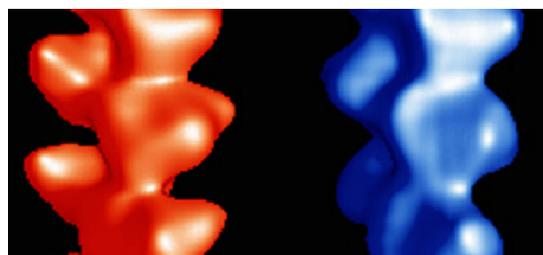
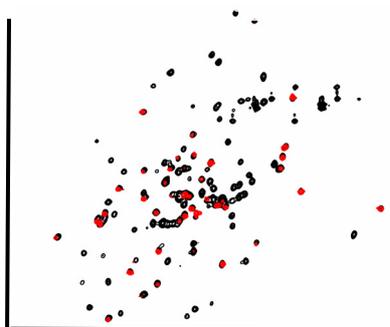
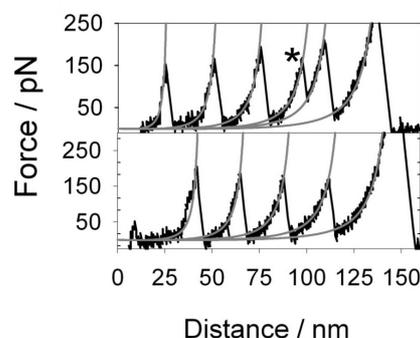
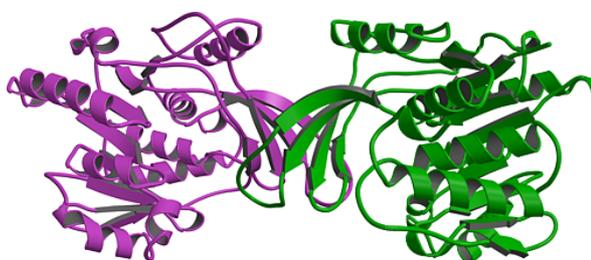


Astbury Centre for Structural Molecular Biology

University of Leeds



Annual Report

2003



Front cover illustration: A collage of pictures illustrating the work of the Astbury Centre. Upper left: The X-ray crystal structure of human ketohexokinase solved at 1.8Å resolution (see page 69); Upper right: Typical data acquired during the mechanical unfolding of a protein using the atomic force microscope (see page 82); Lower right: Three dimensional reconstructions of filaments of arthrin (red) and actin (blue) from cryo-electron microscopy (see page 111); Lower left: Part of the HSQC NMR spectrum of the VMA-7 subunit of the *Saccharomyces cerevisiae* vacuolar H⁺-ATPase (see page 24).

Acknowledgement

The Astbury Centre for Structural Molecular Biology thanks its many sponsors for support of the work and its members for writing these reports. The report is edited by Alan Berry.

This report is also available electronically via <http://www.astbury.leeds.ac.uk>

Mission Statement

The Astbury Centre for Structural Molecular Biology will promote interdisciplinary research of the highest standard on the structure and function of biological molecules, biomolecular assemblies and complexes using physico-chemical, molecular biological and computational approaches.

Introduction

This has been a great year for the Astbury Centre for Structural Molecular Biology (ACSMB). The new facilities supported by the JIF funding from the Wellcome Trust, allowing us to open the Wellcome Trust Centre for Biomolecular Interactions, are finally fully up and running thanks to the efforts of very many staff within ACSMB and in the wider University. Many thanks and congratulations to all those involved. The facilities include state-of-the-art instruments for molecular synthesis and computational analysis in the School of Chemistry; biomolecule production, purification and characterisation in the Faculty of Biological Sciences, and extensive equipment for analysing molecular interactions; the basis of complexity in all living organisms. Our structure determination facilities have also received a major boost with the installation and commissioning of a FEG electron microscope and a 750 MHz NMR spectrometer. As well as excellent facilities, we were also able to get our Wellcome 4 Year PhD programme extended for another four years. Well done to Steve Homans and Alan Berry for their efforts on ACSMB's behalf in this area of our activities. We also welcome Prof Jim Hogle from Harvard Medical School as our first Leverhulme Trust funded sabbatical visitor. Jim's pioneering contributions to the structure and function of polio virus are well known and I'm sure he'll help to stimulate activities amongst many project areas within the Centre during his time here.

In addition to these developments, Astbury PIs secured >£7M in additional grant support over the year and the attached document, containing 63 individual Annual Reports covering work from 28 laboratories, is the largest in our history, reflecting the fact that our staff published ~100 refereed papers in the year (see list at the end of this Report). Never a group to rest on its laurels, major plans were also laid this year to form two University Interdisciplinary Institutes under the Astbury umbrella in Molecular Biophysics and Bionanosciences. The former will be based around the work of Alastair Smith, Sheena Radford and colleagues in protein folding. A wonderful example of the innovative experiments in this area is provided by Dave Brockwell's pioneering work in the mechanical unfolding of proteins, which has attracted international attention. The latter UII will foster the application of a "bottom-up" approach to the manipulation of matter on a molecular scale by developing ways to subvert existing biological complexes, such as virus-like particles, molecular motors etc. to accomplish useful work outside of their natural cellular environments. Funding of ~£5M for these new activities will come from the University. In addition, ACSMB staff (Steve Baldwin, Simon Phillips, Mike McPherson, John Trinick, Steve Homans & Peter Henderson) are part of major UK consortia in bids to the BBSRC Structural Genomics competition. I wish them every success in these endeavours. The Centre continues to host a very successful seminar programme. I would like to thank our industrial sponsors who have contributed financial support over the year. Their logos are shown on page 120 of the Report.

The pages that follow describe some of the highlights of our work over the last year. These reports have largely been written by our younger researchers. Their tremendous enthusiasm for this kind of interdisciplinary work augurs well for our future. As always I am particularly struck by the breadth of activity in the Centre ranging from the sophisticated applications of synthetic organic chemistry to the developments in single molecule biophysics. In between these extremes you will find groundbreaking activity in many traditional areas for structural biology. ACSMB has always been outward looking and this tradition continues with the many external collaborations acknowledged in these pages, from both within the UK and beyond. We would welcome discussions with anyone wishing to collaborate or simply to make use of our facilities, the details of which can be found on our web page (<http://www.astbury.leeds.ac.uk/>).

These brief summaries, however, only scratch the surface of the work of the Centre. I hope you enjoy reading them, and if you wish to learn more please visit our website or contact the Director. This annual report is also available as a ~11.5MB PDF document that can be downloaded from our web site.

Peter G. Stockley

Director, Astbury Centre for Structural Molecular Biology

Leeds, May 2004

Contents

	Pages
Ashcroft, A.	Mass spectrometry facility 1-3
	Investigating the structure of monomeric and fibrillar β_2m by limited proteolysis and mass spectrometry 4-5
Baron, A. J.	Analytical centrifuge facility 6-7
Berry, A.	Enzyme engineering of sialic acid synthesising enzymes 8-9
	Directed evolution of dihydrodipicolinate synthase (DHDPS) 10-11
	Directed evolution of an aldolase to enhance solubility 12-13
Findlay, J.B.C.	Structure/function of G-protein coupled receptors (GPCRs) 14-16
Harris, M.	Functional studies on the HIV-1 Nef protein. 17
	Structural and functional studies on the Hepatitis C virus non-structural NS5A protein. 18-19
Homans, S. W.	Thermodynamics of binding of 2-methoxy-3-isopropylpyrazine and 2-methoxy-3-isobutylpyrazine to the major urinary protein 20-21
	Dissecting the cholera toxin-ganglioside GM1 interaction by isothermal titration calorimetry 22-23
	The NMR solution structure of the VMA-7 subunit of the vacuolar H^+ -ATPase from <i>Saccharomyces cerevisiae</i> 24-25
	Large scale millisecond inter-subunit dynamics in the B subunit homopentamer of the toxin derived from <i>E. coli</i> O157 26-27
Henderson, P. J. F.	Collecting bacterial membrane proteins 28-30

Henderson, P.J.F.H.	Characterisation of ligand binding to membrane transport proteins from <i>Escherichia coli</i> by solid state NMR	31-33
Hiscox, J.	Characterization of the coronavirus RNA binding protein	34-35
Jackson, R.	Molecular modelling of SH2 domain-peptide interactions	36-37
	Flexligdock: A flexible ligand docking tool	38-39
	Searchable database containing comparisons of ligand binding sites at the molecular level for the discovery of similarities in protein function	40-41
	Structural modelling of protein-DNA interactions	42-43
Jäger, J	Structural studies of RNA-directed RNA polymerases from Hepatitis C virus and related viruses	44
	Biochemical and structural studies of Human Rhinovirus 3D polymerase	45
	Functional and structural analysis of motif F in viral RNA polymerases	46-47
	The mechanism of initiation in Hepatitis C virus RNA polymerase	48-49
	Structural and biophysical studies on HIV-1 Nef	50-51
Kalverda, A.	The NMR facility	52-53
McDowall, K. J.	Quaternary structure and catalytic activity of the <i>Escherichia coli</i> ribonuclease E amino-terminal catalytic domain	54-55
	Expanding the use of zymography by the chemical linkage of small, defined substrates to the gel matrix	56-57
McPherson, M. J.	Enhanced enzyme properties by protein engineering of galactose oxidase	58-59
	Mobile TPQ cofactor and metals in <i>E.coli</i> copper amine oxidase	60
Nelson, A. S	Synthesis of screening substrates for the directed evolution of sialic acid aldolase	61

Parker, M.J.	A combinatorial selective labelling method for the assignment of backbone NMR resonances	62-64
	Water as a conformational editor in protein folding	65-66
Phillips, S. E. V.	Crystal structures of Q β RNA stemloop operators complexed with bacteriophage MS2 coat protein mutants; towards an explanation of binding discrimination.	67-68
	Structural determination of human ketohexokinase	69-70
Phillips-Jones, M.K.	Cytolysin gene expression in <i>Enterococcus faecalis</i> is regulated in response to aerobiosis conditions	71-72
	Expression, purification and characterisation of full-length histidine protein kinase RegB (PrrB) from <i>Rhodobacter sphaeroides</i>	73-74
	Solution structure and DNA binding of the effector domain from the global regulator PrrA (RegA) from <i>Rhodobacter sphaeroides</i> : Insights into DNA binding specificity	75-77
Radford, S. E.	Amyloid formation of peptides and point mutants of β_2 -microglobulin	78-79
	Investigating amyloid self-assembly by atomic force microscopy	80-81
	Characterising the factors that describe the mechanical resistance of proteins	82-83
	Mapping the folding landscapes of the four-helix proteins Im7 and Im9	84-85
Ranson, N.	Cryo-EM of macromolecular complexes	86
Rowlands, D.J.	Targeting the Hepatitis C virus ion channel p7 for anti-viral therapy	87-88
	Discrimination among rhinovirus serotypes for a variant ICAM-1 receptor molecule	89-90
Smith, D. A.	Atomic force microscopy of microdomains in lipid bilayers	91-92
	The dynamic response of biomolecules	93-95

Smith, D. A.	Fast folding of B domain protein A by laser induced temperature jump	96-97
	Single molecule fluorescence spectroscopy of RNA hairpin loop folding	98
Stonehouse, N. J.	Investigating the molecular interactions of the ϕ 29 packaging motor	99
	Using <i>in vitro</i> selection to investigate RNA-protein interactions in picornaviruses	100-101
Thomas, C. D.	High throughput techniques as a strategy to overcome crystal twinning	102
	Molecular mechanism of Staphylococcal plasmid transfer	103
	Investigating the interaction of quinolones with topoisomerase IV of <i>Staphylococcus aureus</i>	104
	Dimer specificity in a replication initiator protein	105
Thomson, N.	Atomic force microscopy of DNA-protein interactions	106-107
Trinick, J.	Class VI myosin	108-109
	Titin: studies of synthetic end-filaments	110
	EM image processing of filamentous structures – arthrin	111
	New apparatus for time-resolved electron cryo-microscopy	112
Warriner, S.L.	Application of the Tmb acyl transfer auxiliary to the total chemical synthesis of Im7	113-114
Westhead, D. R.	Design and implementation of a database to archive and aid the analysis of tissue microarray data	115-116
Whitehouse, A.	Herpesviral-host cell interactions which regulate viral gene expression	117-118
Astbury Seminars 2003		119-120
Publications		121-128

Mass spectrometry facility

Alison E. Ashcroft

Overview of the facility

The Mass Spectrometry (MS) Facility has a Q-ToF orthogonal acceleration quadrupole-time-of-flight tandem instrument with nano-electrospray ionisation (ESI) and on-line capillary HPLC, a Platform II (ESI) quadrupole instrument with on-line HPLC and CE, a TSQ 7000 ESI tandem quadrupole instrument, and a surface enhanced laser desorption ionisation/matrix assisted laser desorption ionisation (SELDI/MALDI) ProteinChip mass spectrometer. There is also a MALDI instrument specifically for high-throughput proteomics screening in the department. The Facility runs an analytical service as well as being actively involved in several research areas within the Astbury Centre for Structural Molecular Biology and the Faculty of Biological Sciences, and also with other groups and external collaborators (see publications).

Research

Research in the facility involves the application of MS to the structural elucidation of biomolecules:

i) Protein folding. Protein folding is an intriguing area of biochemistry and protein misfolding is thought to be a contributing factor to several diseases. Working with Sheena Radford's group, ESI-MS is being used to monitor β_2 -microglobulin conformations using charge state distribution analysis, enzymatic digestion and H/D exchange to gain insights into folding intermediates.

ii) Protein-ligand non-covalent interactions and macromolecular assembly. In collaboration with Peter Stockley, Sheena Radford and Nicola Stonehouse, ESI-MS is being used to investigate non-covalently bound macromolecular structures. Such studies include protein-peptide, protein-protein, and protein-RNA complexes. The latter are important in virus assembly, an area we are investigating with respect to the MS2 and Q β systems. Protein-protein macromolecular complexes are critical species in fibrillogenesis and are under investigation as an integral part of our β_2 -microglobulin amyloid studies.

iii) Reaction monitoring. We have used MS to measure the uptake of ATP by the muscle protein myosin (with Howard White (Eastern Virginia Medical School, USA) and John Trinick). Myosin has a motor domain that interacts with filaments of F-actin and splits ATP to generate force and movement (Fig. 1). To increase our understanding of how the free

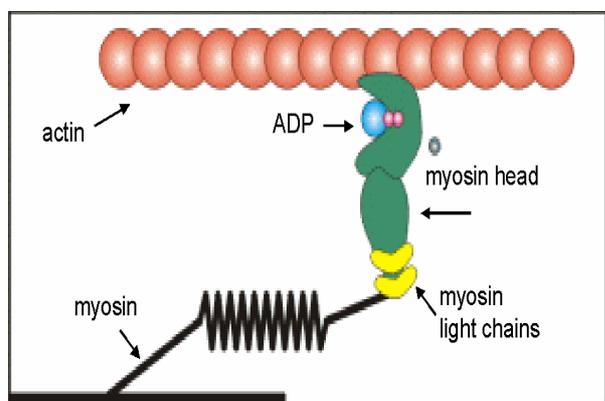


Fig 1. Schematic diagram of myosin and actin in muscle

energy of ATP hydrolysis is coupled to the production of mechanical work in the muscle, we have measured the non-covalent binding of ATP and ADP to the active site of rabbit skeletal myosin-S1 (heavy chain protein mass 92,382 Da) (Fig. 2). We have shown that conversion of the nucleotide complex containing an ATP-myosin-S1 complex to one containing an ADP-myosin-S1 complex occurs at a rate consistent with that of the known steady state rate of ATP hydrolysis.

iv) Structural elucidation and proteomics. Tandem MS (MS/MS) sequencing of proteins and peptides is an important bioanalytical technique. Several proteomics-related projects are in

progress, including the identification of post-translational modifications of kinesin by the generation of mass maps and MS/MS sequence tags from 2D-gel digests (Andy Grierson, University of Sheffield), the functional analysis of proneuropeptide genes from the *Drosophila* genome (with Elwyn Isaac, Biology), and an investigation into parasite manipulation of host sex (with Alison Dunn, Biology). Other structural analysis projects include the characterisation of lipopolysaccharides (with Deidre Devine, Oral Biology).

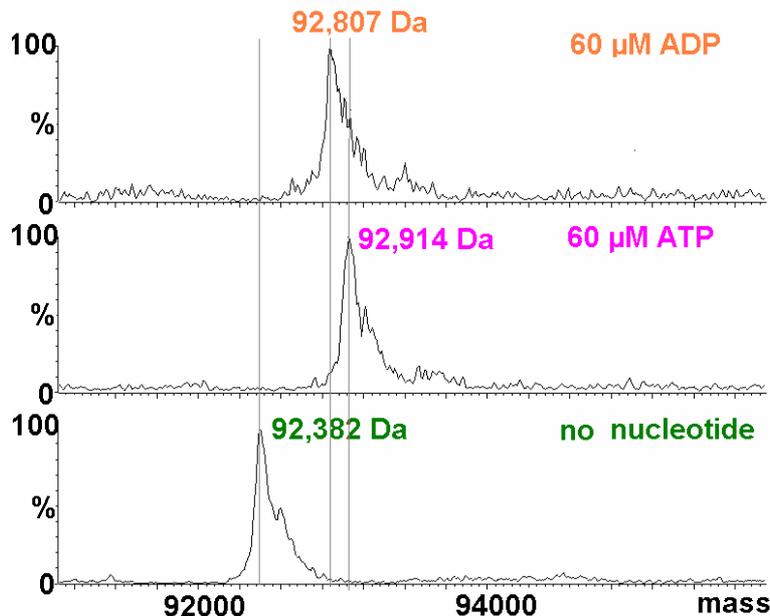


Fig 2. Mass profiles of the heavy chain of myosin-S1 (lower trace) during and after steady state hydrolysis of ATP (60 μ M) showing ATP addition (middle trace) and ADP addition (upper trace).

Publications

Meky, F.A., Turner, P.C., Ashcroft, A.E., Miller, J.D., Qiao, Y-L., Roth, M.J. & Wild, C.P. (2003) Development of a urinary bio-marker of human exposure to deoxynivalenol. *Food & Chemical Toxicology*, **41**, 265-273.

Rodgers-Gray, T.P., Ashcroft, A.E., Smith, J.E. & Dunn, A.M. (2004) Mechanisms of parasite-induced sex reversal in *Gammarus duebeni*. *Int. J. Parasitology*, in press.

Pearson, A.R., Jones, L.H., Higgins, L.A., Ashcroft, A.E., Wilmot, C.M. & Davidson, V.L. (2003) Understanding quinone cofactor biogenesis in methylamine dehydrogenase through novel cofactor generation. *Biochemistry*, **42**, 3224-3230.

Korchazhkina, O.V., Ashcroft, A.E., Croom, J. & Exley, C. (2003) Does either the gastrointestinal peptide PYY or the neuropeptide Y bind aluminium? *J. Inorg. Biochemistry*, **94**, 372-380.

Bailey, S., Sedelnikova, S.E., Mesa, P., Ayora, S., Waltho, J.P., Ashcroft, A.E., Baron, A.J., Alonso, J.C. & Rafferty, J.B. (2003) Structural analysis of *Bacillus subtilis* SPP1 phage helicase loader protein G39P. *J. Biol. Chem*, **278**, 15304-15312.

Ashcroft, A.E., Trinick, J.A. & White, H.D. (2003) Measurement of myosin nucleotide complexes by mass spectrometry, Biophysical Society, 47th Annual Meeting, 1-5th March 2003, San Antonio, Texas, USA.

Ashcroft, A.E. Protein and peptide identification: the role of mass spectrometry in proteomics (2003) *Natural Product Reports*, **20**, 202-215.

Group Members

Antoni Borysic (with Sheena Radford),
Simona Francese (with Nicola Stonehouse & Peter Stockley),
Michelle Morgan (with Deidre Devine),
Andrew Smith (with Sheena Radford & Peter Stockley).

Funding

Financial support from the University of Leeds, the Wellcome Trust, BBSRC, the European Commission, Micromass UK Ltd/Waters, AstraZeneca and Pfizer is gratefully acknowledged.

Investigating the structure of monomeric and fibrillar β_2m by limited proteolysis and mass spectrometry

Toni J.H. Borysik, Sarah L. Myers, Andrew M. Smith,
Sheena E. Radford and Alison E. Ashcroft

Introduction

Beta 2-microglobulin (β_2m) is a small protein, 11.8kDa in mass and 99 amino acids in length, which readily forms amyloid-like fibrils in acidic conditions (pH 2.5) *in vitro* (Fig. 1). *In vivo*, β_2m is found within the plasma at low concentrations. However, in patients suffering from chronic kidney disease the concentration of β_2m can increase up to 60-fold, and this can lead to the formation and deposition of amyloid fibrils within collagen-rich areas of the body, such as the joints, which is marked by chronic inflammation. The mechanism by which β_2m forms amyloid fibrils *in vivo* at a neutral pH where the protein is in its native state fold is currently unknown. We are using a number of mass spectrometric techniques to study the structure and function of β_2m . Using monomeric recombinant β_2m , we have formed amyloid-like fibrils *in vitro* under acidic conditions (pH 2.5) by firstly populating an acid unfolded state, which is a precursor to fibril formation. Limited proteolysis with pepsin was carried out on the acid unfolded monomer and the amyloid-like fibril, as well as seven short, synthetic peptides, which correspond to the seven beta strands comprising the beta sandwich fold of native β_2m . These digests allowed us to determine which protected sites were constant in the monomer, fibril and peptides, and then to map the conformational changes that occur between these states. The digest products were analysed by electrospray ionisation (ESI) tandem mass spectrometry (MS/MS) to generate peptide sequence information so that the protein fragments could be identified unambiguously (Fig. 2).

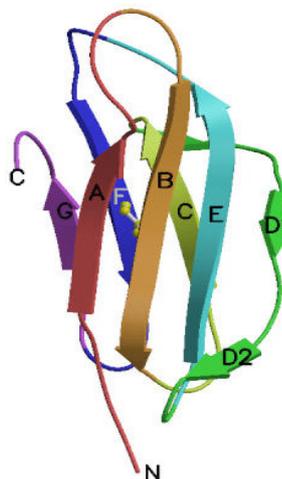
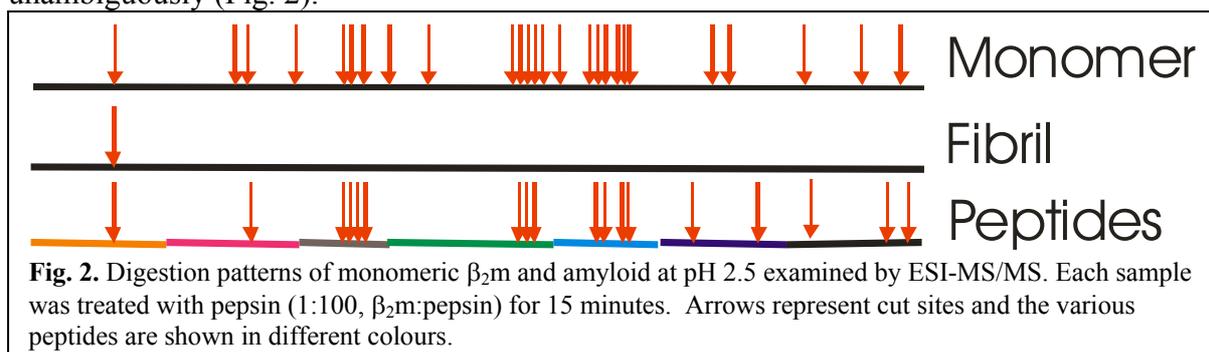


Fig. 1 Ribbon diagram of β_2m



Further experiments using limited proteolysis with different enzymes are underway, and the products will be analysed by MS/MS to give more clues into which key areas of the β_2m sequence and structure are involved in amyloid fibril formation.

Mass spectrometry (MS) is normally a non-quantitative method of protein analysis as it is not possible to relate the observed peak heights to protein concentration. However, we have used an internal standard of known concentration to create a calibration curve across a concentration range for a protein of interest. In the process of amyloid formation, monomeric β_2m undergoes oligomerization, and we have used ESI-MS to follow the loss of monomeric β_2m over time using bradykinin as an internal standard. We have also followed the

appearance of β_2m oligomers using this technique. Data from these experiments have revealed a rapid decrease in monomeric β_2m in conjunction with amyloid formation.

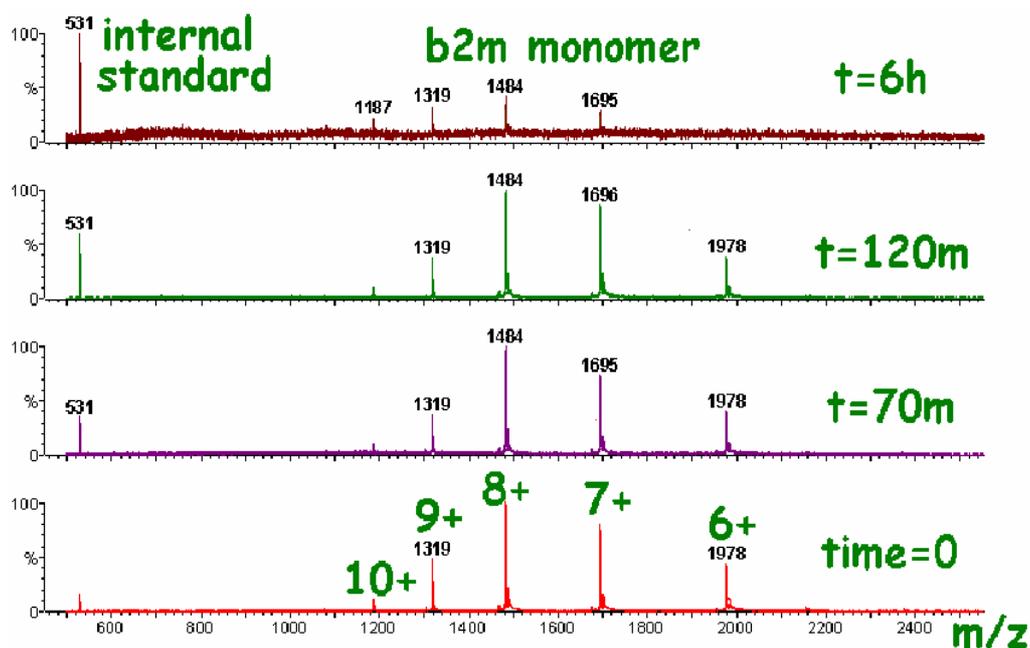


Fig. 3 Time course of amyloid formation monitored using ESI-MS to detect changes in the ratio of the internal standard bradykinin to β_2m over time, as monomeric β_2m forms oligomers.

ESI-MS has also been used to examine the binding of small molecules to β_2m . These molecules could potentially act as inhibitors of amyloid assembly and hence be of great therapeutic value.

Below pH 5.0, β_2m populates a number of different conformational states *in vitro*. We have developed a method to resolve and quantify these co-populated species by applying linear deconvolution methods to the ESI-MS m/z spectra acquired over the pH range 6.0 to 2.0. The method has been tested successfully with mutants of β_2m , and provides the foundation for the identification of the role of each species in the generation of amyloid fibrils. We are now working on novel mass spectrometric methods using ion mobility spectrometry to develop this concept further.

Publications

Borysik, A.J.H., Radford, S.E. & Ashcroft, A.E. (2004) Co-populated conformational ensembles of β_2 -microglobulin uncovered quantitatively by ESI-MS, submitted to J. Biol. Chem..

Ashcroft, A.E., Borysik, A.J.H., Radford, S.E., Read, P., Little, D.R. & Bateman, R.H. (2004) The use of high-field asymmetric waveform ion mobility coupled to ESI-MS to separate protein conformers, in "Advances in Mass Spectrometry 16" Proceedings of the 16th International Mass Spectrometry Conference, Edinburgh, UK, Eds. Ashcroft, A.E., Brenton, A.G & Monaghan, J.J.

Funding

Financial support from the University of Leeds, the Wellcome Trust, BBSRC, Micromass UK Ltd./Waters, AstraZeneca and Pfizer is gratefully acknowledged. SER is a BBSRC Professorial Fellow.

Analytical centrifuge facility

Andy Baron and Peter Stockley

Introduction

The Centre now has two Beckman XL-I analytical ultracentrifuges installed in the Wellcome Trust JIF Centre in Biomolecular Interactions. Both instruments are equipped with absorbance and interference optics, 4-place, and 8-place rotors, and velocity and equilibrium cells with a choice of quartz or sapphire windows. We employ a range of data analysis methods, enabling the determination of properties of macromolecules in free solution including species distribution, mass, degree of asymmetry, and association constants of interacting species.

Work done in 2003

Installation of the second instrument in April enabled more flexible use of the system, including exclusive use of one instrument for a visiting scientist, performance of hands-on experiments by postgraduate students, and simultaneous analysis of more than 3 samples of time sensitive material where interference optics could not be used. In addition, it has been possible to do more of the time-consuming sedimentation equilibrium experiments, while still being able to perform velocity analyses on the other instrument. The facility was used by 38 scientists from 21 different research groups, including 4UK labs outside Leeds and 2 overseas labs.

Example

The subject of the experiments described was coat protein (CP) of the bacteriophage MS2. During the course of on-going research in the laboratories of Prof P.G. Stockley and Dr N.J. Stonehouse, mutant coat proteins have been constructed to study factors affecting assembly into viral capsids. Analytical ultracentrifugation, in conjunction with electron microscopy, has provided qualitative and quantitative evaluation of the structures produced under various conditions. Fig 1 shows the result of incubating wild type MS2 CP with a 19-base RNA (TR).

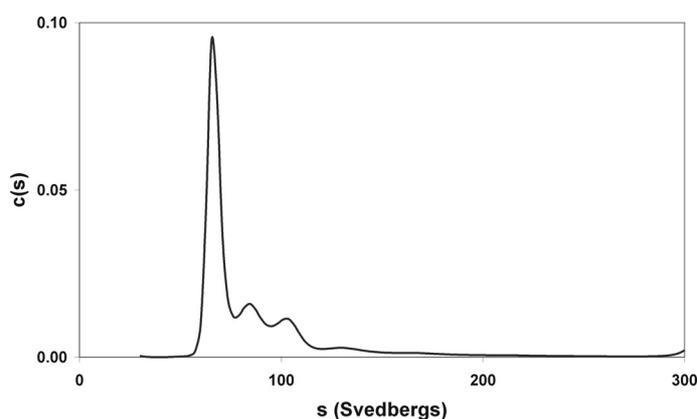


Fig 1. Sedimentation velocity analysis of CP + TR

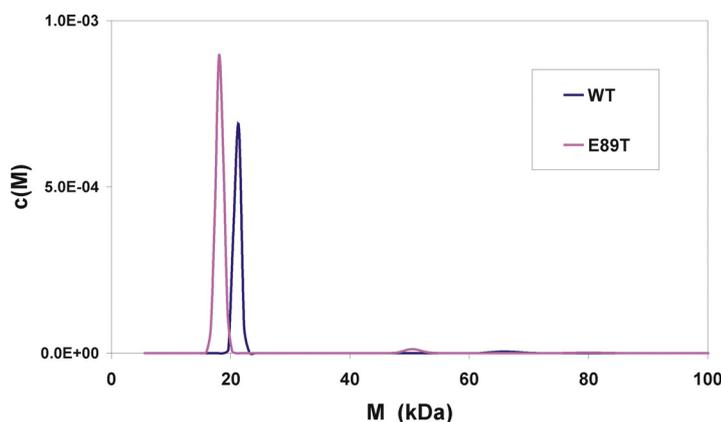
Solute sedimenting at 65 S comprises 58% of the total while the rest of the solute sediments faster. EM analysis showed that most of the sample was in the form of capsids, some of which looked misformed. From many such observations the 65 S component can be confidently identified as capsid.

Properties of acid treated coat protein

The first step in performing CP assembly reactions is to dissociate purified capsids by treatment with acetic acid. It was shown by Sugiyama and Nakada in 1967 that the wild-type forms stable dimers under these conditions. The experiments described below, on material prepared by Simona Francese, were performed to check whether the mutants had similar properties. Results obtained with wild-type and mutant E89T are given as an example. First, sedimentation velocity analysis was performed. Acid treated material was centrifuged at 50000 rpm for 4.5 hours, and boundary movement measured by Rayleigh interference scans. Analysis by the c(s) method using Sedfit (by Peter Schuck) showed that, in both samples, more than 90% of the protein sedimented at 2.1 S. However after least squares fitting, the

Fig 2. Sedimentation velocity analysis of CP in 20mM acetic acid

The weight average masses constituting the major peaks are wild-type: 21.1 ± 0.6 kDa, mutant: 18.2 ± 0.7 kDa. The theoretical mass of monomers is 13.7 kDa so both these values fall between monomer and dimer, suggesting a mixed population of these species in each sample.



frictional ratio of wild-type converged to 1.35, whereas the frictional ratio of the mutant converged to 1.21, and applying these values to $c(M)$ analysis resulted in a small but significant difference between the masses calculated. (Fig 2).

Sedimentation equilibrium analysis, which provides a primary measure of molecular weight, showed that the average mass in a $2\mu\text{M}$ wild-type sample was 17.6 ± 0.8 kDa, and the average mass in a $2\mu\text{M}$ sample of mutant was 15.5 ± 0.9 kDa. However, a reversible monomer-dimer equilibrium model gave a better fit to the results. The wild-type K_d was $74\mu\text{M}$ monomers, and that of mutant $162\mu\text{M}$.

The results from both these techniques show that in 20mM acetic acid the proteins are in a monomer-dimer equilibrium. At the concentration of the sedimentation equilibrium experiment, the average mass in the wild-type sample was between that of monomer and dimer, whereas the average mass in the sample of mutant was only slightly greater than the monomer mass. This observation is reflected in the weaker affinity between mutant monomers. The velocity experiment was performed with much more concentrated protein ($68\mu\text{M}$ wild-type, $36\mu\text{M}$ E89T), so there would be higher proportions of dimer in these samples, thus explaining the higher average masses determined by sedimentation velocity analysis.

These variations in affinity are just one factor affecting the reassembly process which is being studied using a range of techniques including mass spectrometry, gel electrophoresis, analytical ultracentrifugation, X-ray crystallography, and NMR.

Publications

Lago, H., Parrot, A. M., Moss, T., Stonehouse, N. J. and Stockley, P. G. (2001), Probing the kinetics of formation of the bacteriophage MS2 translational operator complex: Identification of a protein conformer unable to bind RNA. *J. Mol. Biol.* **305**, 1131-1144.

Ashcroft, A.E. Lago, H. Stonehouse, N. J. Stockley, P.G. (2000). Q-ToF tackles virus. *Anal Chem*, **2000**, 684-685

Bailey, S., Sedelnikova, S.E., Mesa, P., Ayora S., Waltho J.P., Ashcroft A.E., Baron A.J., Alonso J.C., Rafferty J.B. (2003). Structural analysis of *Bacillus subtilis* SPP1 phage helicase loader protein G39P. *J. Biol. Chem.* **278**, 15304-15312

Funding

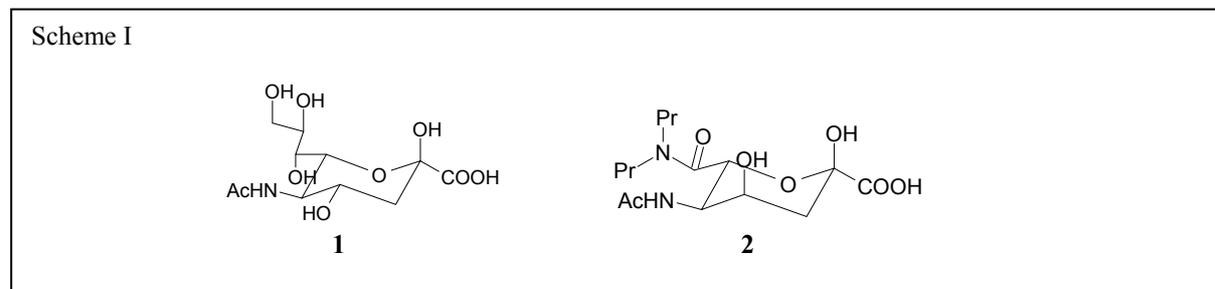
We gratefully acknowledge funding by HEFCE and the Wellcome Trust.

Enzyme engineering of sialic acid synthesising enzymes

Gavin Williams, Thomas Woodall, Adam Nelson and Alan Berry

Introduction

Analogues of sialic acid (**1**) are potent and selective inhibitors of influenza sialidase and therefore represent important therapeutic agents against influenza. In addition, sialic acids play pivotal roles in a wide range of cellular recognition processes.



Sialic acid analogues are however, difficult to chemically synthesise *de novo*. A number have been synthesised by the use of N-acetylneuraminic acid aldolase (NANA). However, the specificity of this enzyme limits its use as a biocatalyst.

As a first step towards overcoming the limited specificity of this enzyme, we aimed to create mutant NANA variants with increased activity towards the synthesis of the 6-dipropylamide (**2**).

Results

Using the known crystal structure of the *E.coli* and *H.influenzae* NANA, three residues were identified that contacted the 6-glycerol moiety of sialic acid. We reasoned that mutagenesis of these residues would produce variants with increased specificity towards analogues of sialic acid which had hydrophobic groups in place of the polar glycerol component. Residues Asp-191, Glu-192 and Ser-208 were mutated to all other 19 amino acids by saturation mutagenesis to create three libraries, D191X, E192X and S208X.

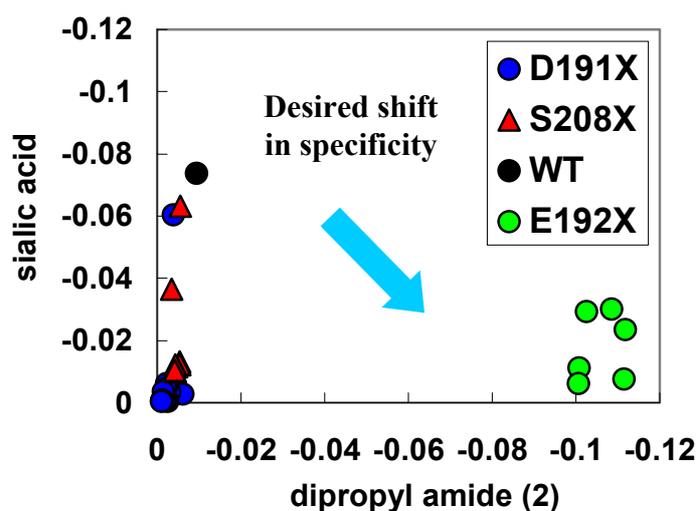


Fig. 1. Activities of mutant sialic acid aldolases expressed in *E.coli*. Colonies expressing mutant aldolase were grown in wells of microtitre plates and crude cell lysates were prepared by lysozyme/freeze thaw treatment. Cleavage of sialic acid and the analogue **2** was monitored via the reduction of pyruvate and the oxidation of NADH using lactate dehydrogenase.

200 members of each library were tested for their ability to cleave the 6-dipropylamide (**2**). The activities of the best clones from these libraries are shown in Fig. 1. Mutations at position 192 produced superior activities. Clearly, variants with greatly improved activity towards the unnatural substrate have been identified. DNA sequencing revealed the wild-type glutamate had been replaced with a range of hydrophobic residues in the most active clones.

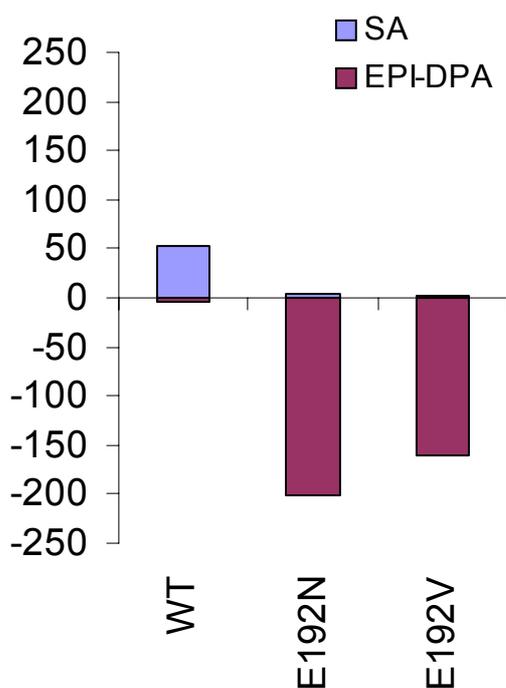


Fig. 2. Specificity constants of wild-type and mutant sialic acid aldolases. The mutant E192N is 4-fold more efficient with the analogue 2 than the wild-type enzyme is with sialic acid.



Fig. 3. Crystal of sialic acid aldolase.

Several variants were chosen for further analysis and were purified and characterised. The best mutant, E192N, showed a 60-fold improvement in catalytic efficiency towards the target compound **2**, and a 170-fold shift in specificity from sialic acid towards the analogue. This resulted in a mutant enzyme that is now 4-fold more efficient with the analogue than the wild-type enzyme with the natural substrate, sialic acid.

In addition, crystals of the wild-type and mutant enzymes have been obtained. Solution of the crystal structure of the evolved biocatalysts will be used to gain insight into the nature of enzyme specificity.

Summary

Several mutant sialic acid aldolase enzymes were created which show greatly improved activity to the dipropylamide **2**, a precursor to an potent and highly selective neuraminidase inhibitor. These novel biocatalysts will be used for the synthesis of a range of clinically relevant sialic acid analogues.

Publications

Suryanti, V., Nelson, A., and Berry, A. (2003) Cloning, over-expression, purification and characterisation of *N*-acetylneuraminase synthase from *Streptococcus agalactiae*. *Protein Expr. Purif.*, **27**, 346-56.

Williams, G.J., Domann, S., Nelson, A., and Berry, A. (2003) Modifying the stereochemical course of an enzyme-catalysed reaction by directed evolution. *Proc. Natl. Acad. Sci. USA*, **100**, 3143-3148.

Funding

This work was funded by the BBSRC and The Wellcome Trust.

Directed evolution of dihydrodipicolinate synthase (DHDPS)

Bernardo Pérez-Zamorano and Alan Berry

Introduction

Dihydrodipicolinate synthase (DHDPS) is a Class I aldolase that is vital for the biosynthesis of the amino acid lysine in prokaryotes, higher plants and some fungi members of the subphylum *Phycomycetes*. The enzyme catalyses the condensation of pyruvate and L-aspartate- β -semialdehyde (L-ASA) to form 2,3-dihydrodipicolinate (DHDP). The immediate product formed from this reaction is not DHDP, but 4-hydroxy-2,3,4,5-tetrahydrodipicolinic acid (HTPA), and the mechanism of the subsequent formation of DHDP under physiological conditions is a matter of conjecture.

The tertiary structure of DHDPS is a $(\beta/\alpha)_8$ barrel structure connected to a α -helical C-terminal domain. The catalytic domain is at the C-terminal end of the barrel, and involves the Schiff-base forming residue Lys-161, located in β strand 6, at the heart of the active center (Fig. 1). DHDPS is a homotetramer, with a subunit molecular mass of 31,372 Da.

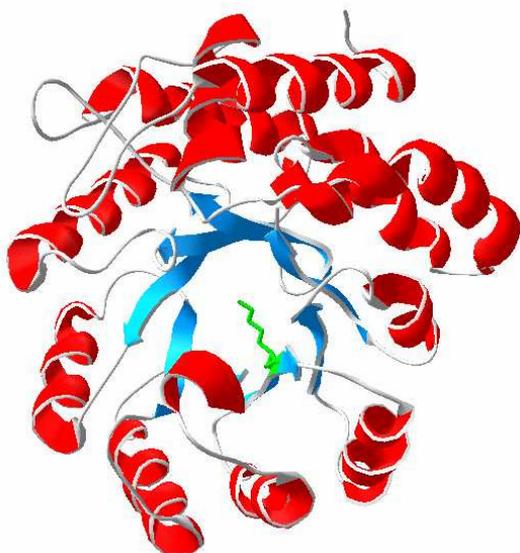


Fig. 1. Tertiary structure of dihydrodipicolinate synthase. The β strands of the $(\beta/\alpha)_8$ barrel are shown in blue and the α helices in red. In green the side chain of Lys 161, the active center, protrudes into the active site cavity from β strand 6.

Evolving DHDPS

Analogues of pipercolinic acid are valuable intermediaries in the synthesis of important compounds in the medical industry such as immunosuppressants and cyclic peptide antibiotics. As an example, a series of hydroxyethylamine-based inhibitors of HIV protease contain a substituted pipercolinic amide. One of the obstacles to overcome in the production of these substances is the fact that *L,D*-pipercolinic acid is very expensive, and the methods known to produce its derivatives are not very efficient. It is the aim of this project to modify DHDPS in order to achieve direct synthesis of analogues of pipercolinic acid. The first target will be to reverse the enantioselectivity of DHDPS towards the use of D-ASA.

Directed evolution involves rounds of random mutagenesis and, in some cases, recombination of the appropriate gene, followed by selection or screening to identify desired changes in the enzymatic activity. The directed evolution process of DHDPS is being carried out by submitting the wild-type DHDPS gene (*dapA*) to error prone polymerase chain reaction (EP-PCR), and the resulting mutant library will be ligated into an expression vector.

Screening for DHDPS activity

The first step in a directed evolution procedure is to validate an assay to monitor the activity of the enzyme to be modified. The screening method used in this case is based on the formation of a characteristic deep red/purple color when *o*-aminobenzaldehyde reacts with a solution containing a DHDPS catalysed reaction; this coloration is believed to be related to the formation of dihydroquinazolium salts. We have shown that the presence of the chromophore is undeniably linked to DHDPS activity (Fig.2). This is a very sensitive method of DHDPS activity detection, and the sensitivity and speed is an advantage for high-throughput screens for new activities.

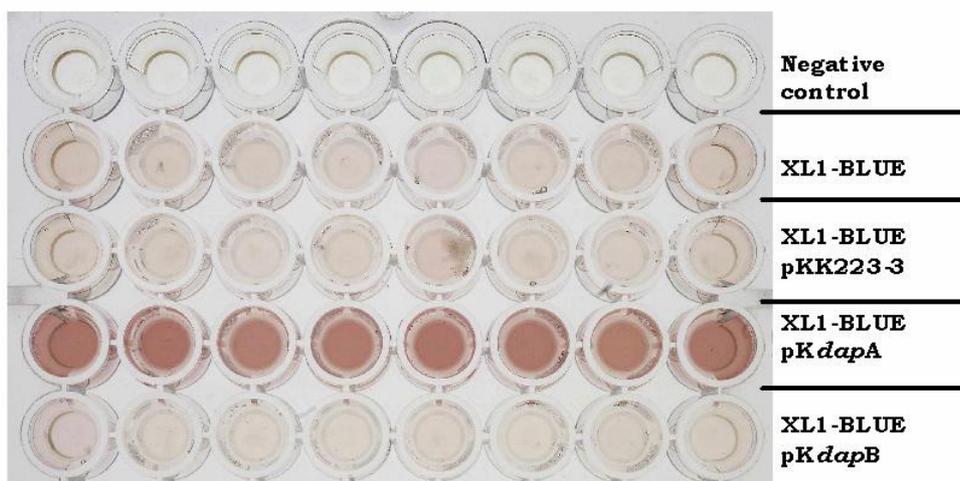


Fig. 2: Colorimetric assay of DHDPS activity. The lysate, substrates and reagents were incubated at 37 °C for ten minutes; and after stopping the reaction the colour was left to develop for 2 hours. Negative control: no cell lysate was used in this reaction; XL1-BLUE: lysate of untransformed cells; XL1-BLUE pKK223-3: lysate of cells transformed with an empty over-expression vector; XL1-BLUE pKdapA: lysate of cells bearing a *dapA* (DHDPS) over-expressing vector; XL1-BLUE pKdapB: lysate of cells bearing a *dapB* (dihydrodipicolinate reductase or DHDPR) over-expressing vector.

The system for the screening of new DHDPS activities is therefore in place and directed evolution of the enzyme is underway.

Collaborators

Adam Nelson and Tom Woodhall (School of Chemistry, University of Leeds).

Funding

This work is supported by the BBSRC and Consejo Nacional de Ciencia y Tecnología (CONACyT), México.

Directed evolution of an aldolase to enhance solubility

Chris Plummer and Alan Berry

Introduction

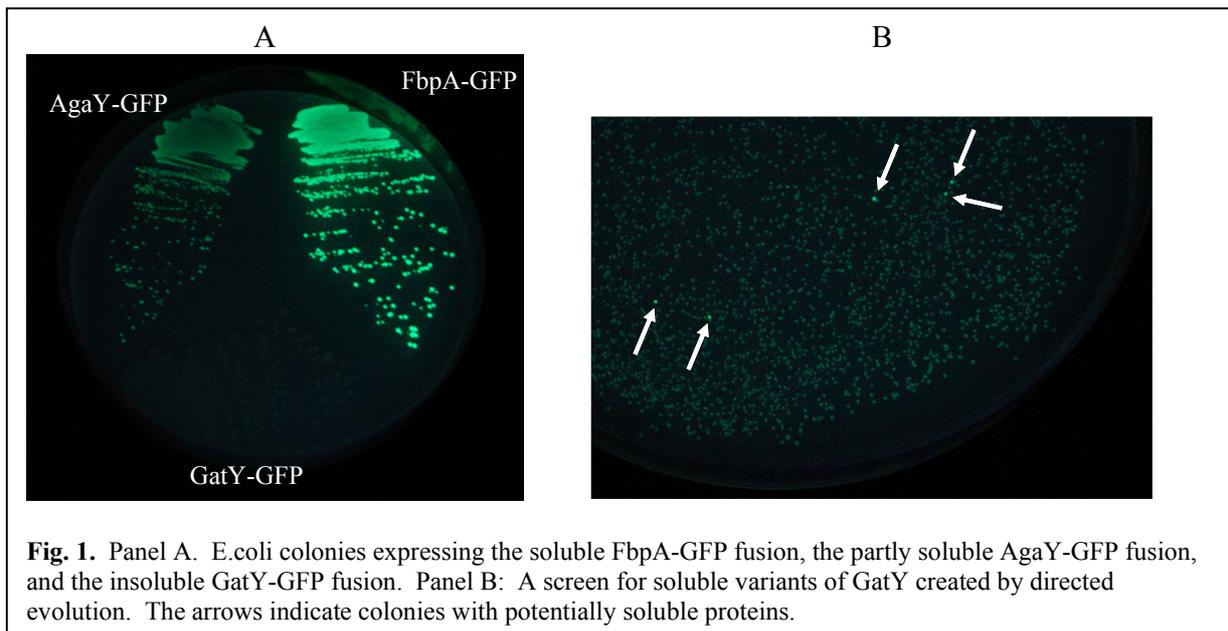
Aldolases are very useful enzymes for generating precursors to complex fine chemicals and pharmaceuticals, which may be very difficult, or impossible, to produce using purely chemical synthesis methods. Indeed many aldolases are already being used for these purposes; however, the capabilities of the natural enzymes are limited. Enzyme engineering can be used to generate new activities within a protein scaffold to broaden our access to other useful reactions. Another problem which may be encountered in the use of enzymes is a limited solubility, especially during the production of the enzyme in an over-expressing bacterial cell. This can lead to the formation of the enzyme as an inactive inclusion body, and these may be difficult or impossible to re-solubilise. The aim of this project is to use directed evolution to modify an insoluble aldolase into a well-expressed, soluble protein.

Directed evolution of a soluble TBP aldolase

The *gat* operon of *E.coli* comprises several genes involved in galactitol metabolism, one of which, *gatY*, appears, by sequence alignment, to encode a tagatose 1,6-bisphosphate aldolase. However, over-expression of protein from the cloned gene yields inclusion bodies, preventing access to this gene product as a biocatalyst. We are using directed evolution to produce a novel, soluble variant of the GatY TBP-aldolase to determine its potential as a biocatalyst. Directed evolution is a technique in which a library of mutant proteins, created by error prone PCR or DNA shuffling, can be screened for variants with improved characteristics. The new characteristic is 'evolved' because the selected mutants can then be improved further by more rounds of mutagenesis and screening. Directed evolution is an ideal approach for the generation of a soluble protein because no prior knowledge of the factors influencing solubility is needed.

A simple screening method for detecting soluble proteins has been developed whereby the protein of interest is expressed as a fusion protein with green fluorescent protein, GFP. Insoluble proteins form inclusion bodies, within which the GFP is not fluorescent; however, soluble proteins allow the correct folding of GFP and so are detectable by intense green fluorescence. This is a particularly useful system when coupled with directed evolution because large libraries ($>10^4$) of mutants can rapidly be visually screened because correctly folded GFP will cause the whole expressing bacterial colony to be fluorescent.

We have cloned the genes encoding a number of soluble aldolases as a fusion to GFP to verify the screening method (Fig. 1). A mutant library generated by error prone PCR has also been screened, and a number of potential positives have been identified. These will be sequenced and then subjected to further rounds of mutagenesis and screening to further increase the solubility. Work can then begin to study and characterize this enzyme, which should not only provide details of its catalytic properties, but also provide some insight into the factors which determine solubility of recombinant proteins.



Publications

Hall, D.R., Kemp, L.E., Leonard, G.A., Marshall, K., Berry, A. & Hunter, W.N. (2003) The organization of divalent cations in the active site of cadmium *Escherichia coli* fructose-1,6-bisphosphate aldolase. *Acta Cryst. D*, **59**, 611-614.

Funding

This work is funded by the BBSRC, Wellcome Trust and the EPSRC.

Structure/function of G-protein coupled receptors (GPCRs)

Nirmala Bhogal and John Findlay

As part of our on-going investigations into the structure and function of GPCRs, we report a series of papers which help to define the location of ligands in the NK₂ tachykinin receptor. The first study describes the use of thiol chemistry to define specific disulphide interactions between Cys-substituted receptor mutants and peptide analogues containing single cysteine replacements (Fig 1A).

N-biotinyl-[Tyr¹, Cys⁹]-NKA and *N*-biotinyl-[Tyr¹, Cys¹⁰]-NKA were both found to reversibly disulphide bond to the NK₂ receptor mutant Met²⁹⁷-Cys (Fig 1B). This is consistent with the improved affinities of these particular analogues for the Met²⁹⁷-Cys receptor as compared to those for the wild-type and Met²⁹⁷-Leu receptors. In our 3D model, Met²⁹⁷ occupies the equivalent position in helix 7 to the retinal binding Lys²⁹⁶ in rhodopsin. Binding of the NK₂ receptor antagonist ³H-SR 48968 and of ¹²⁵I-NKA was used to characterise additional receptor mutants. It seems that the aromatic residues Trp⁹⁹ (helix 3), His¹⁹⁸ (helix 5), Tyr²⁶⁶, His²⁶⁷ and Phe²⁷⁰ play an important role in NKA binding as structural determinants (Fig 1C). These data suggest that the peptide binding site of the NK₂R is, at least in part, formed by residues buried deep within the transmembrane bundle, and that this intramembranous binding domain may correspond to the binding sites for substantially smaller endogenous GPCR ligands.

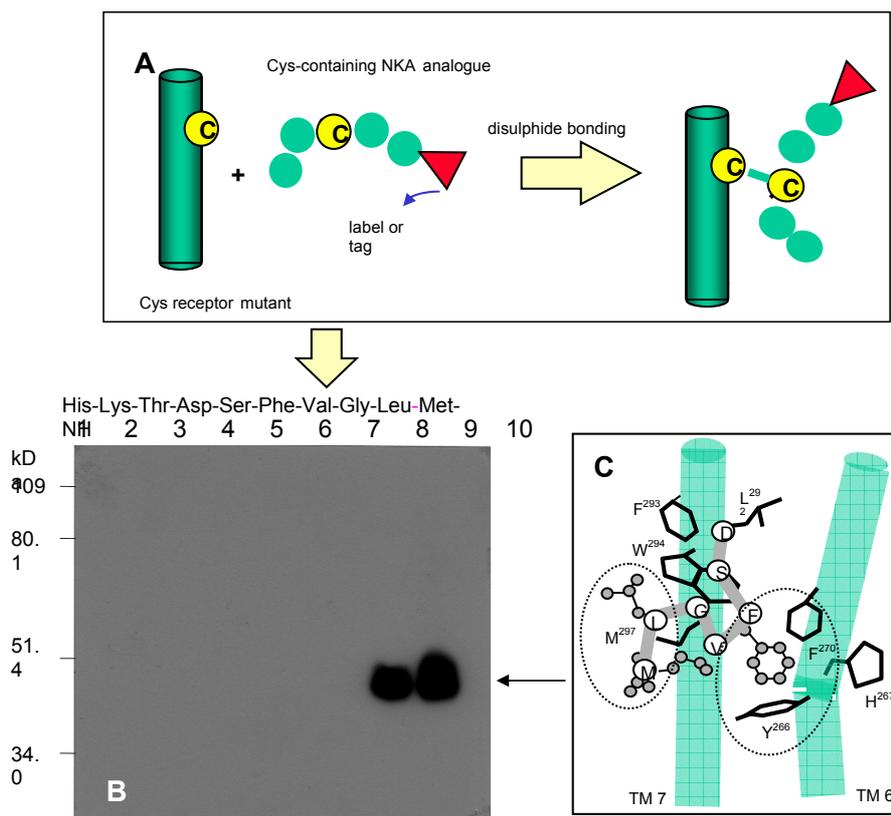


Fig. 1. Thiol chemical approach to define peptide-receptor contacts. A shows the approach taken. B shows the specific thiol cross linking of Cys analogues of NKA to the Met²⁹⁷-Cys receptor mutant as detected by biotin-streptavidin affinity. Only two peptides cross-linked: Those containing Cys substitutions in positions 9 and 10. C shows the predicted disposition of NKA within the NK₂ receptor. Reproduced from Labrou *et al.* 2001.

The second study involves the identification of the binding site for [³H]-SR48968, a piperidinyll antagonist. This binding is inhibited by methanethiosulfonate ethylammonium (MTSEA) in a time-, and concentration-dependent manner. By the systematic alanine replacement of putative loop and transmembrane region cysteine residues (Cys⁴, Cys⁸¹, Cys¹⁶⁷, Cys²⁶², Cys²⁸¹, Cys³⁰⁸ and Cys³⁰⁹), we have determined that MTSEA perturbs [³H]-SR48968 binding by modifying Cys¹⁶⁷ in transmembrane helix 4. Data were substantiated using Gly, Ser, and Thr substitutions of Cys¹⁶⁷. MTSEA preferentially modifies cysteine residues that are in proximity to a negatively charged environment. Hence, aspartate and glutamate residues were systematically substituted with leucine or valine, respectively and the inhibitory effects of MTSEA on [³H]-SR48968 binding re-evaluated in order to determine those acidic residues close to the MTSEA binding crevice. Most significantly, substitution of Asp⁵ in the receptor's extreme N-terminus abolished the effects of MTSEA on [³H]-SR48968 binding. Therefore, our data would suggest close association of the extreme N-terminus with the extracellular surfaces of helices 4 and 3 in the NK₂ receptor in forming a binding crevice for MTSEA. The inhibition of SR 48968 binding appears to result from alteration of the binding conformation of Gln¹⁶⁶ by MTSEA (Figs. 2-4).

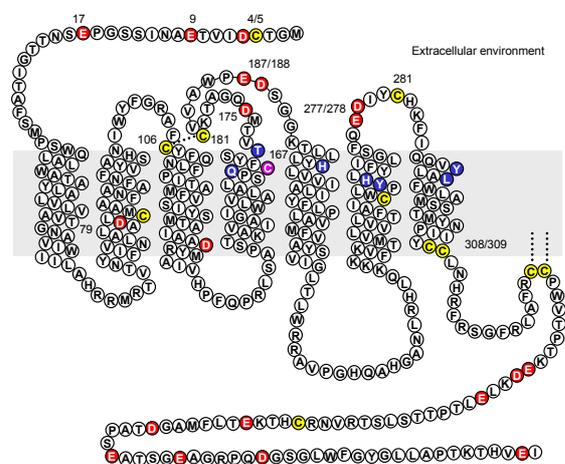


Fig. 2 Representation of the NK₂R. Native cysteine (yellow) and negatively charged residues (red) and residues that form part of the SR48968 binding site (blue) are shown. Cys¹⁶⁷ (purple), the putative S-S bond between Cys¹⁰⁶ and Cys¹⁸¹ (hashed line) and potential palmitoylation sites (Cys³²⁴ and Cys³²⁵) are also indicated.

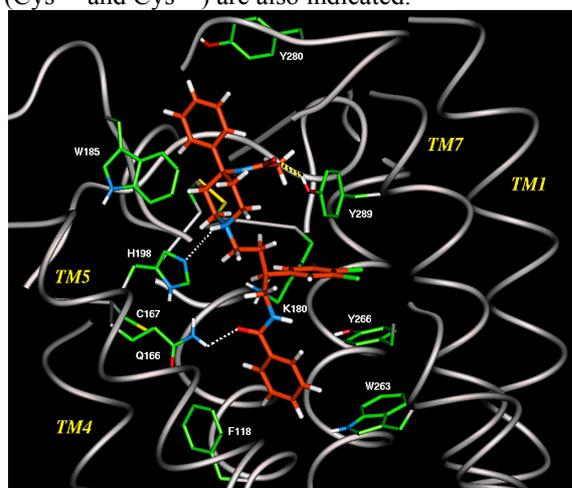


Fig. 3. Proposed binding site for SR48968. Dotted lines represent strong hydrogen bonds, viz. Gln¹⁶⁶ with an amide carbonyl, Tyr²⁸⁹ with the 2nd amide SR48968 carbonyl (orange).

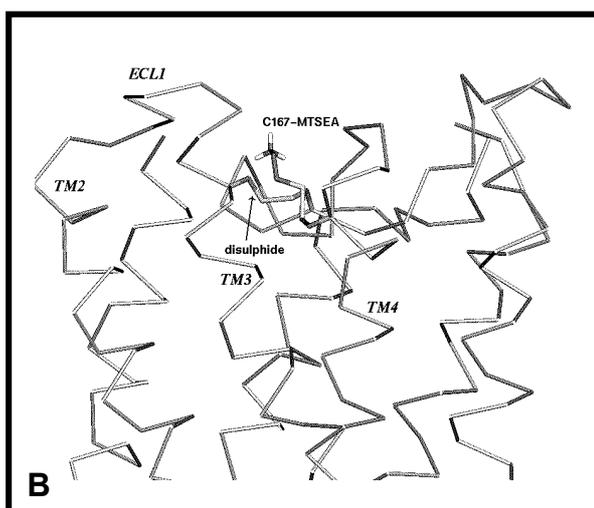
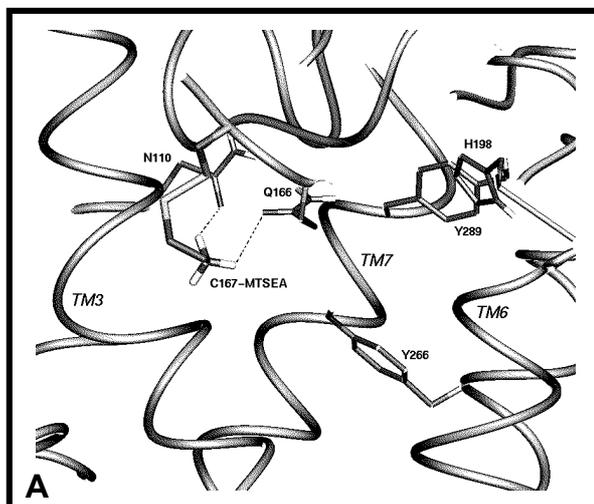


Fig. 4. Alternative conformations of MTSEA cross-linked to Cys¹⁶⁷. (A) the ammonium headgroup of MTSEA interacts with Asn¹¹⁰ and Gln¹⁶⁶, breaking the interaction of the latter residue with SR48968; (B) the headgroup extends to the exterior of the TM bundle between TM3 and TM4. In this conformation MTSEA is believed to interact with Asp⁵.

Publications

Ali, M.A., Bhogal, N., Fishwick, C.W.G. & Findlay, J.B.C. (2001) Spatial requirements of the antagonist binding site of the NK2 receptor. *Bioorg Med Chem Lett* **11**, 819-822.

Labrou, N.E., Bhogal, N., Hurrell, C.R. & Findlay, J.B.C. (2001) Interaction of Met (297) in the seventh transmembrane segment of the tachykinin NK2 receptor with neurokinin A. *J. Biol. Chem.* **276**, 37944-37949.

Bhogal, N., Blaney, F.E., Ingley, P.M., Rees, & Findlay, J.B.C. (2003) The proximity of the extreme N-terminus of the NK₂ tachykinin receptor to Cys¹⁶⁷ in the putative fourth transmembrane helix. (In Press)

Funding

We acknowledge support from BBSRC, the EC and GSK.

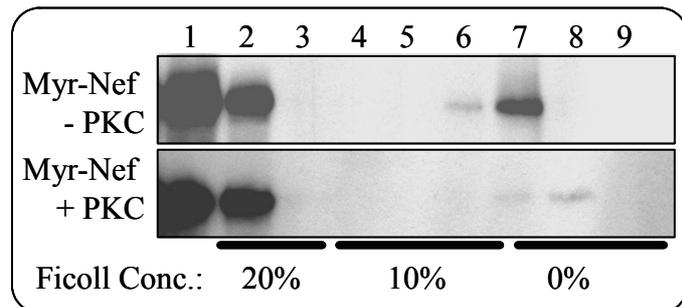
Functional studies on the HIV-1 Nef protein.

Matthew Bentham, Sabine Mazaleyrat, Gemma Dixon, Caitriona Dennis,
Joachim Jäger and Mark Harris.

HIV-1 Nef is a 205 amino acid, N-terminally myristoylated protein that plays a critical role in viral pathogenesis. Myristoylation is a eukaryotic specific, co-translational modification that is catalysed by a ribosomal associated enzyme - N-myristoyltransferase (NMT). As yet there is no information about the structure of the full-length myristoylated form of Nef. To address this issue we have generated large amounts of myristoylated Nef by co-expression with human NMT in *E.coli*. As detailed elsewhere in this report, a range of biophysical and biochemical techniques have been used to determine the effects of myristoylation on the physical properties of Nef. This protein has also been used to analyse the interactions between Nef and cellular factors. In one set of experiments, we have examined the binding of purified Nef to synthetic liposomes using a membrane flotation assay. The data demonstrate that the Nef-membrane interaction, as for many cellular myristoylproteins, involves not only myristate but additional electrostatic interactions between basic amino acids and acidic phospholipids. Fig. 1 shows that the Nef-membrane interaction can be abolished by phosphorylation. We postulate that phosphorylation of Nef by PKC (previously shown by our group and others) will induce the dissociation of Nef from the plasma membrane – this mechanism of regulating membrane association is termed a myristoyl-switch.

Fig.1 Phosphorylation abrogates Nef-membrane association *in vitro*.

Purified myristoylated Nef was mixed with liposomes and layered at the bottom of a discontinuous Ficoll gradient. After ultracentrifugation, fractions were collected, methanol precipitated and analysed for the presence of Nef by western blotting. Lane 1: positive control, lanes 2,3 soluble protein, lanes 6,7: liposome fractions.



One of the functions of Nef is to down-modulate the cell surface glycoprotein CD4 (which also functions as the viral receptor). This involves a direct physical interaction between myristoylated Nef and the cytoplasmic domain of CD4. Previously this interaction had only been demonstrated in insect cells or *in vitro*. We have now been able to show, using a CD4 mutant that is not endocytosed that this interaction also occurs in mammalian cells. As part of this work we have developed both ELISA and SPR based *in vitro* assays to analyse the interaction between Nef and CD4. Ultimately these may be of utility in screening for inhibitors of the interaction.

Publication

Bentham, M., Mazaleyrat, S. & Harris, M. (2003) The di-leucine motif in the cytoplasmic tail of CD4 is not required for binding to HIV-1 Nef, but is critical for CD4 down-modulation. *J. Gen. Virol.* **84**, 2705-2713.

Collaborators

EU Framework Five 'Targeting Nef' (URL: <http://www.uta.fi/~ltkasa/eu/index.html>)

Funding

This work was funded by the BBSRC, MRC and EU Framework Five.

Structural and functional studies on the Hepatitis C Virus non-structural NS5A protein.

Andrew Macdonald, Andrew Street, Katherine Crowder, Virginie Cazeaux, Holly Shelton, Chris McCormick and Mark Harris.

Hepatitis C virus (HCV) is an increasingly important cause of liver disease. The virus has a single stranded positive sense RNA genome of 9.5Kb that contains a long open reading frame encoding a single polyprotein of 3000 amino acids which is cleaved into 10 individual polypeptides by a combination of host cell and virus specific proteases. The molecular mechanisms of pathogenesis remain to be elucidated. To this end, our laboratory is interested in understanding the interactions between the NS5A protein and host cell signal transduction pathways.

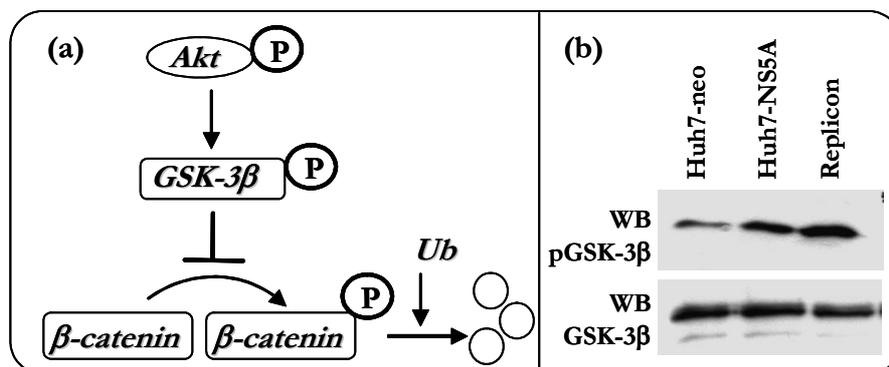
Our data have shown that NS5A perturbs two key mitogenic pathways within the cell. Firstly, NS5A prevents the activation of the AP1 transcription factor by blocking the Ras-ERK MAP kinase pathway. Secondly, NS5A interacts with the SH3 domain of the phosphatidylinositol 3-kinase (PI3K) p85 regulatory subunit via a novel SH3-binding motif and stimulates PI3K activity 10-fold. This results in the activation of PI3K signalling pathways. In particular we have demonstrated activation of the downstream kinase Akt (PKB) with a subsequent increase in cell survival and stimulation of the activity of the proto-oncogene β -catenin.

Fig 1: NS5A stimulates the Akt- β -catenin pathway.

(a) Phosphorylated Akt inactivates GSK-3 β . This stabilises β -catenin in the unphosphorylated form which can enter the nucleus and activate transcription of genes such as cyclin D1 and c-myc.

(b) Huh7 (hepatoma) cells expressing either NS5A alone or the HCV sub-

genomic replicon (NS3-5B) show enhanced GSK-3 β phosphorylation, as shown by western blotting using a phospho-specific antibody (top panel). Overall levels of GSK-3 β are unaltered (lower panel).



We have also shown that two closely spaced poly-proline motifs in NS5A interact with the SH3 domains of members of the Src family of protein tyrosine kinases. We are currently using phage display and *in vitro* and *in vivo* binding assays to understand more precisely these interactions. Current work is also focussed on identifying the functional consequences of these interactions. In particular, we are using the BacMAM system - baculovirus vectors with mammalian specific promoters that are able to efficiently enter hepatic cells and drive expression of foreign genes – to express NS5A both alone, and in the context of the complete HCV genome in hepatocyte derived cell lines.

Publications

Macdonald, A., Crowder, K., Street, A., McCormick, C., Saksela, K. & Harris, M. The Hepatitis C Virus NS5A protein inhibits Activating Protein-1 (AP1) function by perturbing Ras-ERK pathway signalling (2003) *J. Biol. Chem.* **278**, 17775-17784.

Street, A., Macdonald, A., Crowder, K. & Harris, M. (2004) The Hepatitis C Virus NS5A Protein Activates a Phosphoinositide 3-kinase Dependent Survival Signalling Cascade. *J. Biol. Chem.*, in press.

Macdonald, A., Crowder, K., Street, A., McCormick, C. & Harris, M. (2004) The Hepatitis C Virus NS5A protein binds to members of the Src family of tyrosine kinases and regulates kinase activity. *J. Gen. Virol.*, in press.

Collaborators:

Derek Mann, University of Southampton

Kalle Saksela, University of Tampere, Finland

Funding:

This work is funded by the BBSRC, MRC and Wellcome Trust.

Thermodynamics of binding of 2-methoxy-3-isopropylpyrazine and 2-methoxy-3-isobutylpyrazine to the major urinary protein

Richard J. Bingham, John B. C. Findlay, Shih-Yang Hsieh, Arnout P. Kalverda, Alexandra Kjellberg, Chiara Perazzolo, Simon E. V. Phillips, Kothandaraman Seshadri, Chi H. Trinh, W. Bruce Turnbull, Geoffrey Bodenhausen and Steve W. Homans

Introduction

Protein-ligand interactions are of fundamental importance in a great many biological processes. However, despite enormous advances in the speed and accuracy of the three dimensional structure determination of proteins and their complexes, our ability to predict binding affinity from structure remains severely limited. One reason for this dilemma is that affinities are governed not only by energetic considerations concerning the precise spatial disposition of interacting groups, but also by the dynamics of these groups, in addition to solvent effects. Thus, in order to predict accurately the affinity of a protein for a given ligand, it is essential to have prior knowledge of both the enthalpy of binding, ΔH_b° and the entropy of binding ΔS_b° . A quantitative measure of the elusive ΔS_b° component is notoriously difficult, since it depends on the dynamics of the complex (including solvent) over all degrees of freedom of the system. Isothermal titration microcalorimetry experiments offer the possibility to measure thermodynamic binding parameters including ΔS_b° , but since the derived parameters are global in nature, it is difficult to separate contributions from protein, ligand and solvent. In principle, characterisation of the internal dynamics of a protein in the absence and presence of ligand should enable measurement of ΔS_b° values associated with the internal degrees of freedom of the protein. In particular, NMR relaxation-time measurements offer scope for the measurement of ΔS_b° on a per-residue basis, and this approach has been pioneered by a number of workers. As a model study, we examine the entropies of binding of two related ligands, namely 2-methoxy-3-isopropylpyrazine (IPMP) and 2-methoxy-3-isobutylpyrazine (IBMP), to the major urinary protein MUP-I, using a combination of isothermal titration calorimetry (ITC), X-ray crystallography and NMR backbone ^{15}N and methyl side-chain ^2H relaxation measurements.

Binding thermodynamics

Global thermodynamic data derived from ITC indicate that binding is driven by favourable enthalpic contributions, rather than the classical entropy-driven hydrophobic effect. Unfavourable entropic contributions from the protein backbone and side-chain residues in the vicinity of the binding pocket are partially offset by favourable entropic contributions at adjacent positions, suggesting a 'conformational relay' mechanism, whereby increased rigidity of residues on ligand binding are accompanied by increased conformational freedom of side-chains in adjacent positions. The principal driving force governing ligand affinity can be attributed to solvent-driven enthalpic effects from desolvation of the protein binding pocket. Interestingly, the specificity of MUP-I for the two ligands cannot be explained by a difference in solvation of the binding site in the complex, but can be explained in terms of an entropic contribution from ligand desolvation.

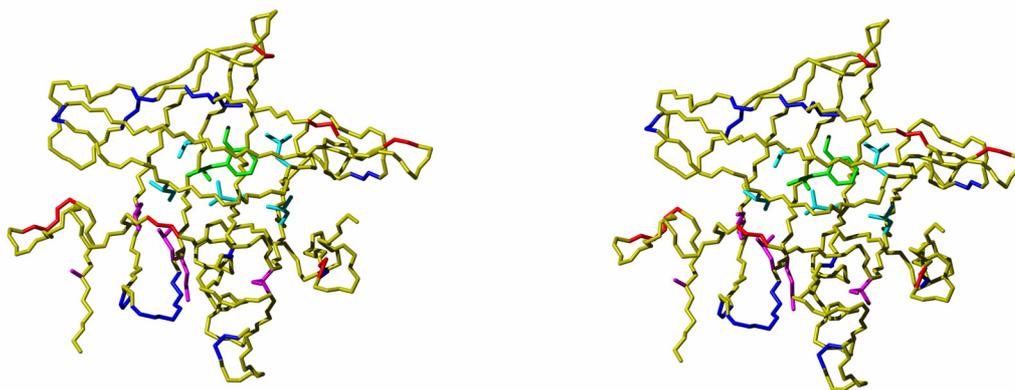


Fig. 1. Stereo view of structural details of residues that contribute to the entropy of binding of 2-methoxy-3-isobutylpyrazine (coloured green) to MUP-I. Backbone residues that exhibit an unfavourable entropic contribution to binding are coloured blue, while those that exhibit a favourable contribution are coloured red. Similarly, residues whose methyl-containing side-chains exhibit an unfavourable contribution are coloured light blue, whereas those that exhibit a favourable contribution are coloured magenta.

Collaborators

Geoffrey Bodenhausen and Chiara Perazzolo, ENS, Paris and EPFL, Lausanne; ProSpect Pharma Inc. (USA).

Publications

Chaykovski, M.M., Bae L.C., Cheng M-C., Murray J.H., Tortolani K.E., Zhang R., Seshadri, K., Findlay, J.B.C., Hsieh, S-Y, Kalverda A.P., Homans, S.W. & Miles Brown, J. (2003) Methyl side chain dynamics in proteins using selective enrichment with a single isotopomer. *J. Am. Chem. Soc.* **125**, 15767-15771.

Bingham, R., Bodenhausen, G., Findlay, J.B.C., Hsieh, S-Y., Kalverda, A.P., Kjellberg, A., Perazzolo, C., Phillips, S.E.V., Seshadri, K., Turnbull, W.B. & Homans, S.W. (2003) Thermodynamics of binding of 2-methoxy-3-isopropylpyrazine and 2-methoxy-3-isobutylpyrazine to the major urinary protein. *J. Am. Chem. Soc.* in press.

Funding

This work has been funded by the Wellcome Trust and BBSRC

Dissecting the cholera toxin-ganglioside GM1 interaction by isothermal titration calorimetry

Bruce Turnbull and Steve Homans

Introduction

Cell surface protein-carbohydrate interactions provide mechanisms for the invasion and colonisation of a wide range of pathogenic viruses and bacteria. Furthermore, a number of protein-based bacterial toxins, e.g., the cholera and heat-labile toxins derived from *Vibrio cholerae* and *Escherichia coli*, respectively, also enter their target cells following an initial adhesion to glycolipids at the cell surface. Both toxins have an AB₅-type structure — a single A-subunit which possesses a toxic ADP-ribosyltransferase activity and five identical B-subunits which are lectins for the cell surface glycolipid, ganglioside GM1. The interaction of the cholera toxin B-pentamer (CTB) with the pentasaccharide portion of GM1 (GM1os, Fig. 1) is one of the highest affinity protein-carbohydrate interactions known, and under multivalent conditions at a cell surface, the system achieves sub-nanomolar avidity. Yet, related oligosaccharide ligands show little, if any, affinity for the toxin. The crystal structure of the CTB-GM1os complex reveals that only the three residues at the non-reducing terminus make significant interactions with the protein. The high affinity of the CTB-GM1os complex has provided us with an opportunity to dissect the intrinsic contributions to the binding affinity of each of the three key monosaccharide residues, thus providing insight into the origin of the remarkable specificity of CTB for GM1.

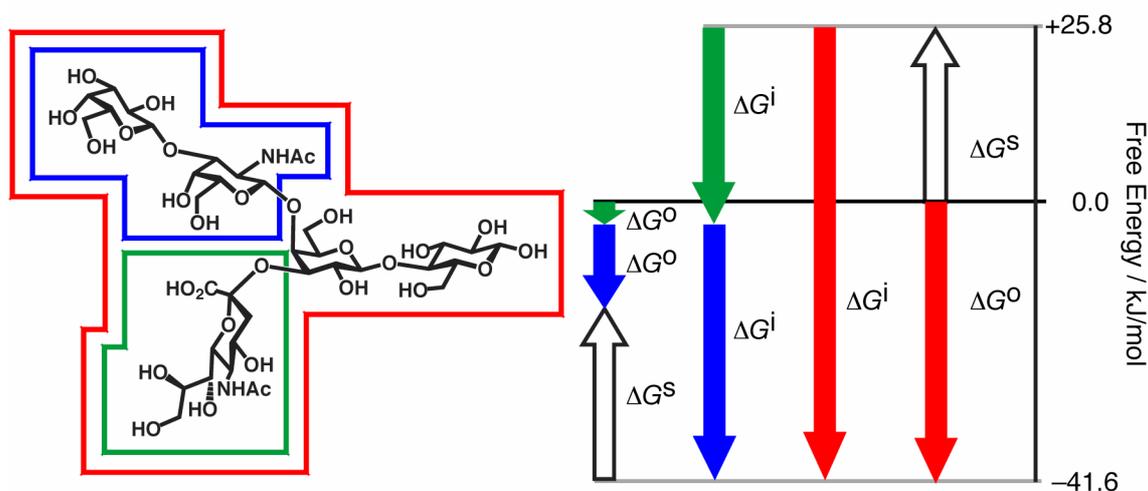


Fig. 1. Breakdown of the standard (ΔG°), intrinsic (ΔG^i) and connection (ΔG^s) free energy changes for the GM1os pentasaccharide (in red), and two of its fragments (in green and blue). The sialic acid residue, shown in green, has a very poor ΔG° , but still contributes approximately half of the total intrinsic free energy for the interaction with CTB.

Synthesis and binding studies of GM1 fragments

Whereas smaller mono and disaccharide fragments of ganglioside GM1 are either commercially available or could be accessed readily by chemical synthesis, larger fragments — including the full pentasaccharide — were most easily produced by stepwise enzymatic degradation of the natural ligand. Isothermal titration calorimetry (ITC) is a technique that exploits small stepwise changes in heat that is released during the course of a ligand-receptor titration, to allow derivation of all thermodynamic parameters (free energy, enthalpy and entropy changes) in a single experiment. Whereas the pentasaccharide was found to bind very tightly with a nanomolar dissociation constant, all of the fragment ligands interacted very weakly with the toxin. We recently demonstrated that such low affinity systems may be readily studied by standard ITC methods (in contradiction to many reports in the literature).

However, in cases such as this one, where a high affinity ligand is available, it is preferable to use a displacement assay, which allows the experiments to be conducted using lower protein concentrations. Analysis of the data using Jencks' concept of intrinsic free energies has demonstrated that the terminal galactosyl and sialosyl residue each contribute just under 50% of the intrinsic CTB-GM1os interaction, but paradoxically, sialic acid itself has little appreciable affinity for the receptor. Our results have allowed us to reconcile the high selectivity of GM1os over its fragments with the crystal structure for the complex that was reported previously. Furthermore, the analysis suggests a different binding mode for a tetrasaccharide analogue (GM2os) and suggests that ligand design for the system would best be served by using galactosyl derivatives which could extend into the space that is occupied by sialic acid in the GM1os-CTB complex. Finally, this study gives an estimate of ca. 25 kJ mol⁻¹ for the largely entropic penalty that must be made on bringing two molecules together to form a complex in aqueous solution.

Collaborators

Bernie Precious, Centre for Biomolecular Science, University of St Andrews, UK.

Publications

Turnbull, W.B. & Daranas, A.H. (2003) On the value of *c*: can low affinity systems be studied by isothermal titration calorimetry? *J. Am. Chem. Soc.* **125**, 14859-14866.

Turnbull, W.B., Precious, B.L. & Homans, S.W. (2004) Dissecting the cholera toxin-ganglioside GM1 interaction by isothermal titration calorimetry. *J. Am. Chem. Soc.* **126**, 1047-1054.

Funding

This work has been funded by the Wellcome Trust through the award of an International Prize Travelling Research Fellowship to B.T.

The NMR solution structure of the VMA-7 subunit of the vacuolar H⁺-ATPase from *Saccharomyces cerevisiae*

Bruce Turnbull, Gary Thompson, Liz Barratt, Lyndsey Durose,
Mike Harrison, John Findlay and Steve Homans

Introduction

Vacuolar H⁺-ATPase (VATPase) is a membrane spanning proton pump which is found in virtually all eukaryotic cells. The action of the pump results in translocation of protons across vacuolar membranes, allowing the pH of intracellular compartments to be regulated. Errors in the function of the pump have been implicated in the pathology of a number of diseases including osteoporosis, diabetes and several common cancers including cervical cancer. A structural understanding of this complex at the atomic level will open up paths to new therapeutics for these diseases by control of its activity. Though the structure of the VATPase enzyme complex has been shown to be related to that of the F₁F₀ ATPase, there are major differences in the subunits and the composition of the two enzymes. Therefore, a significant effort is underway at Leeds under the MASIF scheme to determine the structure of the enzyme subunits and the complete complex using modern structure determination methods: Xray crystallography, NMR spectroscopy and electron microscopy. The VATPase complex contains at least ten distinct subunit types and four of these are of a size amenable to modern NMR techniques. The current target, vma7p, is a 14 kDa subunit of the complex, which has unknown function and fold, and has been over expressed in *E. coli* as a GST fusion protein incorporating the PreScission Protease cleavage site.

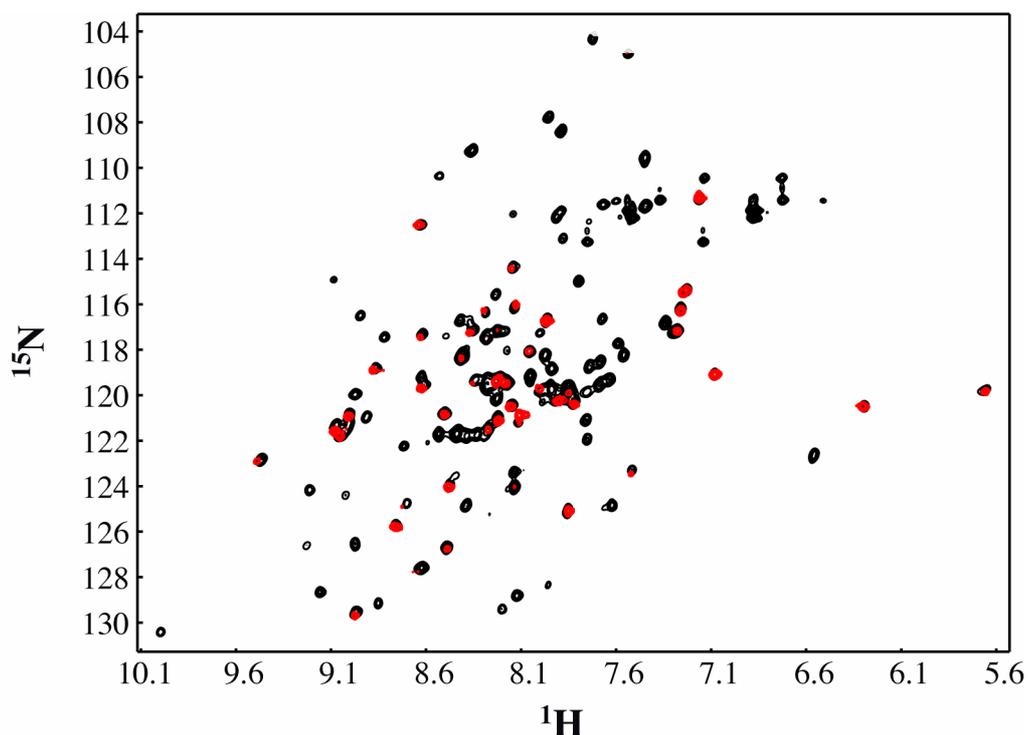


Fig. 1. HSQC spectrum of the amide resonances in vma7p at 750 MHz (in black). Signals arising from only the Ile, Asp and Asn residues of a selectively labeled sample, are shown in red.

Production of a triple labeled sample of vma7p (¹⁵N / ¹³C / 50% ²H) has allowed us to complete the backbone assignment, using a combination of HNCA, HN(CO)CA, HNC(O), HN(CA)CO and HNCACB experiments. Partial deuteration of the protein was found to be necessary to obtain adequate levels of signal-to-noise in the less sensitive triple resonance experiments. This assignment has been confirmed by comparison with a vma7p sample

selectively ^{15}N -enriched in isoleucine, aspartate and asparagine (Fig. 1). Secondary chemical shift analysis has been used to predict the secondary structure of the protein. Residual dipolar coupling measurements for the backbone atoms are currently being used to calculate the 3D structure of the protein in conjunction with NOESY distance restraints.

Funding

We wish to thank the BBSRC and EPSRC for funding.

Reference

Jones, R.P.O., Hunt, I.E., Jaeger, J., Ward, A., O'Reilly, J., Barratt, E.A., Findlay, J.B.C. & Harrison, M.A. (2001), Expression, purification and secondary structure analysis of *Saccharomyces cerevisiae* vacuolar H⁺-ATPase subunit F (VMA7p), *Molecular Membrane Biology*, **18**, 283-290.

Large scale millisecond inter-subunit dynamics in the B subunit homopentamer of the toxin derived from *E. coli* O157

Anna Yung, W. Bruce Turnbull, Arnout P. Kalverda, Gary S. Thompson, Steve W. Homans, Pavel Kitov and David R. Bundle

Introduction

The toxin derived from *E. coli* O157 is a member of the AB₅ class of cytotoxins, comprising a catalytically active A subunit, and a torus-shaped homopentameric B subunit that binds to the cell-surface glycolipid globotriaosylceramide (Gb₃) in a multivalent manner. The crystal structure of the B subunit shows similar inter-monomer β -sheet interactions between β 2 of monomer *n* and β 6 of monomer *n*+1, except between two monomers, resulting from a screw component of about 0.13 nm in the 5-fold rotation axis of the pentamer. In contrast, NMR measurements suggest a dominant conformer that is a symmetric homopentamer in solution, consistent with a single set of crosspeaks in ¹H-¹⁵N correlation spectra. However, ¹H-¹⁵N correlation spectra recorded on the protein in the presence of a five-fold molar excess of the bivalent inhibitor P^k-dimer, which is a bridged dimer analogue of the Gb₃ carbohydrate, are characterised by a reduction of resonance line-widths for a number of resonances. These data suggest an exchange contribution to the line-widths of the B subunit in the absence of P^k-dimer, which accordingly was probed by use of NMR relaxation dispersion measurements.

Slow conformational exchange between subunits

NMR relaxation dispersion experiments allow us to detect slow time-scale (millisecond) motions in proteins on a per-residue basis, and provide, *inter alia*, a quantitative measure of this time-scale. Our measurements showed significant motions for residues immediately adjacent to each monomer-monomer interface in the protein, each of which could be characterized by a single exchange rate of 1000/sec. The data can be interpreted structurally in terms of fast exchange between a symmetric conformation and a minor conformer that may be related to that observed in the crystal structure. On binding of P^k-dimer, which straddles binding-sites on adjacent monomers, these motions disappeared, indicating that this novel ligand effectively ‘locks’ the pentamer in a single conformation. Consequently, binding of P^k dimer cannot be interpreted in terms of a simple two-state model, but is consistent with a sequential binding model involving co-operative effects. These effects involve a positive entropic component to co-operativity that has not been observed previously in multivalent systems to our knowledge.

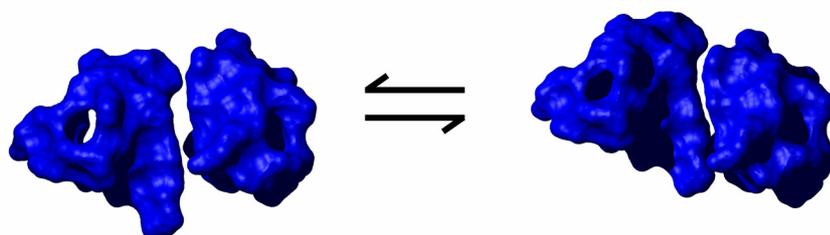


Fig. 1: Side-view of *E. coli* O157 B subunit in the asymmetric ‘excited’ (left) and symmetric ground states. These states exchange in solution on the millisecond time-scale.

Collaborators

David Bundle, Pavel Kitov, University of Alberta

Publications

Kitov, P.I., Shimizu, H., Homans, S.W. & Bundle, D.R. (2003) Optimization of tether length in non-glycosidically linked bivalent ligands that target sites 2 and 1 of a *Shiga*-like toxin. *J. Am. Chem. Soc.* **125**, 3284-3294.

Yung A., Turnbull, W.B., Kalverda, A.P., Thompson, G.S., Homans, S.W., Kitov, P. & Bundle, D.R. (2003) Large scale millisecond inter-subunit dynamics in the B subunit homopentamer of the toxin derived from *E. coli* O157. *J. Am. Chem. Soc.* **125**, 13058-13062.

Funding

This work has been funded by the Wellcome Trust and BBSRC

Collecting bacterial membrane proteins

Massoud Saidijam, Georgios Psakis, Joanne Clough, Christopher Hoyle, Scott Morrison, Chris Potter, Atif Abu-bakr, Simon Baumberg, Alison Ward, John O'Reilly, Nicholas Rutherford, Mary Phillips-Jones and Peter Henderson

Introduction

A general strategy for the amplified expression in *Escherichia coli* of membrane transport and receptor proteins from other bacteria has been devised. As an illustration, we report the cloning of the putative α -ketoglutarate membrane transport gene from the genome of *Helicobacter pylori*, overexpression of the protein tagged with RGS(His)₆ at the C-terminus and its purification in mg quantities. The retention of structural and functional integrity was verified by circular dichroism spectroscopy and reconstitution of transport activity. This strategy for overexpression and purification is extended to many other membrane proteins from Gram-positive and Gram-negative bacteria, including pathogens. The amounts of protein obtained are sufficient for crystallisation trials, NMR and other biophysical studies.

Introduction of the gene encoding a putative transport protein into the plasmid pTTQ18 vector

In order to clone the 1.3 kb *jhp0334* gene into the pTTQ18 vector (Fig. 1), the plasmid pNorAH₆ (pTTQ18 containing the gene *norAH*₆), was isolated from *E. coli* strain BLR and digested with the restriction endonucleases *EcoRI* and *PstI* to yield two DNA fragments of 4.56kb and 1.2kb. The larger fragment of 4.56kb (pTTQ18 with the RGS(His)₆ coding DNA sequence) was isolated from an agarose gel.

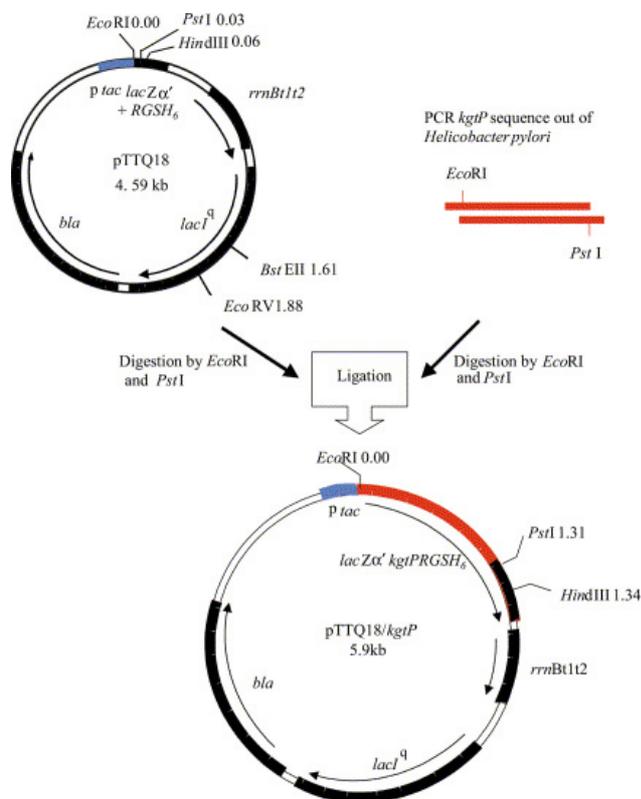


Fig. 1. Cloning strategy for membrane proteins using plasmid pTTQ18.

The gene putatively encoding an α -ketoglutarate transport protein was amplified from *H. pylori* genomic DNA, using mutagenic oligonucleotides designed to introduce an *EcoRI* site at the 5' end and a *PstI* site at the 3' end for subsequent ligation with the 4.6kbp pTTQ18/RGS(His)₆ fragment. The PCR product was isolated and then digested with *EcoRI* and *PstI*.

Ligation reactions were performed using the *EcoRI*-*PstI* digested gene and pTTQ18-RGS(His)₆ fragments at various vector:insert molar ratios. After the ligated product was transformed into *E. coli* XL1-blue, recombinant clones were selected on LB plates containing carbenicillin. Plasmid DNA was prepared from carbenicillin-resistant colonies and subjected to restriction analysis with the uniquely-cutting restriction enzymes *HindIII* and *NcoI*, and automated DNA sequencing of the 5' end to confirm the correct size of the plasmid and the presence of the gene. The size and DNA sequence of

the insert in the new plasmid (Fig. 1) were confirmed. The plasmid was then transformed into *E. coli* strain BL21(DE3) for expression studies.

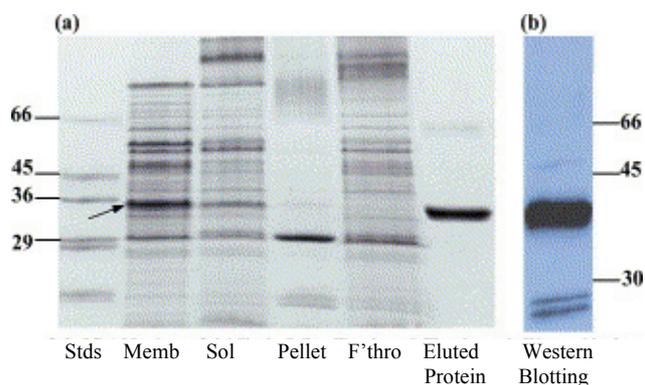


Fig.2. (a) Purification (and identification) of the *H. pylori* KgtP(His)₆ protein. Membrane preparations were solubilised in 1.5% dodecyl- β -D-maltoside. Samples of the original preparation ('Memb'), the soluble ('Sol') and insoluble ('Pellet') material were examined by SDS-PAGE (Coomassie-stained). A sample of the material that failed to adhere to Ni NTA-agarose ('F'thro') and the Ni NTA-bound protein subsequently eluted by 200mM imidazole ('Eluted protein') were processed in the same gel. (b) Confirmation of the identity of the KgtP(His)₆ protein by Western blotting.

Solubilisation and purification of histidine-tagged transport protein

IPTG-induced cells showed transport of [¹⁴C]- α ketoglutarate absent from uninduced cells, and membranes from induced bacteria contained an extra protein of apparent Mr 35k (Fig. 2). In the case of cloned 'Two-component system' (TCS) receptor proteins their ATP-dependent self-phosphorylation activity can similarly be checked in membrane preparations from IPTG-induced cells and uninduced cells. The overexpressed KgtP(His)₆ protein in the membrane (Fig. 2) was purified using Ni-chelate affinity chromatography. The protein migrated at 35 kDa in the eluted fractions (Fig. 2). Its identification as KgtP(His)₆ was reinforced by a positive reaction to an RGS(His)₆ antibody (Fig. 2b) and N-terminal sequencing (MNSHMNPQIQ). Also, the purified protein was reconstituted into liposomes and its transport activity confirmed.

Conclusions. The same strategy, with minor modifications in growth, solubilisation and purification conditions, has been used for overexpression of 37 other membrane proteins from *H. pylori*, *Campylobacter jejuni*, *Neisseria meningitidis*, *Brucella abortus*, *Staphylococcus aureus*, *Bacillus subtilis*, *Rhodobacter sphaeroides*, *Escherichia coli* and *Streptomyces coelicolor*. The yields of purified protein are sufficient for X-ray crystallography, NMR studies and other variety of biophysical techniques.

Publications

Morrison, S., Ward, A., Hoyle, C.J. & Henderson, P.J.F. (2003) Cloning, expression, purification and properties of a putative multidrug resistance efflux protein from *Helicobacter pylori*. *Intern J. Antimicrob. Agents* 22, 242-249.

Patching, S.G., Herbert, R.B., O'Reilly, J., Brough, A.R. & Henderson, P.J.F. (2004) Low ¹³C-background for NMR-based studies of ligand binding using ¹³C-depleted glucose as carbon source for microbial growth: ¹³C-labelled glucose and ¹³C-forskolin binding to the galactose-proton symport protein GalP in *Escherichia coli*. *J. Am. Chem. Soc.* 126, 86-87.

Potter, C.A., Ward, A., Laguri, C., Williamson, M.P., Henderson, P.J.F. & Phillips-Jones, M.K. (2002) Expression, purification and characterisation of full-length histidine protein kinase RegB from *Rhodobacter sphaeroides*. *J. Mol. Biol.* 320, 201-213

Saidijam, M., Psakis, G., Clough, J., Meuller, J.K., Hoyle, C.J., Palmer, S.L., Morrison, S.M., Pos, M.K., Essenberg, R.C., Maiden, M.C.J., Phillips-Jones, M. E., Abu-bakr, A., Baumberg, S.G., Stark, M.J., Ward, A., O'Reilly, J., Rutherford, N.G. & Henderson, P.J.F. Collection and characterisation of bacterial membrane proteins (2003) *FEBS Lett.* 555, 170-175.

Collaborators Shun' Suzuki, Ajinomoto Co., Japan; Johan Meuller, Goteborg, Sweden; Johannes Stegmeier, Wurzburg, Germany; Xu Zhiqiang, Sydney, Australia; Martin Pos, Zurich, Switzerland.

Acknowledgements This work was funded by BBSRC, MRC, EU, SmithKline Beecham (SB, now GlaxoSmithKline, GSK) and by equipment grants from the Wellcome Trust and BBSRC. MS is grateful to the Iranian Ministry of Health for financial support.

Characterisation of ligand binding to membrane transport proteins from *Escherichia coli* by solid state NMR

Simon Patching, Adrian Brough, Richard Herbert, Steve Baldwin and Peter Henderson

Transport proteins from *E. coli*

Bacterial membranes contain transport proteins that mediate the uptake of specific nutrients into the organism from the environment, or that expel waste products, antibiotics and toxins. In some cases these proteins are homologous with human transporters that provide a route of entry for drugs into cells and tissues in addition to their natural substrates. The more accessible bacterial proteins can therefore serve as a model for their human counterparts where knowledge about substrate specificities and binding characteristics has implications for therapeutic development. Unlike the human transporters, the bacterial proteins can be overexpressed to provide sufficient quantities of protein that enables the use of techniques such as NMR for analysis. We have been using solid state NMR to characterise ligand binding to a range of transporters present in the inner bacterial membrane of *E. coli*, including a glucuronide transporter (GusB), two nucleoside transporters (NupC and NupG) and the galactose transporter (GalP).

Observation of ligand binding to membrane proteins by CP-MAS solid state NMR

The cross-polarisation magic-angle spinning (CP-MAS) solid state NMR experiment allows the direct observation of ligand binding to membrane proteins overexpressed in their native membranes without the need for their purification. The approach requires a substrate for the protein of interest that is labelled with an NMR active isotope, usually ^{13}C . In our experiments the substrates are equilibrated with non-energised *E. coli* inner membrane preparations in which a specific transporter has been expressed at levels of 20-50% of total protein. The CP experiment only produces a signal for substrate that has bound in the membranes and does not detect substrate that remains in solution; for example, the binding of methyl $[1-^{13}\text{C}]\text{-}\beta\text{-D-glucuronide}$ to GusB (Fig. 1A).

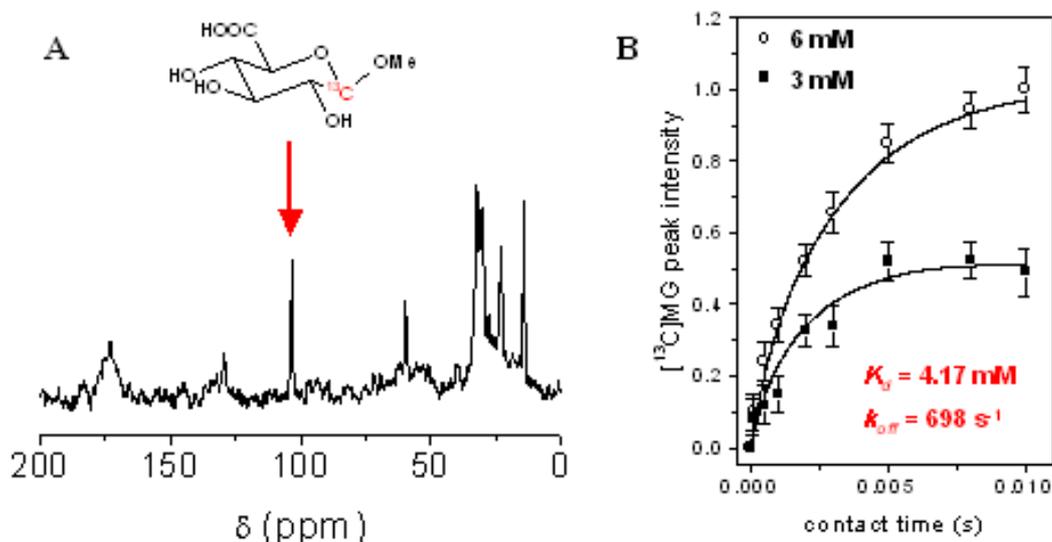


Fig. 1. (A) A ^{13}C CP-MAS NMR spectrum of *E. coli* inner membranes overexpressing the glucuronide transporter GusB and containing 6 mM methyl $[1-^{13}\text{C}]\text{-}\beta\text{-D-glucuronide}$ ($[^{13}\text{C}]\text{MG}$). (B) Peak intensities for two different concentrations of $[^{13}\text{C}]\text{MG}$ in GusB membranes from spectra recorded with a range of contact times and showing the best fitting simulations of CP profiles for the given values of K_d and k_{off} .

Methods to quantify substrate affinities for membrane proteins by CP-MAS NMR

When the CP-MAS NMR spectra are recorded over a range of CP contact times, we have shown that the shape of the resultant substrate peak intensity profile is sensitive to changes in binding affinity. Simulations of CP intensity profiles can be fitted to the experimental data to characterise substrate binding in terms of the binding constant (K_d) and of the rate constant for dissociation from the binding site (k_{off}) (Fig. 1B). In some cases this is the only method that can be used to obtain such information for a particular transporter. We have also developed a competitive displacement approach that allows the determination of the binding constants for unlabelled ligands when at least one suitable labelled substrate has been identified, therefore allowing the screening of an unlimited number of compounds.

Elimination of non-specific signals and reduction of obtrusive backgrounds

When using more hydrophobic substrates in these experiments, *e.g.* [1'- ^{13}C]uridine with the nucleoside transporters, non-specific interactions may contribute to the substrate signal. This complicates the measurement of substrate affinities and so we have applied a spectral editing experiment to selectively eliminate the non-specific component of the signal.

The ^{13}C label introduced into the chosen ligand may produce a signal that overlaps with the natural abundance ^{13}C background from the membranes in the ^{13}C NMR spectrum, which interferes with its detection. We have overcome this problem for observing the binding of the inhibitor [22- ^{13}C]forskolin to the galactose transporter, GalP, by growth of the producing organism in medium containing ^{13}C -depleted D-glucose as the carbon source (Fig. 2).

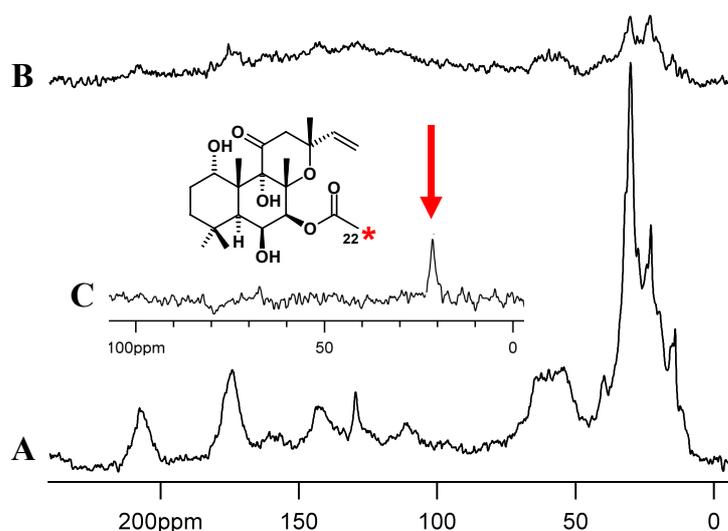


Fig. 2. ^{13}C CP-MAS NMR spectra of *E. coli* inner membranes overexpressing the galactose transporter GalP from cultures grown in medium containing normal D-glucose (A) and ^{13}C -depleted D-glucose (B) as the carbon source. (C) A spectrum of the ^{13}C -depleted membranes with the addition of 167 μM [22- ^{13}C]forskolin and following subtraction of the spectrum in B.

Publications

Patching, S. G., Middleton, D. A., Henderson, P. J. F. and Herbert, R. B. (2003) The economical synthesis of [2'- ^{13}C , 1,3- $^{15}\text{N}_2$]uridine; preliminary conformational studies by solid state NMR. *Org. Biomol. Chem.*, **1**, 2057-2062.

Patching, S. G., Herbert, R. B., O'Reilly, J., Brough, A. R. and Henderson, P. J. F. (2004) Low ^{13}C -background for NMR-based studies of ligand binding using ^{13}C -depleted glucose as carbon source for microbial growth: ^{13}C -labelled glucose and ^{13}C -forskolin binding to the galactose- H^+ symport protein GalP in *Escherichia coli*. *J. Am. Chem. Soc.* **126**, 86-87.

Patching, S. G., Brough, A. R., Herbert, R. B., Rajakarier, J. A., Henderson, P. J. F. and Middleton, D. A. (2004) Substrate affinities for membrane transport proteins determined by ^{13}C cross-polarisation magic-angle spinning NMR. *J. Am. Chem. Soc.* (in press).

Collaborators

Parts of this work were performed in collaboration with Dr David Middleton at UMIST, who we thank for valuable contributions to the project. Thanks also to Mr John O'Reilly for culturing bacteria in a fermenter and for membrane preparations.

Funding

This work was funded by the BBSRC, EPSRC and MRC.

Characterization of the coronavirus RNA binding protein

Brian Dove, Steve Emmett, Mark Reed, Kelly Anne Spencer, Jae Hwan You, and Julian Alexander Hiscox.

Introduction

Coronaviruses are a group of positive strand RNA viruses, which replicate in the cytoplasm of the infected cell and cause a variety of respiratory and gastrointestinal illnesses. In 2003 coronaviruses were identified as the causative agent of severe acute respiratory syndrome (SARS). The death rate following infection approached almost 10%. Work in our laboratory has focused on characterizing the coronavirus nucleoprotein (N protein), which is involved in binding viral RNA, the regulation of viral replication and also signaling in the host cell. With regard to the latter point, our work has indicated that coronaviruses interact with the nucleolus of the host cell and also control the cell cycle (Fig. 1). Whilst this report summarizes our recent work on the avian coronavirus N protein we are now expanding our research to study the SARS-coronavirus N protein.

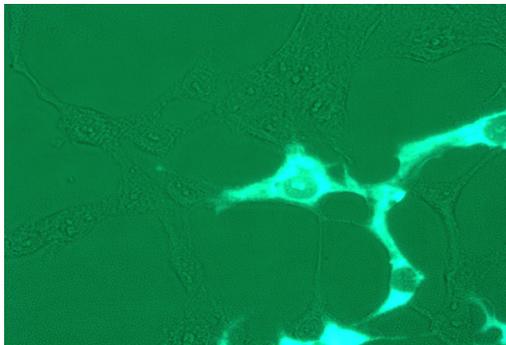
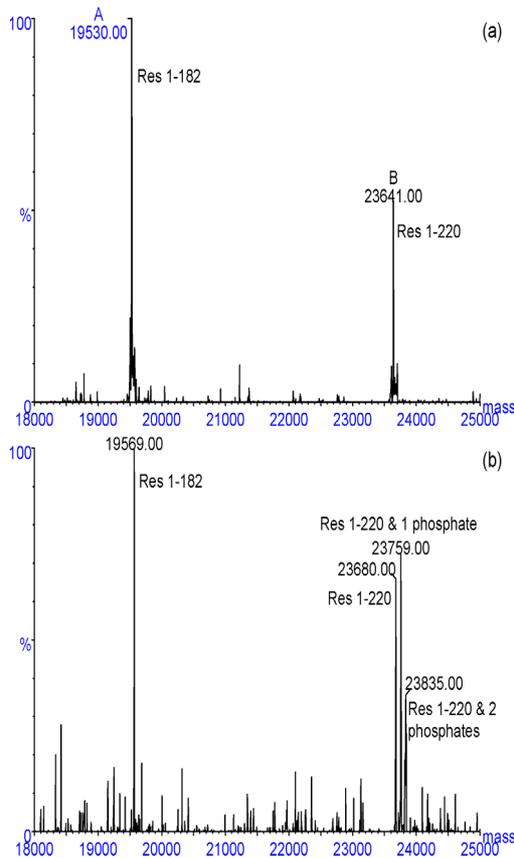


Fig. 1. Combined light and fluorescent image of cells expressing a GFP-coronavirus N fusion protein. The GFP-N fusion protein localizes not only to the cytoplasm, but in many cells also to a sub-nuclear structure called the nucleolus.

Phosphorylation of N protein has been predicted to play a role in RNA binding. To investigate this hypothesis we examined the kinetics of RNA binding between non-phosphorylated and phosphorylated infectious bronchitis virus (IBV) N protein with non-viral and RNA using surface plasmon resonance (BIAcore). Mass spectroscopic analysis of N protein identified phosphorylation sites that were proximal to RNA binding domains (Fig. 2).

Fig. 2. Deconvoluted electrospray mass spectra of polypeptides released by endoproteinase Glu-C digestion of (a) $N_{E.coli}$ and (b) N_{SF9} proteins.



In (a), peaks corresponding to residues 1-182 (theoretical mass 19531.7 Da) and 1-220 (theoretical mass 23641.3 Da) are evident. In (b), peaks corresponding to the same polypeptides have masses that have increased by around 40 Da. This mass difference was measured to be 44 Da from the original peak envelopes. In addition, the polypeptide spanning residues 1-220 is present in three different forms, each separated by around 80 Da. This is consistent with up to two phosphorylations in this polypeptide. The absence of these signals in the peptide corresponding to residues 1-182 implies that the sites of phosphorylation are between residues 183-220.

Kinetic analysis, using surface plasmon resonance, indicated that non-phosphorylated N protein bound with the same affinity to viral RNA as phosphorylated N protein. However, phosphorylated N protein bound to viral RNA with higher binding affinity than non-viral RNA, indicating that phosphorylation of N protein determined the recognition of virus RNA. The data also indicated that a known N protein binding site (and involved in transcriptional regulation) consisting of a conserved core sequence present near the 5' end of the genome (in the leader sequence) functioned by promoting high association rates of N protein binding. Further analysis of the leader sequence indicated that the core element was not the only binding site for N protein and that other regions functioned to promote high affinity binding.

As an immunogen of the coronavirus, the N protein is a potential antigen for the serological monitoring of IBV. The Beaudette strain of IBV was produced and purified from *E.coli* as well as Sf9 (insect) cells, and used for the coating of enzyme-linked immunosorbent assay (ELISA) plates. Our data indicated that N protein purified from *E.coli* was more sensitive to anti-IBV serum than the protein from Sf9 cells. The recombinant N protein did not react with the anti-sera to other avian pathogens, implying that it was specific in the recognition of IBV antibodies. In addition, the data from the detection of field samples and IBV strains indicated that using the recombinant protein as coating antigen could achieve an equivalent performance to an ELISA kit based on infected material extracts as a source of antigen(s). ELISAs based on recombinant proteins are safe (no live virus), clean (only virus antigens are present), specific (single proteins can be used) and rapid (to respond to new viral strains and strains that cannot necessarily be easily cultured). This technology can be adapted for the rapid detection of many different viruses. An ELISA kit is now marketed and manufactured by our industrial collaborator, Guildhay Ltd.

Collaborators

Andrew Gill, Institute for Animal Health.
Guildhay Ltd.

Publications

Chen, H., B. Coote, S. Attree, & Hiscox, J. A. (2003) Evaluation of a nucleoprotein based ELISA for the detection of antibodies against infectious bronchitis virus. *Avian Path.* 32, 519-526.

Hiscox, J. A. (2003) The interaction of animal cytoplasmic RNA viruses with the nucleus to facilitate replication. *Virus Res.* 95, 13-22.

Funding

This work was funded by the BBSRC and Guildhay, Ltd.

Molecular modelling of SH2 domain-peptide interactions

Stephen Campbell and Richard Jackson

Before carrying out docking studies on a potential target, it is important to gain a thorough understanding of the system. We have been studying SH2 domain-peptide interactions as a model system using a variety of molecular modelling techniques.

SH2 domains are highly homologous phosphotyrosine-binding motifs of approximately one hundred amino acids found in a range of signal transduction proteins. Uncontrolled signalling through protein tyrosine pathways involving these SH2 domains can lead to a variety of disease states, meaning that the search for novel agents which can interrupt such pathways has become an intense field of research. Although there has been much evidence to suggest that interactions involving SH2 domains occur in a substrate-specific manner, it has also recently been argued that the family of domains may be too highly homologous and structurally conserved to allow for the design of specific inhibitors lacking undesired effects.

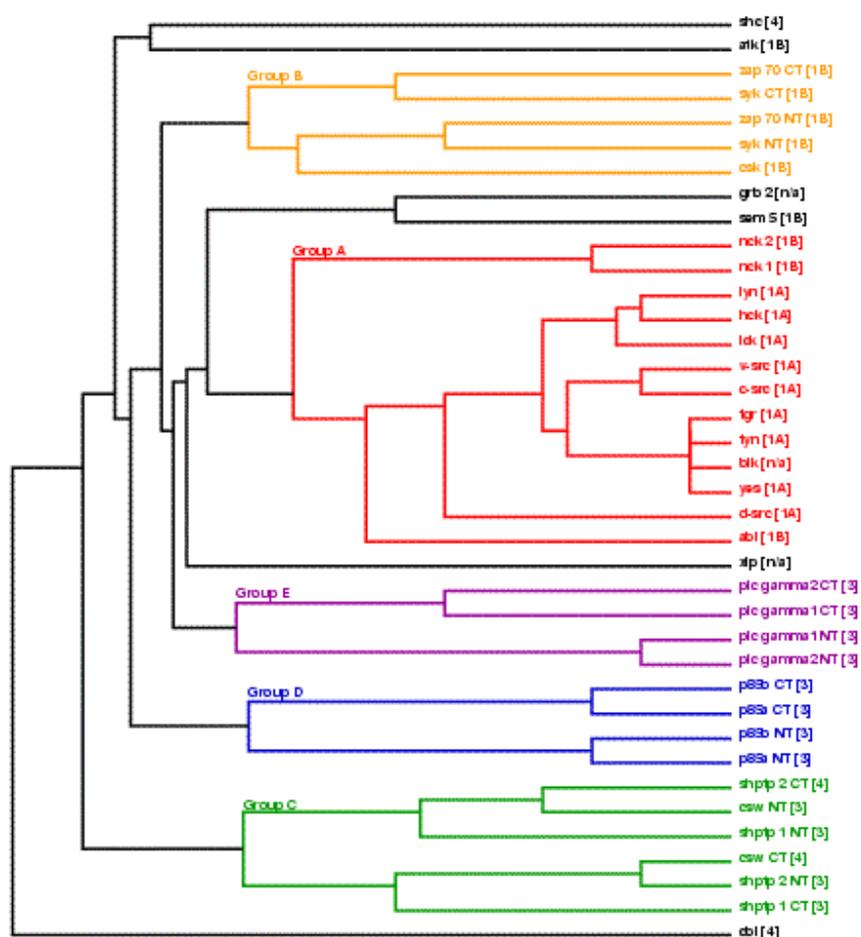


Fig 1: Clustering of Binding Sites. We have used the experimentally defined protein-ligand contacts to define the protein binding site, allowing the clustering of the domains according to amino acid conservation in the binding site.

The aim of our work has been to gain a greater understanding of the SH2 domain system by investigating these similarities and characterising the diversity within the family. This may allow us to determine residues which may be potentially exploitable in terms of structure-based drug design. The main techniques we have been using include structural and sequence

alignment, clustering of binding site residues (See Fig. 1) and conservation studies (see Fig. 2).

This study of conservation and diversity within the SH2 domain allows a greater understanding of the system, revealing regions that may be important in SH2 domain interactions.

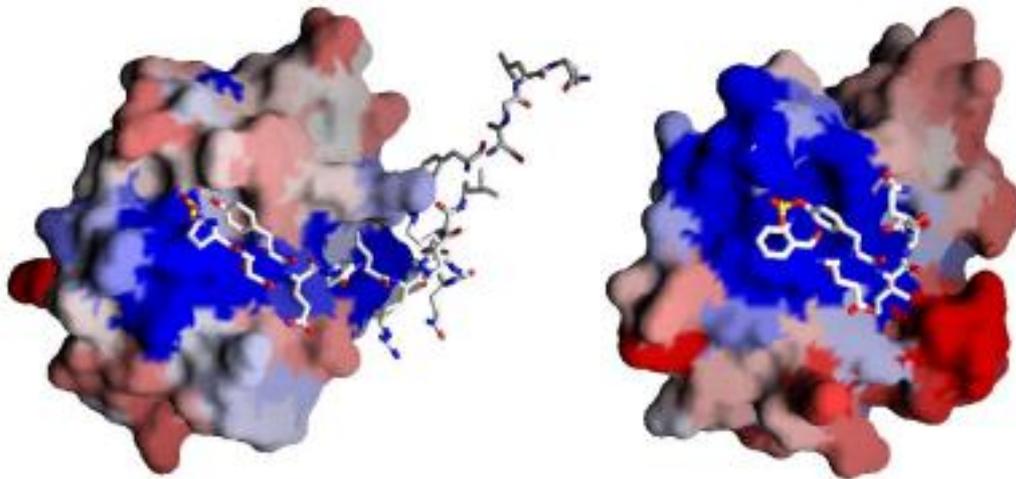


Figure 2: Conservation maps for SH2 domain subgroups Conservation is displayed on a representative SH2 domain structure for each sub-group. Blue indicates a high level of conservation and red indicates low conservation.

Publications

Campbell, S. J. & Jackson, R. M. (2003) Diversity in the SH2 domain family phosphotyrosyl peptide binding site. *Protein Eng.* **16**, 217-227.

Funding

We wish to acknowledge the support of the BBSRC for this work.

Flexligdock: A flexible ligand docking tool

Peter Oledzki, Paul Lyon and Richard Jackson

The algorithm *Flexligdock* is being developed to provide a comprehensive flexible ligand docking tool for small molecule docking to proteins. The program is based on QFIT and uses a probabilistic sampling method in conjunction with a molecular mechanics force field to predict the ligand binding mode (Fig. 1).

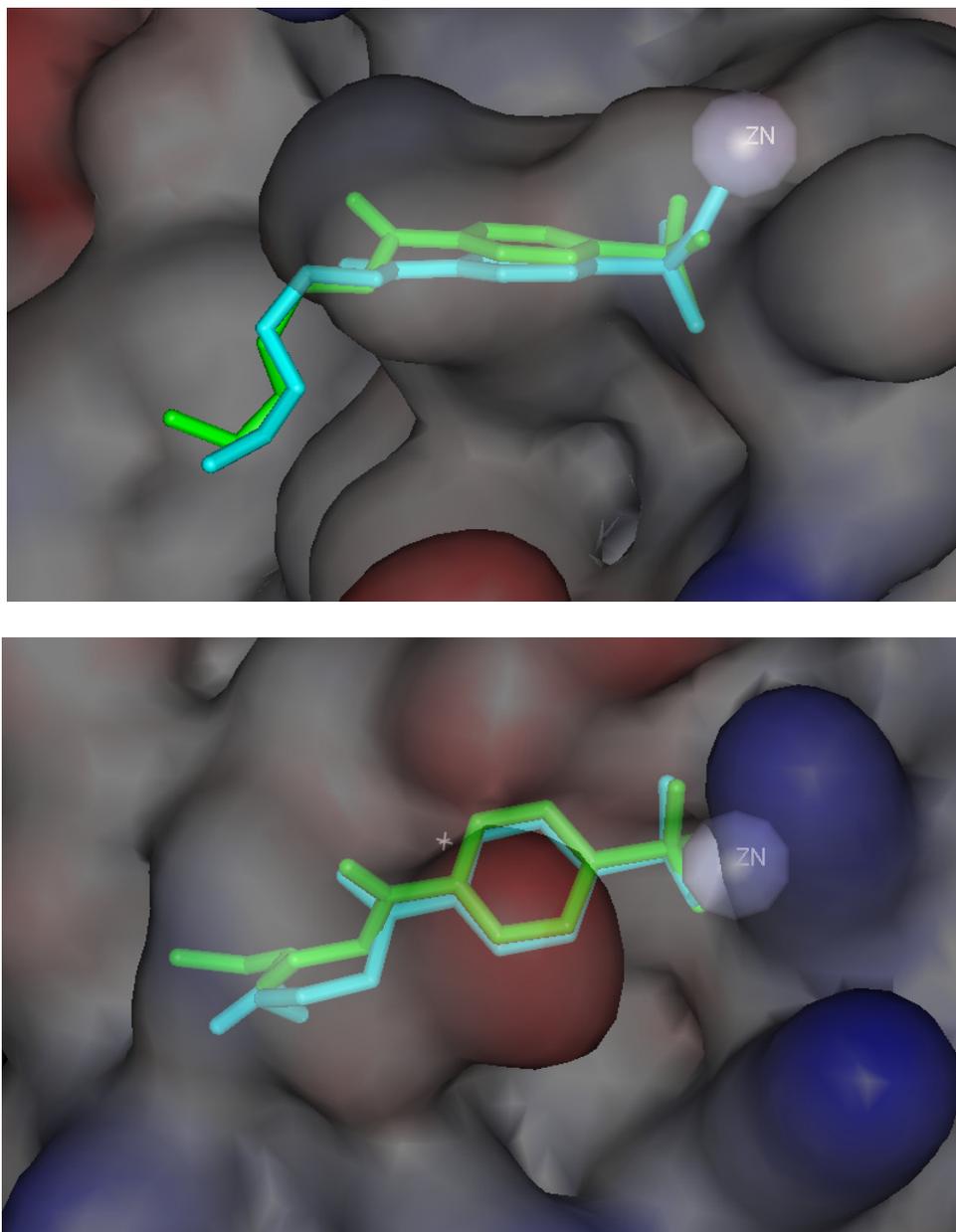


Fig. 1 Ligand conformations of 4-sulphonamide-[1-(4-aminobutane)]benzamide in complex with carbonic anhydrase II. The experimental crystal conformation is shown in light blue and the successfully docked solution is shown in green.

The method fragments a ligand to produce a series of rigid fragments, and utilises an interaction point methodology to map the ligand fragments onto an interaction energy grid map of the protein target (Fig. 2). The method then uses an incremental construction method to build the ligand fragment by fragment in the protein binding site. This stage uses torsion angle sampling to search for low energy ligand conformations.

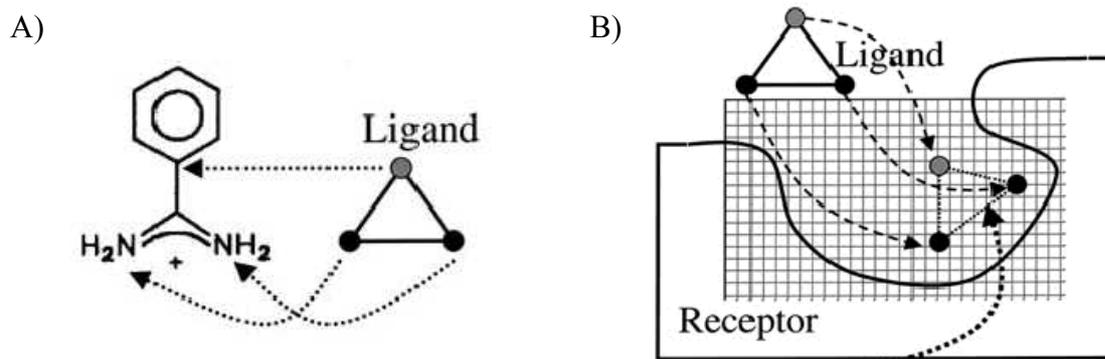


Fig. 2: A) Ligands can be represented by a series of three functional groups that make up triplets. B) Ligand triplets are matched to the receptor triplets and ranked to find the best conformations.

The algorithm has been parameterised on a data set of 46 protein-ligand complexes obtained from a recently released docking data set. The parameterisation data set contained a structurally diverse set of proteins and a variety of ligands that contained between 0-23 torsion angles.

Currently, a validation data set of 200 protein-ligand complexes is being used to validate *Flexligdock* and allow comparison against other existing protein-ligand docking algorithms.

Funding

We wish to acknowledge support of the BBSRC.

Searchable database containing comparisons of ligand binding sites at the molecular level for the discovery of similarities in protein function

Nicola Gold and Richard Jackson

Structural genomics projects produce large amounts of data, some of which are solved structures of hypothetical proteins of unknown function. The aim of this project is to aid the characterisation of these proteins by structure based prediction of protein function.

The current project extends previous work which demonstrated a method based on geometric hashing to compare the structures and properties of ligand binding sites and assess the extent of their similarity. In particular the binding site of the phosphate moiety of the large class of nucleotide ligands (ATP/ADP, GTP/GDP, FAD, NAD) was studied and is now being extended to include the entire ligand binding sites of these proteins as well as others.

This project uses a geometric hashing technique (described previously) which gives a score of similarity, a superposition, RMSD and equivalenced atoms for each pair of compared binding sites. These data are stored in a World Wide Web accessible database. The prototype database is searchable with a PDB code and ligand information (such as ligand name, number and chain). Submission of these data rapidly returns a ranked list of similar ligand binding sites (above a certain similarity score cut-off) with the most similar at the top (Fig. 1).

Family Relative		Super-Family Relative		Fold Relative		Class Relative		Unrelated/Unclassified	
<input type="button" value="Submit"/>					<input type="button" value="Reset"/>				
PDB ID	Ligand Name	Ligand Number	Chain	Score	Score 2	Scop Class	Scop Class		
Query: 1dg1	GDP	1999	0	83	1	c.37.1.8			
<input type="checkbox"/> 1d2e	GDP	1301		79	0.840	c.37.1.8			
<input type="checkbox"/> 1d8t	GDP	999		78	0.839	c.37.1.8			
<input type="checkbox"/> 1got	GDP	355	G	53	0.477	c.37.1.8			
<input type="checkbox"/> 1g16	GDP	202		53	0.596	c.37.1.8		a.66.1.1	
<input type="checkbox"/> 1fqk	GDP	361		53	0.477	c.37.1.8		a.66.1.1	
<input type="checkbox"/> 1doa	GDP	198		50	0.625	c.37.1.8			
<input type="checkbox"/> 1hh4	GDP	1190	B	50	0.667	c.37.1.8		c.10.1.2	
<input type="checkbox"/> 1ng1	GDP	900		49	0.570	c.37.1.10			
<input type="checkbox"/> 1hh4	GDP	1190	A	49	0.598	c.37.1.8			
<input type="checkbox"/> 1i4l	GDP	200		48	0.686	c.37.1.8			
<input type="checkbox"/> 1git	GDP	355		47	0.423	c.37.1.8		a.66.1.1	
<input type="checkbox"/> 1as2	GDP	355		47	0.423	c.37.1.8		a.66.1.1	
<input type="checkbox"/> 1e2f	ADP	302	A	46	0.622	c.37.1.1			
<input type="checkbox"/> 1ii9	ADP	590		46	0.451	c.37.1.10			
<input type="checkbox"/> 1mkj	ADP	600		46	0.648	c.37.1.9			
<input type="checkbox"/> 1ii0	ADP	590		45	0.429	c.37.1.10			
<input type="checkbox"/> 1k3c	ADP	541		41	0.427	c.91.1.1		g.41.2.1	
<input type="checkbox"/> 1p60	ADP	501		40	0.471				
<input type="checkbox"/> 1p60	ADP	401		40	0.494				
<input type="checkbox"/> 1p5z	ADP	301		40	0.500				

Fig. 1. Example ranked hits with the GDP binding site from Elongation factor Tu protein (1dg1) as the query. Family relatives from the SCOP database are coloured red while superfamily and fold relatives are coloured yellow and blue respectively. Unclassified proteins in SCOP 1.61 are coloured purple. The list of hits details the ligand name, number and chain along with the similarity score and the score normalised for binding site size. The final columns contain SCOP codes for the proteins.

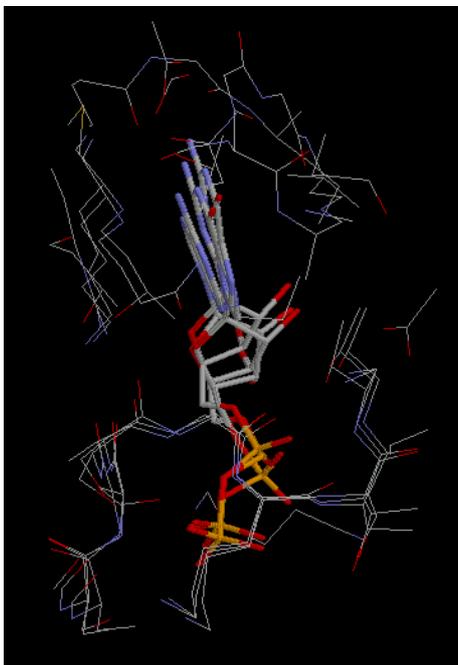


Fig. 2. Superposition of the GDP binding site of Elongation factor Tu protein (1dg1), the GDP binding site of signal sequence recognition protein Ffh (1ng1) and the ADP binding site of phosphoenolpyruvate carboxykinase (1k3c). 1dg1 and 1ng1 are from different SCOP families sharing the same superfamily and fold, whereas 1k3c is classified in a different superfamily and fold group.

The similarity scores and SCOP (Structural Classification Of Proteins) codes of these ‘hits’ are also displayed and each hit is coloured according to its similarity to the query’s overall fold and SCOP family, thus showing the nature and relative interest of the hits in a visual and easily assimilated layout. In the example shown in Fig. 1, a highly similar binding site has been found which has a different overall fold to the query and is easily identifiable as such because of its colour. The next stage of the program calculates optimal superpositions of the binding site and ligand of interesting hits on the query allowing further examination and visualisation with molecular graphics packages (Fig. 2). A multiple alignment of structurally equivalenced atoms is also provided.

Future work will see this method and database extended to include ligand binding sites from other proteins and possibly other functionally useful annotations such as PROSITE patterns (database of protein families and domains) and CATH codes (protein structure classification database). We will also implement a statistical measure of similarity. This database and comparison method can then be used to discover new similarities indicating functional relationships between proteins and may uncover binding site similarities in proteins previously thought to be unrelated. It may also be possible to develop a classification scheme for ligand binding sites useful to researchers attempting to identify possible functional relatives.

Publication

Brakoulias A. & Jackson RM. (2004) Towards a structural classification of phosphate binding sites in protein-nucleotide complexes: an automated all-against-all structural comparison using geometric matching. *Proteins Struct. Funct. and Bioinformatics*. In press.

Funding

We wish to acknowledge the support of the BBSRC.

Structural modelling of protein-DNA interactions

Richard Gamblin and Richard Jackson

A large number of *in silico* models of protein-DNA interactions reflect a simplistic view of DNA recognition, whereby proteins recognise a specific sequence of nucleotides. These models use aligned binding sites to generate a position specific scoring matrix (PSSM), in what is essentially a table giving the prevalence of nucleotides at each position in the section of bound DNA. This can then be used to characterise and identify other such binding sites in DNA sequences.

This representation of recognition is, however, far simpler than present *in vivo*. DNA is a negatively charged bi-polymer projecting into three dimensional space, rather than a one dimensional series of letters, and it is this that proteins must negotiate to distinguish specific regions. The difference between the PSSM models of transcription factor binding and the *in vivo* reality provides an explanation as to why PSSM binding site predictions are plagued with erroneous predictions. Recent studies into protein-DNA complex structures have indicated the presence of trends in amino acid-DNA base interactions, and preliminary results from the latest models of protein-DNA recognition, which utilise this type of information, appear promising.

We have developed a novel method with the aim of quantifying structural features that confer specific binding properties not evident from sequence similarity alone. Using a catalogue of hydrogen bonding and non-bonded contact patterns from a non-redundant set of protein and DNA complex structures, an overall statistical knowledge-based model was developed to represent specific amino acid-DNA base/ backbone interactions. This was applied to create new PSSM-type models, termed structurally derived matrices (SDMs).

Assessment of the functional differences between the SDM and PSSM models revealed that SDM predictions were significantly poorer than the equivalent PSSM predictions. The SDMs correctly predicted binding sites as the top 'hit' in only 2% of cases, compared with 58% of cases by the equivalent PSSM for a diverse set of experimentally characterised binding sites (see Fig. 1).

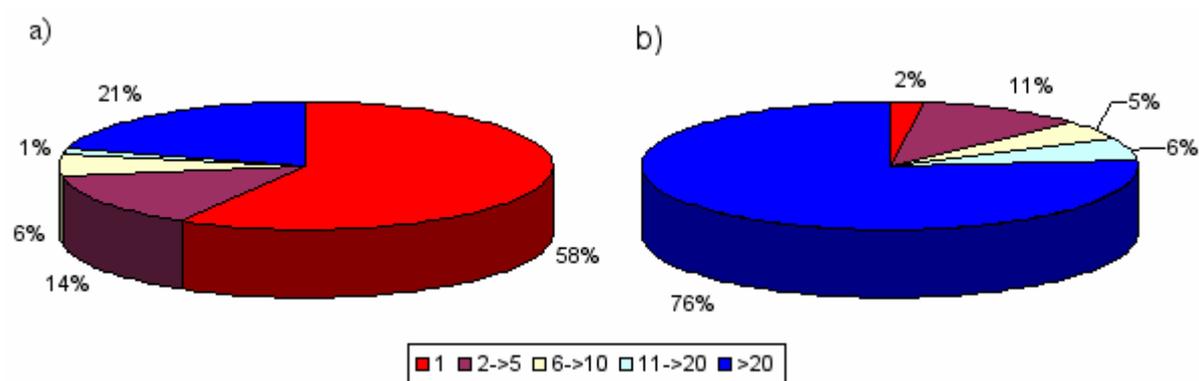


Fig. 1. Binding site predictions by a) the PSSM model and b) the SDM model. Correctly predicted binding sites as the top 'hit' appear in red, correctly predicted binding sites in the top 2 to 5 'hits' appear in maroon, those in the top 6 to 10 appear in yellow, those in the top 11 to 10 appear in light blue and finally binding sites predicted outside the top 20 'hits' appear in blue.

Our findings suggest that, while there is clearly some information to be obtained from analysis of these intermolecular interactions, application at the amino acid–DNA base level to a matrix-type model is much worse than PSSM models currently available. Indeed, although PSSM models have their limitations, they do perform very well on short sections of DNA sequence when representing a well defined binding site.

Recently we have shifted our broad spectrum analysis and focussed our attention on the GATA family of transcription factors in *Arabidopsis thaliana*. The members of this class IV zinc finger proteins, are typified by the GATA motif that they selectively bind. PSSMs representative of this DNA recognition site are available for mammalian systems, however consideration of the length of sequence to be searched in *A. thaliana*, coupled with the abbreviated nature of the motif, mean that binding site predictions made with the PSSM model could never achieve statistical significance.

Our most recent approaches have involved analysis of the *A. thaliana* GATA factor proteins. Preliminary results from multiple sequence alignments of *A. thaliana* protein sequences aligned against a mammalian GATA factor template, suggest they may interact with their cognate DNA in the same manner at the molecular level.

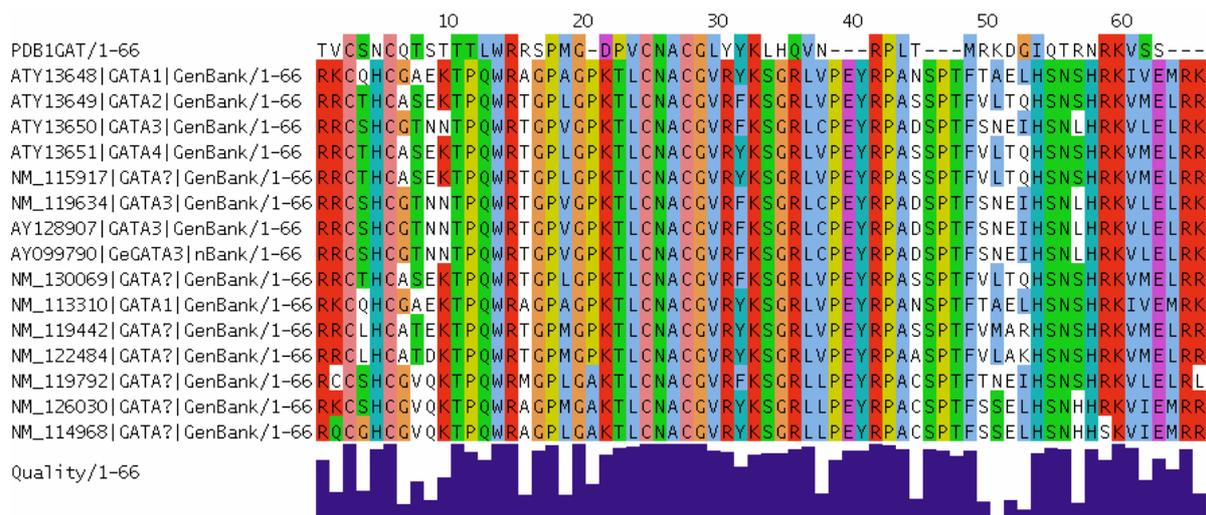


Fig. 2. Multiple sequence alignment of *Arabidopsis thaliana* sequences aligned against a mammalian template. The domain selected from the mammalian sequence is the zinc finger DNA binding domain as found in the 1gat PDB structure.

Collaborators

Professor P. M. Glimartin, Centre for Plant Science, University of Leeds, UK.

Funding

This work was supported by the BBSRC.

Structural studies of RNA-directed RNA polymerases from Hepatitis C virus and related viruses

Caitríona Dennis, Russell Green, Amy Cherry, Mark Saw, Elena Curti, Ben Linford, David Rowlands and Joachim Jäger

Hepatitis C virus and related viruses

Hepatitis C virus (HCV) belongs to the hepaciviruses in the family of *flaviviridae*, which are positive strand RNA viruses. Other members belonging to this family are yellow fever, Dengue fever, GBV-B and bovine viral diarrhoea virus. HCV infection is a major public health threat affecting an estimated 180 million carriers worldwide. About 80% of Hepatitis C infections lead to liver disease including cirrhosis and hepatocellular carcinoma. There is no vaccine for HCV, but the search for antiviral agents is ongoing, mainly focusing on three enzymes encoded by the viral genome, namely the bi-functional NS3 protease-helicase and the NS5B polymerase. Crystallographic studies on NS5B have had a major impact on the research towards the development of antivirals. In the Astbury Centre, we are addressing the structure/function relationships of the polymerase activity, providing a more detailed mechanistic understanding for nucleotide import, initiation, elongation and targeting these events for the development of antiviral inhibitors.

Crystallographic Studies on HCV NS5B

Structural studies have been undertaken on various forms of NS5b polymerase (genotype 1b: HC-J4 and HC-BK; genotype 1a, H77 and genotype 2, HC-J6). Crystals of these proteins (HC-J4 and BK and J4) have been grown in low ionic strength conditions - necessary for the introduction of various ligands into the active site of the enzyme. Crystal soaking experiments with J4 and BK have shown a single binding site for the incoming ribonucleoside-triphosphates (rNTP) into the active site cavity. UTP and GTP have been shown to bind in a similar manner, each interacting with the catalytic manganese metal ions

(Fig. 1). Combining the rNTP soaks with small single-stranded RNA oligonucleotides has provided a structural insight into the mechanism of template binding and initiation mechanism. In order to complete the understanding of the polymerase catalytic cycle, it is necessary to study the elongation mechanism. Co-crystallisation of the active polymerase with a primer:template duplex is currently underway.

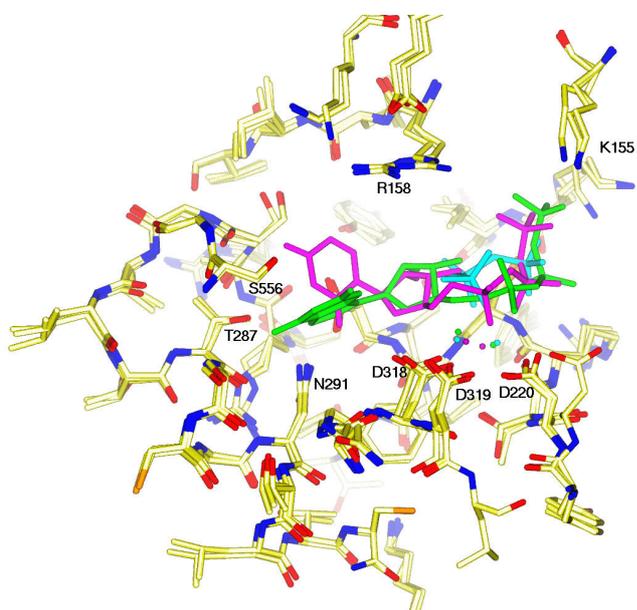


Fig. 1: An overlay of J4 with UTP (pink), GTP (green) and BK with GTP, phosphates only (cyan)

Collaborators

Onkar Singh and Pia Thommes, GlaxoSmithKline, Stevenage, UK.

Funding

This work is supported by GSK, BBSRC, MRC and Replizyme

Biochemical and structural studies of Human Rhinovirus 3D polymerase

Mark Saw, Gill Hart and Joachim Jäger

Human rhinovirus (HRV) is a major etiological cause of colds and upper respiratory tract cold symptoms. In over 50 % of these illnesses HRV can be found. They are also associated with complications in patients with other primary illnesses such as bronchitis, asthma and immunodeficiency conditions.

HRV, a member of the *picornaviridae* family, has a positive-sense, single-stranded RNA genome which is translated to a single, post-translationally cleaved polypeptide. At the C-terminal end of the polyprotein is an RNA dependent RNA polymerase (RdRp) responsible for the replication of the genome. Six conserved sequence motifs that are the hallmark of an RdRp domain are found in this sequence. Motifs A-D are involved in binding of the catalytic magnesium ions and in the recognition of the primer:template RNA. Motif E functions as the 'primer grip', and motif F is thought to be involved in ribonucleotide import.

In order to study the structure/function relationships of the HRV 3D polymerase protein in more detail, a cDNA clone of rhinovirus serotype 16 was obtained and the polymerase cloned into two vectors (pGEX 4, pET23a). Both the N-terminally GST-tagged protein, and the C-terminally hexa-His-tagged proteins are active on a poly-A RNA template using a Oligo U₍₁₂₎ primer. After further purification using gel filtration (purity >98% as determined by mass spectrometry and SDS-PAGE), crystallisation trials using the His-tagged protein were initiated using sparse matrix screens. Small needle-shaped crystals have been observed and conditions are being optimised to produce large diffraction grade crystals.



Single crystal of HRV 3D polymerase

Mutations in this motif F have been introduced and are being investigated using the polyA:oligoU primer extension and misincorporation assays. All mutants retain some activity. Further characterisation under steady state and single-turnover conditions are currently underway.

Funding

This work has been supported by a BBSRC Case Studentship with IDS Ltd. Boldon.

Functional and structural analysis of motif F in viral RNA polymerases

Elena Curti, Catherine Joyce and Joachim Jäger

In recent years, several lines of evidence have led to the proposal that all nucleic acid polymerases exhibit fundamental similarities in structure and mechanism of catalysis. This apparent similarity is preserved across the division of polymerases into classes based on whether the template and the product strand are DNA or RNA. Recent studies have shown that beyond several conserved motifs, there is no primary sequence conservation between polymerases. Multiple sequence alignments of viral polymerases, such as HIV reverse transcriptase, hepatitis C, bovine diarrhoea (BVDV), poliovirus and rhinovirus RNA dependent RNA polymerases (RdRp) revealed five conserved motifs (A-E) unique to these classes of enzymes. Motifs A to D map into the polymerase catalytic core, coordinate the catalytic metal ions and/or participate in primer:template or NTP binding. Motif E or “primer grip” is involved in positioning the 3'-end of the primer in the active site for an in-line attack of the α -phosphate group of the incoming nucleotide.

More recently, however, through careful sequence alignments and 3D structure comparisons, an additional motif (F) has been identified amongst RdRps and reverse transcriptases (RT). Several basic residues of motif F have been shown by co-crystallisation with nucleoside triphosphates to participate in nucleotide binding in HIV RT (Fig. 1a), Phi6, HCV and reovirus RdRp, leading to the hypothesis of a role in faithfully incorporating the correct NTPs. However the possibility that motif F could control binding of the RNA template and primer can not be excluded. Similarly to HCV and Phi6 polymerases, the guanidinium group of Arg72 in RT interacts with the α -phosphate oxygen atom of the incoming nucleotide and forms tight stacking interactions with the incoming substrate, whereas the side chain of Leu74 creates a platform for the template base which pair with incoming NTP.

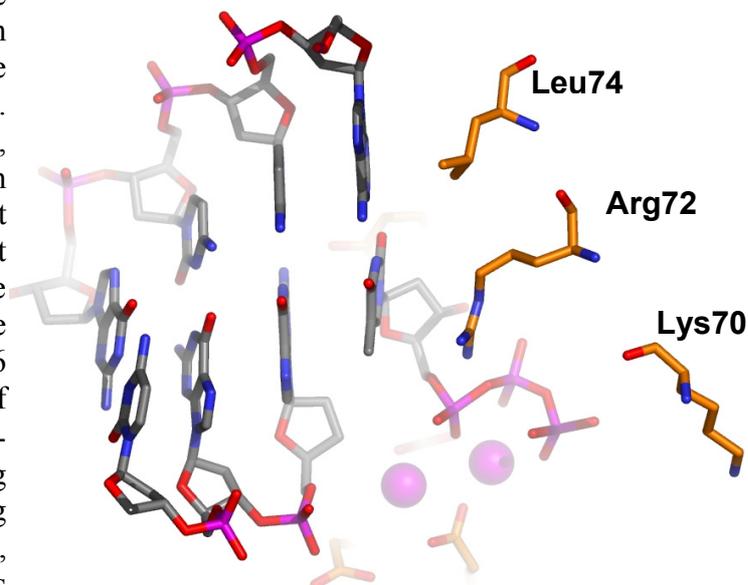


Fig. 1. Primer:template bound to the active site of HIV-1 RT.

The functional role of the conserved residues (Arg/Lys- Nnn_i -Arg- Nnn_i -Ile/Leu, where $i=1$ or 2) in motif F of BVDV RdRp has been probed by site directed mutagenesis. RdRp assays using poly(C)/oligo(G)₁₂ as a substrate (primer-dependent elongation) as well as a 192nt fragment of BVDV 3'UTR (*de novo* initiation) were used to investigate the overall activity of the mutants in comparison to wild-type. Replacement of Arg284 with Lys, Asn or Ala resulted in a substantial decrease in polymerase activity both in primer extension and *de novo* activities. Mutations at position 286 (corresponding to Arg72 in HIV RT) produced a 3-fold decrease in the activity of the enzyme. Less significant was the loss of activity for mutants at position 288. Steady-state and pre-steady-state kinetics were also performed to further investigate the effect of these mutations on the ability to bind rNTPs. Primer elongation experiments are being used to investigate the overall rate of nucleotide misincorporation.

Mutations in the conserved residues Lys155, Arg158 and Ile160 of motif F of HCV NS5B have also been introduced for crystallographic studies. Crystals of Lys155Ala, Lys158Ala and Ile160Ala NS5B have been grown using sitting drop vapour diffusion in crystallisation conditions similar to the wild type. The kinetic and structural characterisation of these mutants is currently underway.

Collaborators:

Joyce/Grindley labs, Yale University, New Haven, USA

Funding:

The work was supported by The Wellcome Trust 4yr PhD Programme in Structural Molecular Biology

The mechanism of initiation in Hepatitis C virus RNA polymerase

Amy Cherry, Pia Thoemmes and Joachim Jäger

Crystallographic structures of native Hepatitis C virus (HCV) polymerase (NS5B) have been reported by many pharmaceutical and academic research groups. In the Astbury Centre, we have determined the structures of 'J4' HCV polymerase alone, and in complex with ribonucleotides and incoming template RNA. Although the architecture of HCV NS5B displays the typical fingers, palm and thumb domains found in most polymerases, it also contains several more distinct elements. Two of these are a long β -hairpin in the thumb domain, and a long C-terminal loop ('arm'), which both block the natural exit from the active site. Comparison with the structure of reovirus and $\phi 6$ polymerases suggests that these elements have a role in positioning the RNA template correctly for *de novo* initiation at the 3'-terminus.

Recombinant proteins containing mutations in either the β -flap or C-terminal arm (see Fig. 1) have been synthesised in order to probe the functional role of these elements in detail. All the mutant proteins are more active than the wild-type enzyme in *in vitro* assays. This is true when using either primed homopolymeric templates, or heteropolymeric templates. Although HCV is believed to use a *de novo* mechanism of initiation *in vivo*, many people have reported heteropolymeric templates folding back on themselves to create a primer, in what is known as the 'copy-back' mechanism. Analysis of products from reactions using the heteropolymeric template shows that there is no difference in the quality of products produced, with all enzymes using both a copy back and a *de novo* mechanism of initiation. Further mutations introduced to disrupt interactions between the C-terminal arm and the fingers domain have a similar effect. Crystals of some of the β -flap mutants have been grown in conditions previously established in our laboratory (300mM NaCl, 16% PEG 4000, 10% glycerol, 5mM DTT). The crystal structure of a Tyr Phe mutant shows that even such a conservative and subtle mutation is sufficient to disrupt the interactions of the β -hairpin with the remainder of the polymerase domain.

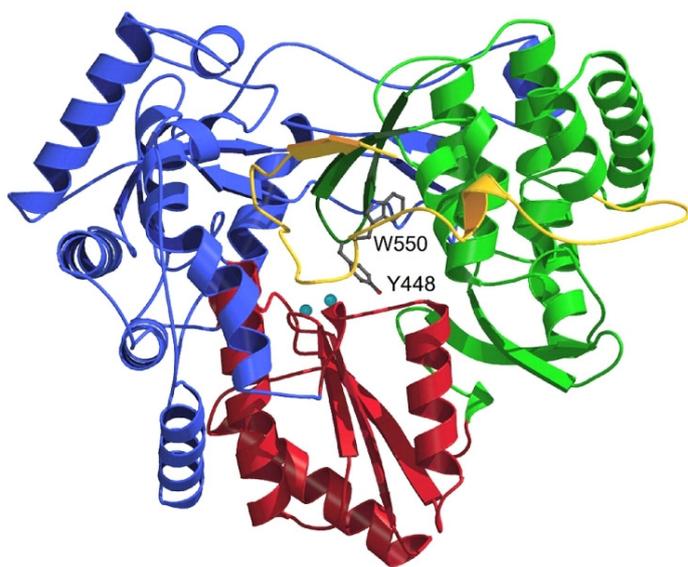


Fig. 1. Backbone structure of NS5B. The structure is coloured according to domain. Blue = Fingers, Red = Palm, Green = Thumb, Dark Green = Flap element, Yellow = C-terminal arm. Two of the mutated residues are indicated.

A recent advance in HCV research has been the development of the replicon assay, which allows viral replication activity to be studied in a cell culture system. When any of these mutations are introduced into the replicon system, replication *in vivo* is substantially decreased, with some mutations eliminating replication altogether.

The combined data suggest that mutations to either the β -hairpin or the C-terminal arm alter the binding of these structures at the exit of the active site, so that the RNA template binds more readily, initiation occurs more efficiently and the switch to elongation is made more rapidly. However, the replicon data show that this

increased rate of replication is not beneficial *in vivo*. This may be because initiation is allowed to begin internally by the mutated proteins, leading to the generation of incomplete genomes.

Collaborators:

Virology labs, GlaxoSmithKline, Stevenage, UK.

Funding:

This work was supported by an MRC Case Studentship with GSK.

Structural and biophysical studies on HIV-1 Nef

Caitríona Dennis, Mark Harris and Joachim Jäger

HIV-1 Nef

In primate immunodeficiency viruses (SIV and HIV), the *nef* gene is situated at the 3' end of the genome and encodes a protein of between 205 and 265 amino acids. Nef has many wide and varied functions within the cell but all contribute to an increased pathogenicity of the virus. It acts primarily as a binding protein/modulator and interacts with proteins such as cell surface receptors (MHC-I and CD4), where it is responsible for their down-regulation; and SH3 domains, where it alters signalling pathways.

Biophysical studies on HIV-1 Nef

Nef consists of two distinct domains – an N-terminal ‘anchor’ domain and a C-terminal ‘core’ domain. The anchor domain is N-terminally myristoylated, binds to the cell membrane and enables Nef to bind its membrane-bound targets. In order to interact with targets in the cytosol, Nef must remain soluble and conceal the highly hydrophobic surfaces and the lipidic modification within the protein. This form has been termed the ‘closed’ form, whereas the membrane-bound form represents the ‘open’ form. In the Astbury Centre, we have been able to express and purify both myristoylated (myr-Nef ‘open’ form) and non-myristoylated Nef (non-myr Nef ‘closed’ form). Initial observations on the two forms indicate differences in solubility - the myristoylated form remains quite insoluble and the non-myristoylated form is very soluble, thus reflecting the physiological properties of the two conformations of Nef. Further studies have shown the two forms also possess different biophysical properties. Circular dichroism (CD) reveals myr-Nef to contain a more helical structure than non-myr Nef, the additional helical region appearing in the anchor domain. Analytical ultracentrifugation (AUC), native SDS-PAGE, dynamic light scattering (DLS) and small angle X-ray scattering (SAXS) have shown different oligomerisation states for the two forms. Myr-

Nef exists as a mixture of monomeric and small order oligomers in equilibrium. Non-myr Nef, on the other hand, exists as a stable tetramer, under reducing conditions [Fig. 1], and a larger hexadecamer under non-reducing conditions. We can deduce from these results that specific oligomerisation of this small protein is essential for its wide and varied functions and this knowledge is an important prerequisite for the ongoing search for novel anti-virals.

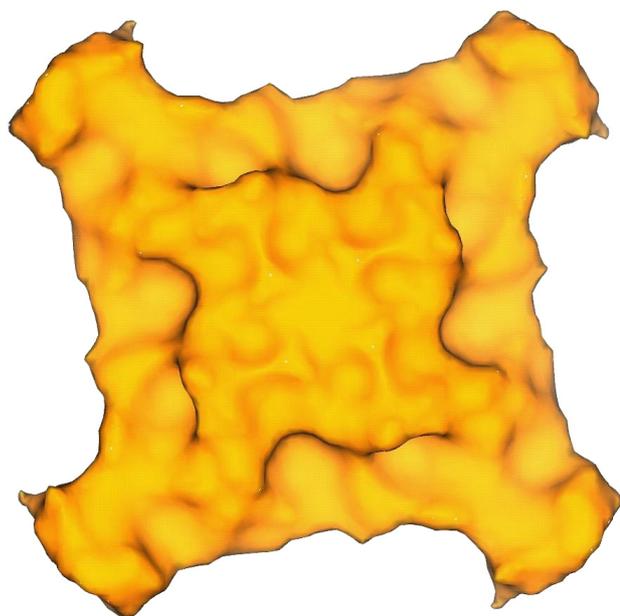


Fig. 1: A surface model of non-myr Nef tetramer obtained from SAXS

Structural studies on HIV-1 Nef

In order to gain crucial information on the function of Nef, it is essential that a three-dimension structure of full-length myristoylated Nef is determined. In the Astbury Centre, we have obtained an orthorhombic crystal form of myr-Nef that diffracts to 1.6Å on a Rigaku RU-300B rotating anode generator equipped with an R-axis IV++ imaging plate detector. Structural determination is currently underway.

Collaborators

Gunther Grossmann, SRS, Daresbury Laboratory, Warrington, UK
Andy Baron, Astbury Centre, Leeds, UK

Funding

This work is supported by the BBSRC.

NMR facility

Arnout Kalverda and Steve Homans

Overview of facility

The NMR facility is equipped with 750 MHz, 600 MHz and two 500 MHz Varian Inova NMR spectrometers. All instruments are setup to use ^1H , ^{13}C , ^{15}N and ^2H during normal operation. The 750 MHz spectrometer was installed in the spring of 2003 to extend the capabilities for NMR research with larger systems. In the coming year a cryoprobe will be added to the 750 MHz NMR spectrometer, to further enhance the sensitivity of the system.

Dynamics and thermodynamics of ligand binding

The binding of a ligand to a protein often results in changes in the dynamics of the protein upon binding. These dynamic changes are reflected in the thermodynamics of binding through their contribution to the entropy of the system. NMR can provide a residue specific view of changes in main chain and side chain dynamics upon ligand binding. An example is the carbohydrate binding to the pentameric Verotoxin B subunit. The apo-protein is found to be interconverting between the major axially symmetric pentamer and a low abundance, higher energy form. Broadened peaks in the ^1H - ^{15}N HSQC spectrum and significant relaxation dispersion are found in the residues that make up the subunit interface. This indicates that the differences between the two forms are found at the subunit interface and the higher energy state may be an asymmetric form of the pentamer. The high energy form is depopulated upon binding a novel bivalent analogue of the natural carbohydrate ligand which has been shown to bind across the subunit interface in adjacent binding sites on two subunits. ITC measurements show that the binding of the bivalent ligand is cooperative and that the first molecule to bind has a substantial unfavourable contribution to the entropy compared to subsequent molecules binding. This entropic cost arises from the suppression of the conformational equilibrium by the first ligand to bind across the subunit interface.

In mouse urinary protein, changes in side chain dynamics upon ligand binding have been probed with ^2H and ^{13}C relaxation methods. The results indicate that the increased rigidity of side chains in the binding pocket upon ligand binding is accompanied by increased mobility of side chains in adjacent positions. This partially compensates for the unfavourable entropic contribution from the side chains in the binding pocket.

Protein folds from residual dipolar couplings

Residual dipolar couplings are a powerful source of conformational restraints for the determination of the solution structure of proteins. An approach utilising only backbone residual dipolar couplings, combined with a minimal number of NH-NH NOE restraints, could result in a fast determination of the global fold, after the completion of backbone assignments. Using ubiquitin as an example, the fold could be determined from three residual dipolar couplings measured in two alignment systems. These couplings can, in principle, be measured from a single HSQC-type experiment. Efforts are underway to demonstrate this approach with a *de novo* fold.

NMR of protein folding

Native state hydrogen exchange has been used with the Colicin Immunity protein Im7 to show that the hydrogen exchange data provide information on the secondary structure of an intermediate state. To demonstrate whether hydrogen exchange occurred from the intermediate state, local fluctuations or global unfolding, data were compared between the native protein and a mutant (I72V). The mutation significantly destabilises the intermediate state relative to the unfolded state, but only slightly destabilises the native state. The

hydrogen exchange rates reflect the free energy difference with the state from which exchange occurs and the hydrogen exchange patterns shift with the changes in free energy upon mutation. Thus, residues that exchange from the intermediate have decreased hydrogen exchange rates, while residues that exchange through global unfolding have increased hydrogen exchange rates in this case. This method, using site directed mutagenesis to identify the state from which hydrogen exchange occurs, could have more widespread use in identifying the presence of secondary structure in intermediate states.

Publications

Giesen, A.W., Homans, S.W. & Brown, J.M. (2003) Determination of protein global folds using backbone residual dipolar couplings and long range NOE restraints. *J. Biomol. NMR* **25**, 63-71.

Kitov, P.I., Shimizu, H., Homans, S.W. & Bundle, D.R. (2003) Optimisation of tether length in non-glycosidically linked bivalent ligands that target sites 2 and 1 of a shiga-like toxin. *J. Am. Chem. Soc.* **125**, 3284-3294.

Yung, A., Turnbull, W.B., Kalverda, A.P., Thompson, G.S., Homans, S.W., Kitov, P. & Bundle, D.R. (2003) Large scale millisecond intersubunit dynamics in the B subunit homopentamer of the toxin derived from *Escherichia coli* O157. *J. Am. Chem. Soc.* **125**, 13058-13062.

Chaykovski, M.M., Bae, L.C., Cheng, M., Murray, J.H., Tortolani, K.E., Zhang, R., Seshadri, K., Findlay, J.H.B.C., Hsieh, S., Kalverda, A.P., Homans, S.W. & Brown, J.M. (2003) Methyl side-chain dynamics in proteins using selective enrichment with a single isotopomer. *J. Am. Chem. Soc.* **125**, 15767-15771.

Funding

We thank the Wellcome Trust and BBSRC for funding

Quaternary structure and catalytic activity of the *Escherichia coli* ribonuclease E amino-terminal catalytic domain

Yulia Redko and Kenneth McDowall

Introduction

RNase E is an essential endoribonuclease that plays a central role in the processing and degradation of RNA in *Escherichia coli* and other bacteria. With limited sequence specificity, RNase E tends to cut transcripts within single-stranded segments that are rich in A/U nucleotides. In many cases, the cleavage by RNase E is '5' end-dependent', in that the enzyme will efficiently cut substrates having a monophosphate group at the 5' end, but not transcripts that have a 5'-triphosphate group, or lack an accessible 5' end. The RNase E gene encodes one of the largest polypeptides in *E. coli* (1061 residues). The N-terminal domain of RNase E (residues 1-529) possesses the site of catalytic activity, and mutants that express this domain alone are viable. Such mutants can process ribosomal RNA satisfactorily, although they degrade bulk RNA more slowly, and are outgrown by cells expressing the full-length wild-type RNase E. The C-terminal domain of *E. coli* RNase E functions as a scaffold in the assembly of RhlB, an ATP-dependent RNA helicase, PNPase, a phosphate-dependent 3'-exonuclease, and enolase, a glycolytic enzyme. The association of these (and other) proteins forms a complex known as the RNA degradosome. It is likely that the selective advantage provided by the full-length protein, in comparison with the shorter N-terminal domain, is related to the co-ordination of RNA degrading and processing activities within the degradosome. Homologues of the *E. coli* RNase E N-terminal catalytic domain have been identified in other bacterial species and in archae, plastid genomes, and the nuclear genomes of several higher plants. *E. coli* itself contains a paralogue of RNase E, known as RNase G (the product of the *cafA* gene), which is closely homologous to the RNase E N-terminal domain. RNase G appears to overlap functionally somewhat with RNase E. Genome sequencing has identified RNase G-type enzymes in almost all known bacteria, with some organisms additionally possessing an RNase E-type enzyme.

The results of recent structural studies

Our collaborators in Cambridge have shown using a variety of biophysical and functional techniques that the N-terminal domain of *E. coli* RNase E forms a tetramer. In Leeds, studies using a new fluorescence-based assay have shown this to be a catalytically active form. Non-dissociating nanoflow-electrospray mass spectrometry suggests that the tetramer binds up to four molecules of a specific substrate RNA analogue. The tetrameric assembly of the N-terminal domain of RNase E is consistent with crystallographic analyses, which indicate that the tetramer possesses approximate D(2) dihedral symmetry. Using X-ray solution scattering data and symmetry restraints, a solution shape was calculated for the tetramer. This shape, together with limited proteolysis data, suggests that the S1-RNA binding domains of RNase E lie on the periphery of the tetramer (see Fig. 1).

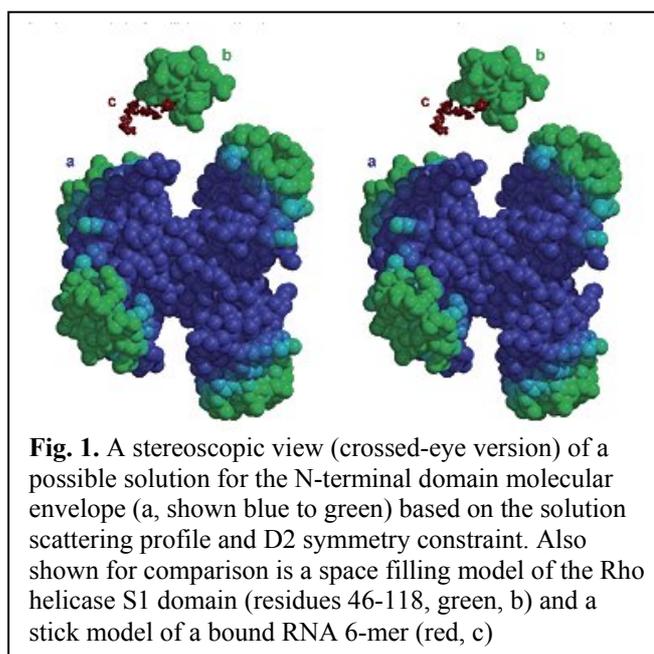


Fig. 1. A stereoscopic view (crossed-eye version) of a possible solution for the N-terminal domain molecular envelope (a, shown blue to green) based on the solution scattering profile and D2 symmetry constraint. Also shown for comparison is a space filling model of the Rho helicase S1 domain (residues 46-118, green, b) and a stick model of a bound RNA 6-mer (red, c)

As RNase E is the central component of the degradosome, and its C-terminal domain is the scaffold for the assembly of the remaining components of the complex, we suggest that these other components will be organised around a tetrameric RNase E core. Reports vary with regard to the compositional stoichiometry of the degradosome, and it is possible that the assembly around the RNase E tetramer is heterogeneous so that each C-terminal domain may not have the same number of components bound.

The tetrameric organisation of the RNase E N-terminal domain suggests that longer RNA substrates could be bound simultaneously to neighbouring protomers, while smaller RNA species, like the 10-mer studied here, might be bound independently. Recently, results from the group of Stanley Cohen (Stanford) provided the first evidence of RNase E acting via a scanning mechanism. Previously, George Mackie and coworkers (British Columbia) had proposed that a hypothetical dimeric organisation of RNase E could permit the interplay of binding and catalytic sites to affect a processive mode of substrate cutting. In their model, the binding of the 5' end of RNA (namely, the 5'-monophosphate) could occur at one protomer, with cleavage of the tethered substrate proceeding at the neighbouring protomer. This tethered, processive model was supported by evidence that segments of RNase E N- and C-terminal domains are sites of self-interaction. It is now clear, however, that the tetrameric assembly of the N-terminal domain would also facilitate this processive activity.

Publications

Callaghan, A.J., Grossmann, J.G., Redko, Y.U., Ilag, L.L., Moncrieffe, M.C., Symmons, M.F., Robinson, C.V., McDowall, K.J. & Luisi, B.F. (2003) Quaternary structure and catalytic activity of the *Escherichia coli* ribonuclease E amino-terminal catalytic domain. *Biochemistry* **42**, 13848-13855.

Collaboration:

This work was carried out in collaboration with Anastasia Callaghan, Günter Grossmann, Leopold Ilag, Martin Moncrieffe, Martyn Symmons, Carol Robinson, and Ben Luisi at the University of Cambridge.

Funding

Much of this work was supported by the Wellcome Trust. K.J.M. is the recipient of a fellowship from the Royal Society.

Expanding the use of zymography by the chemical linkage of small, defined substrates to the gel matrix

Vladimir Kaberdin and Kenneth McDowall

Introduction

In the postgenomic era, the comprehensive proteomic analysis of metabolic and signalling pathways is inevitably faced with the challenge of large-scale identification and characterisation of polypeptides with a particular enzymatic activity. Previous work has shown that a wide variety of enzymatic activities of microbial, plant, and animal origin can be assigned to individual polypeptides using zymography – a two-stage technique that involves protein separation by electrophoresis, followed by *in situ* assay of enzymatic activities. Certain advantages of zymography over conventional assays, such as the ability to assess the repertoire of enzymes that have a particular activity in non-fractionated cell extracts, and to estimate the molecular weight and isoelectric point of the corresponding polypeptides and their isoforms, can be invaluable in identifying and monitoring specific and non-specific activities in complex biological and clinical samples, and developing purification schemes. The most straightforward and widely used zymographic assays are based on the degradation or modification of high-molecular-weight substrates (e.g., gelatin, RNA transcripts) that are retained within the gel and remained unchangeable except within the immediate vicinity of the enzyme(s) being assayed. Suitable substrates, however, are only available for a limited number of classes of enzymes. A significant obstacle in developing assays for specific enzymes using high-molecular-weight substrates can be the existence of multiple sites that are recognisable by enzymes with different specificities. In this regard, small defined substrates, which are currently being synthesised in great variety and used in enzymological studies, have obvious advantages over macromolecules, but cannot be used directly in zymography as they are not retained efficiently within gels due to their small size. We have overcome this barrier, however, by covalently linking small, defined substrates to the gel matrix prior to electrophoresis.

Proof of principle

The effectiveness of using synthetic molecules that are covalently attached to the gel matrix has been shown using a short synthetic decaribonucleotide as a model substrate. The oligonucleotide was synthesised with the sequence ACAGUAAUUUG linked at the 3' end via an 18-link spacer to an acrylamide group, which was incorporated using acrydite phosphoramidite. This acrylamide–oligonucleotide was radiolabelled at the 5' end when required, co-polymerised into standard SDS-polyacrylamide gels and used for the in-gel activity staining of several evolutionarily unrelated enzymes; mammalian RNase A, bacterial alkaline phosphatase, and T4 polynucleotide kinase. As shown in Fig. 1 (panels A–C), we were able to detect the release of radiolabel from gel areas that corresponded to RNase A and alkaline phosphatase, and the attachment of radiolabel to the 5' end of hydroxylated substrate by polynucleotide kinase. The [$\gamma^{32}\text{P}$] ATP for the latter reaction was supplied in buffer after

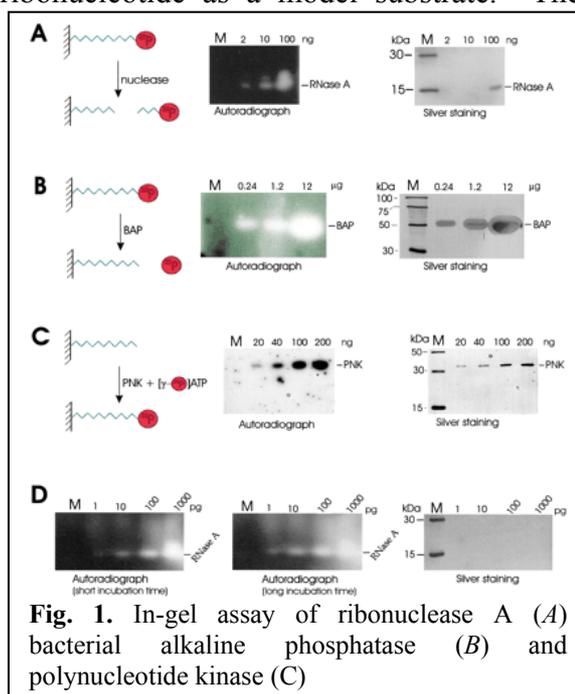


Fig. 1. In-gel assay of ribonuclease A (A) bacterial alkaline phosphatase (B) and polynucleotide kinase (C)

electrophoresis and the removal of SDS. Remarkably, these assays were at least as sensitive as silver staining without modification of standard gel electrophoresis and washing conditions. Further increases in sensitivity can, however, be obtained by optimising the conditions for each enzyme. We were able, for example, to increase the sensitivity of our RNase A assay by 1,000-fold. This was achieved simply by omitting reducing reagent and increasing the amount of substrate by fivefold (Fig. 1D).

The future

The above approach can be adopted readily and expanded to assay a wide variety of activities that modify or cleave nucleic acids as acrydite-containing oligonucleotides can be purchased (e.g., from Sigma-Genosys) or synthesised with relative ease by laboratories that have access to phosphoramidite chemistry. For example, enzymes such as helicases, polymerases, transcriptases, methylases, and certain endo-nucleases that have substrates with double-stranded segments can be assayed by hybridising complementary segments to acrydite-linked oligonucleotides prior to polymerisation of the gel. Peptides can also be chemically synthesised, in most cases, by relatively straightforward approaches, and procedures have been developed recently that allow peptides to be linked efficiently to mono- and oligonucleotides. Combined, these technologies provide a route for linking small, defined peptide substrates to the matrix of polyacrylamide gels thereby expanding our development of zymography to include protein-modifying enzymes such as proteases, phosphatases, and protein kinases. The latter should be particularly advantageous in the identification and analysis of activities that have important roles in apoptosis and signalling pathways. In principle, the chemical linkage of any type of small substrate to the gel matrix could facilitate the identification of the corresponding enzyme activity by zymography.

Mass spectrometric techniques are being developed that are able to analyse femtomole to attomole amounts of protein within 2D gels. These levels are below that detectable using protein stains such as Coomassie Blue, silver nitrate, zinc imidazole, or fluorescent dyes such as SYBRO Ruby. However, zymography can offer much higher levels of sensitivity without necessitating the use of radio-labelled substrates. Fluorescent markers (and quenchers) can be incorporated readily during the synthesis of, for example, oligonucleotide and peptide substrates and used in conjunction with plate readers that quantitatively measure specific fluorescence signals within gels. In combination with the continuing development of mass spectrometric techniques, zymography may lead to the identification and characterisation of enzymes that are below the current limits of detection by non-enzymatic staining. Increased sensitivity will be particularly beneficial in the analysis of biological material that can only be obtained in limited amounts and that is difficult to culture; for example, that from biopsies and swaps, and may lead, not only to new research tools, but also to assays for monitoring disease and detecting infection and biological contamination.

Publications

Kaberdin, V.R. & McDowall, K.J. (2003) Expanding the use of zymography by the chemical linkage of small, defined substrates to the gel matrix. *Genome Research* **13**, 1961-1965.

Funding

This work was supported by a grant from the Austrian Science Foundation to V.R.K. and by a Royal Society University Research Fellowship to K.J.M

Enhanced enzyme properties by protein engineering of galactose oxidase

Diane Wilkinson, Sarah Deacon, Ramon Hurtado Guerrero, Nana Akumanyi, Peter Knowles, Simon Phillips and Mike McPherson.

Galactose oxidase (GO; E.C. 1.1.3.9) is a 68kDa mononuclear copper-containing enzyme that oxidises primary alcohols to the corresponding aldehyde with coupled reduction of molecular oxygen to hydrogen peroxide according to the reaction scheme:



One of the copper ligands, Tyr 272, is covalently bonded through C_ε to the sulphur of Cys 228 and is the site of the free radical. The (Cu(II)-Tyr^{*}) centre can therefore catalyse the two electron oxidation of substrate. Recently both expression in *E. coli* and improvement in kinetic parameters have been achieved by directed evolution of GO.

We have investigated a series of single and double mutants and a triple mutant comprising V494A, C383S and Y436H identified by directed evolution. The locations of the mutated residues are shown in Fig. 1. The most interesting effect is due to the C383S mutation that leads to a significant reduction in *K_M*. Cys 383 lies in a pocket at the back of the active site behind the copper atom and therefore is not able to make direct interactions with the alcohol substrate. It is one of five cysteines that can be titrated in the denatured, reduced enzyme, but no role has yet been assigned to this residue. To explore the structural basis for improvement in *K_M* we have determined the 3-dimensional structure of the C383S variant by X-ray crystallography (Fig. 2).

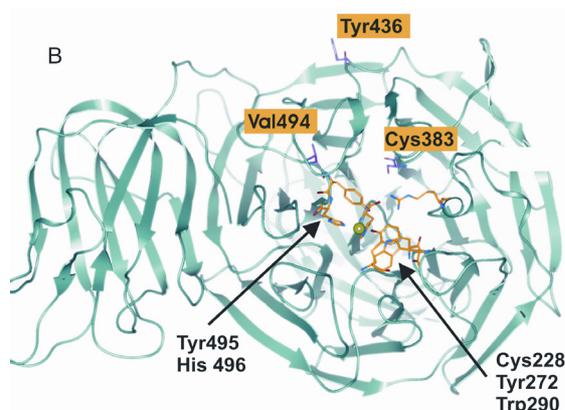


Fig. 1. Target residues in GO

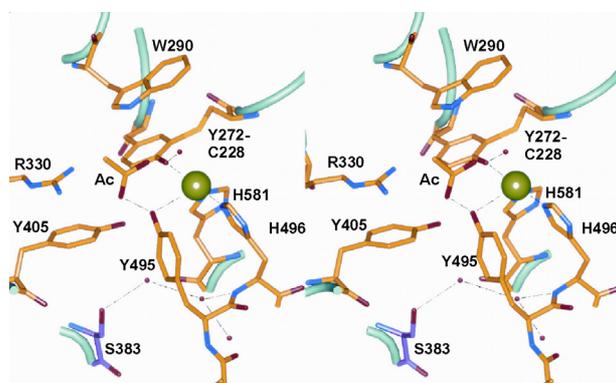


Fig. 2. Stereo of the active site of C383S. Ac is acetate and represents the substrate binding site.

Examination of the 3D crystal structure of the C383S variant provides little understanding of the reason for the beneficial effect of this substitution. Substitution by Ser replaces a weak H-bonding group with a strong H-bonding group and would lead to a stronger interaction with neighbouring water molecules. Such a change could be transmitted through the H-bond network to the substrate binding site, subtly altering its structure or flexibility. Perhaps more obvious differences within the active site would become apparent upon substrate binding. The reasons for beneficial effects mediated by mutations can often be difficult to discern, presumably as a consequence of the subtle manner in which they lead to concerted effects on protein structure and H-bond networks. Recent studies have identified substrate bound to the C383S variant and will allow analysis of this question. The detailed characterisation of these mutational variants of GO that display improvements in catalytic properties provides an excellent starting point for further selection strategies to optimise the enzyme to act on new polysaccharide substrates.

We have also mutated residues implicated in substrate binding, such as Arg330, and the R330K variant represents a more effective fructose oxidase than wild type GO. Combining various mutations may further enhance Go activity towards important substrates such as fructose.

Publication

Wilkinson, D., Akumanyi, N., Hurtado-Guerrero, R., Dawkes, H., Knowles, P.F., Phillips, S.E.V. & McPherson, M.J. (2004) Structural and kinetic studies of a series of mutants of galactose oxidase identified by directed evolution. *PEDS* In Press.

Funding

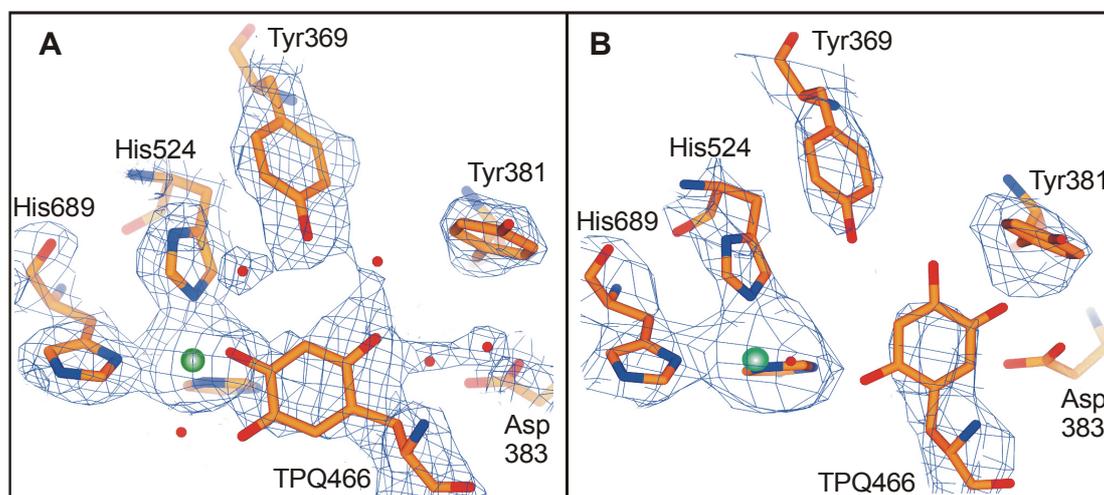
This work was supported by Hercules Inc. and the BBSRC.

Mobile TPQ cofactor and metals in *E.coli* copper amine oxidase

Christian Kurtis, Mark Parsons, Simon Phillips, Peter Knowles and Mike McPherson

Escherichia coli amine oxidase (ECAO) is a member of the copper amine oxidase enzyme family. These enzymes are homodimeric and contain a single cupric ion and a post-translationally modified tyrosine cofactor; 2,4,5-trihydroxyphenylalanine quinone (TPQ). TPQ biogenesis is an autocatalytic event requiring copper and oxygen. The role of the copper atom in cofactor formation is widely accepted; however, the role of the metal centre during catalysis is less widely agreed upon. There is a long held view that the copper centre is involved in catalytic redox processes based primarily on the observation of a semiquinone radical-Cu(I) species in the anaerobically reduced form of the enzyme. In support of this view, substitution of the copper with divalent metals such as zinc, cobalt and nickel in ECAO or amine oxidase from *Arthrobacter globiformis*, significantly reduces the catalytic velocity. However, previous studies with the amine oxidase from *Hansenula polymorpha* in which the cobalt substituted enzyme was essentially active, indicated that the metal centre is not directly involved in redox processes during catalysis.

We have determined the crystal structures of all the metal substituted forms of the ECAO in a low temperature nitrogen stream (100 K) and this revealed a tendency of the TPQ to coordinate the metal centre in a non productive conformation. Spectroscopic properties of these metal substituted enzymes in solution, and determination of the room temperature structure of the nickel substituted enzyme, indicated that the TPQ was not coordinating the metal centre. In ECAO the TPQ appears to readily adopt alternate conformations and in some cases cryo-crystallography leads to the trapping of energetically favourable but non-productive conformations that do not represent the major species in solution at room temperature. An in-depth study of the transient kinetics of TPQ reduction and re-oxidation in the metal substituted enzymes is now underway to demonstrate definitively how the metal centre is involved in catalysis.



(A) Active site of the nickel substituted form of ECAO determined by X-ray cryocrystallography to a resolution of 2.3 Å. The $2F_o - F_c$ electron density map contoured at 1σ reveals how the TPQ ring oxygen in the para position is coordinating the nickel atom directly. (B) X-ray crystal structure of the nickel substituted enzyme determined at room temperature. Data processed to 3.4 Å were sufficient to show the TPQ ring in a 'productive' conformation in the $2F_o - F_c$ electron density map contoured at 1σ .

Funding:

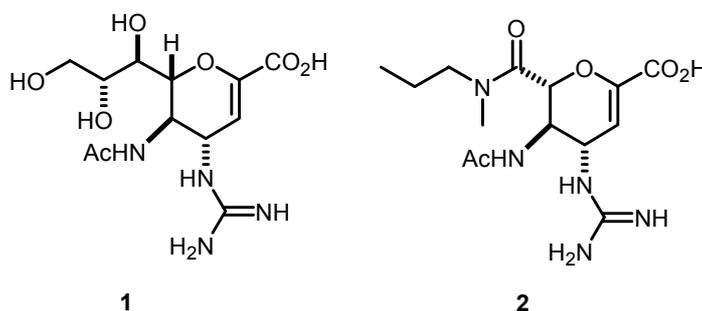
This work was supported by the BBSRC.

Synthesis of screening substrates for the directed evolution of sialic acid aldolase

Thomas Woodhall, Alan Berry, Gavin Williams and Adam Nelson

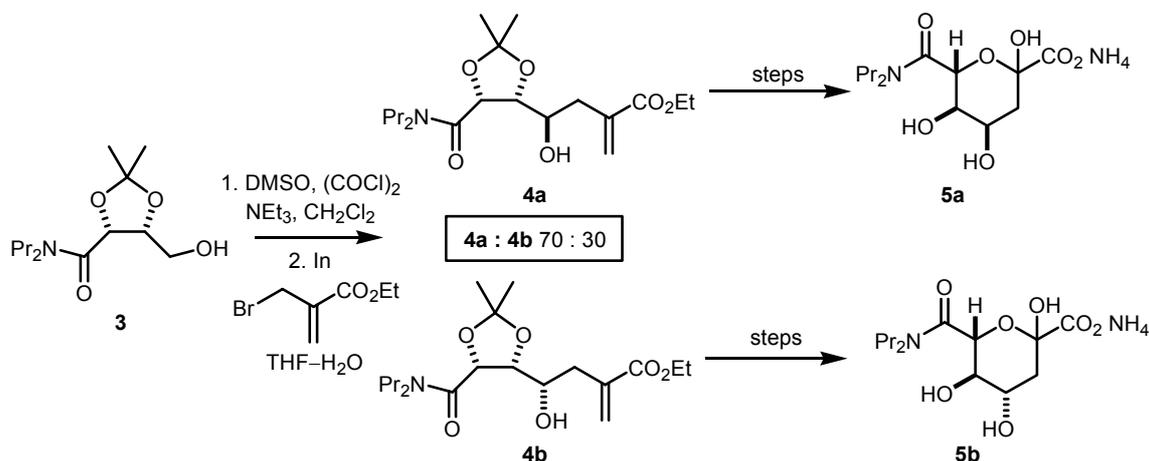
Introduction

Sialidase inhibitors have tremendous value as therapeutic agents for the treatment of influenza. For example, Relenza (**1**) is in clinical use, and the sialic acid mimetic **2** is a potent and selective inhibitor of influenza sialidase A. We are engaged in a research programme which involves the directed evolution of tailored enzymes for exploitation in organic synthesis. Here, we describe the synthesis of screening substrates for the directed evolution of sialic acid aldolase. The resulting enzymes will be of value in the preparation of a wide range of sialic acid mimetics (e.g. **2**).



Synthesis of screening substrates

The screening substrates **5a** and **5b** were prepared using the approach outlined below. An indium-mediated allylation of the aldehyde derived from the alcohol **3** gave the homoallylic alcohols **4a** and **4b** as a 70:30 mixture of diastereoisomers. Ozonolysis and deprotection of each of these diastereoisomers gave the screening substrates **5a** and **5b**.



Summary

Tailored enzymes which have been generated using directed evolution are an excited addition to the armoury of the synthetic chemist. The synthesis of screening substrates for the directed evolution of sialic acid aldolase has been described. The results of these evolution experiments are discussed elsewhere in this report.

Acknowledgements

We thank EPSRC and BBSRC for funding.

A combinatorial selective labelling method for the assignment of backbone NMR resonances

Marc Aulton-Jones and Martin J. Parker

Introduction

NMR spectroscopy is an enormously powerful technique with which to study a wide range of biochemical problems in solution. One of its unique features is that it can provide a set of residue specific probes with which to study protein interactions. The prime source for these probes is the ^1H - ^{15}N HSQC spectrum. Interaction sites are inferred from perturbations of crosspeaks upon titration of ligand or unlabelled protein into the sample. These are then mapped on to the experimental structure. The method is particularly suited for studying weakly interacting systems, and has led to an active area called 'SARS by NMR', where lead pharmaceutical compounds with affinities as low as 10mM can be screened rapidly for interaction with the target protein.

Application of the method obviously demands the assignment of crosspeaks to specific residues in the protein. Despite many technological developments, this remains a significant challenge, especially for large systems. Recently, several groups have presented methods using selective amino acid labelling to accelerate the identification of protein interaction sites. These techniques only probe a very limited number of residues however. We have developed a Combinatorial Selective Labelling (CSL) method for the efficient assignment of the majority of HSQC crosspeaks using a small number of samples that can be rapidly and cost effectively produced in parallel in a commercially available cell-free system.

The CSL method

The CSL method is based upon the dual amino acid-selective $^{13}\text{C}/^{15}\text{N}$ labelling technique, which utilises protein samples in which the main chain carbonyl carbons of one amino acid type (A) are labelled with ^{13}C , and the amide nitrogens of another amino acid type (B) are labelled with ^{15}N . The NMR signals of the amino acid residues that possess a ^{13}CO - ^{15}N linkage can be extracted on the basis of the ^{13}C - ^{15}N spin coupling, i.e. if an (A)B pair exists only once in the sequence then a unique crosspeak will appear in the ^1H - ^{15}N 2D HNCOC spectrum, and the NH group of the residue type B can be unambiguously assigned. To date these selective labelling methods have only been applied to identify single peaks in ^1H - ^{15}N correlation spectra. In a simple approach, the identification of each different AB pair would require a separate sample, demanding a prohibitively large total number of such samples. Our novel approach is to use a much smaller number of samples labelled with different combinations of amino acids, using the resulting patterns of crosspeak intensities across these samples to differentiate each AB pair.

The CSL method requires five protein samples, each containing a different combination of 16 labelled amino acid types (see Fig. 1 (a)). For each sample, a ^1H - ^{15}N HSQC spectrum and a ^1H - ^{15}N 2D plane of an HNCOC spectrum are acquired. Comparison of the relative peak intensities in the HSQC spectra yields the amino acid type of each peak. The 16 amino acid types chosen here can be assigned in the 4 samples as there are 2^4 (= 16) such patterns (sample 1 is the reference). For a particular crosspeak, one can tell the amino acid type of the preceding residue in the sequence by examining the presence or absence of peaks in the five 2D HNCOC spectra. Therefore, all 16x16 possible amino acid pairs are identifiable simultaneously from these five samples. If a pair appears n times in the sequence then n peaks will appear in these spectra with the same intensity pattern, and the assignment will be n -fold degenerate.

We have demonstrated the feasibility of the method using the 27 kDa, beta barrel protein GFP as a model system. The samples were labelled in the pattern depicted in Fig. 1 (a), using the Rapid Translation System (RTS) 500 *E. coli* HY kit. The HSQC spectrum of GFP with all 16 of the chosen set of amino acids fully $^{13}\text{C}/^{15}\text{N}$ labelled (sample 1) is shown in Fig. 1 (b). The majority of the resolved peaks in the HSQC/HNCO spectra display intensity patterns conforming to those expected for the published assignments (green peaks in Fig. 1 (b)), demonstrating that the method is practicable.

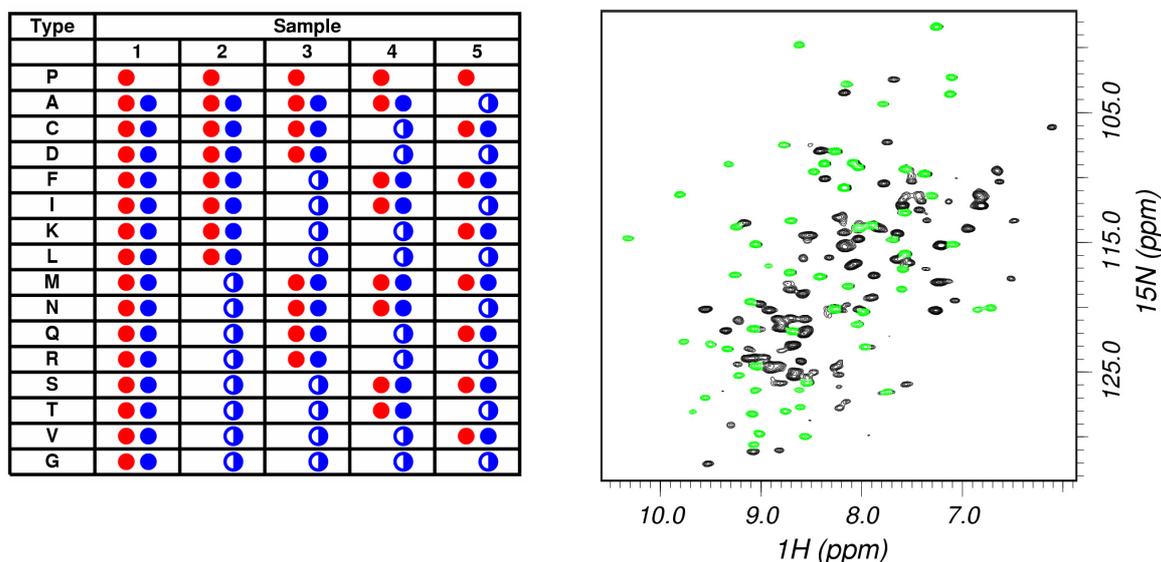


Fig. 1: LHS (a) CSL labelling method. Red and blue filled circles denote 100% ^{13}C and ^{15}N labelling. Blue half filled circles denote samples in which a 50:50 mix of ^{15}N and ^{14}N labelled amino acids is used. RHS (b) ^1H - ^{15}N HSQC spectrum of GFP labelled according to sample 1 in (a). The peaks coloured in green show the correct pattern of intensities, based on cutoffs relative to sample 1 of 0.75 and 0.25 for the HSQC and HNCO spectra, respectively.

Traditional assignment methods are sensitive to the completeness of the data for all residues, since incorrect or incomplete information about one residue can confound assignment of another residue. The assignment of a particular crosspeak in our method to a particular (A)B pair depends solely on information about that single crosspeak, which makes our method much less sensitive to crosspeak overlap or missing crosspeaks. Furthermore, and crucially, our method uses two of the most rapid/sensitive NMR experiments, which make it applicable to proteins suffering from poor solubility, tumbling or sample lifetime characteristics. The analysis of the data in our technique is also much less demanding and could be carried out by someone with limited NMR experience. Thus the method opens up opportunities for applying NMR to systems that give less than ideal spectra, to studying groups of related proteins rapidly in parallel, and to making NMR more accessible to non-NMR specialists.

Further studies

We are currently developing and refining the CSL method. In particular, we are combining it with protein ligation technology to increase the completeness of the assignments and to reduce spectral complexity. Together with recent developments in NMR (specifically the ‘TROSY’ method), this will greatly increase the range of applicability. These developments are best achieved by exploring their application to challenging and medically important problems. We are applying the methods to: (i) the interaction of bactericidal/permeability increasing protein with lipopolysaccharide (LPS; endotoxin), an important part of the innate immune response to bacteria, and (ii) the interactions among those members of the Bcl-2 family of proteins involved in apoptosis; potential targets for cancer therapeutics. The

experiments will identify those residues comprising the protein interaction sites in these systems, providing important insights on their biochemistry and aiding in the design and development of drugs.

Collaborators

Dr C. Jeremy Craven, University of Sheffield.

References

Parker, M.J., Aulton-Jones, M., Hounslow, A. & Craven, C.J. (2004). A combinatorial selective labelling method for the assignment of backbone NMR resonances. *J. Am. Chem. Soc.*, in press.

Funding

This work was funded by the BBSRC. M.J.P. is a BBSRC David Phillips and University Research Fellow.

Water as a conformational editor in protein folding

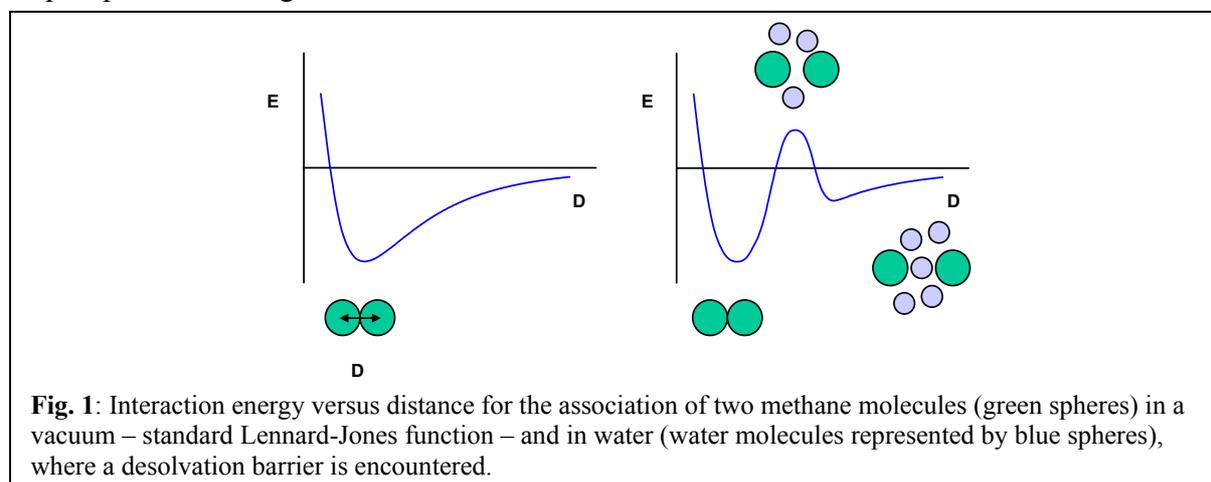
Geraint Thomas and Martin J. Parker

Introduction

Currently, even the most extensive all-atom, explicit-solvent simulations of proteins do not provide sufficient coverage to tackle many equilibrium and long-timescale kinetic properties under physiological conditions. A complementary approach is to use simplified lattice or continuum models, trading high structural resolution for the extensive conformational sampling that is often necessary for conceptual advances. Most studies using such simplified representations use unrealistic, native-centric force fields, i.e. they are based on the native structure and not on the chain's sequence of amino acids. As such, these models are incomplete and capture only the physics arising from chain connectivity and excluded volume; competing non-native interactions are not considered in the energy spectrum.

We have developed an algorithm for protein structure prediction called 'RAFT' (Reduced Ramachandran Angle and Folding Forcefield for Tertiary Structure Prediction). The method aims to keep complexity of the geometric description to a minimum without seriously compromising the accuracy of the structural representation. RAFT uses 6 optimised ϕ - ψ angles and fixed side chains approximating rotationally averaged real side chains. The energy of a particular conformation is calculated using a simple physico-chemical force field, based on hydrophobic, hydrophilic, steric and hydrogen bonding potentials derived from experimental data. A search technique based on simulated annealing is then employed to find the conformation with the lowest energy. RAFT has been used successfully to predict the structures of peptides up to ~ 40 residues in length, and has been found to be similarly effective as a more complex all-atom model at predicting the structures of proposed independent folding units.

RAFT provides a reasonably realistic, non-native-centric model with which to explore fundamental questions related to protein folding thermodynamics and kinetics. One such question relates to the influence water molecules have on protein folding reactions. As a protein's constituent groups cluster together in aqueous solvent they encounter a desolvation free energy barrier. Calculations of the potential of mean force for the association of methane in explicit water, for example, reveal two minima: a pronounced minimum at the optimal contact distance, and a second solvent-separated minimum. Between these two minima water cannot penetrate; interactions with the solvent are lost and an energetically unfavourable cavity is formed. Experiments suggest that such a barrier may contribute to the rate-limiting step in protein folding.



Thermodynamic and kinetic effects

We incorporated desolvation barriers in to the RAFT force field to explore their influence on the energy landscape of protein folding. Monte Carlo sampling techniques were used to examine the effect on the conformational energy spectra of a number of different peptide/protein systems. In each case, desolvation barriers increase the stability of the native conformation and the cooperativity of the major folding/unfolding transition (Fig. 2). The influence on folding kinetics was also explored by calculating median first passage times (MFPT). Interestingly, at relatively low temperatures, where the numbers of admissible moves are low, desolvation barriers speed up the search for the native fold (Fig. 2). By ‘editing out’ alternative conformations that have interaction distances in the barrier zone, water desolvation engenders a more defined route to the native fold.

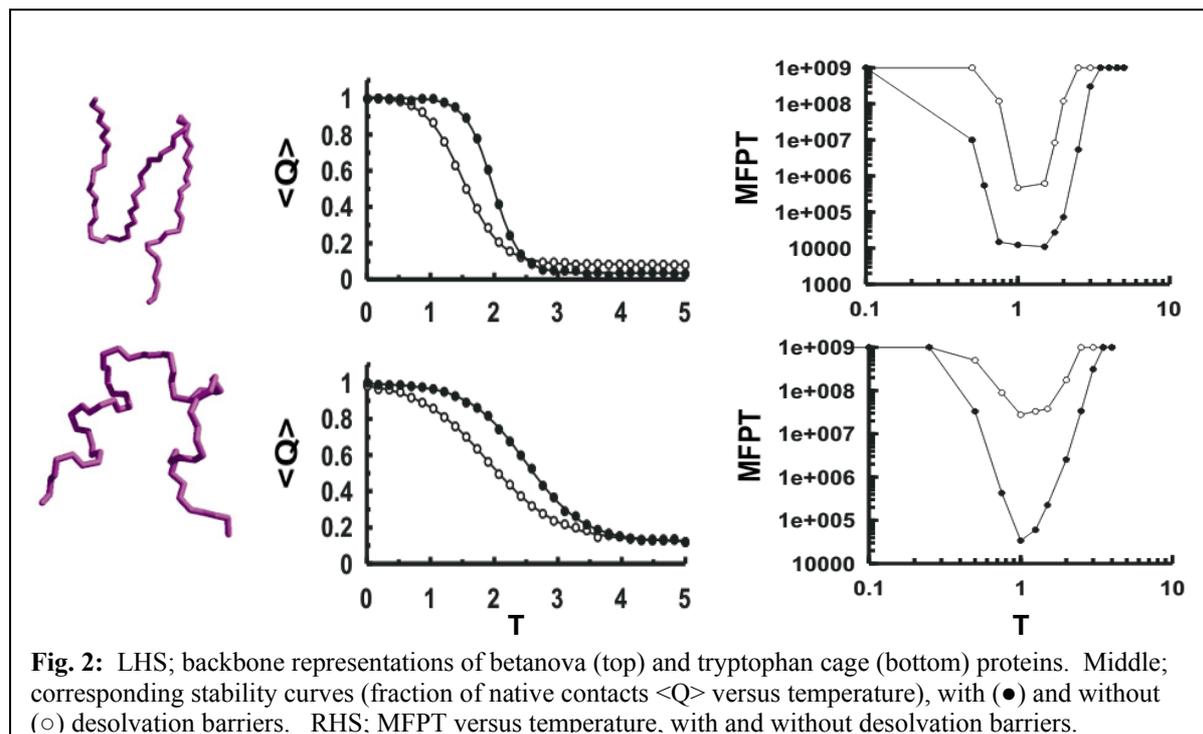


Fig. 2: LHS; backbone representations of betanova (top) and tryptophan cage (bottom) proteins. Middle; corresponding stability curves (fraction of native contacts $\langle Q \rangle$ versus temperature), with (●) and without (○) desolvation barriers. RHS: MFPT versus temperature, with and without desolvation barriers.

Collaborators

Dr Richard Sessions, University of Bristol.

References

Sessions, R. B., Thomas, G. & Parker, M. J. (2004). Water as a conformational editor in protein folding. *J. Mol. Biol.*, in press.

Funding

This work was funded by the BBSRC and Wellcome Trust. M.J.P. is a BBSRC David Phillips and University Research Fellow.

Crystal structures of Q β RNA stemloop operators complexed with bacteriophage MS2 coat protein mutants; towards an explanation of binding discrimination.

Wilf T. Horn, Nicola J. Stonehouse, Peter G. Stockley and Simon E.V. Phillips

Introduction

MS2 and Q β are evolutionarily related $T=3$ icosahedral bacteriophages with single stranded RNA genomes that infect *E. coli*. Although the protein subunits of the MS2 and Q β capsids share less than 25% sequence identity, the structures of the Q β subunits are very similar to those of MS2. Subunits of both the MS2 and Q β capsid shells exist as three distinct conformers (A, B and C) that associate to form AB and CC dimers which comprise the basic building blocks of both capsid shells. The structures of the MS2 and Q β capsids have been determined via X-ray crystallography by our collaborators in Uppsala, Sweden.

The two bacteriophages both utilise a similar mechanism of translational repression. *In vivo*, a small RNA stemloop within the viral genomes binds to a specific site on a coat protein dimer, acting to inhibit viral replicase gene translation. The translational complex of MS2 has, for many years, been the paradigm for studying RNA/protein interactions at the atomic level.

The RNA stemloop operator binding site is located on a 10 stranded β sheet formed by AB and CC dimers within the capsid shells of both MS2 and Q β . Many of the amino acid residues that have previously been shown to be

important for high affinity binding of MS2 and Q β stemloops are conserved between the two bacteriophages (Fig. 1). Although the protein surfaces of the stemloop binding site of the two bacteriophages display considerable similarity, profound differences exist in the sequence and

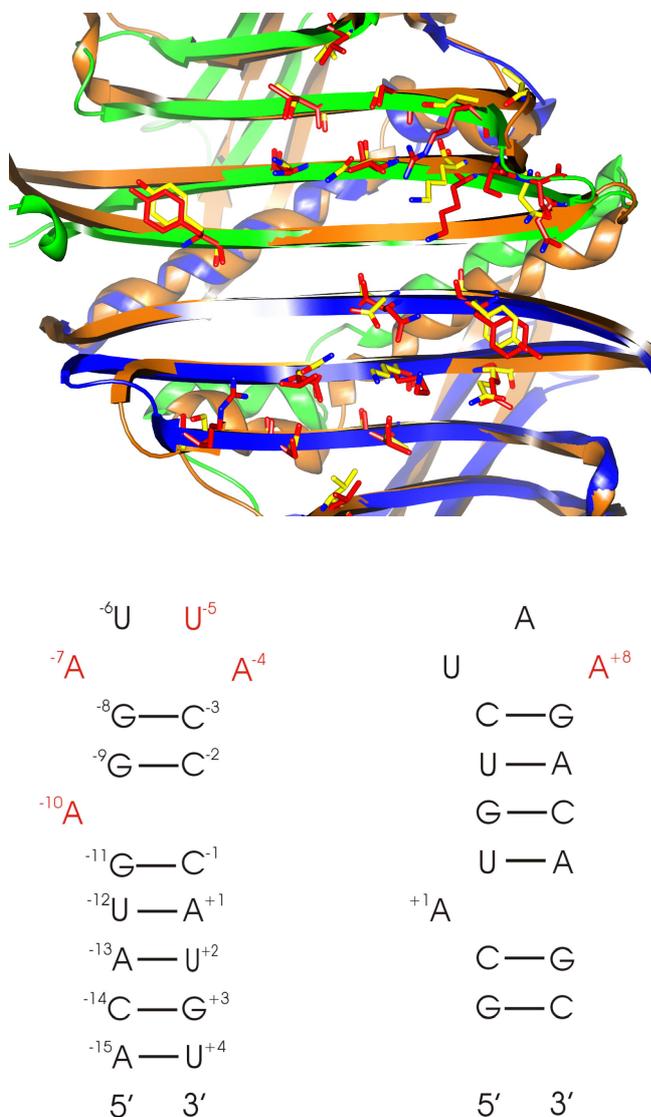


Fig. 1;

(Top) Ribbon diagram of the ten-stranded β sheet of an AB coat protein dimer of the Q β phage dimer (in orange) superimposed on an AB dimer from the MS2 capsid (subunit A in blue, B in green). Amino acid residues that are involved in protein-RNA interactions are shown in stick format, with MS2 residues in red and their Q β equivalents in yellow.

(Bottom) Diagram of the secondary structure of the MS2 (left) and Q β (right) stemloop operators. Residues that have been shown to be important for high affinity binding to capsid protein are highlighted in red.

secondary structures of the two stemloop operators (Fig. 1). *In vivo*, each bacteriophage preferentially discriminates against binding the stemloop operator of the other. Affinity binding studies have, however, identified specific coat protein mutants of MS2 that overcome this discrimination mechanism, some of the mutations allowing the binding of the Q β RNA operator to MS2 mutant capsids with an affinity comparable to that of the wild type MS2 operator. In order to gain new insight into this discrimination mechanism, Q β RNA stemloop operators were soaked into pre-crystallised MS2 mutant capsids and the structure of the capsid/RNA complexes determined via X-ray crystallography.

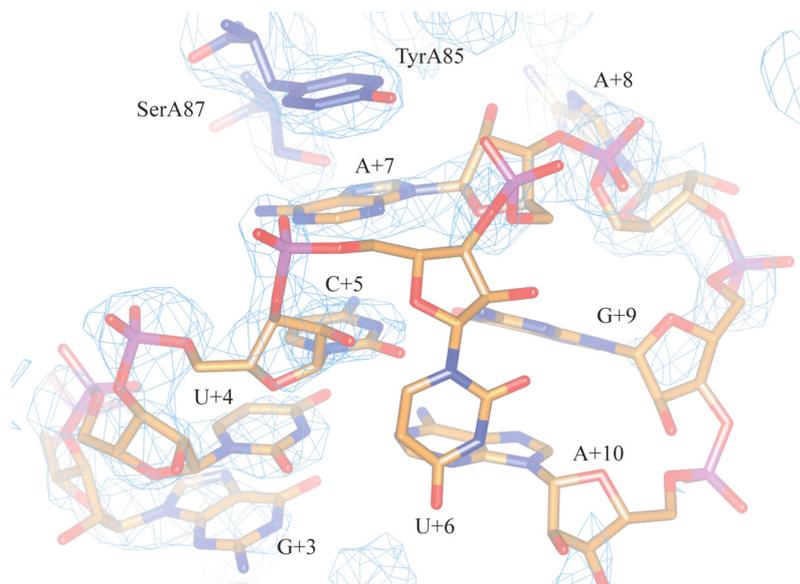


Fig. 2; Diagram showing the $2Fo-Fc$ electron density (blue) associated with the Q β RNA complexed at the AB binding site of the N87S MS2 mutant. The loop and upper stem region of the Q β RNA is modelled into the electron density.

Results

Diffraction data were collected at the SRS, Daresbury, UK for three different MS2 mutants (N87S, E89K/N87S, E89K) complexed with Q β RNA stemloop operators. Although the electron density for the lower stem region of the RNA is weak in each of the complexes, unambiguous modelling of the loop and upper stem region of the RNA is possible (Fig. 2). In contrast to the four base loop observed in the MS2 stemloop-coat protein complex, the Q β RNA maintains its three base loop topology on complex formation. The structures of both the N87S and the N87S/E89K MS2 mutants complexed with Q β RNA demonstrate that their increase in affinity for Q β RNA can partly be explained by the identity of the nucleotide that stacks onto the underside of TyrA85. In the MS2 RNA – coat protein complex this nucleotide is a uracil, whereas in the two mutant complexes it is an adenine that makes this important stacking interaction. The mutation of the bulky Asn87 residue to the smaller Ser87 mediates the stacking of the adenine onto TyrA85, an interaction that would not be as favourable if Asn87 was present. Further crystallographic studies are continuing in the expectation of explaining the role played by the E89K mutation on binding affinity.

Collaborator

Lars Liljas and Kaspar Tars, Department of Cell and Molecular Biology, Uppsala University, Sweden

Funding

This work was funded by the BBSRC.

Structural determination of human ketohexokinase.

Chi H. Trinh, Aruna Asipu, Bruce E. Hayward, David T. Bonthron and Simon E.V. Phillips

Introduction

In mammals, dietary fructose is primarily metabolised through a pathway distinct from that responsible for glucose metabolism. Ketohexokinase (fructokinase; KHK), is the first enzyme in this pathway, catalysing the phosphorylation of fructose to fructose 1-phosphate. Alternative splicing of the KHK gene generates a "central", predominantly hepatic, isoform (ketohexokinase-C) and a more widely distributed ketohexokinase-A. Hepatic KHK deficiency (caused by mutations that inactivate KHK) underlies essential fructosuria, a benign inborn error of metabolism. This is characterised by a large and persistent rise in blood fructose after the ingestion of fructose, sucrose or sorbitol and the excretion of 10-20% of the ingested load in the urine.

It has been shown that both ketohexokinase isoforms are active. Ketohexokinase-A has much poorer affinity for fructose than ketohexokinase-C but is considerably more thermostable. At physiological temperature, mutations that cause essential fructosuria consequently result in significant loss of ketohexokinase-C activity but not of ketohexokinase-A. Affected individuals therefore probably have a selective deficiency of hepatic KHK, with peripheral ketohexokinase-A being preserved. No function is presently defined for the low levels of ketohexokinase-A that are widely distributed in many tissues. Ketohexokinase has no significant primary structure similarity to other mammalian hexokinases, and is a member of the family of prokaryotic ribokinases and furanose sugar kinases. Here, we report the structure of the human ketohexokinase-A, and in addition its complexes with the sugar molecules fructose and xylulose, and a nucleotide analogue.

Crystallographic studies

Ketohexokinase-A has been crystallised using the vapour diffusion technique. Complexes of ketohexokinase-A with different substrate ligands have been co-crystallised. Data for both the native apo-enzyme and complexes were collected using synchrotron radiation at Daresbury Laboratory. All X-ray diffraction data were collected at a temperature of 100 K. Heavy-atom derivatised crystals were prepared by soaking the crystals in the stabilisation solution containing heavy atom compounds for various periods of time. Screening for a derivatised crystal and subsequent data collection were performed using the X-ray facility within the Astbury Centre for Structural Molecular Biology.

The ketohexokinase-A structure was determined by the method of single isomorphous replacement. The density modified SIR map was interpretable and the structure of the ribosekinase was positioned into the density. Rigid body refinement using the program CNS was carried out on the structure of ribosekinase to give a starting model for ketohexokinase-A. Model building and refinement was accomplished at 1.8 Å resolution using the programs O and CNS, respectively.

Structure of ketohexokinase-A

There is one monomer of ketohexokinase-A in the asymmetric unit, with the other monomer of the dimer related by a crystallographic twofold axis. Each ketohexokinase-A monomer consists of 298 residues. The monomer has two distinct secondary structural elements; a central α/β fold and a 4-stranded β -sheet. The α/β fold consists of a 9 β -stranded sheet flanked on both sides by α -helices. The dimer interface between the two monomers is formed from the 4-stranded β -sheet (Fig. 1). The alternative splicing of the KHK gene results in a difference of a single region between the two isoforms of the protein. This region contains 44 residues and translates structurally to 2 of the 4 β -strands that form the dimer interface. There is one active site per monomer and this is located between the α/β fold and

the β -sheet forming the dimer interface. For the structures of the KHK complexes electron density are observed for both the sugar molecules and the nucleotide analogue.

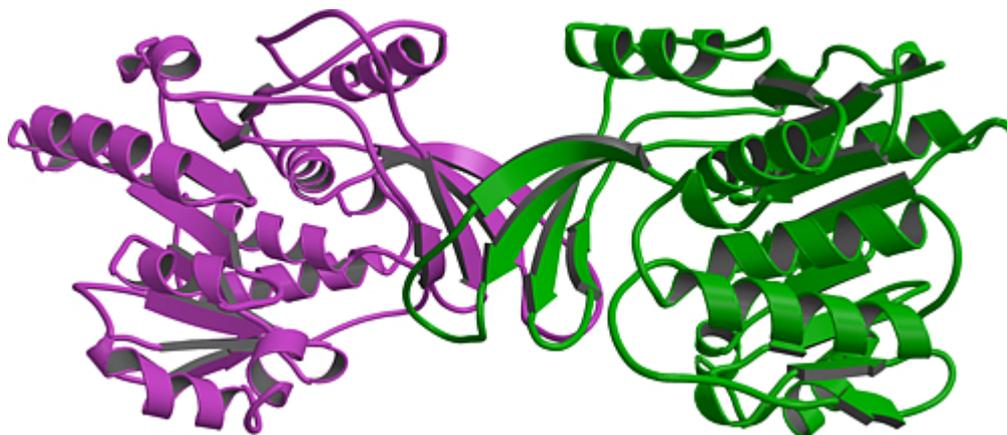


Fig. 1. Structure of the ketohexokinase-A, each monomer of the protein is coloured in a different colour.

Work is currently being undertaken to elucidate the structure of the human ketohexokinase-C. Our findings will hopefully provide an insight into the molecular basis for the differences in function between the hepatic and peripheral isoforms of KHK.

Collaborators

Aruna Asipu, Bruce E. Hayward and David T. Bonthron, Molecular Medicine Unit, University of Leeds, St. James's University Hospital, Leeds, UK

Funding

This work was funded by the BBSRC.

Cytolysin gene expression in *Enterococcus faecalis* is regulated in response to aerobiosis conditions

Alison. Day, Jonathan Cove and Mary Phillips-Jones

Enterococcus faecalis is a facultatively anaerobic, biofilm-forming bacterium that is a significant causative agent of hospital-acquired infections. It has added significance because of the ready emergence amongst some strains of resistance to the last-line antibiotic vancomycin and other glycopeptides. It is generally thought to be a homolactic fermentative bacterium, but evidence of oxidative metabolism has also emerged. The ability to switch between aerobic and anaerobic metabolism is evidenced by the global changes in gene expression that occur in response to changing oxygen conditions, and at least one oxidative stress protein homologue (OxyR) has already been identified.

In a recent study, we investigated expression of *cyLL_L* and *cyLL_S*, encoding the structural subunits of the essential virulence factor cytolysin/haemolysin of *Enterococcus faecalis* in response to aerobiosis conditions. Haemolysis assays of *E. faecalis* strains cultured under aerobic and anaerobic conditions revealed three different haemolytic phenotypes, one of which showed greater haemolysis under anaerobic conditions than under aerobic conditions, and we demonstrated that this was associated with the presence of the *cyL* genes. Reporter studies using a green fluorescent protein reporter (that was permitted to attain full maturation prior to assays) revealed significantly greater (up to 8.6-fold) *cyLL_LcyLL_S* promoter activity anaerobically compared to aerobically throughout batch growth, demonstrating that these genes are regulated in response to aerobic/anaerobic conditions (Fig. 1).

Gel retardation assays confirmed the binding of a protein factor between 202 and 37bp upstream of the *cyLL_L* start codon, with greater binding observed using anaerobically-derived cell-free extracts compared with aerobic extracts. This is the first report of an oxygen-regulated virulence factor in *E. faecalis* (that is distinct from the quorum-sensing regulatory system reported by others in 2002), and may be of relevance *in vivo* to the bacterium within biofilm and other oxygen-gradient environments.

We are currently identifying the oxygen-responsive pathway involved and other pathways that regulate antibiotic and essential functions in this pathogenic bacterium.

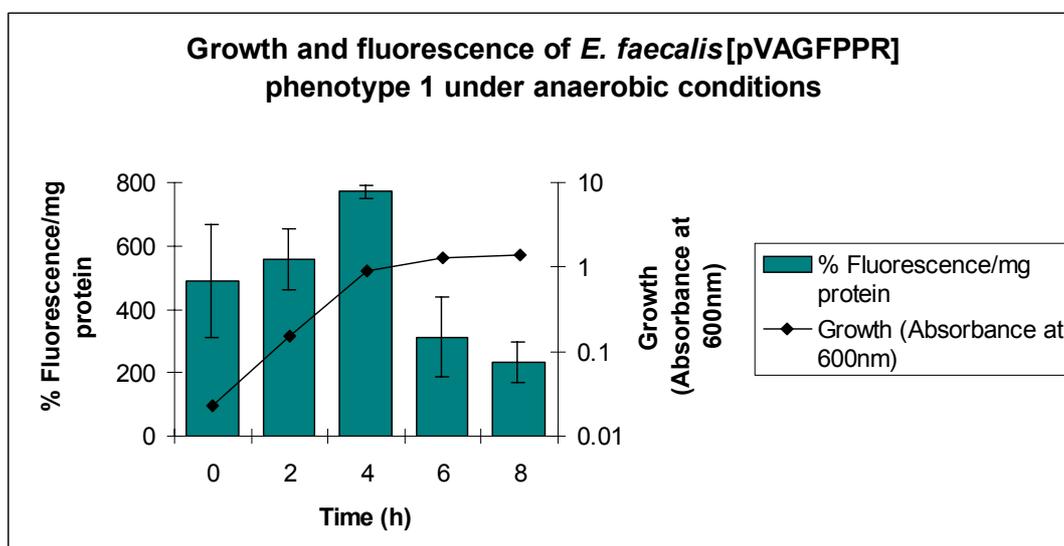
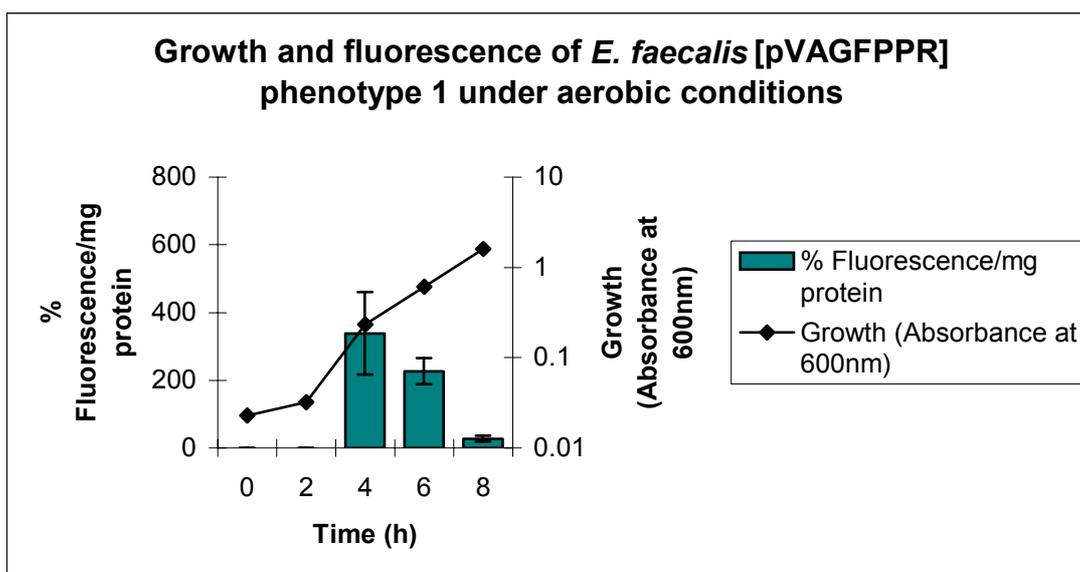


Fig. 1. Green fluorescent protein (GFP) reporter study to examine activity of the *cylL_L* promoter throughout batch growth of *E. faecalis* under aerobic and anaerobic conditions.

GFP reporter levels are expressed as % fluorescence using *E. coli* DH5 α [pGFPuv] as 100% fluorescence standard and *E. faecalis* *cyl*-minus strain 30[pVAGFP] in the appropriate growth phase as 0% fluorescence. Growth (filled squares) was measured by absorbance at 600nm. Experiments were performed in triplicate and the results of one typical experiment are shown. Values are the means of 3 replicate assays per cell sample. *E. faecalis* haemolysis *cyl*⁺ phenotype 1 strains harbouring pVAGFPPR were cultured aerobically (upper panel), or (B) anaerobically (lower panel).

Publications

Day, A.M., Cove, J.H. and Phillips-Jones, M.K. (2003) Cytolysin gene expression in *Enterococcus faecalis* is regulated in response to aerobiosis conditions. *Mol. Genet. Genom.* **269**, 31-39.

Funding

AMD was supported by a Emma and Leslie Reid Scholarship from the University of Leeds.

Expression, purification and characterisation of full-length histidine protein kinase RegB (PrrB) from *Rhodobacter sphaeroides*

Christopher A. Potter, Alison Ward, Peter J.F. Henderson and Mary K. Phillips-Jones

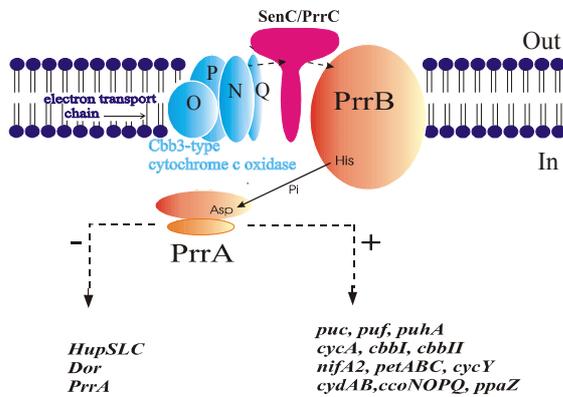


Fig.1. The Reg (Prr) pathway of *R. sphaeroides*

The global redox switch between aerobic and anaerobic growth in *Rhodobacter sphaeroides* is controlled by the RegA/RegB two-component system (also known as the PrrA/PrrB), in which RegB (PrrB) is the integral membrane histidine protein kinase, and RegA (PrrA) is the cytosolic response regulator. Despite the global regulatory importance of this system and its many homologues, there have been no reported examples, to date, of heterologous expression of full-length RegB or indeed any membrane histidine protein kinases (HPKs).

Yet knowledge of membrane HPKs such as RegB are important because most members of this protein family are membrane-bound, transferring signal information across the membrane and into the cell, and little is currently known about the mechanisms of signal transduction across the membrane. This lack of information is mainly due to the technical difficulties of overexpressing and purifying these bacterial membrane proteins. In the case of RegB, the signal may arrive at the protein from within the membrane itself, from the *cbb*₃-type cytochrome *c* oxidase of the electron transport chain via another membrane protein SenC (PrrC).

In order to address this problem, we utilised plasmid pTTQ18His, a membrane protein expression plasmid that has been used previously for the successful overexpression of 25 membrane proteins to amplify expression, and subsequently isolate and purify, intact RegB. Using this vector, RegB accounts for 2-10 % of *E. coli* membrane proteins and yields of ca. 1 mg of purified protein per litre of culture have been obtained (Fig. 2). This is the first successful overexpression of an HPK in a heterologous *E. coli* host.

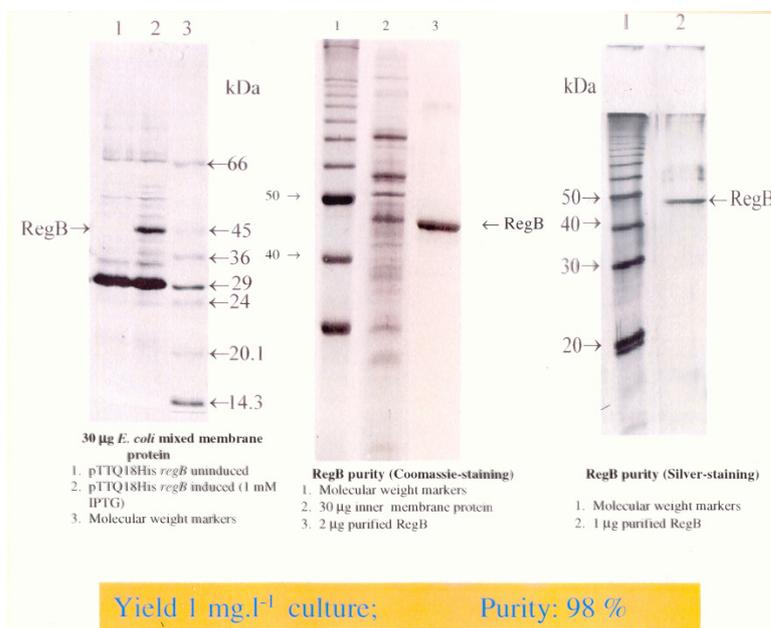


Fig. 2. SDS-PAGE analysis of overexpressed and purified RegB. Left panel: 30 µg mixed membrane proteins. Lane 1, pTTQ18HisregB uninduced; lane 2, pTTQ18HisregB induced (expressed); lane 3, molecular mass markers (MMM). Middle panel: Lane 1, MMM; lane 2, 30 µg inner membrane protein expressing RegB; lane 3, 2 µg purified RegB. Right panel: lane 1, MMM; lane 2, 1 µg purified RegB stained with silver. Left and middle panels show Coomassie-stained gels.

The overexpressed protein is functional in *E. coli* inner membranes, as shown by its autophosphorylation activity and its ability to be dephosphorylated by RegA. Importantly, it is also functional following purification, since autophosphorylation, phosphotransfer and RegA-dephosphorylation activities were all demonstrable (e.g. Fig. 3).

Our kinetic data obtained for this intact protein reveal important differences compared with the truncated version of soluble RegB from *R. sphaeroides*, demonstrating that the transmembrane region has important regulatory activity.

The system has been developed for *in vitro* studies to elucidate the signal sensing mechanism of RegB, by examining RegB autophosphorylation in combination with candidate interacting proteins such as SenC and components of the cytochrome *c* oxidase complex. Targetted mutagenesis will facilitate elucidation of the structure-activity relationships of the single RegB, RegA and regulatory proteins as well as their complexes. The ability to produce milligram quantities of highly purified RegB protein is also enabling us to undertake 2D/3D crystallisation in order to elucidate the 3D structure of this sensor kinase by electron or X-ray diffraction.

Publications

Saidijam M., Psakis, G., Clough, J.L., Meuller, J., Suzuki, S., Hoyle, C.J., Palmer, S.L., Morrison, S.M., Pos, M.K., Essenberg, R.C., Maiden, M.C.J., Abu-bakr, A., Baumberg, S.G., Stark, M.J., Ward, A., O'Reilly, J., Rutherford, N.G., Phillips-Jones, M.K. & Henderson, P.J.F. (2003) Collection and characterisation of bacterial membrane proteins. *FEBS Lett.* **555**, 170-175.

Potter, C.A., Ward, A., Laguri, C., Williamson, M.P., Henderson, P.J.F. & Phillips-Jones, M.K. (2002) Expression, purification and characterisation of full-length heterologously expressed histidine protein kinase RegB from *Rhodobacter sphaeroides*. *J. Mol. Biol.* **320**, 201-213.

Collaborators

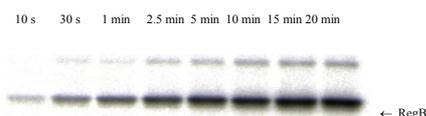
Mike P. Williamson and Cedric Laguri, Krebs Institute Structural Studies Group, University of Sheffield.

Funding

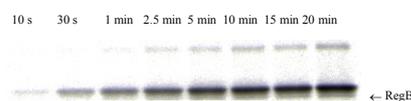
We are grateful to the BBSRC for financial support.

RegB autophosphorylation

A. Aerobic



B. Anaerobic



Phosphotransfer from RegB to RegA

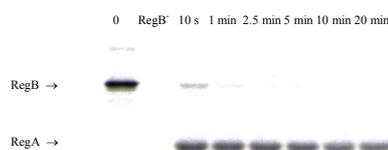


Fig.3. Purified RegB is functional.

Autophosphorylation activity of purified RegB in the presence and absence of oxygen, and phosphotransfer from RegB to RegA.

Solution structure and DNA binding of the effector domain from the global regulator PrrA (RegA) from *Rhodobacter sphaeroides*: Insights into DNA binding specificity

Mary K. Phillips-Jones

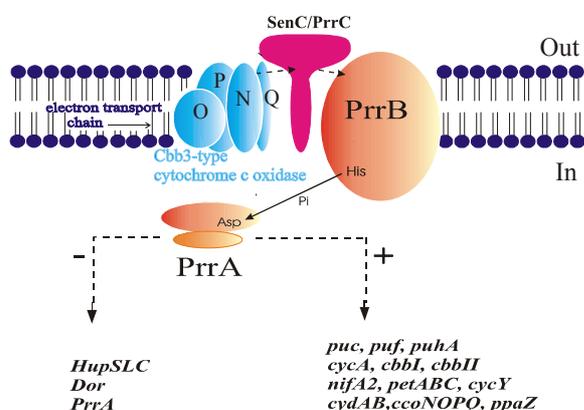


Fig. 1. The Prr pathway in *Rhodobacter sphaeroides*.

strong conservation in several other bacteria, and because of the relative lack of structural studies on the response regulator (RR) protein family compared with the considerable biochemical and genetic information. Furthermore, PrrAC (the C-terminal DNA-binding domain of PrrA) is predicted to be one of the simplest folds so far for RR effector domains and does not belong to any of the three main effector domain families (OmpR/PhoB, NtrC/Dcdd, NarL/FixJ).

We recently presented the solution structure of PrrAC (residues 125-184). It forms a three-helix bundle, each helix forming about a 30° angle with the next one. Helices $\alpha 7$ and $\alpha 8$ form the helix-turn-helix DNA binding motif, of which the recognition helix, $\alpha 8$, is expected to insert into the major groove of DNA. PrrAC belongs to the abundant family of three-helix bundle HTH DNA-binding domains, and it presents structural homologies with prokaryotic and eukaryotic DNA-binding domains such as DNA polymerase and many transcription factors, but with less than 20% identity on average. The PrrAC fold is most similar to Fis (Factor for Inversion Stimulation) protein; indeed, the domain architecture is different from previously characterised response regulator effector domains, as it is shorter than any characterised so far.

We wanted to determine the nature of any interactions between PrrAC and PrrAN (the N-terminal domain) and to find out if PrrAN ‘blocked’ the helix-turn-helix motif in helix $\alpha 8$, as is the case for some response regulators in their inactive state. Comparison of chemical shifts in the full-length protein and PrrAC (facilitated by an unusual inability to see signals from the N-terminal domain, perhaps due to partial unfolding or chemical exchange process) suggested that the region of PrrAC in contact with the N-terminal domain in the inactive full-length PrrA protein is the $\alpha 6$ helix. This area of the protein having interactions with the regulatory domain would not support a direct blockage of the DNA-binding recognition helix by the N-terminal domain. Importantly for PrrA function, the phosphorylation of PrrA by a phosphate analogue, BeF_3^- , shows that PrrA dimerises upon phosphorylation, as observed for many RRs. A complete loss of NMR signal from the PrrAC domain occurs upon phosphorylation, suggesting a drastic conformational and/or dynamic change upon dimerisation.

The Prr/RegA response regulator is a global transcription regulator in purple bacteria *Rhodobacter sphaeroides* and *R. capsulatus*, and is essential for controlling the metabolic changes between aerobic and anaerobic environments in these extremely versatile bacteria. It is part of the Prr signal transduction pathway that senses and responds to redox potential (Fig. 1). The study of PrrA is important because of its essential role in oxygen-mediated regulation of gene expression, the unusually large variation in DNA sequences it recognises and to which it binds, its occurrence and

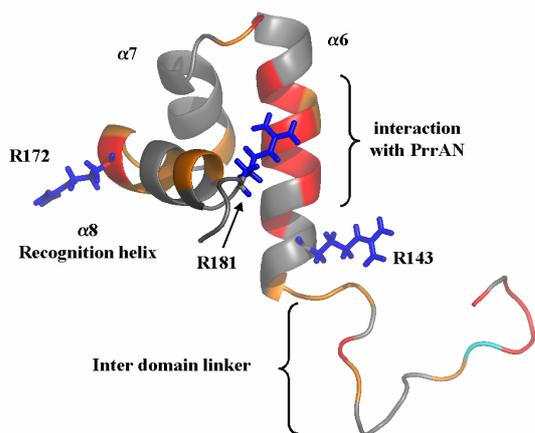


Fig.2. Differences in chemical shift between PrrAC in isolation and in full length PrrA. In red are represented large HN backbone chemical shifts changes and peaks that could not be found in the PrrA spectrum. Medium backbone chemical shift changes are shown in orange. Arg143, Arg172 and Arg181 H ϵ experience large chemical shift changes and their side chains are represented in blue. The end of PrrA N-terminal domain is shown in cyan.

Alignment of Prr/RegA DNA targets permitted a refinement of the consensus sequence, which contains two GCGNC inverted repeats with variable half-site spacings. We noticed that the number of bases between GCG and CGC motifs range from 3 (for the PrrA cluster sequence) to 9 nucleotides. This variable distance puts the recognition elements at different relative positions on the B-DNA helix and suggests that the PrrA dimer and/or the DNA itself would have to adopt different conformations to adapt for different spacings, probably with different affinities. An ability to bind differently spaced, poorly conserved sequences but with some similar structural features might be an advantage for PrrA activity as a global regulator, controlling many different genes and highly conserved in several even non-related, organisms.

Our NMR titrations of PrrAC with specific and non-specific DNA have revealed which surfaces are involved in DNA binding and suggest residues important for binding specificity. We constructed a model of the PrrAC/DNA complex in which two PrrAC molecules are bound to DNA in a symmetrical manner.

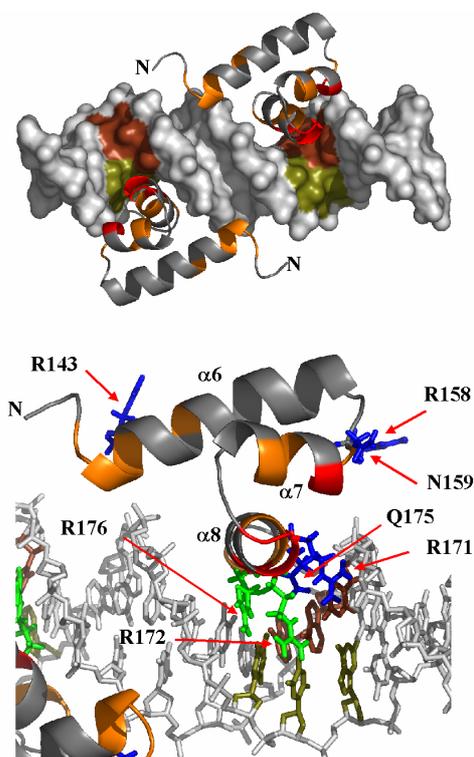


Fig.3. Model of two PrrAC monomers bound to a 20bp DNA fragment. Top: a view of both monomers. Bottom: a detail of PrrAC/DNA interactions. PrrAC monomers were fitted with respect to current published rules for protein/DNA contacts, the DNA titrations and the DNA and protein surfaces. R171, R172, Q175 and R176 sidechain positions are at a distance where the predicted contacts would be possible. The recognised GCG inverted repeats have been coloured by DNA strand (brown and olive). The PrrAC colour code is red and orange for large and medium backbone HN chemical shift variations respectively. Sidechains experiencing chemical shift variations are shown, in green for the ones affected only by puc (specific) binding and in blue the ones affected by both specific and non-specific DNA binding. R166 and R181 have been omitted for clarity.

A careful study of DNA/protein complexes showed that in most such complexes, the DNA is bent. Furthermore, the bending is not usually continuous but shows kinks at discrete sites. The kink sites are generally formed by pyrimidine-purine (YR) steps [CA (=TG), TA or CG], a particularly flexible combination, which, for proteins that bind DNA in adjacent major grooves as many dimers do, are usually found about one helix turn apart (8 to 10 bp). In the PrrA binding sequences, whatever the distance between the right and left repeats there are always two YR steps 8 to 11 bases apart. We therefore suggest that the binding of PrrA to DNA fits this model and is accompanied by kinking. In support of the kink hypothesis, we note that GC-rich regions and AT-rich regions favour compression of the major and the minor groove respectively, therefore likely to encourage the bending process.

An investigation of how the full-length PrrA dimer binds DNA targets with drastically different half-site spacings, and the influence on DNA local structure, are now being investigated. Such investigations could provide important information on the strategies a transcription activator can adopt to bind different DNA sequences and its influence on gene regulation exerted by the PrrA family of response regulators in bacteria.

Publications

Laguri, C., Phillips-Jones, M.K. & Williamson, M.P. (2003) Solution structure and DNA binding of the effector domain from the global regulator PrrA (RegA) from *Rhodobacter sphaeroides*: insights into DNA binding specificity. *Nucl. Acids Res.* **31**, 6778-6787.

Collaborations

This work was carried out in collaboration with Cédric Laguri and Michael P. Williamson of the Krebs Institute Structural Studies Group, University of Sheffield. We are grateful to Peter Henderson for helpful discussions.

Funding

We thank the BBSRC for financial support.

Amyloid formation of peptides and point mutants of β_2 -microglobulin

Susan Jones, David Smith, Clemens Stilling, Geoff Platt, Thomas Jahn, Isobel Morton, Steve Homans and Sheena Radford

Introduction

There are approximately twenty proteins with unrelated amino acid sequences and native structures that aggregate to form highly ordered amyloid fibrils *in vivo*, and result in specific disease states. More recently, several of these proteins, and some that are not disease-related, have been shown to form amyloid *in vitro* by manipulation of solution conditions. In all cases, the normally soluble, monomeric proteins deposit as insoluble fibres with characteristic cross- β structure. Our research aims to elucidate the mechanism of fibril formation for human beta-2-microglobulin (β_2 m), which causes haemodialysis related amyloidosis in all patients with renal failure.

Fibril formation by peptides

We have analysed the role of each individual β -strand of β_2 m in the formation of amyloid-like fibrils *in vitro* by utilising peptides that correspond to each of the seven strands in the native protein. Only two sequences, both of which encompass strand E of intact β_2 m (residues 59-71 (peptide E) and 59-79 (peptide E')) form fibrils *in vitro*. These peptides form fibrils under the conditions previously shown to promote amyloid formation by the intact protein, but can also associate to form fibrils at neutral pH. In addition, fibrils formed from these two peptides enhance fibrillogenesis of the intact protein. No correlation was found between secondary structure propensity, peptide length, pI or hydrophobicity and the ability of the peptides to associate into amyloid. However, only the amyloidogenic peptides contain a relatively high content of aromatic side-chains. We propose that residues 59-79 may be one of the important regions involved in the self-association of partially folded β_2 m into amyloid fibrils and studies of this sequence could shed light on the assembly mechanism of the intact protein *in vitro*.

Fibril formation by edge-strand mutants

We have also examined the role of edge strand fraying in the initiation of β_2 m fibrillogenesis, by creating mutants in the N- and C-terminal (A and G) strands that specifically remove hydrophobic packing between these strands and the rest of the protein. By contrast with similar mutations elsewhere in the protein, we have shown that mutations in strands A and G

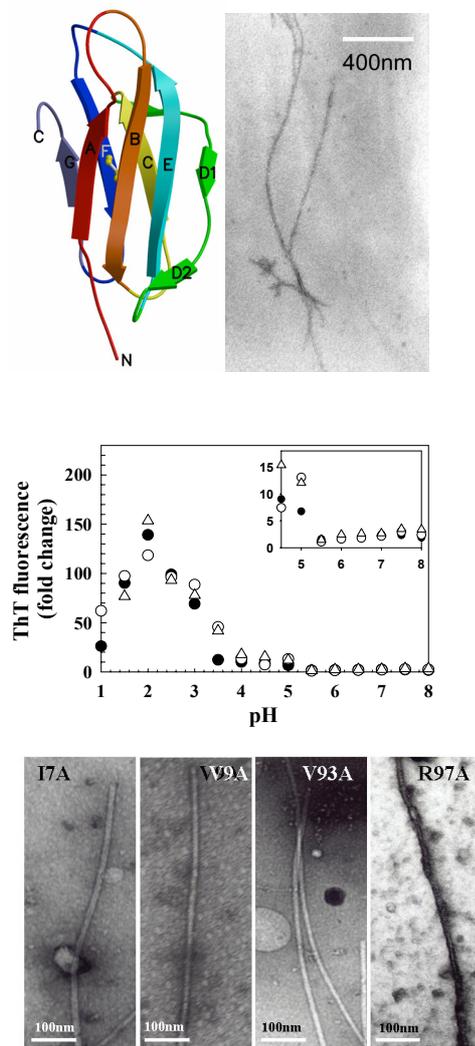


Fig. 1. Fibril formation by peptides and mutants. Ribbon diagram of β_2 m rainbow coloured according to the position of the peptides (A). Electron micrograph of peptide E fibrils at pH 3.5 (B). Fibril formation of the strand A mutant I7A monitored by Thioflavin-T (C) and electron micrograph of all strand A and G mutant fibrils at pH 7.0 (D)

not only increase the rate of fibril formation at acidic pH values, but also result in fibril formation at neutral pH. This is the first time that β_2m has been shown to form fibrils *ab initio* at physiological pH *in vitro*. The data are consistent with the view that perturbation of the N- and C-terminal strands is an important feature in the generation of assembly-competent states of β_2m .

Structural properties of the amyloidogenic intermediates at different pHs revealed by 2D NMR

Structural characterisation of amyloidogenic precursor conformations is vital to further our understanding of the mechanisms of fibril formation. We are currently using a variety of NMR techniques to study the structural and dynamic properties of monomeric β_2m at a range of pH values. By comparing wild-type β_2m and specific mutants, we aim to identify characteristics that make proteins like β_2m intrinsically amyloidogenic.

Publications

Jones, S., Manning, J., Kad N.M., and Radford, S.E. (2003) Amyloid-forming peptides from β_2 -Microglobulin—Insights into the mechanism of fibril formation *in vitro*. *J. Mol. Biol.* **325**, 249-257.

Jones, S., Smith, D.P., and Radford, S.E. (2003). Role of the N and C-terminal strands of β_2 -microglobulin in amyloid formation at neutral pH. *J. Mol. Biol.* **330**, 935–941

Smith D.P., Jones, S., Serpell, L.C., Sunde, M. and Radford, S.E. (2003). A systematic investigation into the effect of protein destabilisation on beta 2-microglobulin amyloid formation. *J. Mol. Biol.* **330**, 943–954

Funding

We gratefully acknowledge the University of Leeds, BBSRC, EPSRC, The National Kidney Research Fund and The Wellcome Trust for financial support. SER is a BBSRC Professorial Research Fellow.

Investigating amyloid self-assembly by atomic force microscopy

Walraj Gosal, Sarah Myers, Sheena Radford, Alastair Smith and Neil Thomson.

Introduction

The clarification of protein self-assembly mechanisms are critical to our understanding of many normal and pathological conditions in biology. One such example is amyloid formation, which is associated with a number of diseases, as well as protein-only ('prion') inheritance. We have utilised atomic force microscopy (AFM) to study the self-association of β_2 -microglobulin, apoC-II and poly-Q containing sequences. This technique involves the use of a very fine probe that scans across samples bound to very flat surfaces, producing three-dimensional images on the nanometre scale.

Complexities in the β_2 -microglobulin self-assembly process

The deposition of β_2 -microglobulin, a 99 residue all- β -sheeted protein, into amyloid fibrils is associated with the condition 'haemodialysis-related amyloidosis'. Amyloid fibrils are formed *in vitro* by the association of partially unfolded molecules formed at low pH. Using AFM, we have identified fibrils various morphologies that are formed during this process, many of which form in a hierarchical manner. This self-assembly is achieved through the lateral association of elementary filaments (or protofilaments), producing a variety of higher-order helical structures.

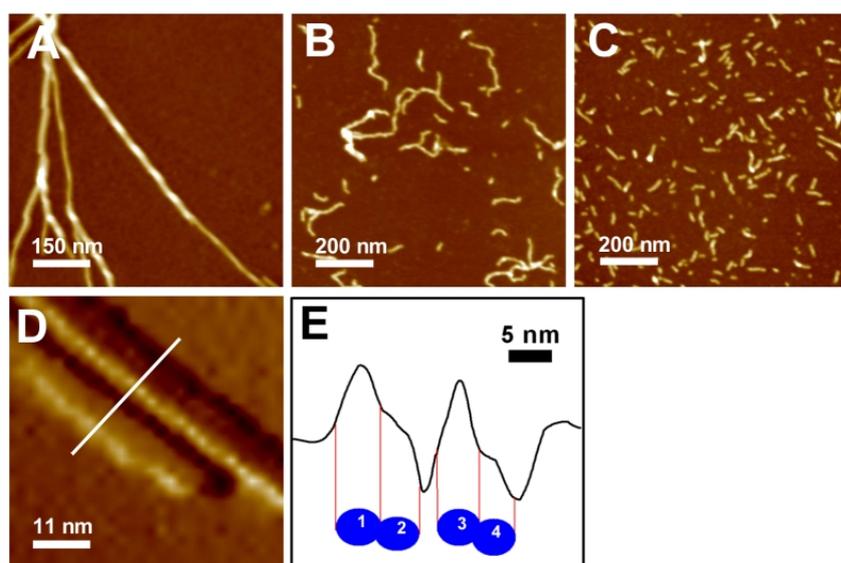


Fig. 1. (A-C) AFM images of the various morphologies of β_2 -microglobulin amyloid fibrils obtained under varying conditions of pH and ionic strength. (D) AFM amplitude image obtained under fluid, and the resulting amplitude profile (E) suggests a fine filamentous elementary subunit (protofilament), of width $\sim 3-4$ nm.

In addition, we have identified other fibril morphologies that form when conditions such as pH, ionic strength and protein concentration are varied (Fig. 1A-C). This raises the question, how do these morphologies arise, and what is the underlying controlling mechanism? One explanation could be changes in the relative ratio between nucleation and growth rate constants. An alternative explanation necessitates the introduction of more than one self-assembly pathway, regulated at the molecular level by at least two distinct β_2 -microglobulin conformations. We are currently investigating whether morphologies can act as self-replicating templates under conditions where they would not usually form, as has been shown to occur in prion self-assembly. Additionally, AFM experiments under fluid have provided us with detailed results regarding the substructure of amyloid fibrils. In particular, amplitude analysis under fluid partially resolves the dimensions of an elementary protofilament, which we suggest is cylindrical with a diameter of $\sim 3-4$ nm (Fig. 1D & 1E).

Early stages in the β_2 -microglobulin fibril assembly pathway

The pathway of fibril formation can be followed by fluorescence measurements using an amyloid specific binding fluorescent dye, Thioflavin T. Under certain conditions, these kinetic data show a lag and elongation phase. Within the lag phase, pre-fibrillar structures are formed which presumably facilitate fibril growth, and these have been analysed using AFM (Fig. 2). These data demonstrate that even as early as ~5 minutes into the lag phase, small aggregated structures form, and these precede the formation of fibrils, which first appear ~25 minutes after the pH is reduced (Fig. 2). Future work will involve the use of biological molecules, especially those implicated in dialysis-related amyloidosis, to determine their effect on fibril growth. Also many types of *ex vivo* amyloid fibrils from different sources will be imaged by AFM.

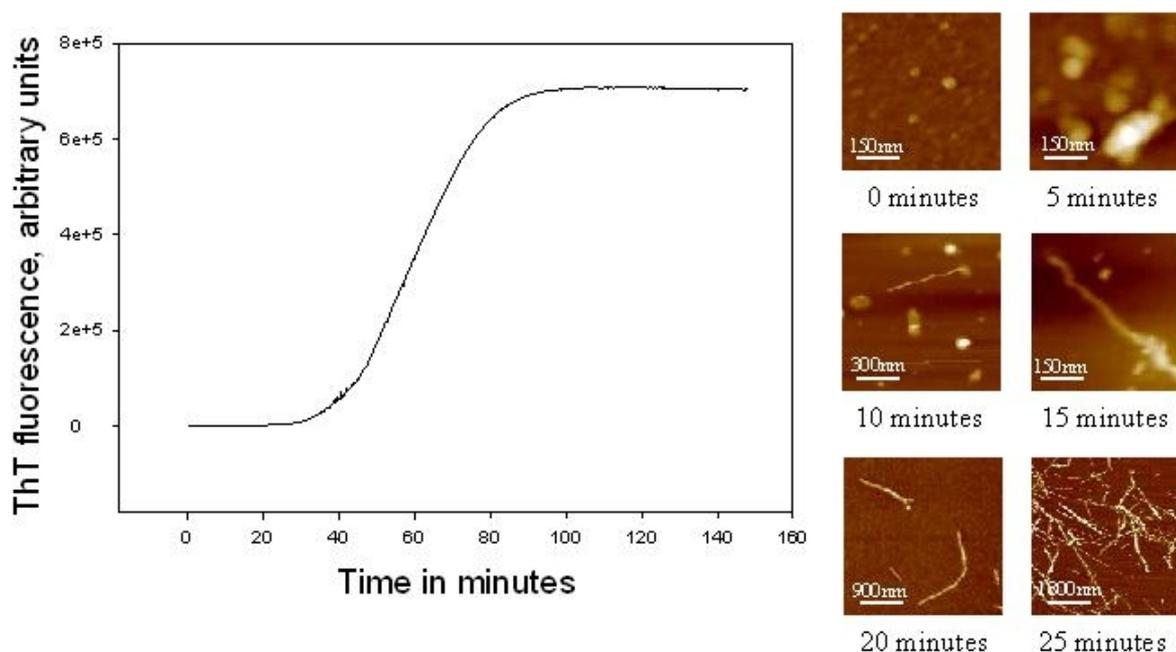


Fig. 2. Thioflavin T fluorescence assay for fibril formation, and corresponding AFM images of structures formed during the 'lag-phase' of fibril formation.

Publications

Kad, N.M., Smith, D.P., Myers, S.L, Smith, D.A., Radford, S.E. & Thomson, N.H. (2003). Hierarchical assembly of β_2 -microglobulin amyloid *in vitro* revealed by atomic force microscopy. *Journal of Molecular Biology*. **330**, 785-797.

Hatters, D.M., MacRaid, C.A., Daniels, R., Gosal, W.S., Thomson, N.H., Jones, J.A., Davis, J.J., MacPhee, C.E., Dobson, C.M. & Howlett, G.J.(2003) The circularisation of amyloid fibrils formed by apolipoprotein C-II. *Biophysical Journal*. **85**, 3979-3990.

Collaborators

Geoff Howlett, University of Melbourne, Australia.

Erich Wanker, Max Delbrück Center for Molecular Medicine, Germany.

Acknowledgements

We gratefully acknowledge the University of Leeds, BBSRC and EPSRC for financial support. SER is a BBSRC Professorial Research Fellow. NHT is an EPSRC Advanced Research Fellow.

Characterising the factors that describe the mechanical resistance of proteins

Anthony Blake, David Brockwell, David Sadler, Tomoko Tezuka-Kawakami, Dan West, Rebecca Zinober, Peter Olmsted, Alastair Smith and Sheena Radford

Introduction

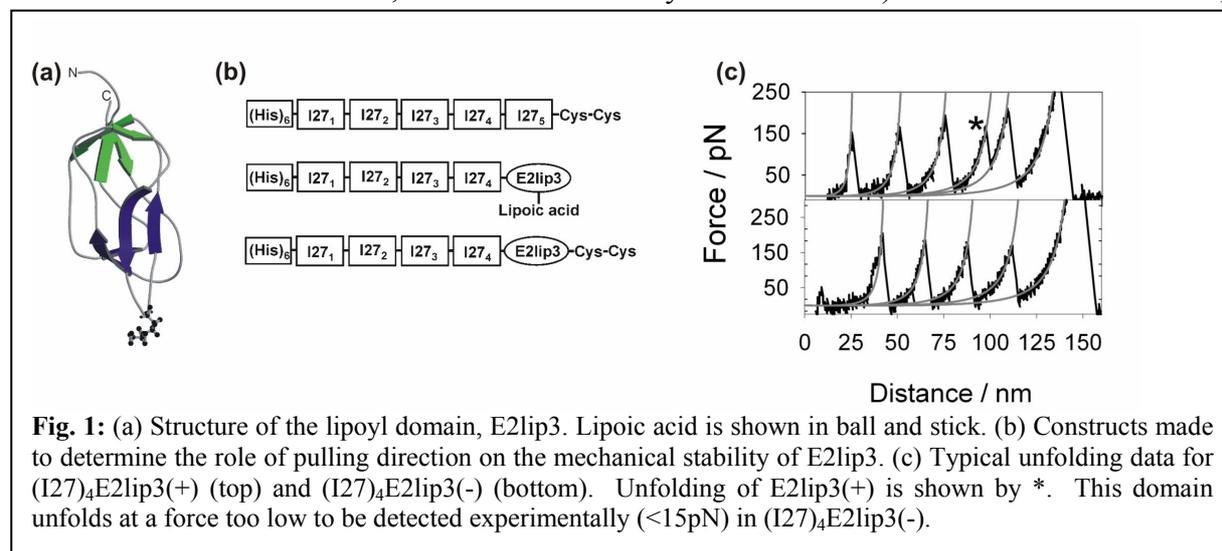
Exciting new experiments using the atomic force microscope (AFM) permit measurements of the mechanical unfolding of proteins at the single molecule level. Using this method, piconewton forces can be measured with nanometer resolution. Whilst such experiments initially focussed on naturally occurring poly-proteins, such as those found in muscle fibres, more recently a number of different protein structures have been measured using this technique, by creating novel polymers of concatenated multiple copies of a gene product joined by short linker regions. Today, mechanical unfolding of the giant muscle-protein titin and one of its domains, I27, have become the paradigms of mechanical unfolding experiments. Other proteins studied using this method include tenascin, T4 lysozyme and barnase, demonstrating the versatility and scope of these experiments.

Ground state perturbations do not affect mechanical stability

Although early experiments suggested that proteins unfold *via* similar pathways under mechanical stress and in the presence of chemical denaturant, more recent results from both our laboratory and elsewhere have used site-directed mutagenesis to show that the chemical and mechanical unfolding landscapes of I27 differ significantly. To determine the properties of the mechanical and chemical unfolding transition states of I27 in more detail, we have investigated the effect of alterations in the solution conditions on the chemical and mechanical unfolding rate constants. We have found that whilst alteration of the pH or the introduction of 8% (v/v) trifluoroethanol have similar effects on the rate constants for chemical and mechanical unfolding, adding 0.4M sodium sulphate decreases the chemical unfolding rate constant significantly, but has little effect on mechanical unfolding. These data support the view that mechanical unfolding probes local rather than global stability and demonstrate that mechanical unfolding offers a new method of mapping the shape of an unfolding energy landscape.

Pulling geometry determines mechanical resistance

One method of exploring in more detail the extent to which mechanical unfolding probes local stability is to pull the *same* protein in *different* directions (rather than simply applying force to the N- and C-termini, as has been routinely utilised to date). In order to achieve this,



we have been investigating the mechanical unfolding properties of the inner lipoyl domain of the dihydrolipoyl acetyltransferase subunit of the pyruvate dehydrogenase multienzyme complex from *E. coli* (E2lip3) (Fig. 1). The natural lipoylation of this domain allows force to be applied to the protein between the N-terminus and lipoyl-lysine at residue 41, as well as between the N- and C-termini (in the absence of lipoylation, but in the presence of C-terminal Cys introduced by mutagenesis) (Fig. 1b). Using these constructs we have shown that extending E2lip3 *via* the N- and C-termini results in very low unfolding forces (<15pN), whilst extension between the N-terminus and the lipoyl-lysine moiety at residue 41 results in a high unfolding force, similar to that observed during mechanical unfolding of I27 (Fig. 1c). These data suggest that application of a peeling force to directly hydrogen bonded strands results in weak mechanical resistance, whilst shearing hydrogen bonded strands results in significant mechanical resistance. These data suggest that mechanical unfolding is intimately related to the geometry of the applied extension. Altering the points of attachment of a protein to substrate and tip thus offers a unique way of probing different barriers in protein unfolding. Current work is directed at pulling I27 in different directions relative to its secondary structure elements to map the mechanical unfolding landscape of this domain and to assess the shape of the energy landscape for unfolding of this classic immunoglobulin domain.

Protein L is mechanically resistant

Most mechanical unfolding studies performed to date have focused on immunoglobulin and fibronectin type domains. Based on the results presented above, we have predicted that other proteins that contain parallel hydrogen bonded terminal strands should be mechanically resistant, regardless of whether they play a mechanical role *in vivo*. To test this hypothesis we have commenced a new study of protein L, a small 62 amino acid domain with a simple α/β topology from *Peptostreptococcus magnus*. Preliminary results show that protein L shows remarkable resistance to force, supporting the hypothesis that mechanical unfolding monitors local stability of hydrogen bonded strands, and opening the door to dissection of the origins of mechanical resistance for a simple protein domain.

Collaborators

Godfrey Beddard, School of Chemistry, University of Leeds.

Emanuele Paci, University of Zürich, Switzerland.

Richard Perham, Department of Biochemistry, University of Cambridge.

Publications

Best, R.B., Brockwell, D.J., Toca-Herrera, J.L., Blake, A.W., Smith, D.A., Radford, S.E. and Clarke, J. (2003) Force mode microscopy as a tool for protein folding studies. *Analytica Chimica Acta* **479**, 87-105

Brockwell, D.J., Paci, E., Zinober, R.C., Beddard, G.S., Olmsted, P.D., Smith D.A., Perham, R.N. and Radford, S.E. (2003) Pulling geometry defines the mechanical resistance of a β -sheet protein. *Nature Structural Biology* **10**, 731-737

Smith, D.A., Brockwell, D.J., Zinober, R.C., Blake, A.W., Beddard, G.S., Olmsted, P.D. and Radford, S.E. (2003) Unfolding dynamics of proteins under applied force. *Phil. Trans. R. Soc. Lond. A* **361**, 713-730

Acknowledgements

We thank Keith Ainley for technical support and the Wellcome Trust, the University of Leeds, the BBSRC and EPSRC for funding. SER is a BBSRC Professorial Fellow.

Mapping the folding landscapes of the four-helix proteins Im7 and Im9

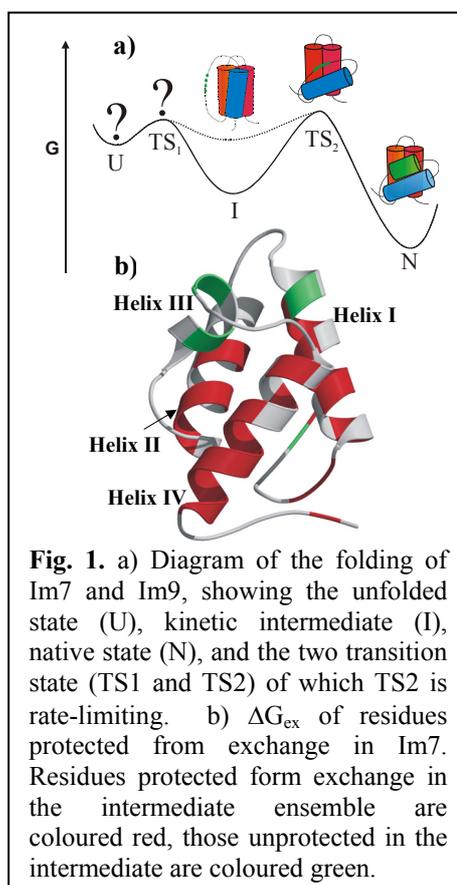
Claire Friel, Graham Spence, Stuart Knowling, Victoria Morton, Eva Cobos Sanchez, Susanne Craz and Sheena Radford.

Introduction

Most single domain proteins have the ability to fold spontaneously into a precise, functional three-dimensional structure in seconds or less. Understanding how this transition occurs will not only help in uncovering the way in which amino acid sequence encodes protein structure, but also is likely to provide insight into the folding/unfolding transitions that many proteins undergo as a part of their normal functioning, as well as the misfolding processes that underlie a number of disease states. We use the four-helix bundle proteins, Im7 and Im9, as a model system for the study of protein folding. These proteins have the same native structure and 50% sequence identity and yet under identical conditions (pH7.0, 10°C) can fold to the native state with mechanisms of different kinetic complexity. Im7 folds through a rapidly populated on-pathway intermediate whereas, Im9 folds without populating an intermediate.

Structure of the Im7 intermediate

Im7 folds through an on-pathway intermediate (Fig. 1a), which has been shown, using Φ -analysis, to have a specific, but partially misfolded hydrophobic core. Using equilibrium hydrogen exchange, monitored site-specifically using NMR, we have gained further insights into the structure of this species. By measuring the amide exchange rates (under EX2 conditions), the backbone hydrogen bonds formed in the intermediate have been identified. The data show that amides in helices I, II and IV exchange slowly, with a free energy similar to that of global unfolding, suggesting that these helices form highly protected hydrogen bonded helical structure in the intermediate (Fig. 1b). By contrast, amides in helix III exchange rapidly, confirming that helix III does not form stable secondary structure in the intermediate. These results are in agreement with the results previously obtained from Φ -analysis, and show that the on-pathway folding intermediate of Im7 contains extensive, stable hydrogen bonded structure in helices I, II and IV (Fig. 1b), and that this structure is stabilised by both native and non-native interactions involving amino acid side-chains in these helices.



In order to gain more detailed structural information about the Im7 folding intermediate, experiments are currently underway to trap this species at equilibrium so that high-resolution structural information can be obtained. A number of mutants have been successfully created that prevent the binding of helix III to the developing structure, such that the intermediate is trapped at equilibrium. These mutants now offer a unique opportunity to determine the structure of a transiently populated intermediate at high resolution.

Tailoring the folding kinetics of Im9

Determination of the structures of the rate-limiting transition states for folding of Im7 and Im9, using Φ -value analysis, has shown that these species are structurally similar ensembles. This suggests that the different kinetic mechanisms for folding of these homologous proteins

do not result from radically different folding landscapes. Instead, these differences may result from the differing stability of a common intermediate species (Fig. 1a). Using knowledge of the structure of the intermediate formed during Im7 folding, gained from Φ -value analysis, we have identified a number of sequence differences between Im7 and Im9 that may be responsible for the instability of the putative high-energy intermediate in Im9 folding. Consequently we have been able to engineer a populated intermediate into the folding of Im9 at pH7.0, by the introduction of specific substitutions at positions known to stabilise the Im7 intermediate (Fig. 2). We have also shown that the intermediate populated during the folding of the redesigned Im9 structurally resembles the misfolded intermediate formed during Im7 folding.

Other work currently ongoing includes investigation into the nature of the reorganisation that occurs on crossing the rate-limiting transition state (TS2), investigating the role of helical propensity in immunity protein folding, and determination of the structures of TS1 and the unfolded state.

Collaborators

Martin Karplus, Harvard University
 Colin Kleanthous, University of York
 Geoff Moore, University of East Anglia
 Emmanuele Paci, University of Zurich
 Michele Vendruscolo, Cambridge University

Publications

Friel, C.T., Capaldi, A.P. & Radford, S.E. (2003) Structural analysis of the rate-limiting transition states in the folding of Im7 and Im9: similarities and differences in the folding of homologous proteins. *J. Mol. Biol.*, **326**, 293-305.

Paci, E., Friel, C.T., Lindorff-Larsen, K., Radford, S.E., Karplus, M. & Vendruscolo, M. (2004) Comparison of the transition state ensembles for folding of Im7 and Im9 determined using all-atom molecular dynamics simulations with value restraints. *Proteins: Structure, Function and Genetics*, **54**, 513-525.

Gorski, S.A., Le Duff, C.S., Capaldi, A.P., Kalverda, A.P., Beddard, G.S., Moore, G.R. & Radford, S.E. (2004) Equilibrium hydrogen exchange reveals extensive hydrogen bonded secondary structure in the on-pathway intermediate of Im7. *J. Mol. Biol.*, in press

Friel, C.T., Beddard, G.S., & Radford, S.E. (2004) Switching two-state to three-state kinetics in the helical protein Im9 by rational redesign. *J. Mol. Biol.*, submitted.

Acknowledgements

We thank the BBSRC and the Wellcome trust for funding and Keith Ainley for technical support.

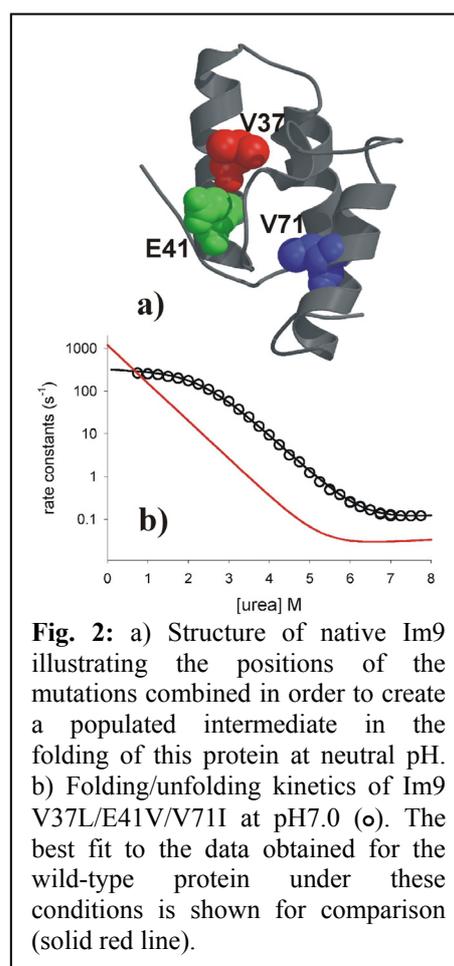


Fig. 2: a) Structure of native Im9 illustrating the positions of the mutations combined in order to create a populated intermediate in the folding of this protein at neutral pH. b) Folding/unfolding kinetics of Im9 V37L/E41V/V71I at pH7.0 (○). The best fit to the data obtained for the wild-type protein under these conditions is shown for comparison (solid red line).

Cryo-EM of macromolecular complexes

Neil Ranson

Introduction

The molecular chaperone GroEL, together with its co-protein GroES, mediates an ATP-dependent enhancement in the folding of a range of proteins both *in vitro* and *in vivo*. ATP binding promotes large conformational changes in GroEL, which have been visualised using cryo-EM, but not in crystallographic studies that show ATP γ S-bound GroEL in an unliganded conformation. Upon ATP binding to one of its two available rings, GroEL undergoes global conformational change that primes the ring for binding of the co-protein GroES, forming a large proteinaceous cage in which substrate proteins become encapsulated. This GroEL-GroES-ATP complex is the “folding-active” state: substrate proteins can fold to reach their native state *inside* the GroEL-GroES cage. The structure of this “folding-active” state has been probed using cryo-EM and single particle image processing, and compared to the X-ray structure of the “non-folding active” GroEL-GroES-ADP complex.

Results & Future Work

The GroEL(D398A)-GroES-ATP complex was imaged in unsupported vitreous ice, using the FEI Tecnai F20 microscope at Birkbeck College. A range of defocus values was used during data collection to allow full CTF correction to be carried out during image processing. Images were refined against an initial starting model created by filtering the crystal structure of the related GroEL-GroES-ADP complex (1AON) to $\sim 35\text{\AA}$ resolution, removing all fine structural detail. Refinement was carried out using an angular refinement strategy implemented in the SPIDER image processing software, together with MSA classification methods using IMAGIC. This refinement has currently reached a resolution of 7.7\AA (measured using a Fourier Shell Correlation of 50%). Fitting of the atomic coordinates of the individual domains is providing a detailed picture of conformational rearrangements within the complex (*manuscripts in preparation*).

With the recently installed Tecnai F20 Field Emission Gun microscope, fitted with a Gatan Ultrascan4000 4k x 4k CCD camera, in Leeds, we are also beginning structural studies on a wide range of macromolecular targets including viruses, viral protein assemblies, and a range of molecular chaperone systems.

Collaborators

Helen Saibil & Dan Clare, Birkbeck College London
Arthur Horwich, George Farr & Wayne Fenton, Yale University, New Haven, Connecticut, USA.
Sabine Rospert, University of Freiburg.

Funding

This work is being funded by the University of Leeds, The Royal Society, and by the BBSRC.

Targeting the Hepatitis C virus ion channel p7 for anti-viral therapy

Stephen Griffin, Dean Clarke, Rachel Trowbridge, Martin Higgs, Steve Evans,
Alastair Smith, Joachim Jäger, Mark Harris and David Rowlands

Introduction

Hepatitis C virus (HCV) currently infects over 3% of the world population and is the major indicator for liver transplant surgery in the west. Acute infection is usually asymptomatic but leads to persistence in the majority of cases, causing chronic liver disease. Treatment of the virus is currently limited to the use of type 1 interferon either alone, or in combination with the guanosine analogue, ribavirin. The therapeutic regime is expensive, poorly tolerated, and effective in only 40% of cases world-wide. Furthermore, resistance to this treatment is common in the viral genotypes found in the west. The search for a vaccine and alternative therapies has been hampered by the current inability to successfully culture the virus *in vitro*, making the identification of new anti-viral drug targets paramount.

The HCV p7 protein is a small hydrophobic protein of 63 amino acids comprised of two *trans*-membrane alpha helices separated by a short positively charged cytoplasmic loop. It has now been shown to be essential for HCV infectivity. Deletion of the cytoplasmic loop of the p7 protein of a related virus, bovine viral diarrhoea virus (BVDV) has also been shown to abolish viral infectivity. p7 is proposed to belong to a family of proteins known as viroporins, which homo-oligomerise to form aqueous pores in cellular membranes. Perhaps the best characterised of these proteins is the M2 channel of Influenza A virus: the target of the first anti-viral drug, amantadine.

Biophysical studies of p7 oligomers

Recombinant expression of p7 was achieved by fusion with a glutathione-S-transferase (GST) in *E.coli*. Oligomeric complexes were visualised in liposomal membranes by TEM in collaboration with Dr. Lucy Beales (University of Texas Medical Branch). Both GST-HIS-p7 and near-native HIS-p7 were seen to form ordered ring structures with consistent dimensions. The same proteins have also been visualised using Atomic Force Microscopy (AFM) in collaboration with Dr. Alastair Smith and Professor Jennifer Kirkham. Oligomeric ring complexes of similar size to those in TEM were observed using AFM. These oligomeric complexes were seen to accumulate in a fluidic environment over time and were excluded from cholesterol-rich lipid regions. Averaging software for analysis of TEM images has been applied to micrographs of GST-FLAG-p7 fusions in liposomal membranes, in collaboration with Dr. Stan Burgess. The software revealed a heptameric complex with 7-fold symmetry (Fig. 1).

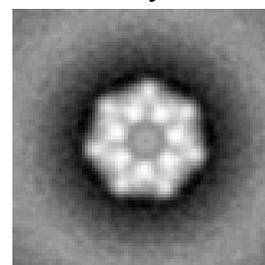


Fig. 1
Averaged TEM image
of GST-FLAG-p7 in
liposomes

p7 ion channel activity is abrogated by amantadine in cells and *in vitro*

Previously, we were the first to demonstrate an ion channel activity for p7 using black lipid membranes (BLM). We also reported that this ion channel activity could be blocked with the addition of 1 μ M amantadine. In support of this data, a cell based assay has been employed as an indirect assessment of ion channel activity, in collaboration with Dr. Wendy Barclay (University of Reading). Mutations of the positive residues within the cytoplasmic loop of p7 abrogated function, as did the inclusion of amantadine. Additionally, we were able to show that the p7 protein of the related virus, BVDV has ion channel activity in this system.

p7 localisation

Co-expression of a green fluorescent protein fused p7 with fluorescent cellular markers in mammalian cells, showed that p7 localises to various membranous organelles, notably mitochondria. Distribution of p7 throughout the cell was also variable over time. Mutations to the positive charge on the p7 cytoplasmic loop had no effect on localisation.

GBVB as a surrogate model for HCV p7 function

The inability to grow HCV in culture highlights a demand for model systems in which individual viral proteins can be analysed in the context of a whole virus. One such model is the related virus, hepatitis GB virus B (GBVB), the closest homologue to HCV, which can be cultured in primary tamarin cells. We are working with GSK to identify the size and position of the p7 protein within the GBVB polyprotein.

Collaborations

Dr. Lucy Beales (University of Texas Medical Branch)

Dr. Wendy Barclay (University of Reading)

Publications

Griffin, S.D., Beales, L.P., Clarke, D.S., Worsfold, O., Evans, S.D., Jager, J., Harris, M.P. & Rowlands, D.J. (2003) The p7 protein of hepatitis C virus forms an ion channel that is blocked by the antiviral drug, amantadine. *FEBS Lett*, **30**, 34-8

Griffin, S.D., Harvey, R., Clarke, D.S., Barclay, W.S., Harris, M. & Rowlands, D.J. (2003) A conserved basic loop in the hepatitis C virus p7 protein is required for amantadine-sensitive ion channel activity in mammalian cells but is dispensable for localisation to mitochondria. *J. Gen. Virol.* **85**, 451-461.

Funding

This work was funded by the MRC “Enabling Technologies” Co-operative, University of Leeds and by GSK.

Discrimination among rhinovirus serotypes for a variant ICAM-1 receptor molecule

Tobias Tuthill, Lisa Challinor, Richard Killington and David Rowlands.

Introduction

Intercellular Adhesion Molecule 1 (ICAM-1) is the cellular receptor for the major group of human rhinovirus serotypes, including human rhinovirus 14 (HRV14) and human rhinovirus 16 (HRV16). A naturally occurring variant of ICAM-1, ICAM-1^{Kilifi}, in which Lys 29 is mutated to Met, binds to these serotypes of HRV with altered characteristics relative to ICAM-1.

Biochemical analysis of receptor binding *in vitro*

HRV14 binds to ICAM-1 only transiently at physiological temperatures but forms a stable complex with ICAM-1^{Kilifi}. Conversely HRV16 forms a stable complex with ICAM-1 but does not bind to ICAM-1^{Kilifi} (Fig. 1).

The structure of the virus-receptor complexes

The three dimensional structures of HRV14 and HRV16 complexed with ICAM-1 and the structure of HRV14 complexed with ICAM-1^{Kilifi} have been determined by cryo-electron microscopy (cryo-EM) image reconstruction to approximately 13Å resolution. The structures of both viruses and of ICAM-1, as determined by X-ray crystallography, were fitted into the cryo-EM density maps. The interfaces between the viruses and receptor contain extensive ionic networks, however the interactions between the viruses and ICAM-1^{Kilifi} contain one less salt bridge than with ICAM-1. As HRV16 has fewer overall interactions with ICAM-1 than HRV14, the absence of this charge interaction will have a greater impact on the binding of ICAM-1^{Kilifi} to HRV16 than to HRV14.

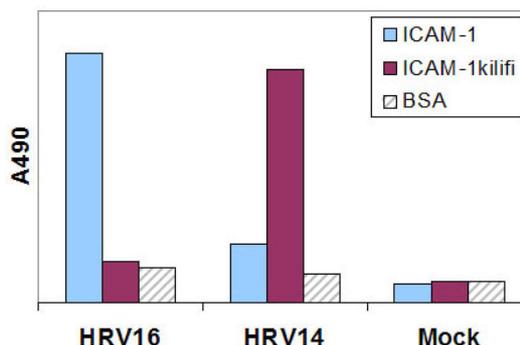


Fig 1. Virus-receptor binding by capture ELISA using immobilized soluble receptor molecules.

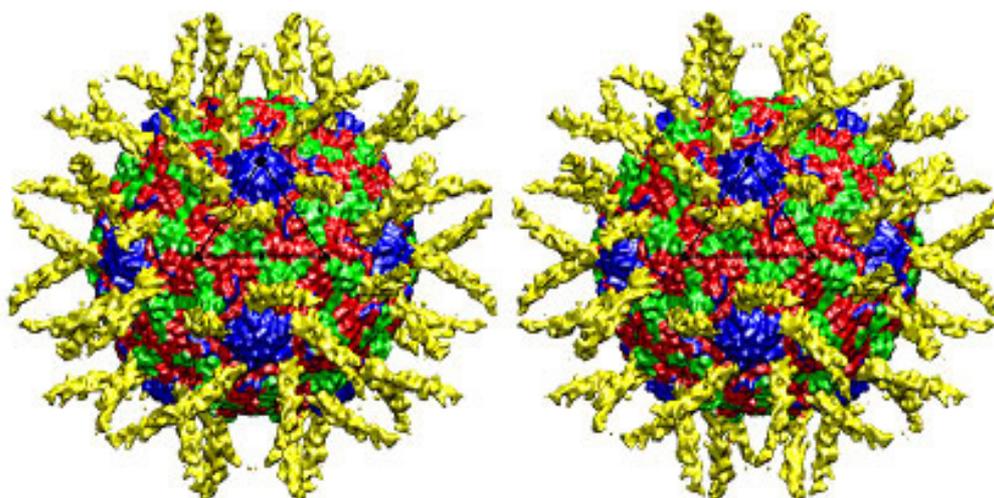


Fig 2. Surface shaded Cryo-EM map of HRV16 complexed with ICAM-1^{wt}. VP1, VP2, VP3 and ICAM-1^{wt} are coloured blue, green, red and yellow, respectively.

Collaborations

Structural aspects of this work were carried out by Chuan Xiao, Carol Bator, Paul Chipman and Michael Rossmann of the Department of Biological Sciences, Purdue University, USA. The study was done in collaboration with Alister Craig of the Liverpool School of Tropical Medicine, UK.

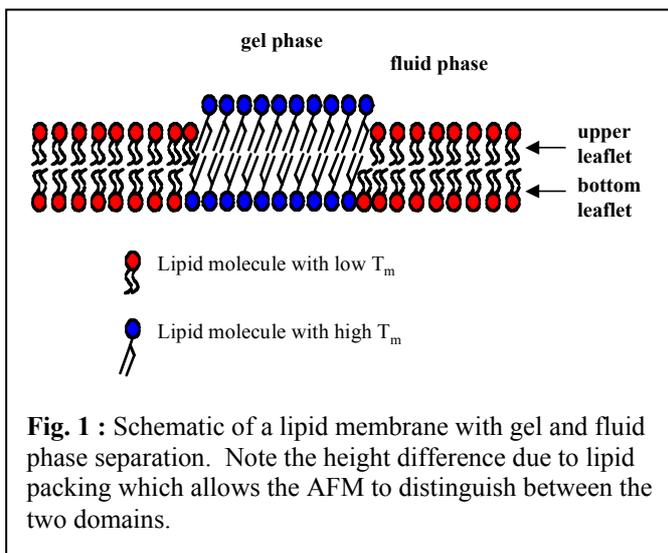
Funding

The work was funded by the MRC.

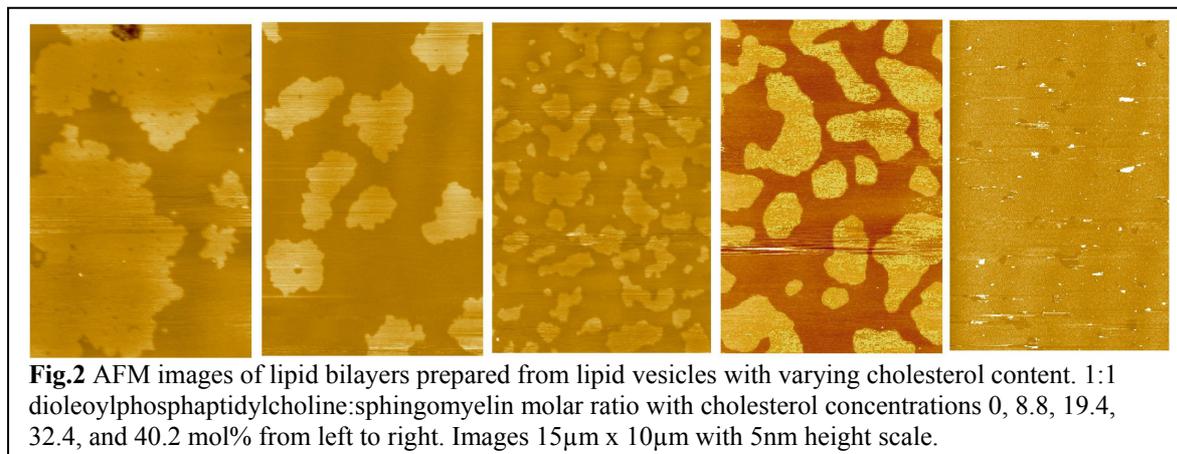
Atomic force microscopy of microdomains in lipid bilayers

Ashley Garner, Simon Connell, Guanbin Li, Alastair Smith, John Colyer and Nigel Hooper

The reported existence of microdomains in plasma membranes has prompted considerable interest in recent years. These domains, often referred to as lipid rafts, are enriched in sphingomyelin and cholesterol and are proposed to be involved in many functions, including membrane trafficking, protein sorting and signal transduction. Despite a wealth of literature concerning lipid rafts and their associated proteins, there is still little direct evidence of their existence, due to difficulties visualizing such small, possibly transient features, in living cells.



Microdomains of a similar composition to lipid rafts have been shown to form in lipid bilayers (Fig. 1). These bilayers can be prepared on solid supports from lipid vesicles and provide a model to investigate raft formation.



Using AFM these lipid bilayers can be imaged under physiological conditions with nanometre resolution. We have been using AFM to investigate the physical properties of microdomains in bilayers of varying lipid compositions (Fig. 2).

Triton X-100 insolubility

Lipid rafts and their associated proteins are frequently defined by their insolubility in the non-ionic detergent Triton X-100 at 4°C. We have investigated Triton X-100 insolubility of supported bilayers in real time by AFM for a variety of lipid compositions. Our results indicate that treatment with Triton X-100 can affect the size of existing microdomains and even induce their formation under certain conditions.

Future work

Future studies will involve using AFM to investigate the distribution of raft associated proteins in bilayer membrane models.

Publications

Li, G., Connell, S.D., Olmsted, P.D., Colyer, J., Hooper, N.M., Smith, D.A. (2004) The nature of detergent resistant domains in supported lipid bilayers. Submitted to *J Biol Chem*.

Funding

We gratefully acknowledge the BBSRC and MRC for financial support.

The dynamic response of biomolecules

Katherine Byrne, Masaru Kawakami, Bhavin Khatri, Alastair Smith, Tony Blake, David Brockwell, Sheena Radford, Peter Olmstead and Tom McLeish.

Introduction

Atomic force microscopy has become a useful tool for examining the mechanical response of single molecules to an applied force.

A commercial microscope has been modified to measure the dynamic response of single molecules to oscillations under tension. This will enable determination of the viscoelastic behaviour of single molecules, and provide extra parameters for modelling the energy landscape associated with the molecule. In addition, the method may be used to explore interesting sections of the energy landscape in detail. The work may also be extended to assess the change in dynamic response of molecules undergoing functional structural change.

There are still some obstacles to be overcome in developing the instrument, but we have already been able to measure the dynamic response of a polysaccharide as it undergoes a structural transition under increasing tension.

Single molecules under tension

The response of single molecules, suspended between the end of a cantilever and a substrate, and subjected to a constant rate of extension, can be measured. Numerous studies have concentrated on the response of multi-domain concatamers of protein domains like I27, a domain of the giant muscle protein, titin. Each peak in the force/extension curve for the concatamer corresponds to the unfolding of a single I27 domain and is modelled using a worm-like chain model to reveal the unfolded length of the monomer. Pyranose ring sugars, such as dextran (α -1-6-polyglucose), have also been investigated. The monomer rings undergo a force-induced structural transition (chair to boat), resulting in molecular extension that can be observed in the force/extension curve (Fig. 1).

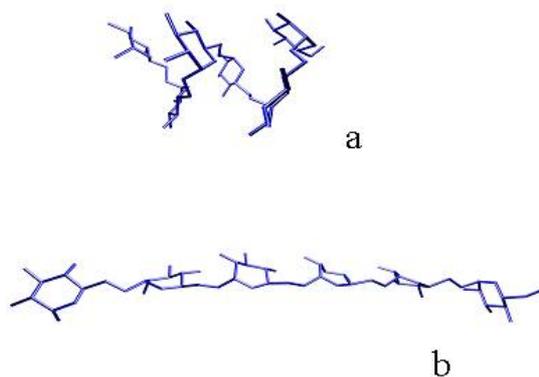


Fig. 1: Structure of dextran before (a), and during (b), application of force along the length of the molecule. Initially, pyranose rings are in the chair conformation and steric restrictions favour a loosely coiled chain. Application of force straightens the helix and forces pyranose rings to adopt the more extended boat conformation. Images taken from simulated pulling of a hexamer using Gaussian 98.

Thermally driven oscillation of single dextran molecules

Using the new apparatus, a single dextran molecule can be extended with gradual increments of force. At each step, thermally driven Brownian motion of the cantilever-molecule system is measured. The time-series data is Fourier transformed to give a power spectrum, and a simple harmonic oscillator model (or more complex model) can be used to fit the data:

$$|u|^2(\omega) = \frac{2k_B T \zeta}{(k - m\omega^2)^2 + \zeta^2 \omega^2}$$

Since the cantilever and molecule are in series, their spring constants, k , and damping constants ζ add. The molecular constants can be found by subtracting the constants of the free cantilever at the same height above the substrate.

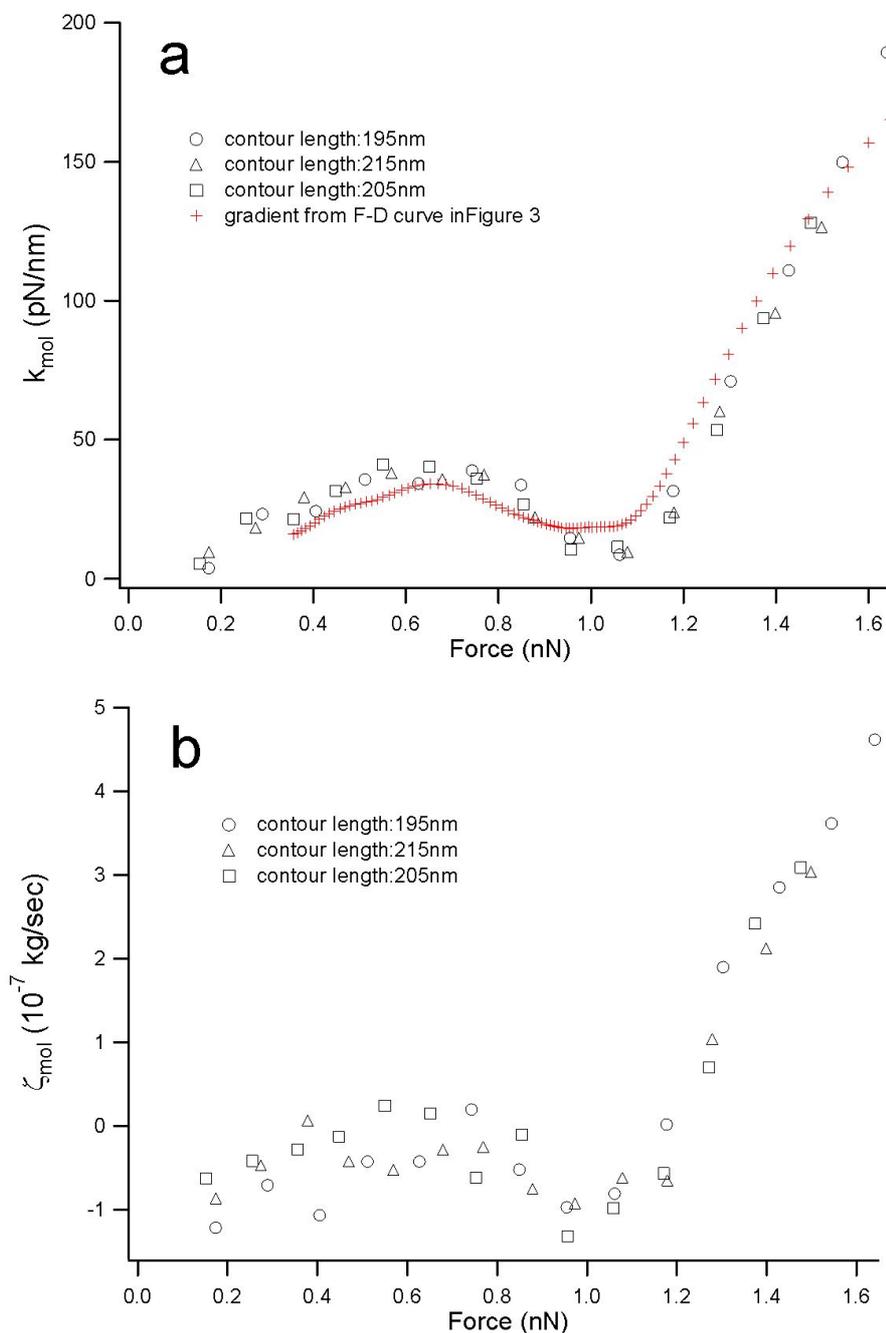


Fig. 2: “Molecular” damping and spring constants as a function of tip-substrate separation for three different molecules. Molecular constants are taken to be the difference between constants found for a free cantilever and the same cantilever with an attached single molecule. Also shown is the molecular spring constant estimated from the slope of the force extension curve for a single molecule. This is in good agreement with the values estimated from thermal noise power spectra.

The molecular spring constant rises with increasing chain extension and falls during the chair-boat transition as extra molecular length is created. Damping constants show that in the

unextended chair and boat states the molecule reduces system dissipation, but increases it in the extended state (Fig. 2).

There are several hundred independent pyranose rings in each molecule. So, despite its single molecule nature, this is an ensemble measurement. The energy landscape for the transition can be modelled as a force perturbed two-state pathway, with equilibrium populations of chair and boat monomers. This model leads to a simple relationship between equilibrium rate constant, λ , and macromolecular properties: $\lambda=k/\zeta$. We hope to use the model to estimate rate constants for the transition.

Note that the simple harmonic oscillator (SHO) model results in a negative value of $\Delta\zeta$ implying that the molecule-cantilever system is less dissipative than the cantilever alone. This counter intuitive result demonstrates that the SHO model is not sufficient to describe the system. We have developed a more complex model which adequately describes the response of the system to thermal fluctuations.

Future work

This will involve application of the thermal oscillation method to I27 concatamers and development of the apparatus to permit forced oscillation of single molecules at a specified frequency using a magnetic drive.

Acknowledgements

We would like to thank Stuart Warriner for useful discussions on dextran structure and Gaussian 98 dextran images. We are also grateful to Jen Struckmeier at Veeco Metrology for continued assistance with technical enquiries regarding the PicoForce instrument.

Funding

This work was funded by the EPSRC. MK was a JSPS Visiting Research Fellow and is now supported by the EPSRC.

Fast folding of B domain protein A by laser induced temperature jump

George Dimitriadis, Kay Sadler, Adam Drysdale, Sheena Radford and Alastair Smith

Introduction

Theoretical approaches to protein folding have created, in the last twenty years, a need for experimental data of the kinetics of increasingly smaller and faster proteins. Experiments like stop flow, or time resolved, CD with dead times of the order of tens of milliseconds are unable to provide data for proteins that fold within hundreds of microseconds, which are now the paradigm for theoretical approaches like MD or Monte Carlo simulations. In our lab, we have built an experiment to measure microsecond folding dynamics using a fluorescence probe following an infra-red pulse which causes a rapid temperature jump. The extremely fast temperature rise (8ns) and the ability to record kinetic data at a time resolution of 130ns allows the measurement of the kinetics of even the fastest folding proteins known today. At the same time, a fully automated experimental setup allows for the collection of kinetic data both against rising temperature and against increasing denaturant concentration, within a reasonable amount of time and with excellent signal to noise. This allows for the creation of kinetic rate (or free energy) surfaces (plotted against temperature and denaturant) that carry much more information than the usual Eyring or chevron plots, and which can be extrapolated with a substantial increase in certainty to regions which are not experimentally accessible.

Setup and Results

Experimental Setup. The temperature jump apparatus utilizes two lasers to increase the temperature of the sample, and to follow the folding of the protein by fluorescence spectroscopy. The temperature jump is achieved by an Nd:YAG Surelite laser giving an 8ns pulse of 800mJ at 1064 nm that is Raman shifted by the use of methane at 30 atm. The 80mJ pulse at ~1550 nm is propagated twice through the sample, producing temperature jumps in the order of 20°C. A NdYAG pumped titanium sapphire laser is used in combination with a pulse picker and frequency tripler to produce a train of UV pulses 130ns apart at 290 nm to excite protein fluorescence. Tryptophan fluorescence (emitting at 350nm) is used to follow protein dynamics following the T jump. The fluorescence is collected by a photomultiplier tube, customized for high linearity and efficiency for short pulses. The signal is then collected by an 8Gsamples/s oscilloscope and finally transferred to a computer for analysis. The setup is constantly monitored and controlled by computer, using a custom program written in Igor 4.0. This allows the automatic collection and analysis of a number of T-jumps (usually between 50 and 100) at a predefined series of final temperatures.

Results of fast protein folding – BdpA. The F13W/G29A double mutant of the beta domain of protein A was known to fold at a much faster rate than the wild-type protein, previously measured at $120,000\text{s}^{-1}$. Using denaturant concentrations between 0.8M and 3M of GuHCl, and final temperatures between 20°C and 60°C, observed kinetic rates between $150,000\text{s}^{-1}$ and 400s^{-1} were measured. The surface that these data represent in the phase space of denaturant concentration and temperature was fitted using a two state kinetic model. The accurate fit allowed the extrapolation (to 0M GuHCl and to temperatures between 0°C and 100°C) and the calculation of the Gibbs free energy of the transition state for all values of denaturant concentration and temperature (Fig. 1). These results have shown that the F13W/G29A mutant of BdpA has a folding rate constant of $248,000\text{s}^{-1}$ at 45°C, i.e. a rate of folding of 4 μs . This makes BdpA the fastest folding protein presently known. At the same time, assuming that BdpA folds following a two state mechanism, this suggests that the pre-

exponential folding factor (considered as the rate of the backbone movement) has a lower limit of $2 \times 10^6 \text{ s}^{-1}$ for this small protein.

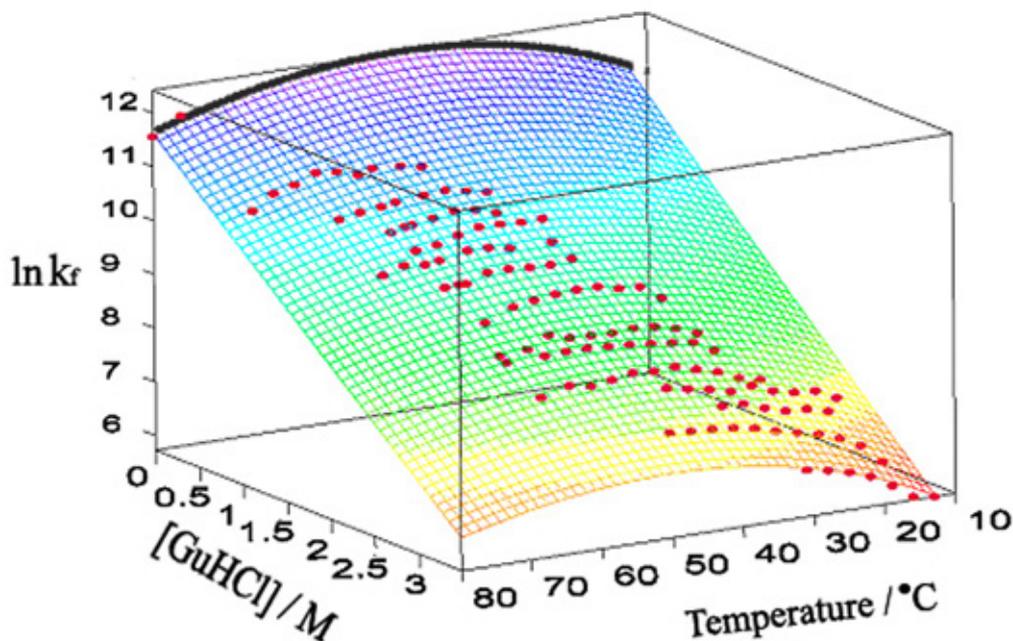


Fig. 1. Folding rates of F13W/G29A BdpA at different denaturant concentrations and temperatures. The two points at 0M GuHCl were not included in the surface fit, but were collected later to demonstrate the accuracy of the extrapolation to 0M GuHCl which our surface fitting procedure allows.

Future Work

The ability to measure fast kinetics at the limit of what is possible, together with the accuracy of any extrapolations both in denaturant concentration and temperature, make our T-jump system a versatile tool for measuring protein folding kinetics. At the moment, two more protein systems are under study. The first is a series of mutants of BdpA designed to test prediction of current folding theory. The second is a series of mutants based on the L53A/I54A double mutant of Im7. This double mutant is trapped in a “native” state that is believed to be a transiently populated intermediate of the wild-type protein. The kinetics of the point mutations on this double mutant will help explain in detail the folding pathway of the wild-type Im7.

Publications

Dimitriadis, G., Drysdale, A., Myers, J. K., Arora, P., Radford, S. E., Oas, T. G. & Smith, D. A. (2004) Microsecond folding dynamics of the F13W G29A mutant of the B domain of Staphylococcal protein A by laser induced temperature jump. *Proc. Natl. Acad. Sciences. USA*. In press.

Acknowledgements

The B domain folding study is a collaboration with Prof Terry Oas and Pooja Arora at Duke University, and Prof Jeff Myers at Vanderbilt University, USA.

Funding

This work was funded by the BBSRC, EPSRC and the University of Leeds.

Single molecule fluorescence spectroscopy of RNA hairpin loop folding

Chris Gell, Rob Leach, Tara Sabir, Sara Pugh, Sheena Radford,
Peter Stockley and Alastair Smith

Introduction

Single molecule methods are now reaching some level of maturity in a range of biological applications. In particular, the use of single molecule fluorescence resonance energy transfer (FRET) and fluorescence correlation spectroscopy (FCS) for structural, dynamic and ligand binding studies is becoming commonplace. In our laboratory in the School of Physics and Astronomy, we have continued to develop an instrumentation suite capable of a range of these studies. Most recently we have constructed a total internal reflection fluorescence (TIRF) microscope for the study of single molecules attached to surfaces, incorporated into surface supported membranes or in living cells (see Fig 1). As an example of our ongoing work we highlight, below, recent results demonstrating the use of single molecule FRET to study RNA hairpin folding kinetics.

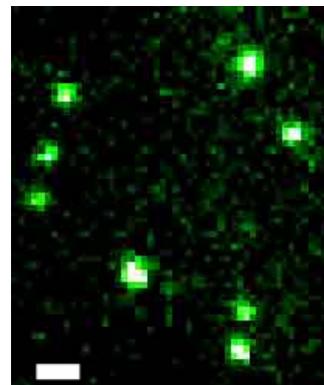


Fig 1. Image of single fluorescein molecules on a glass surface taken with our new TIRF microscope. Bar is approximately 1 μm .

FRET of an RNA hairpin

We have used single molecule methods to probe the thermally induced unfolding of an RNA hairpin. We intend to use this system as a model to assess the potential of these methods in the investigation of more complex systems. The RNA has been labelled with the dye pair fluorescein (FL) and tetramethylrhodamine (TMR): TMR-5'-ACA-UGA-GGA-UUA-CCC-AUG-U-3'-FL, which results in strong FRET when the hairpin is folded. Using diffusion FRET we have monitored multiple single molecule events at various temperatures. Fig. 2 shows the measured histograms of the FRET efficiencies of the single molecules at three temperatures along the denaturation profile of the RNA (ensemble FRET experiment, black circles, blue line). The histograms reveal three peaks corresponding to (from left to right) donor only labelled, unfolded and folded molecules. The relative areas under the folded and unfolded peaks clearly reflect the populations of folded and unfolded hairpins in agreement with the bulk measurements (arrows show “positions” of the measurements).

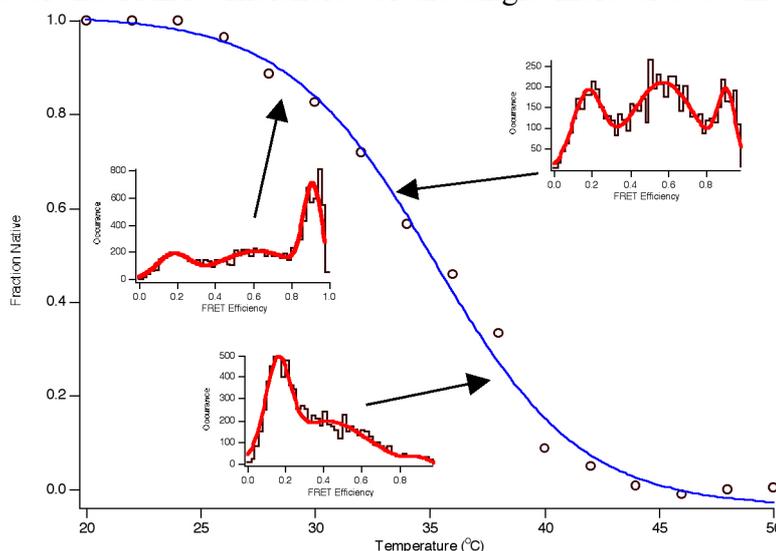


Fig 2. Thermal unfolding of RNA hairpin (circles, blue line) showing the fraction of formed hairpins. The inset shows histograms of single molecule FRET measurements that show three peaks (red lines are Gaussian fits) corresponding to single labelled molecules (artefact), unfolded and folded molecules. Arrows indicate “positions” of single molecule measurements.

Funding

This work was funded by the Wellcome Trust, University of Leeds and SRIF II initiative.

Investigating the molecular interactions of the $\phi 29$ packaging motor

Mark A. Robinson, Dawn Baird, Holly Dunsby and Nicola J. Stonehouse

Introduction

The bacteriophage $\phi 29$ infects *Bacillus subtilis* by the transfer of viral genomic dsDNA. Before infection, this dsDNA is packaged into preformed procapsids via the action of a protein-RNA molecular motor (Fig. 1). A novel RNA-RNA multimerisation event is thought to be involved in motor assembly, resulting, in combination with ATPase activity, in rotation of the molecular motor and concomitant translation of genomic DNA into the procapsid. Packaging of the 19.3 kbp dsDNA has been determined at a rate of 100 bp per second, generating a final internal pressure of 6 Mpa. $\phi 29$ may act as a model system for the investigation of systems with potentially analogous modes of DNA packaging.

Results

The multimerisation of the RNA implicated in the molecular motor function, packaging RNA (pRNA), has been previously investigated in this laboratory by analytical ultracentrifugation (AUC) and light scattering experiments. More recently, dynamic light scattering (DLS) experiments indicate the Mg^{2+} -dependent formation of monomeric, dimeric and trimeric species, which may act as components of the ultimate higher order species.

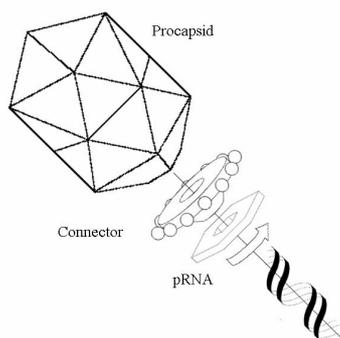


Fig. 1: Proposed rotational movement of the molecular motor results in linear translation of the genomic DNA

Current work is now targeting the chemical and enzymatic synthesis of pRNA species – hairpin loop domains and full length molecules, both wild-type and mutant. This library of RNA species allows a varied approach towards understanding the interactions involved in the $\phi 29$ molecular motor. In particular, chemical syntheses that allow the orthogonal incorporation of molecular probes for use in surface plasmon resonance and fluorescence experiments have been completed. Crystallisation screening trials to allow investigation of intermolecular interactions in the pRNA multimerisation event are underway.

Collaborators

Neil Thomson

Peixuan Guo, Purdue University

Funding

Funding from the BBSRC and NESBiC is gratefully acknowledged.

Using *in vitro* selection to investigate RNA-protein interactions in picornaviruses

Mark Ellingham, David Rowlands and Nicola J. Stonehouse

Introduction

The *Picornaviridae* is a family of single-stranded, negative sense RNA viruses that includes many important pathogens, such as poliovirus, rhinovirus and foot-and-mouth disease virus (FMDV). Although the life-cycle of these viruses has been described in great detail, a comprehensive model must include detail of the interactions that occur between viral proteins and RNA, for example during genome replication, the exact nature of which has yet to be elucidated. Using *in vitro* selection, or SELEX, to generate high affinity RNA ligands against picornaviral proteins, it may be possible to gain more insight into these events and further the understanding of this important group of viruses.

Results

SELEX is a powerful technique that works on the principle that nucleic acid sequences isolated on the basis of the exhibition of a desired property, for example binding affinity, can be preferentially amplified over competing sequences. Repetition of this process results in a population with a high proportion of sequences that behave in the desired manner. A schematic diagram of this process is illustrated in Fig. 1.

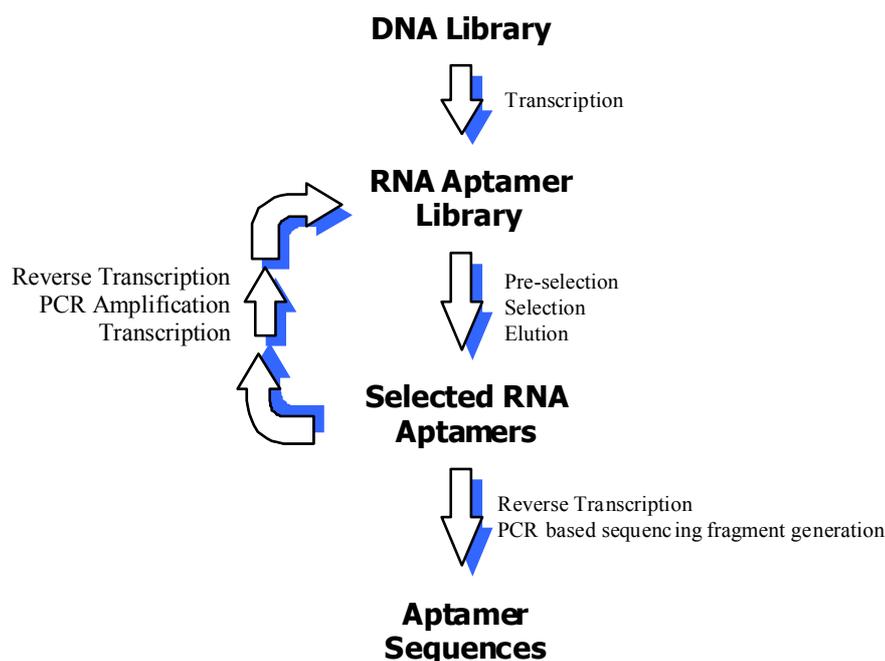


Fig. 1: Schematic of the process of SELEX

15 rounds of selection have been performed against the 3D protein, an RNA polymerase, of FMDV-C using a random sequence library of RNA as a starting pool and with the intention of generating high affinity RNA ligands against the protein. Binding studies have been carried out, which indicated that the final RNA product of the process had a higher affinity for the 3D protein than the starting pool of random RNA sequences. In addition, a number of RNAs from the final product were sequenced and subjected to sequence homology searches against the genome of the virus. Several regions of significant homology were obtained

between the virus genome and sequences from the selected aptamer pool, indicating possible binding sites of the 3D protein. Current work aims to elucidate further the nature of the binding between the aptamer pool and the target protein. In addition, other proteins involved in the replication of FMDV are being investigated, as it is believed that they may form a complex with 3D during genome synthesis.

Collaborators

Esteban Domingo, Madrid

Funding

With thanks to the BBSRC for funding.

High throughput techniques as a strategy to overcome crystal twinning

Stephen Carr, Simon Phillips and Chris Thomas

Background

Staphylococcal plasmids of the pT181 family use a “rolling circle” mode of replication. This requires a plasmid-encoded initiator protein, Rep. These proteins share over 80% sequence identity, yet, *in vivo*, the Rep proteins are specific for their cognate plasmid: thus RepD protein initiates replication for pC221, RepC for pT181 and RepN for pCW7.

Previously crystallisation trials have been performed with RepD protein and chimeric constructs in which the C-terminal domain of RepD has been exchanged with that of either RepC or RepN. The latter constructs produced diffraction quality crystals, but structure solution has been hampered by the presence of near-perfect merohedral twinning in crystals of both proteins.

Recent Findings

In an attempt to produce non-twinned crystals suitable for diffraction studies, the choice of potential target proteins has been expanded to encompass the entire protein family. These can be considered point mutations on a larger scale as multiple residues will be altered between constructs.

Rep proteins C, D, E, I, J and N have all been cloned into plasmid pET11a. Truncated variants of the above proteins, which lack a highly variable region of 30 amino acids at the extreme N-terminus, have also been produced. The truncated constructs remove what is likely to be a highly flexible region of the proteins which could hamper crystallisation. These have been shown to be active in the case of RepD.

The novel Rep targets have all been over-expressed and can be readily purified in milligram quantities. These targets are currently undergoing automated crystallisation trials using a robot produced by Douglas Instruments.

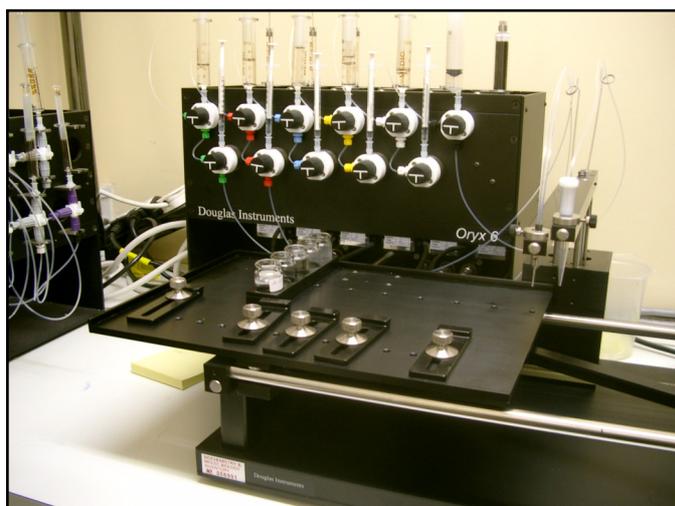


Fig 1. Crystallisation robot (Douglas Instruments)

Acknowledgements

We thank Val sergeant for technical support. This work is funded by the Wellcome Trust.

Molecular mechanism of Staphylococcal plasmid transfer

Jamie Caryl and Chris Thomas

Background

Horizontal gene transfer in bacteria results in genetic diversity with important medical consequences. Small non-self transmissible, mobilisable, Staphylococcal plasmids such as pC221 offer a simple system that embodies the initial events in plasmid mobilisation. pC221 is a 4.6kb chloramphenicol resistance plasmid of *Staphylococcus aureus*. It is a non-self transmissible plasmid that can be mobilised by a co-resident self-transmissible plasmid such as pGO1. Being a small plasmid, pC221 contains only those genes required for its own DNA processing and contains four such loci: an origin of transfer (*oriT*); a DNA relaxase, MobA; and the putative accessory proteins MobB and MobC.

The nicking reaction has been reconstituted *in vitro* and has demonstrated the requirement for MobA and MobC proteins, in the presence of Mg^{2+} or Mn^{2+} , for site- and substrate-specific nicking.

Recent findings

DNaseI footprinting by primer extension has been performed with MobC and MobA on supercoiled pC221cop903 *oriT* DNA. MobC binds to two regions within the *oriT* (Fig. 1), and additionally to a region within the *mobC* gene. A consensus binding sequence has been identified at all three protected regions. Between the *oriT* binding sites are regions of periodic MobC induced structural perturbation.

The addition of MobA to the complex induces modified protection footprints to the left of the nick site and increased hypersensitivity immediately downstream from the nick site.

The requirement for each of the binding sites for nicking and mobilisation has been investigated by selective cloning. Only the binding site adjacent to the nick site was required for both processes. The role of these binding sites in regulation of Mob protein expression is currently being investigated.

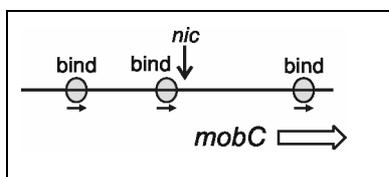


Fig. 2. Relative positions of MobC binding sites, MobA *nic* site, and *mobC* reading frame.

Publications

Smith, M.C.A. & Thomas, C.D. (2004) An accessory protein is required for relaxosome formation by small staphylococcal plasmids. *J. Bacteriol.* (in press)

Caryl, J.A., Smith, M.C.A. & Thomas, C.D. (2004) Reconstitution of a staphylococcal plasmid-protein relaxation complex *in vitro*. *J. Bacteriol.* (in press)

This work has also been presented at the UK Mobile Genetic Elements Workshop 2003, Birmingham, UK.

Acknowledgements

We thank Val Sergeant for technical support. This work is funded by the BBSRC.

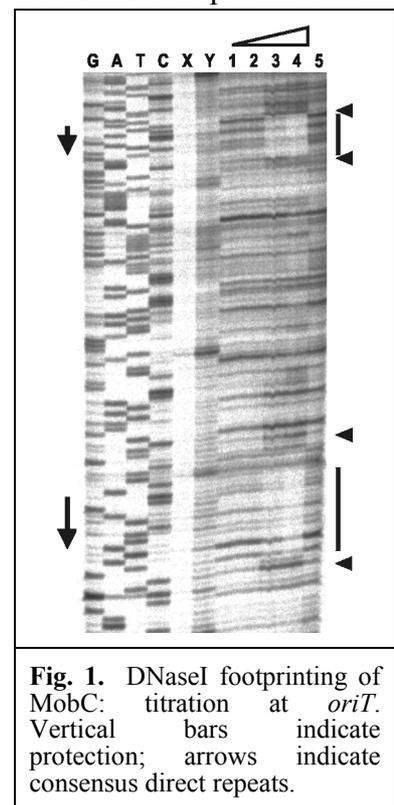


Fig. 1. DNaseI footprinting of MobC: titration at *oriT*. Vertical bars indicate protection; arrows indicate consensus direct repeats.

Investigating the interaction of quinolones with topoisomerase IV of *Staphylococcus aureus*

George Makris, Jon Cove and Chris Thomas

Background

DNA topoisomerases are ubiquitous enzymes responsible for resolving topological problems arising due to DNA transcription, recombination, replication and chromosome partitioning. Topoisomerase IV (topo IV) of *Staphylococcus aureus* is a heterotetrameric protein composed of two homodimeric subunits: GrlA, which is responsible for DNA binding, the cleavage of both strands (type II class enzyme) and the religation reaction, and GrlB, which binds and hydrolyses ATP allowing enzyme turnover. The action of topo IV results in reduced superhelical density, via the conversion of negatively supercoiled (SC) DNA into the relaxed, covalently closed form (RC).

We have developed an *in vitro* assay that enables us to measure the relaxation of SC DNA substrate (staphylococcal plasmid pC221*cop903*) over time using in-house expressed and purified topo IV. In the context of this assay, we have verified that quinolone drugs such as ciprofloxacin cause significant reduction in the catalytic efficiency of topo IV of *S. aureus in vitro*, leading to the accumulation of cleaved DNA complexes, such as nicked open circular (OC) and linear (Lin) forms of DNA (Fig. 1).

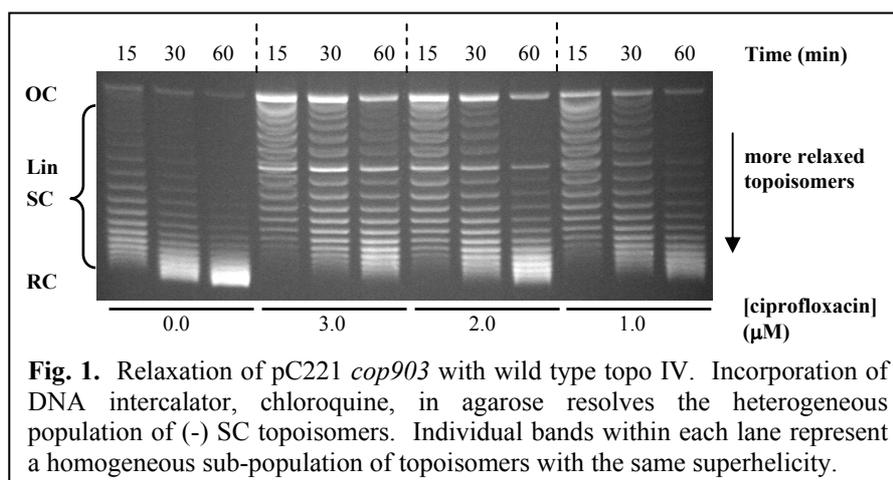


Fig. 1. Relaxation of pC221 *cop903* with wild type topo IV. Incorporation of DNA intercalator, chloroquine, in agarose resolves the heterogeneous population of (-) SC topoisomers. Individual bands within each lane represent a homogeneous sub-population of topoisomers with the same superhelicity.

Recent Findings

In order to identify the mechanism by which quinolone drugs exert their effect, we set out to identify domains with discrete functions within GrlA. Partial proteolysis of GrlA demonstrated the existence of a stable fragment, which we have cloned, expressed and purified.

Sedimentation velocity analysis demonstrated that the GrlA fragment exists in a monomer-dimer equilibrium. We have also shown that the fragment retains part of the topo IV activity by being able to convert supercoiled and covalently-closed, relaxed DNA substrates to nicked open circular and linear forms independent of GrlB. Crystallisation trials with the fragment have been performed yielding a number of potential targets for X-ray diffraction analysis.

Acknowledgements

We thank Val Sergeant for technical support. This work is funded by Aventis.

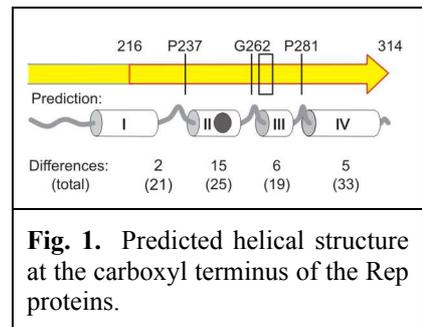
Dimer specificity in a replication initiator protein

Chris Thomas

Background

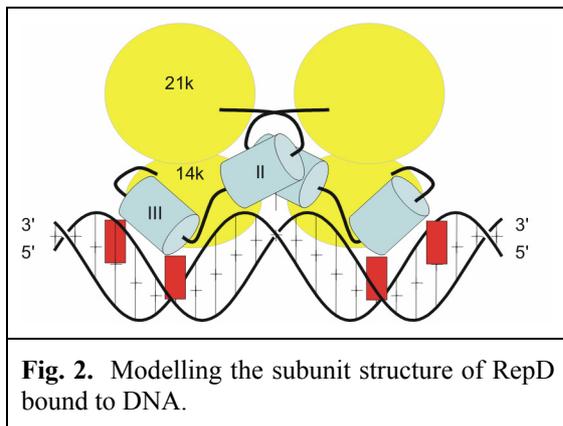
Replication of staphylococcal plasmids such as pT181 and pC221 requires the plasmid-encoded Rep proteins RepC and RepD. These proteins are over 80% identical in primary sequence, yet retain DNA binding specificity for their cognate plasmid replication origins due to a small divergent region of six amino acids within the carboxyl terminus.

The 34kDa fusion protein RepDC contains residues 35-216 of RepD and 217-314 of RepC, corresponding to the 21kDa and 14kDa fragments obtained by partial proteolysis of these proteins. RepDC displays DNA sequence specificity for pT181 but cannot form heterodimers with RepD, despite sharing over 80% amino acid sequence identity. Dissection of the C-terminal domain previously correlated this dimerisation specificity with residues 237-262, which is adjacent to (but not overlapping) the DNA binding determinant at 265-270. This dimerisation determinant includes the second of four predicted alpha-helical regions (I - IV) as shown in Fig. 1.



Recent findings

Dimerisation studies have been conducted using a wide range of 34kDa variants, with sequences exchanged in each of the four indicated regions. Each has now been tested for dimerisation against both RepD and RepC. In each case, dimerisation specificity is conferred solely by the sequence at region II. No additional specificity is contributed by the rest of the protein: for example, a RepD variant altered only by possessing the 15 RepC-specific amino acids in region II will form heterodimers with RepC protein.



Dimerisation trials have also been conducted using fusion proteins based on other Rep proteins of the pT181 family. For example, RepDN (which substitutes residues 217-314 with those of the pCW7 RepN protein) demonstrates dimerisation against RepC but not RepD. From such results the contribution of individual residues within region II can be deduced, and the likely arrangement of helices and loops at the dimer interface modelled as in Fig. 2.

Based on this model, point mutations are currently being designed to create novel dimerisation interfaces. These will allow the construction of obligate heterodimers, with asymmetric DNA binding and nicking characteristics.

Acknowledgements

Thanks to Val Sergeant for technical support. This work is funded by the Wellcome Trust.

Atomic force microscopy of DNA-protein interactions

Neal Crampton and Neil Thomson

Introduction

Atomic force microscopy (AFM) can be used to distinguish proteins bound to nucleic acid templates. Imaging can be conducted in air or under fluid, the only requirement being that the protein-DNA complexes must be bound to a flat substrate. By scanning across the support with a sharp AFM tip we are able to produce a topographic image of the support and the bound DNA-protein complexes. By measuring quantities such as the position of the protein on the template, its height, or its width, we are able to infer changes in the molecular conformation of the protein under different conditions. Also, the amount of template wrapping by the protein can be discerned, by measuring the protein induced bend angle, and the apparent shortening of the template. A central aim of our work is to study such systems under fluid in real time. In this setup, we can potentially study the transition of single molecules between different conformational states. Diffusion controlled processes such as a promoter search occur on time scales of seconds to minutes and are readily studied with current AFM technology. However enzymatically controlled processes such as transcription occur at much faster time scales, and in order to observe enzymatic processes, imaging times must be reduced. Fast scan technologies are being developed by a number of research groups worldwide. As these technologies become more widely available, AFM will play a larger role in understanding of mechanisms of interaction between proteins and DNA

Nested-gene transcription model

“Nested gene” is a term coined for a gene that lies completely within the sequence of another gene. Such genes are known to exist in the human genome and eukaryotes such as *Saccharomyces cerevisiae*. Whether these genes come about as an unavoidable consequence of their compressed genetic arrangement, or whether such genes have an intrinsic effect on transcriptional regulation is unclear. We aim to study the effect two convergently aligned genes have on the expression of these gene’s mRNAs, either through collision of polymerases, RNA interference of sense and antisense transcripts, or topological constraints introduced by the template.

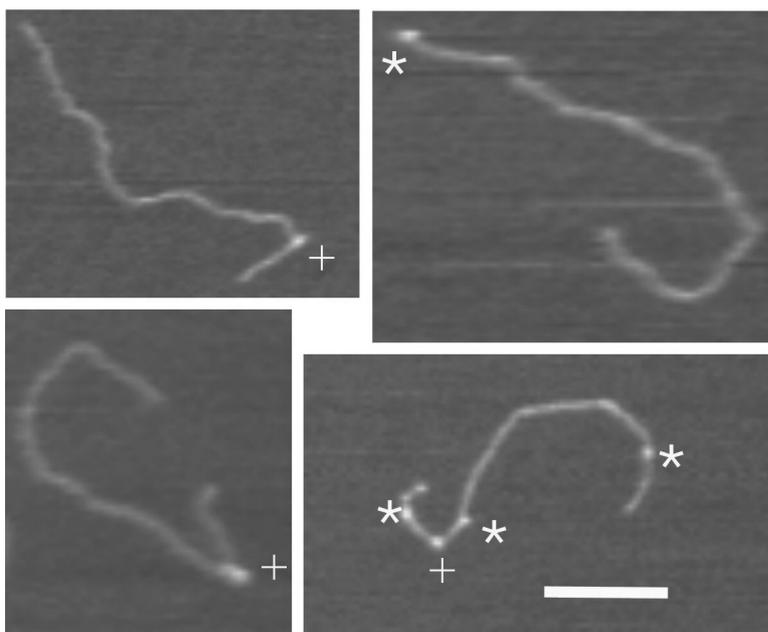


Fig. 1: Gallery of representative images of a DNA template with T7 polymerases bound. Polymerases bound specifically, as judged by their position, are marked by +, and those non-specifically bound are marked *. The scale bar is 200nm.

Present work has focused on bacteriophage T3 and T7 polymerases due to their high specificities for their promoters. Using this system, we have been able to bind the polymerases to the template DNA, and, by measuring their position along the template, we are able to determine if they are bound at the promoter. We have studied the effect of omission of each nucleoside triphosphate (NTP) and their respective affinity to form stalled complexes downstream of the promoter. Future work will investigate different methods of polymerase stalling, in an attempt to decouple promoter binding and polymerase stalling. Different templates are currently being constructed with the aim to move to a true nested gene, when single molecule real time imaging will be undertaken.

DNA gyrase

DNA gyrase is a bacterial motor protein in a class known as topoisomerases, which are responsible for controlling the topological properties of DNA (i.e. amount of supercoiling or catenation). Most topoisomerases can relax supercoiled DNA, which is an energetically favourable process. DNA gyrase is unique amongst this class, because it can introduce supercoils as well as remove them. To wind or unwind DNA, it must break both strands of DNA, capture another segment of the same DNA molecule and pass this through the double-strand break before resealing. We have used AFM to study the length of linear DNA wrapped around the gyrase in the presence, and absence of 5'-adenylyl β,γ -imidodiphosphate (ADPNP). There is a shortening of the DNA when the gyrase binds consistent with a full wrap of DNA around the protein. On binding of ADPNP this wrap is completely lost. These AFM data have been compared with previous foot-printing data from the Maxwell group in the presence and absence of ADPNP. Both data sets are consistent and have allowed a new model for DNA wrapping in gyrase to be proposed.

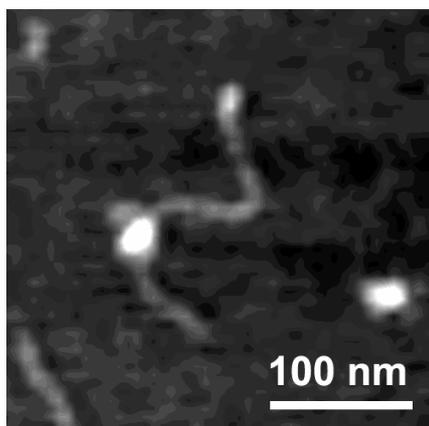


Fig. 2: DNA gyrase bound to a 1070bp linear DNA template. This DNA-protein complex is deposited on a mica surface and imaged using tapping-mode AFM in air.

Collaborators

Nested-gene Project:

Jennifer Kirkham and William Bonass (Oral Biology, Dental School, Leeds)

DNA gyrase Project:

Sylvain Mittelheiser and Anthony Maxwell (John Innes Centre); Fiona Hurrell (Dental School, Leeds)

Publication

Heddle, J.G., Mittelheiser, S., Maxwell, A. & Thomson, N.H. (2004) Nucleotide binding to DNA gyrase causes loss of DNA wrap. *J. Mol. Biol.* In press.

Acknowledgements

We gratefully acknowledge the EPSRC for financial support. Neil Thomson is an EPSRC Advanced Research Fellow.

Class VI myosin

Matthew Walker and John Trinick

Myosins are identified on the basis of sequence similarities in their motor domain regions, which binds actin and have ATPase activity. Extending from the motor domain is a single α -helix, which binds light chains of the calmodulin class. This acts as a lever arm and produces motion by changing its angle relative to the motor domain. At the end of the lever arm the molecule tail binds cargo. The vast majority of work on myosin has focussed on the type found in muscle, which is called class II. Muscle myosin is a dimeric molecule, with two heads, each with a motor domain and a lever arm. The heads are dimerised through the tail, which is a coiled-coil α -helix.

Recent years have seen the discovery of 17 other, non-muscle myosin classes, most of which have been identified only on the basis of sequence similarities and have not been characterised experimentally. Of the myosin classes that have been characterised, all move in the same direction along actin filament tracks except one, class VI myosin. Class VI myosin has been implicated in the development and maintenance of the stereocilia in the inner ear in mammals and thus in hearing and balance mechanisms. Other roles involve endocytosis at the plasma membrane, maintenance of Golgi complex morphology and secretion. Class VI myosin was thought to be a dimeric molecule, because “heptad” patterns of hydrophobic residues of the form HxxHxxx were found in its tail region. These were predicted to form a coiled-coil tail, as in myosin II. Published optical trap studies of a myosin VI construct, in which dimerisation was enforced by a leucine zipper, showed processive hand-over-hand walking of single molecules.

We have studied myosin VI in collaboration with the groups of Dr John Kendrick-Jones (MRC Laboratory of Molecular Biology (Cambridge) and with Dr Claudia Veigel (National Institute for Medical Research, Mill Hill). Surprisingly, our electron microscopy of purified myosin revealed a monomeric molecule. Cross-linking studies of myosin VI *in situ* done in Cambridge confirmed this conclusion. Optical trap studies in Mill Hill show myosin VI has a disproportionately large step size for its predicted lever arm length, which contains only one calmodulin light chain. Recent evidence from Germany suggests myosin VI’s reverse directionality may result from a $\sim 180^\circ$ bend in the light chain. Our single molecule EM studies appear compatible with this conclusion.



Fig. 1. Images averages of myosin VI molecules in the presence (upper panels) and absence of ATP.

Publication

Lister, I., Schmitz, S., Walker, M., Trinick, J., Veigel, C. & Kendrick-Jones, J. (2004)
Myosin VI is a non-processive monomer with a large working stroke. *EMBO J* **in press**.

Collaborators

John Kendrick-Jones and Ida Lister (MRC-LMB, Cambridge), Claudia Veigel and Stephan Schmitz (MRC-NIMR, Mill Hill).

Funding

NIH (USA) and BBSRC

Titin: studies of synthetic end-filaments

Ahmed Houmeida, Beatrix Thompson, Peter Knight, Kavitha Thirumurugan, Andy Baron,
Larissa Tskhovrebova and John Trinick

Titin is the largest protein yet described (chain weight 3-3.7 MDa) and the third most abundant protein in muscle. More than half the titin molecule is bound to muscle thick filaments in the sarcomere, where we have suggested it regulates exact assembly of the 294 myosin molecules known to comprise the filament. The remainder of the titin molecule forms an elastic connection between the end of the thick filament and the Z-line. These connections are the main origin of the passive elasticity of muscle. They also ensure that thick filaments stay in the middle of the sarcomere, which ensures even forces are developed by myosin in each half of the filament.

We previously discovered distinctive structures ~85 nm long at the tips of thick filaments, which we called end-filaments. We proposed that end-filaments are the in-register aggregate of a section of the 6 titin molecules that emerge from each end of the thick filament. We suggested that end-filaments are composed of 22 conserved immunoglobulin domains, I20 to I41, in the titin structure. We are studying a large proteolytic fragment of titin, which we can now purify. The fragment has N-terminal sequence beginning at the start of I20. Its molecular weight has been measured as 286 kDa in the analytical ultracentrifuge. In the electron microscope, the peptide forms aggregates appear very similar to native end-filaments. These data are consistent with the view that titin molecules do not act independently in the region thought to be elastic.

Publications

Tskhovrebova, L. & Trinick, J. (2003) Titin: properties and family relations. *Nature Reviews Cell Molecular Biology* **4**, 679-689.

Trinick, J. & Tskhovrebova, L. (2003) Giant protein titin: structural and functional aspects. *Elastomeric Proteins (eds P R Shewry, A S Tatham and A J Bailey)* Cambridge University Press, pp242-258.

Mutungi, G., Trinick, J. & Ranatunga, K. W. (2003) Resting tension characteristics in differentiating intact rat fast and slow twitch muscle fibers. *J. Appl. Physiol.* **95**, 2241-2247.

Collaborator

Walter Stafford (Boston Biomedical Research Inst. Boston, USA).

Funding

The British Heart Foundation.

EM image processing of filamentous structures – arthrin

Stan Burgess, Matt Walker, Peter Knight and John Trinick

Many proteins exist *in vivo* in the form of filamentous aggregates. In general, these are helical. Because all views of a helical filament are seen at different axial points, it is possible to generate a 3D model from the 2D projections in a single electron micrograph. As a consequence, 3D helical reconstruction has been widely applied. In the case of highly ordered filaments, such as bacterial flagellae, this can lead to atomic resolution. Often, however, filaments are not exactly helical but exhibit disorder, which severely limits the resolution attainable. This is the case in actin filaments, where there is $\sim 6^\circ$ azimuthal disorder between successive subunits in the filament. We have developed an alternative method that does not rely on helical symmetry, but windows out filament segments and uses these as the particles in single particle image processing.

As proof of concept, we have used the new method to reconstruct filaments of arthrin (arthropod actin), which is the mono-ubiquitinated actin found in insect flight muscle and malaria parasite. Cryo-EM images of frozen-hydrated arthrin filaments made at York University were windowed into $\sim 15,000$ segments, each containing ~ 10 subunits. The segments were aligned, averaged and used in the reconstruction. Filaments of pure actin were used to make a control reconstruction and the position of the ubiquitin was revealed in a difference map. This showed that the ubiquitin is located on the opposite side of sub-domain 1 in actin from where myosin binds. Our collaborator at Aventis, John Holt, independently determined the position of the ubiquitin linkage by peptide mapping and mass spectrometry. This showed the ubiquitin to be on Lys118 in actin, which is consistent with the EM data. Molecular modelling by our collaborator in Bristol was used to show that the ubiquitin adopts a few conformers, stabilised with the interface with actin. Our collaborators in Heidelberg, Belinda Bullard and Kevin Leonard, used antibodies to show that every seventh actin subunit in insect flight muscle thin filaments is ubiquitinated. This suggests that ubiquitination is involved in regulating the muscle contractile activity.

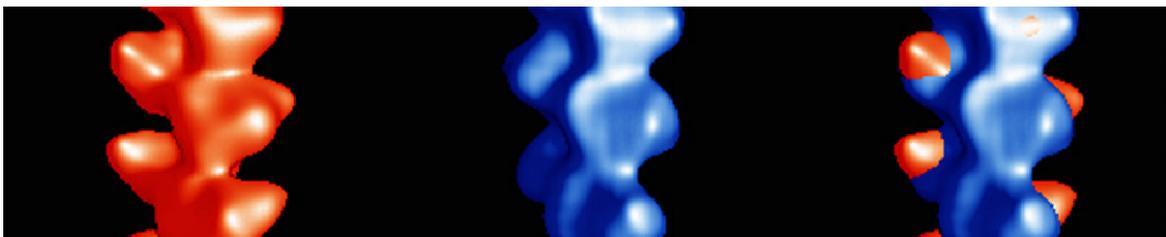


Fig. 1. The figure shows the arthrin map (left), control actin (middle) and the difference map (right). Note the extra mass on sub-domain 1 at high radius due to ubiquitin (9 kDa).

Publication

Burgess, S. A., Walker, M. L., Knight, P. J., Sparrow, J. C., Schmitz, S., Offer, G., Bullard, B., Leonard, K., Holt, J. C. & Trinick, J. (2004) Structural studies of arthrin: mono-ubiquitinated actin. *J. Mol. Biol. in press*,

Collaborators

John Sparrow and Stephan Schmitz (York University); Belinda Bullard and Kevin Leonard (EMBL-Heidelberg); Gerald Offer (Bristol University); John Holt (Aventis, Swiftwater Pennsylvania)

Funding

NIH (USA) and BBSRC

New apparatus for time-resolved electron cryo-microscopy

Matt Walker, Kavitha Thirumurugan, Howard White and John Trinick

Understanding the reaction mechanisms of proteins and large complexes is one of the major goals of modern biology. The way to accomplish this task is through the identification and characterization of key reaction intermediates. Time-resolved electron cryo-microscopy provides the only general-purpose means available to accomplish this and is capable of ~1 nm spatial resolution and millisecond time resolution.

Cryo-EM specimens require ultra-fast cooling ($\sim 10^{6-7} \text{ }^\circ\text{C/sec}$) in order to retain the vitreous state of water and avoid ice crystal damage. It was clear that such rapid cooling might provide excellent time resolution, but it was only through introduction of a spraying method that rapid mixing of reactants, while retaining a water layer thin enough for microscopy, became feasible. This technique involves spraying a solution containing one reactant onto an EM grid covered with a thin layer of another reactant immediately (~ 5 msec) before freezing. Through this approach it became possible to correlate directly structural work with bulk solution fast kinetics.

We previously designed and built apparatus for this purpose and have now built a second generation apparatus. Reactions run for controlled times ranging from a few milliseconds to tens of seconds after initiation are trapped by rapid freezing. Blotting of the electron microscope grid and freezing it in liquid ethane uses computer-controlled microstepping motors. For the fastest time resolution, ~ 5 ms, a blotted grid containing a thin film of one reactant is sprayed with small droplets containing a second reactant just before freezing. In the new apparatus, the spray is, for the first time, produced electrically (electrospray), which gives a dense cloud of droplets < 1 μm in diameter from the 1-2 μl of solution required per grid. A second method in which two solutions are first mixed by turbulent flow and then blotted prior to freezing is used for reactions with time courses > 1 second.

The spraying approach is important, because sprays can be used to deliver anything from small molecules to proteins or even protein assemblies. Its wide applicability makes it the only general purpose method to visualize fast conformational changes in large protein complexes. As such, it is uniquely powerful for the study of mechanism in protein complexes generally.

Publication

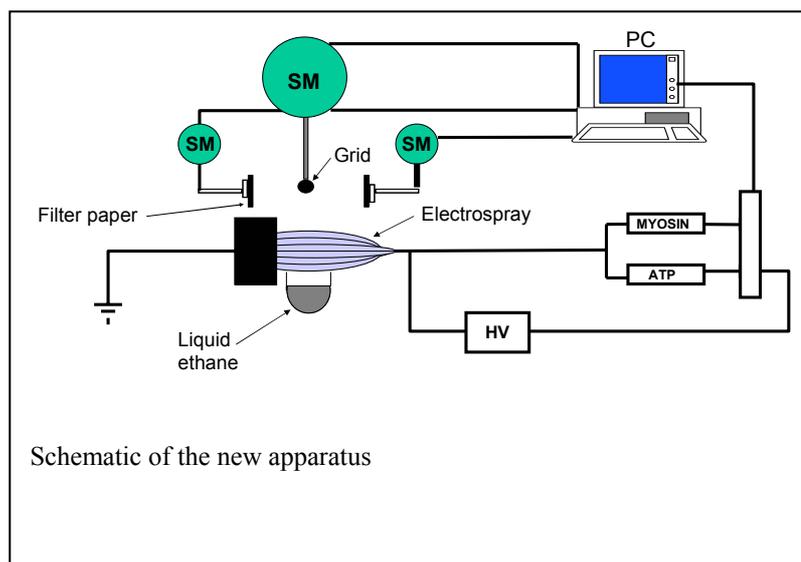
White, H. D., Thirumurugan, K., Walker, M. L. & Trinick, J. (2003) A second generation apparatus for time-resolved electron cryo-microscopy using stepper motors and electrospray. *J. Struct. Biol.* **144**, 246-252.

Collaborator

Howard White, Eastern Virginia Medical School

Funding

NIH (USA)

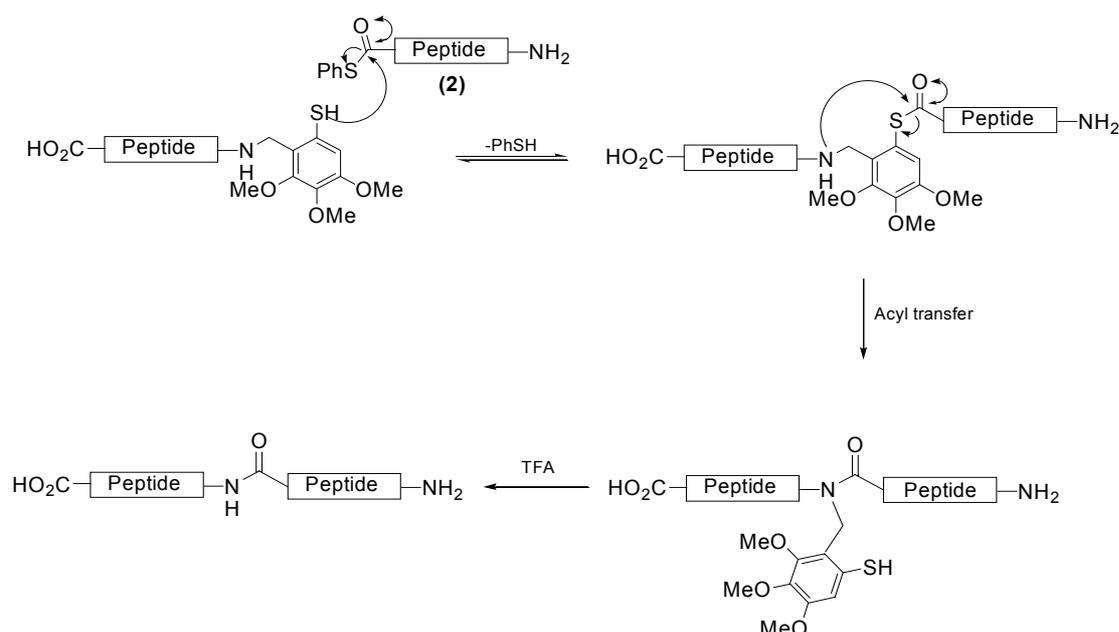


Application of the Tmb acyl transfer auxiliary to the total chemical synthesis of Im7

Anil Nagalingam, Sheena Radford and Stuart Warriner

Introduction

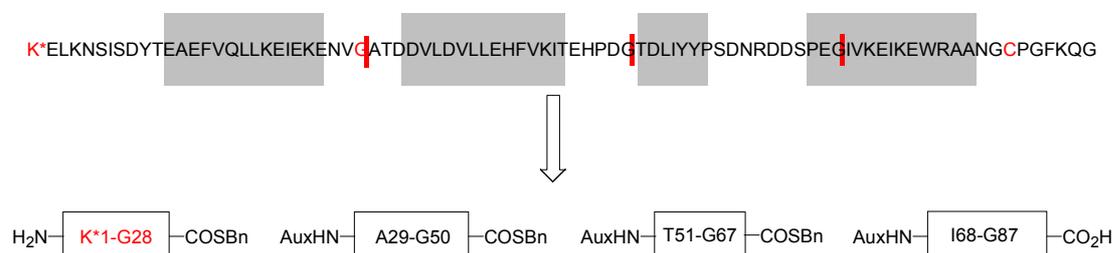
Total chemical synthesis of proteins allows easy access to a wide variety of modified species, which can facilitate folding studies. Limitations of SPPS for large polypeptides have been overcome by the development of chemical ligation methods of unprotected peptides. The most recent of these developments employs a removable acyl-transfer auxiliary to carry out the regioselective coupling of the unprotected peptides. Subsequent cleavage of the directing group yields the native protein (Scheme 1).



Scheme 1

Synthesis of Im7

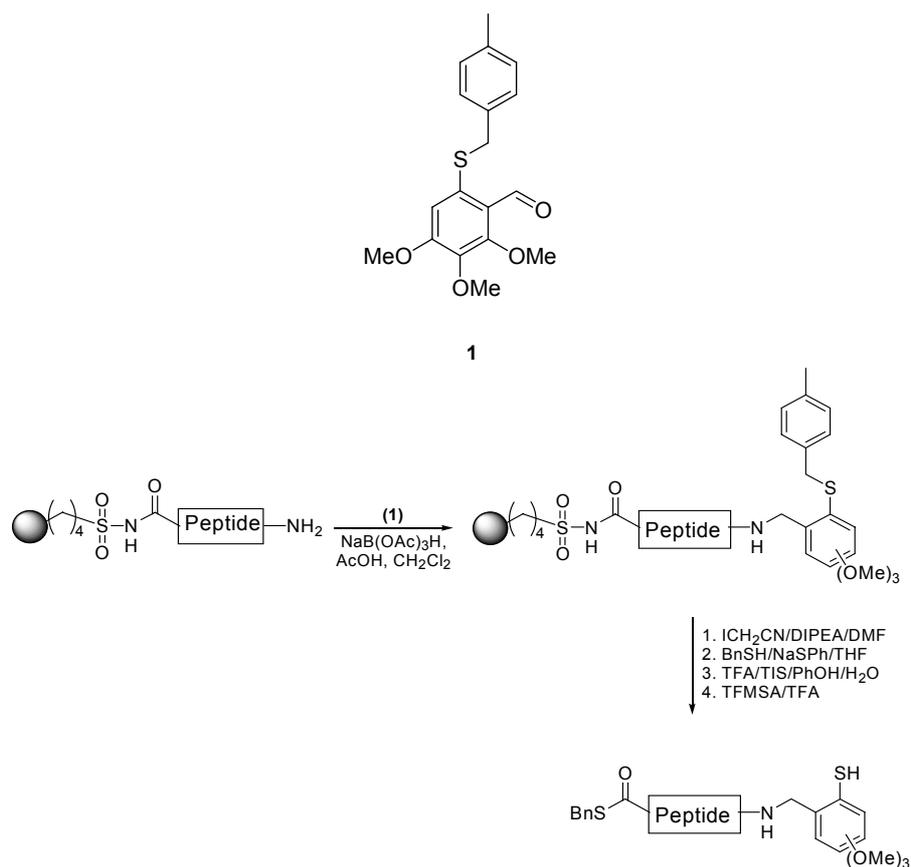
We have divided the Im7 protein into four fragments which will be synthesised using SPPS (Fig. 1).



Red = Mutation, * = TMR Labelled

Fig. 1

The C-terminal thio-esters are synthesised through the use of a 4-sulfamyl butyryl safety-catch resin, while treatment of the resin bound peptide with aldehyde (1) introduces the acyl-transfer auxiliary to the N-terminus via reductive amination (Scheme 2).



Scheme 2

Ligation reactions are carried out in neutral Gn.HCl conditions in the presence of thiophenol to form the more reactive SPh thio-ester (**2**).

Summary

Native chemical ligation will give efficient access to Im7, allowing residual modifications to be easily incorporated. Mutations have been made to allow selective labelling of the protein with various fluorophores. The completed structure will facilitate single molecule FRET studies to probe the folding kinetic of Im7.

Funding

We thank MRC for a project studentship.

Design and implementation of a database to archive and aid the analysis of tissue microarray data

Archana Sharma-Oates, Phil Quirke and David Westhead

Introduction

Tissue microarrays (TMA) provide a high-throughput method of analysing the prognostic benefit of a number of potential (protein) targets on a large cohort of tumour samples (Fig. 1). The technique enables the analysis of molecular alterations in thousands of tissue specimens in parallel at the DNA, RNA, or protein level. The TMA technology was developed to enable genome-scale molecular pathology studies. With conventional methods (whole tissue sections) only three hundred 5 mm sections can be cut from an average sized clinical tissue specimen for use in molecular analyses (such as *in situ* hybridisation (ISH) and immunostaining). Analysis of 300 molecular targets correspond to approximately 0.85% of the estimated ~35 000 genes in the human genome. TMAs have the advantage of assaying all the predicted genes in the human genome in parallel as a single experiment.

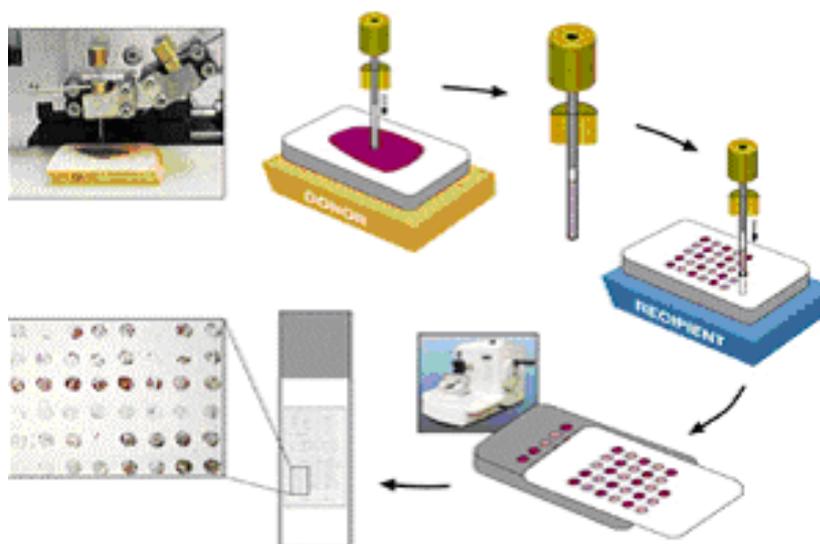


Fig. 1: demonstrates the construction of a tissue microarray. Arrays are assembled by extracting cylindrical cores of 0.6mm diameter from specific locations in “donor” paraffin-embedded tissue blocks and re-embedding them in an arrayed “recipient” block. (Figure from DL Rimm’s website [<http://www.yalepath.org/DEPT/research/YCCTMA/tisarray.htm>].)

Project aims

TMAs are used in the laboratory of Phil Quirke to assess, on a large-scale, the diagnostic and therapeutic significance of various genes and proteins in colorectal tumour samples. The initial aim of the project was to design a database to collate all aspects of data relating to TMA. These data include the pathology reports associated with each tumour sample, tumour grade, TMA construction protocol, tissue staining protocols and results, including images scanned from the microscope. Additional information includes experiment authors, dates and the storage location of each TMA in the laboratory. Once created, the next goal was to interface the database with the World Wide Web (WWW) thereby enabling users to query and assimilate their own data into the database.

Current progress and future directions

A relational database has been designed and implemented in MySQL and has been interfaced with World Wide Web. Currently users can submit their data for incorporation into the

database via the WWW. The next step is to develop query statements for fast and informative interrogation of the database.

The long-term aim of the project is to analyse the data amassed within the database using machine learning tools.

Collaborations: This work was carried out in collaboration with Prof P. Quirke, Faculty of Medicine, University of Leeds.

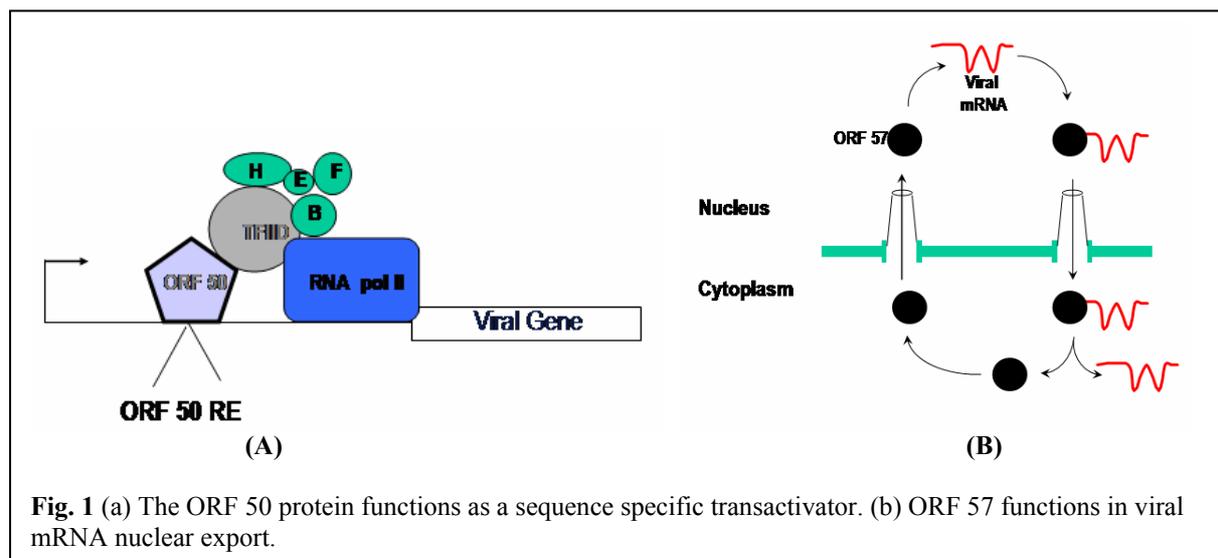
Herpesviral-host cell interactions which regulate viral gene expression

Matthew Walters, Michael Calderwood and Adrian Whitehouse

Gamma-2 herpesviruses are an increasingly important sub-family of herpesviruses with oncogenic potential, particularly as a result of the identification of the first human gamma-2 herpesvirus, Kaposi's sarcoma-associated herpesvirus (KSHV). KSHV has rapidly become the focus of intensive research as epidemiological studies suggest it is the etiologic agent of Kaposi's sarcoma, the most common AIDS-related malignancy. In addition, the presence of the virus has been detected in a variety of lymphoproliferative disorders including primary effusion lymphoma and multicentric Castleman's disease. However at present, analysis of KSHV gene function is hampered by the lack of a permissive cell culture system. Therefore, the ability to easily grow and manipulate the prototype gamma-2 herpesvirus, HVS *in vitro*, has made this virus an attractive model for the analysis of gamma-2 herpesviruses in general. Therefore, we have a major research focus investigating the virus-host cell interactions which regulate the early events in gamma-2 herpesvirus replication cycles, in particular HVS and more recently KSHV.

The interaction of the major transcription control protein, ORF 50 and viral promoters.

The ORF 50 protein is the latent-lytic switch gene in gamma- herpesviruses and transactivates delayed-early gene expression. It functions as a sequence specific transactivator, binding to an A/T rich ORF 50 response element with DE promoters (Fig. 1a).



We have demonstrated that ORF 50 contains a DNA binding domain that has homology to an AT-hook DNA binding motif. The AT-hook is a small DNA-binding protein motif that was first described in the non-histone chromosomal protein HMGA, and allows binding to the minor groove of short stretches of AT-rich DNA. The AT-hook has a core consensus sequence of Pro-Arg-Gly-Arg-Pro (with R-G-R-P being invariant), flanked on either side by a number of positively charged lysine/arginine residues. The core of the AT-hook peptide motif is highly conserved in evolution from bacteria to humans and is found in one or more copies in a large number of other, HMGA proteins, many of which are transcription factors or components of chromatin remodelling complexes. Deletion analysis of this domain reduces ORF 50-mediated transactivation of the DE ORF 6 and ORF 57 promoters by 100% and 90%, respectively. Furthermore, gel retardation experiments demonstrated that the AT-hook motif was required for binding the ORF 50 response element in the promoters of DE genes. Single site-directed mutagenesis of the AT-hook revealed that mutation of the glycine residue

at position 408 to an alanine reduced ORF 50 transactivation of the ORF 57 promoter by 40%. Moreover, mutation of multiple basic residues in conjunction with the glycine residue within the core element of the AT-Hook abolishes ORF 50-mediated transactivation. In addition, the p50GFP Δ AT-hook mutant was capable of functioning as a *trans*-dominant mutant leading to a reduction in virus production of approximately 50% compared to wild-type ORF 50. We are presently characterising these protein-DNA interactions using surface plasmon resonance (SPR) and fluorescence resonance energy transfer (FRET).

The interaction of the nucleocytoplasmic shuttle protein, ORF 57 and viral mRNA.

The ORF 57 protein encodes a nuclear cytoplasmic shuttle protein which mediates the nuclear export of viral mRNAs. We have recent analysis demonstrating that ORF 57 has the ability to bind viral RNA, shuttle between the nucleus and cytoplasm and is required for efficient nuclear export of viral transcripts (Fig. 1b). Moreover, we have shown that ORF57 shuttles between the nucleus and cytoplasm in an CRM-1 independent manner. ORF 57 interacts with the mRNA export factor REF and two other components of the exon-junction complex, Y14 and Magoh. The association of ORF57 with REF stimulates recruitment of the cellular mRNA export factor TAP, and HVS infection triggers the relocalisation of REF and TAP from the nuclear speckles to several large clumps within the cell. Using a dominant negative form of TAP and RNA interference to deplete TAP, we show that it is essential for bulk mRNA export in mammalian cells and is required for ORF57 mediated viral RNA export. Furthermore, we show that disruption of TAP reduces viral replication. These data indicate that γ -2 herpesviruses utilises ORF57 to recruit components of the exon-junction complex and subsequently TAP to promote viral RNA export via the cellular mRNA export pathway. We now aim to analyse the domains required for these interactions in more detail using structural analysis.

Publications.

Williams, B., Goodwin, D.J., Roaden, L.R., Wilson, S.A. & Whitehouse, A. (2004) The prototype gamma-2 herpesvirus nucleocytoplasmic shuttle protein, ORF 57, transports viral RNA via the cellular mRNA export pathway. Submitted to *J. Biol. Chem.*

Walters, M.S., Hall, K.T. & Whitehouse, A. (2004). The Herpesvirus Saimiri open reading frame 50 (Rta) protein encodes an AT-Hook required for binding to the ORF 50 response element in delayed early promoters. *J. Virology*, In Press.

Collaborators.

Stuart Wilson, University of Sheffield.

Funding.

This work has been funded in parts by the BBSRC, MRC, YCR and Royal Society.

Astbury Seminars 2003

Thursday, 23th January 2003

Prof. Joel Gottesfeld, The Scripps Research Institute, La Jolla
"Probing Genes and Chromosomes with Synthetic DNA Ligands"

Thursday, 20th February 2003

Prof. Mike Ferguson, Biological Chemistry & Molecular Microbiology, Dundee
"Glycosylphosphatidylinositol (GPI) biosynthesis: a target for anti-parasite drug development"

Thursday, 6th March 2003

John Ladbury, Department of Biochemistry and Molecular Biology, UCL
"Experimental Thermodynamic and Structural Correlation: A Route to Enhance Drug Design"

Thursday, 13th March 2003

Prof Chris Dobson, FRS, University of Cambridge. The Royal Society Bakerian award lecture.
"Protein folding and misfolding: from theory to therapy"

Thursday, 10th April 2003

Prof. So Iwata, Imperial College, London
"Crystallisation of membrane proteins"

Thursday, 8th May 2003

Prof. Stephen Neidle, Biomolecular Structure Group, UCL
"DNA quadruplexes as selective anticancer targets for rational drug design"

Thursday, 22nd May 2003

Dr. Adam Nelson, Chemistry, Leeds
"Exploring conformational space: Chemical biology of libraries of stereoisomeric carbohydrate mimetics"

Monday 8th September, 2003

Dr. Kurt L Krause, Department of Biol. & Biochem., University of Houston
"Structure aided drug design for anti-tuberculosis agents using alanine racemase"

Wednesday 1st October, 2003

Prof. Venki Ramakrishnan, Structural Studies Division, LMB
"Structure of the translational apparatus"

Thursday 9th October, 2003

Dr. Stan Burgess, Biomedical Sciences, University of Leeds
"Structure and power stroke of the molecular motor Dynein"

Thursday 6th November, 2003

Dr. Michele Vendruscolo, Department of Chemistry, University of Cambridge
"Determination of protein folding pathways at atomic resolution"

Thursday 27th November, 2003

Dr Michael Koonce, Wadsworth Center, New York State Department of Health
"Heads or tails: structure function relationships of the microtubule-based motor, dynein"

Thursday 4th December, 2003

Dr. Laura Itzhaki, MRC Cancer Cell Unit, Cambridge
"Integrating biophysics into medicine: cell cycle regulation and protein degradation"

We would like to acknowledge our corporate sponsors for 2003:



BELLE TECHNOLOGY

Waters



Publications by Astbury Centre Members 2003

- Ackley, M. A., R. J. M. Governo, C. E. Cass, J. D. Young, S. A. Baldwin and A. E. King (2003). "Control of glutamatergic neurotransmission in the rat spinal dorsal horn by the nucleoside transporter ENT1." *Journal of Physiology-London* **548**: 507-517.
- Aggeli, A., M. Bell, L. M. Carrick, C. W. G. Fishwick, R. Harding, P. J. Mawer, S. E. Radford, A. E. Strong and N. Boden (2003). "pH as a trigger of peptide beta-sheet self-assembly and reversible switching between nematic and isotropic phases." *Journal of the American Chemical Society* **125**: 9619-9628.
- Al-Sabah, S. and D. Donnelly (2003). "A model for receptor-peptide binding at the glucagon-like peptide-1 (GLP-1) receptor through the analysis of truncated ligands and receptors." *British Journal of Pharmacology* **140**: 339-346.
- Al-Sabah, S. and D. Donnelly (2003). "The positive charge at Lys-288 of the glucagon-like peptide-1 (GLP-1) receptor is important for binding the N-terminus of peptide agonists." *Febs Letters* **553**: 342-346.
- Anstiss, M., J. M. Holland, A. Nelson and J. R. Titchmarsh (2003). "Beyond breaking the mirror plane: The desymmetrisation of centrosymmetric molecules as an efficient strategy for asymmetric synthesis." *Synlett*: 1213-1220.
- Ashcroft, A. E. (2003). "Protein and peptide identification: the role of mass spectrometry in proteomics." *Natural Product Reports* **20**: 202-215.
- Atkinson, H. J., P. E. Urwin and M. J. McPherson (2003). "Engineering plants for nematode resistance." *Annual Review of Phytopathology* **41**: 615-639.
- Bailey, S., S. E. Sedelnikova, P. Mesa, S. Ayora, J. P. Waltho, A. E. Ashcroft, A. J. Baron, J. C. Alonso and J. B. Rafferty (2003). "Structural analysis of *Bacillus subtilis* SPP1 phage helicase loader protein G39P." *Journal of Biological Chemistry* **278**: 15304-15312.
- Bartlett, S., R. Hodgson, J. M. Holland, M. Jones, C. Kilner, A. Nelson and S. Warriner (2003). "Exploiting predisposition in the stereoselective synthesis of mono-, bi- and tetracyclic oxygen heterocycles: Equilibration between, and trapping of, alternative di- and tetraacetals." *Organic & Biomolecular Chemistry* **1**: 2393-2402.
- Beales, L. P., A. Holzenburg and D. J. Rowlands (2003). "Viral internal ribosome entry site structures segregate into two distinct morphologies." *Journal of Virology* **77**: 6574-6579.
- Beaumont, N. J., V. O. Skinner, T. M. M. Tan, B. S. Ramesh, D. J. Byrne, G. S. MacColl, J. N. Keen, P. M. Bouloux, D. P. Mikhailidis, K. R. Bruckdorfer, M. P. Vanderpump and K. S. Srai (2003). "Ghrelin can bind to a species of high density lipoprotein associated with paraoxonase." *Journal of Biological Chemistry* **278**: 8877-8880.
- Bentham, M., S. Mazaleyrat and M. Harris (2003). "The di-leucine motif in the cytoplasmic tail of CD4 is not required for binding to human immunodeficiency virus type 1 Nef, but is critical for CD4 down-modulation." *J Gen Virol* **84**: 2705-13.
- Bentley, J., D. Itchayanan, K. Barnes, E. McIntosh, X. W. Tang, C. P. Downes, G. D. Holman, A. D. Whetton, P. J. Owen-Lynch and S. A. Baldwin (2003). "Interleukin-3-mediated cell survival signals include phosphatidylinositol 3-kinase-dependent translocation of the glucose transporter GLUT1 to the cell surface." *Journal of Biological Chemistry* **278**: 39337-39348.

- Best, R. B., D. J. Brockwell, J. L. Toca-Herrera, A. W. Blake, D. A. Smith, S. E. Radford and J. Clarke (2003). "Force mode atomic force microscopy as a tool for protein folding studies." *Analytica Chimica Acta* **479**: 87-105.
- Bradford, J. R. and D. R. Westhead (2003). "Asymmetric mutation rates at enzyme-inhibitor interfaces: Implications for the protein-protein docking problem." *Protein Science* **12**: 2099-2103.
- Brockwell, D. J., E. Paci, R. C. Zinober, G. S. Beddard, P. D. Olmsted, D. A. Smith, R. N. Perham and S. E. Radford (2003). "Pulling geometry defines the mechanical resistance of a beta- sheet protein." *Nature Structural Biology* **10**: 731-737.
- Burchmore, R. J. S., L. J. M. Wallace, D. Candlish, M. I. Al-Salabi, P. R. Beal, M. P. Barrett, S. A. Baldwin and H. P. de Koning (2003). "Cloning, heterologous expression, and *in situ* characterization of the first high affinity nucleobase transporter from a protozoan." *Journal of Biological Chemistry* **278**: 23502-23507.
- Burgess, S., M. Walker, P. J. Knight, J. Sparrow, S. Schmitz, G. Offer, B. Bullard, K. Leonard, J. Holt and J. Trinick (2003). "Structural studies of arthrin: Mono-ubiquitinated actin." *Biophysical Journal* **84**: 480A-480A.
- Burgess, S., M. Walker, H. Sakakibara, K. Oiwa and P. Knight (2003). "Dynein structure and power stroke." *Biophysical Journal* **84**: 443A-443A.
- Burgess, S. A., M. L. Walker, H. Sakakibara, P. J. Knight and K. Oiwa (2003). "Dynein structure and power stroke." *Nature* **421**: 715-718.
- Callaghan, A. J., J. G. Grossmann, Y. U. Redko, L. L. Ilag, M. C. Moncrieffe, M. F. Symmons, C. V. Robinson, K. J. McDowall and B. F. Luisi (2003). "Quaternary structure and catalytic activity of the *Escherichia coli* ribonuclease E amino-terminal catalytic domain." *Biochemistry* **42**: 13848-13855.
- Campbell, S. J., N. D. Gold, R. M. Jackson and D. R. Westhead (2003). "Ligand binding: functional site location, similarity and docking." *Current Opinion in Structural Biology* **13**: 389-395.
- Campbell, S. J. and R. M. Jackson (2003). "Diversity in the SH2 domain family phosphotyrosyl peptide binding site." *Protein Engineering* **16**: 217-227.
- Carter, C., S. Fletcher and A. Nelson (2003). "Towards phase-transfer catalysts with a chiral anion: inducing asymmetry in the reactions of cations." *Tetrahedron-Asymmetry* **14**: 1995-2004.
- Chaykovski, M. M., L. C. Bae, M. C. Cheng, J. H. Murray, K. E. Tortolani, R. Zhang, K. Seshadri, J. H. Findlay, S. Y. Hsieh, A. P. Kalverda, S. W. Homans and J. M. Brown (2003). "Methyl side-chain dynamics in proteins using selective enrichment with a single isotopomer." *J Am Chem Soc* **125**: 15767-71.
- Chen, H. Y., B. Coote, S. Attree and J. A. Hiscox (2003). "Evaluation of a nucleoprotein-based enzyme-linked immunosorbent assay for the detection of antibodies against infectious bronchitis virus." *Avian Pathology* **32**: 519-526.
- Cobbold, C., J. Coventry, S. Ponnambalam and A. P. Monaco (2003). "The Menkes disease ATPase (ATP7A) is internalized via a Rac1- regulated, clathrin- and caveolae-independent pathway." *Human Molecular Genetics* **12**: 1523-1533.
- Cobbold, C., A. P. Monaco, A. Sivaprasadarao and S. Ponnambalam (2003). "Aberrant trafficking of transmembrane proteins in human disease." *Trends in Cell Biology* **13**: 639-647.

- Dalton, J. A. R., I. Michalopoulos and D. R. Westhead (2003). "Calculation of helix packing angles in protein structures." *Bioinformatics* **19**: 1298-1299.
- Damaraju, V. L., S. Damaraju, J. D. Young, S. A. Baldwin, J. Mackey, M. B. Sawyer and C. E. Cass (2003). "Nucleoside anticancer drugs: the role of nucleoside transporters in resistance to cancer chemotherapy." *Oncogene* **22**: 7524-7536.
- Day, A. M., J. H. Cove and M. K. Phillips-Jones (2003). "Cytolysin gene expression in *Enterococcus faecalis* is regulated in response to aerobiosis conditions." *Molecular Genetics and Genomics* **269**: 31-39.
- de Maturana, R. L., A. Willshaw, A. Kuntzsch, R. Rudolph and D. Donnelly (2003). "The isolated N-terminal domain of the glucagon-like peptide-1 (GLP-1) receptor binds exendin peptides with much higher affinity than GLP-1." *Journal of Biological Chemistry* **278**: 10195-10200.
- Declais, A. C., J. M. Fogg, A. D. J. Freeman, F. Coste, J. M. Hadden, S. E. V. Phillips and D. M. J. Lilley (2003). "The complex between a four-way DNA junction and T7 endonuclease I." *Embo Journal* **22**: 1398-1409.
- Dellisantia, C. D., S. Spinelli, C. Cambillau, J. B. C. Findlay, P. F. Zagalsky, S. Finet and V. Receveur-Brechot (2003). "Quaternary structure of alpha-crustacyanin from lobster as seen by small-angle X-ray scattering." *Febs Letters* **544**: 189-193.
- Dixon, N., T. Pali, S. Ball, M. A. Harrison, D. Marsh, J. B. C. Findlay and T. P. Kee (2003). "New biophysical probes for structure-activity analyses of vacuolar-H⁺-ATPase enzymes." *Organic & Biomolecular Chemistry* **1**: 4361-4363.
- Firbank, S. J., M. Rogers, R. Hurtado-Guerrero, D. M. Dooley, M. A. Halcrow, S. E. V. Phillips, P. F. Knowles and M. J. McPherson (2003). "Cofactor processing in galactose oxidase." *Biochemical Society Transactions* **31**: 506-509.
- Fournier, L., A. Gaudel-Siri, P. J. Kocienski and J. M. Pons (2003). "The beta-lactone route to beta-hydroxy or alpha,beta-unsaturated gamma- and delta-lactones. Syntheses of (+/-)-massoialactone and (+/-)-prelactone B." *Synlett*: 107-111.
- Friel, C. T., A. P. Capaldi and S. E. Radford (2003). "Structural analysis of the rate-limiting transition states in the folding of Im7 and Im9: Similarities and differences in the folding of homologous proteins." *Journal of Molecular Biology* **326**: 293-305.
- Giesen, A. W., S. W. Homans and J. M. Brown (2003). "Determination of protein global folds using backbone residual dipolar coupling and long-range NOE restraints." *Journal of Biomolecular Nmr* **25**: 63-71.
- Giles, M. S., P. G. Smith, P. L. Coletta, K. T. Hall and A. Whitehouse (2003). "The herpesvirus saimiri ORF 73 regulatory region provides long-term transgene expression in human carcinoma cell lines." *Cancer Gene Therapy* **10**: 49-56.
- Grewal, S., S. Ponnambalam and J. H. Walker (2003). "Association of cPLA(2)-alpha and COX-1 with the Golgi apparatus of A549 human lung epithelial cells." *Journal of Cell Science* **116**: 2303-2310.
- Griffin, S. D. C., L. P. Beales, D. S. Clarke, O. Worsfold, S. D. Evans, J. Jaeger, M. P. G. Harris and D. J. Rowlands (2003). "The p7 protein of hepatitis C virus forms an ion channel that is blocked by the antiviral drug, Amantadine." *Febs Letters* **535**: 34-38.
- Guy, J. L., R. M. Jackson, K. R. Acharya, E. D. Sturrock, N. M. Hooper and A. J. Turner (2003). "Angiotensin-converting enzyme-2 (ACE2): Comparative modeling of the active site, specificity requirements, and chloride dependence." *Biochemistry* **42**: 13185-13192.

- Hall, D. R., L. E. Kemp, G. A. Leonard, K. Marshall, A. Berry and W. N. Hunter (2003). "The organization of divalent cations in the active site of cadmium *Escherichia coli* fructose-1,6-bisphosphate aldolase." *Acta Crystallographica Section D-Biological Crystallography* **59**: 611-614.
- Harding, M., R. Hodgson, T. Majid, K. J. McDowall and A. Nelson (2003). "A stereodivergent, two-directional synthesis of stereoisomeric C-linked disaccharide mimetics." *Organic & Biomolecular Chemistry* **1**: 338-349.
- Harrison, M., L. Durose, C. F. Song, E. Barratt, J. Trinick, R. Jones and J. B. C. Findlay (2003). "Structure and function of the vacuolar H⁺-ATPase: Moving from low-resolution models to high-resolution structures." *Journal of Bioenergetics and Biomembranes* **35**: 337-345.
- Hatters, D. M., C. A. MacRaid, R. Daniels, W. S. Gosal, N. H. Thomson, J. A. Jones, J. J. Davis, C. E. MacPhee, C. M. Dobson and G. J. Howlett (2003). "The circularization of amyloid fibrils formed by apolipoprotein C-II." *Biophysical Journal* **85**: 3979-3990.
- Hiscox, J. A. (2003). "The interaction of animal cytoplasmic RNA viruses with the nucleus to facilitate replication." *Virus Research* **95**: 13-22.
- Horton, J. R., J. M. Bostock, I. Chopra, L. Hesse, S. E. V. Phillips, D. J. Adams, A. P. Johnson and C. W. G. Fishwick (2003). "Macrocyclic inhibitors of the bacterial cell wall biosynthesis enzyme Mur D." *Bioorganic & Medicinal Chemistry Letters* **13**: 1557-1560.
- Houmeida, A., B. Thompson, S. Burgess, J. Keen, K. Thirumurugan, L. Tskhovrebova, P. J. Knight and J. Trinick (2003). "Preparation of synthetic titin end-filaments." *Biophysical Journal* **84**: 563A-563A.
- Hubert, A., P. J. F. Henderson and D. Marsh (2003). "Lipid-protein interactions in *Escherichia coli* membranes overexpressing the sugar-H⁺ symporter, GalP EPR of spin-labelled lipids." *Biochimica Et Biophysica Acta-Biomembranes* **1611**: 243-248.
- Jones, S., J. Manning, N. M. Kad and S. E. Radford (2003). "Amyloid-forming peptides from beta(2)-microglobulin - Insights into the mechanism of fibril formation *in vitro*." *Journal of Molecular Biology* **325**: 249-257.
- Jones, S., D. P. Smith and S. E. Radford (2003). "Role of the N and C-terminal strands of beta 2-microglobulin in amyloid formation at neutral pH." *Journal of Molecular Biology* **330**: 935-941.
- Kaberdin, V. R. and K. J. McDowall (2003). "Expanding the use of zymography by the chemical linkage of small, defined substrates to the gel matrix." *Genome Research* **13**: 1961-1965.
- Kad, N. M., S. L. Myers, D. P. Smith, D. A. Smith, S. E. Radford and N. H. Thomson (2003). "Hierarchical assembly of beta(2)-microglobulin amyloid *in vitro* revealed by atomic force microscopy." *Journal of Molecular Biology* **330**: 785-797.
- Kitov, P. I., H. Shimizu, S. W. Homans and D. R. Bundle (2003). "Optimization of tether length in nonglycosidically linked bivalent ligands that target sites 2 and 1 of a Shiga-like toxin." *Journal of the American Chemical Society* **125**: 3284-3294.
- Knight, P. J., S. Burgess, M. Walker, F. Wang, J. R. Sellers, H. D. White and J. Trinick (2003). "Motor domain strain and the origin of force in myosin." *Biophysical Journal* **84**: 328A-328A.

- Korchazhkina, O. V., A. E. Ashcroft, J. Croom and C. Exley (2003). "Does either the gastrointestinal peptide PYY or the neuropeptide NPY bind aluminium?" *Journal of Inorganic Biochemistry* **94**: 372-380.
- Krishnan, V. G. and D. R. Westhead (2003). "A comparative study of machine-learning methods to predict the effects of single nucleotide polymorphisms on protein function." *Bioinformatics* **19**: 2199-2209.
- Laguri, C., M. K. Phillips-Jones and M. P. Williamson (2003). "Solution structure and DNA binding of the effector domain from the global regulator PrrA (RegA) from *Rhodobacter sphaeroides*: insights into DNA binding specificity." *Nucleic Acids Research* **31**: 6778-6787.
- Li, G. Y., K. F. Liu, S. A. Baldwin and D. W. Wang (2003). "Equilibrative nucleoside transporters of *Arabidopsis thaliana* - cDNA cloning, expression pattern, and analysis of transport activities." *Journal of Biological Chemistry* **278**: 35732-35742.
- Loewen, S. K., A. M. L. Ng, N. N. Mohabir, S. A. Baldwin, C. E. Cass and J. D. Young (2003). "Functional characterization of a H⁺/nucleoside co-transporter (CaCNT) from *Candida albicans*, a fungal member of the concentrative nucleoside transporter (CNT) family of membrane proteins." *Yeast* **20**: 661-675.
- Macdonald, A., K. Crowder, A. Street, C. McCormick, K. Saksela and M. Harris (2003). "The hepatitis C virus non-structural NS5A protein inhibits activating protein-1 function by perturbing Ras-ERK pathway signaling." *Journal of Biological Chemistry* **278**: 17775-17784.
- Meky, F. A., P. C. Turner, A. E. Ashcroft, J. D. Miller, Y. L. Qiao, M. J. Roth and C. P. Wild (2003). "Development of a urinary biomarker of human exposure to deoxynivalenol." *Food and Chemical Toxicology* **41**: 265-273.
- Milne, J. E. and P. J. Kocienski (2003). "A synthesis of alpha-lithiated enol ethers from alpha- arenesulfinyl enol ethers: Ni(0)- and Pd(0)-catalysed coupling of enol triflates and phosphates derived from lactones with sodium arenethiolates gives alpha- arenesulfanyl enol ethers." *Synthesis-Stuttgart*: 584-592.
- Morrison, S., A. Ward, C. J. Hoyle and P. J. F. Henderson (2003). "Cloning, expression, purification and properties of a putative multidrug resistance efflux protein from *Helicobacter pylori*." *International Journal of Antimicrobial Agents* **22**: 242-249.
- Mutungi, G., J. Trinick and K. W. Ranatunga (2003). "Resting tension characteristics in differentiating intact rat fast- and slow-twitch muscle fibers." *Journal of Applied Physiology* **95**: 2241-2247.
- O'Farrell, D., R. Trowbridge, D. Rowlands and J. Jager (2003). "Substrate complexes of hepatitis C virus RNA polymerase (HC- J4): Structural evidence for nucleotide import and De-novo initiation." *Journal of Molecular Biology* **326**: 1025-1035.
- Oke, O., S. Burgess, F. Wang, J. R. Sellers, P. J. Knight and J. Trinick (2003). "Structural studies of mutant myosin V molecules." *Biophysical Journal* **84**: 116A-116A.
- O'Leary, R., S. Ponnambalam and E. J. Wood (2003). "Pioglitazone-induced myofibroblast cell death: implications for cutaneous scarring." *British Journal of Dermatology* **149**: 665-667.
- Olivetta, E., Z. Percario, G. Fiorucci, G. Mattia, A. Schiavoni, C. Dennis, J. Jager, M. Harris, G. Romeo, E. Affabris and M. Federico (2003). "HIV-1 Nef induces the release of inflammatory factors from human monocyte/macrophages: Involvement of nef endocytotic signals and NF-kappa B activation." *Journal of Immunology* **170**: 1716-1727.

- Parkinson, N. M., C. M. Conyers, J. N. Keen, A. D. MacNicoll, I. Smith and R. J. Weaver (2003). "cDNAs encoding large venom proteins from the parasitoid wasp *Pimpla hypochondriaca* identified by random sequence analysis." *Comparative Biochemistry and Physiology C-Toxicology & Pharmacology* **134**: 513-520.
- Patching, S. G., D. A. Middleton, P. J. F. Henderson and R. B. Herbert (2003). "The economical synthesis of 2'-C-13, 1,3-N-15(2) uridine; preliminary conformational studies by solid state NMR." *Organic & Biomolecular Chemistry* **1**: 2057-2062.
- Pearson, A. R., L. H. Jones, L. A. Higgins, A. E. Ashcroft, C. M. Wilmot and V. L. Davidson (2003). "Understanding quinone cofactor biogenesis in methylamine dehydrogenase through novel cofactor generation." *Biochemistry* **42**: 3224-3230.
- Pommier, A., V. Stepanenko, K. Jarowicki and P. J. Kocienski (2003). "Synthesis of (+)-manoalide via a copper(I)-mediated 1,2- metalate rearrangement." *Journal of Organic Chemistry* **68**: 4008-4013.
- Ponnambalam, S. (2003). "Foreword: Protein secretion and the Golgi apparatus." *Molecular Membrane Biology* **20**: 97-98.
- Ponnambalam, S. and S. A. Baldwin (2003). "Constitutive protein secretion from the trans-Golgi network to the plasma membrane." *Molecular Membrane Biology* **20**: 129-139.
- Radford, S. E. (2003). "Co-translocational misfolding in the ER of living cells." *Nature Structural Biology* **10**: 153-154.
- Redko, Y., M. R. Tock, C. J. Adams, V. R. Kaberdin, J. A. Grasby and K. J. McDowall (2003). "Determination of the catalytic parameters of the N-terminal half of *Escherichia coli* ribonuclease E and the identification of critical functional groups in RNA substrates." *Journal of Biological Chemistry* **278**: 44001-44008.
- Robinson, C., R. C. Shore, S. R. Wood, S. J. Brookes, D. A. M. Smith, J. T. Wright, S. Connell and J. Kirkham (2003). "Subunit structures in hydroxyapatite crystal development in enamel: Implications for amelogenesis imperfecta." *Connective Tissue Research* **44**: 65-71.
- Rowlands, D. J. (2003). "Special issue - Foot-and-mouth disease." *Virus Research* **91**: 1-1.
- Saidijam, M., G. Psakis, J. L. Clough, J. Mueller, S. Suzuki, C. J. Hoyle, S. L. Palmer, S. M. Morrison, M. K. Pos, R. C. Essenberg, M. C. J. Maiden, A. Abu-Bakr, S. G. Baumberg, A. A. Neyfakh, J. K. Griffith, M. J. Stark, A. Ward, J. O'Reilly, N. G. Rutherford, M. K. Phillips-Jones and P. J. F. Henderson (2003). "Collection and characterisation of bacterial membrane proteins." *Febs Letters* **555**: 170-175.
- Saunders, S. M., S. Pascoe, A. P. Johnson, M. J. Pilling and M. E. Jenkin (2003). "Development and preliminary test results of an expert system for the automatic generation of tropospheric VOC degradation mechanisms." *Atmospheric Environment* **37**: 1723-1735.
- Siepen, J. A., S. E. Radford and D. R. Westhead (2003). "β-Edge strands in protein structure prediction and aggregation." *Protein Science* **12**: 2348-2359.
- Smith, D. A., D. J. Brockwell, R. C. Zinober, A. W. Blake, G. S. Beddard, P. D. Olmsted and S. E. Radford (2003). "Unfolding dynamics of proteins under applied force." *Philosophical Transactions of the Royal Society of London Series A-Mathematical Physical and Engineering Sciences* **361**: 713-728.
- Smith, D. P., S. Jones, L. C. Serpell, M. Sunde and S. E. Radford (2003). "A systematic investigation into the effect of protein destabilisation on beta 2-microglobulin amyloid formation." *Journal of Molecular Biology* **330**: 943-954.

- Suryanti, V., A. Nelson and A. Berry (2003). "Cloning, over-expression, purification, and characterisation of N-acetylneuraminase synthase from *Streptococcus agalactiae*." *Protein Expression and Purification* **27**: 346-356.
- Szkotak, A. J., A. M. L. Ng, S. F. P. Man, S. A. Baldwin, C. E. Cass, J. D. Young and M. Duszyk (2003). "Coupling of CFTR-mediated anion secretion to nucleoside transporters and adenosine homeostasis in Calu-3 cells." *Journal of Membrane Biology* **192**: 169-179.
- Thirumurugan, K., S. A. Burgess, J. Trinick, F. Wang, J. R. Sellers, H. D. White and P. J. Knight (2003). "Temperature effects on myosin V head conformation." *Biophysical Journal* **84**: 116A-116A.
- Thomson, N. H., S. Kasas, B. M. Riederer, S. Catsicas, G. Dietler, A. J. Kulik and L. Forro (2003). "Large fluctuations in the disassembly rate of microtubules revealed by atomic force microscopy." *Ultramicroscopy* **97**: 239-247.
- Tskhovrebova, L. and J. Trinick (2003). "Titin: Properties and family relationships." *Nature Reviews Molecular Cell Biology* **4**: 679-689.
- Turnbull, W. B. and A. H. Daranas (2003). "On the value of c: can low affinity systems be studied by isothermal titration calorimetry?" *J Am Chem Soc* **125**: 14859-66.
- Tuthill, T. J., N. G. Papadopoulos, P. Jourdan, L. J. Challinor, N. A. Sharp, C. Plumpton, K. Shah, S. Barnard, L. Dash, J. Burnet, R. A. Killington, D. J. Rowlands, N. J. Clarke, E. D. Blair and S. L. Johnston (2003). "Mouse respiratory epithelial cells support efficient replication of human rhinovirus." *J Gen Virol* **84**: 2829-36.
- White, H. D., K. Thirumurugan, M. L. Walker and J. Trinick (2003). "A second generation apparatus for time-resolved electron cryo- microscopy using stepper motors and electrospray." *Journal of Structural Biology* **144**: 246-252.
- Whitehouse, A. (2003). "Herpesvirus saimiri: A potential gene delivery vector." *International Journal of Molecular Medicine* **11**: 139-148.
- Williams, G. J., S. Domann, A. Nelson and A. Berry (2003). "Modifying the stereochemistry of an enzyme-catalyzed reaction by directed evolution." *Proceedings of the National Academy of Sciences of the United States of America* **100**: 3143-3148.
- Wong, C. K. W., L. D. Jewell, L. Dabbagh, S. A. Baldwin, J. D. Young, J. R. Mackey and C. E. Cass (2003). "Localization of human equilibrative nucleoside transporter-1 (HENT1) protein in normal human gastrointestinal tissues." *Gastroenterology* **124**: A285-A286.
- Yoburn, J. C., S. Deb, I. W. Manfield, P. G. Stockley and D. L. Van Vranken (2003). "A biaryl peptide crosslink in a MetJ peptide model confers cooperative, nonspecific binding to DNA that ablates both repressor binding and in vitro transcription." *Bioorganic & Medicinal Chemistry* **11**: 811-816.
- Yung, A., W. B. Turnbull, A. P. Kalverda, G. S. Thompson, S. W. Homans, P. Kitov and D. R. Bundle (2003). "Large-scale millisecond intersubunit dynamics in the B subunit homopentamer of the toxin derived from *Escherichia coli* O157." *Journal of the American Chemical Society* **125**: 13058-13062.
- Zhang, J., F. Visser, M. F. Vickers, T. Lang, M. J. Robins, L. P. C. Nielsen, I. Nowak, S. A. Baldwin, J. D. Young and C. E. Cass (2003). "Uridine binding motifs of human concentrative nucleoside transporters 1 and 3 produced in *Saccharomyces cerevisiae*." *Molecular Pharmacology* **64**: 1512-1520.