '*Chemical Applications of Atomic and Molecular Electrostatic Potentials. Reactivity, Structure, Scattering and Energetics of Organic, Inorganic and Biological Systems*'; Plenum Press; New York; (1981)
 3. Fricke; '*On the correlation between electric polarisabilities and the ionisation potential of atoms*'; *J. Chem. Phys.* **84** (1986), p2554
-

Docking and direct design in the binding pocket -libraries for serine protease inhibitors

Gerlach, C.¹, Steuber H.M.¹, Velec, H.F.G¹, Smolinski, M.², Hangauer, D.² Heine, A.¹ and Klebe, G¹

¹Philipps-Universität Marburg, DE-35032 Marburg, Germany

²Department of Chemistry, State University of New York at Buffalo, Buffalo NY 14260-3000, USA

The successful design of potential inhibitors by means of structure-based drug design strongly relies on the correct prediction of the putative binding geometries along with a reliable estimation of the binding affinity. Hence, the success of any virtual screening and automated ligand or library design in the computer for modern drug discovery depends intimately on this crucial step and requires rigorous validation.

To assess the predictive power of our docking and library design tools we selected trypsin and thrombin as well-established model cases. In particular in combination with the thermodynamic analysis of ligand binding through isothermal titration calorimetry we try to factorize binding properties in enthalpic and entropic contributions. Such data will help to select the most appropriate strategy for lead optimization.

We start with a privileged ligand scaffold well-suited to address the key interactions of the conserved recognition pattern shared by the members of the serine protease family. Binding modes of ligands with such a privileged scaffold are generated by computer docking. The obtained solutions are ranked using different scoring functions recently developed in our group [1]. In a consecutive step, the best ranked solutions are further optimized by means of local search techniques based on knowledge-based potentials derived from the CSD [2]. Intermolecular interactions are modeled by distance-dependent pair potentials whereas torsional energies are represented by empirical potentials as implemented into MIMUMBA [3].

Subsequently for individual members of the series the crystal structure in complex with trypsin and thrombin as well as the binding affinity with respect to both enzymes are determined. Facing the obtained experimental data with the computer-generated

binding geometries enables us to assess the accuracy of our predictions. In a proximate step, based on the experimentally determined binding geometry of the privileged scaffold and searches by means of docking routines for novel side-chain decorations to reveal ligands with optimal complementarity with the target protein a targeted library of potential leads is generated. Through syntheses, enzymatic assay and crystal structure analysis the most promising library members are further characterized. Based on this thorough investigation, the design of a structure-based combinatorial library directed towards highly selective thrombin inhibitors is currently performed in our laboratories.

References:

- [1] Velec, H.F.G., Gohlke, H., Klebe, G., DrugScore^(CSD)-Knowledge-Based Scoring Function Derived from Small Molecule Crystal Data with Superior Recognition Rate of Near-Native Ligand Poses and Better Affinity Prediction, *J. Med. Chem.* (2005) **48(20)**, 6296-6303
 - [2] Allen F.H., The Cambridge Structural Database: a quarter of a million crystal structures and rising, *Acta Cryst.* (2002), **B58**, 380-388
 - [3] Klebe G., Mietzner T., A fast and efficient method to generate biologically relevant conformations, *J. Comput Aided Mol Des* (1994) **8**, 583-606
-

QM/MM modelling of reactions in citrate synthase

Marc van der Kamp¹, Francesca Perruccio², Adrian Mulholland¹

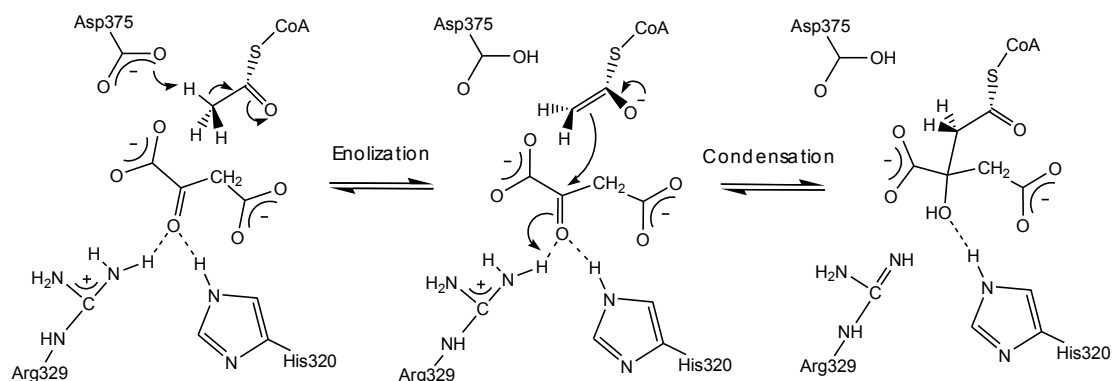
¹*School of Chemistry, University of Bristol, UK*

²*Current address: Medicinal Informatics, Structure and Design, Pfizer Global R&D, UK.*

Citrate synthase catalyses the first step in the citric acid cycle: the conversion of oxaloacetate to citrate. The first reaction in this conversion is the Claisen condensation of acetyl-CoA with the carbonyl of oxaloacetate. This reaction comprises two steps: proton abstraction from acetyl-CoA to form an enolate (enolization) and the subsequent nucleophilic attack of the enolate on the carbonyl carbon of oxaloacetate (condensation). An important issue is the stabilization of the enolate by the enzyme, which is thought to be essential for the enzymatic catalysis. Hydrogen bonding in the active site has shown to be important for this stabilization. Free energy profiles of the enolization reaction, obtained by QM/MM (AM1/CHARMM27) umbrella sampling simulations, show that the polarization of the oxaloacetate carbonyl, important for the nucleophilic attack, also plays a role in the stabilization of the enolate.

Comparison of AM1/CHARMM27 and B3LYP/6-31+G*/CHARMM27 potential energy profiles of the enolization reaction however, indicate that the AM1 method significantly overestimates the stability of the enolate. Specific AM1 parameters were adjusted to get relative energy values closer to B3LYP/6-31+G* results. Including a single point energy correction, the AM1/CHARMM27 profile using these specific reaction parameters (AM1-SRP) is almost identical to the B3LYP/6-31+G*/CHARMM27 profile. Preliminary QM/MM simulations on the next step of the condensation reaction, nucleophilic attack on the oxaloacetate carbonyl carbon,

show that this step is most likely to occur concerted with proton transfer from an arginine residue to the carbonyl oxygen.



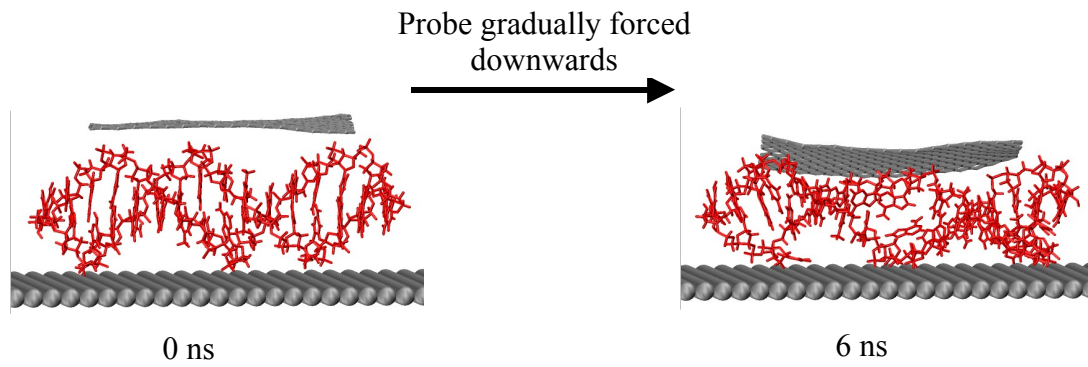
The enolization and condensation reaction steps catalysed by citrate synthase.

Molecular Modelling of DNA Compression Caused by Physical Probing Techniques

Verity G Hudson and Charles A Laughton

*School of Pharmacy, Centre for Biomolecular Sciences, University of Nottingham
NG7 2RD*

The Atomic Force Microscope (AFM) is a surface analysis tool that is now well established in the biomolecular field. DNA has been particularly well studied by this technique as it can be easily immobilised on a surface and imaged whilst still in solution, representing near biological conditions (Hansma et al, 2004). The AFM consists of a tiny probe mounted on a flexible cantilever. The probe is scanned across the sample surface and the deflections of the cantilever are recorded, providing a topographical image of the surface. Single DNA molecules on a surface are resolved and this method has been used to observe DNA supercoiling, DNA microcircles and DNA-protein complexes to name a few. However, there is still some uncertainty surrounding the interpretation of AFM data. For example it is believed that soft biological molecules are likely to be compressed by the tip during imaging but there has been little theoretical investigation of this. The evidence for compression comes from the fact that DNA molecules always appear much flatter in AFM images. In this project we use molecular modelling to study the radial compressibility of DNA under mechanical forces to aid in the interpretation of AFM experiments. We have performed a fully atomistic steered Molecular Dynamics simulation of DNA being compressed by an AFM tip. We have also run extended simulations at fixed points along the simulation path to extract the entropy using the method of Schlitter (Schlitter, 1993). This has enabled us to plot a free energy curve for the pathway. This shows that there is a small free energy barrier to compression up to a compression depth of $\sim 2\text{\AA}$, and that after this barrier is overcome the DNA deforms very easily.



References:

Hansma H G, Kasuya K and Oroudjev E. *Current Opinion in Structural Biology* (2004) **14**, 380-385.

Schlitter J. J. *Chem. Phys. Lett.* (1993) **215**, 617-621.
