

Martin Noble.

Adhesive interactions, the cell cycle, and NAT enzymes.

Introduction

Work in this group addresses structure-function relationships of medically important proteins from three different areas: adhesive cellular interactions, the eukaryotic cell cycle, and the enzyme arylamine N-acetyltransferase (NAT). We study these proteins by both experimental and theoretical approaches. Experimentally, proteins are subject to biochemical and biophysical characterisation, as well as structural analysis by X-ray crystallography and nuclear magnetic resonance (in collaboration with Dr. J. Werner and Prof I. Campbell). Development of theoretical methods has centred on developing the molecular graphics program AESOP. AESOP interfaces with the program GRID to allow quantitative and qualitative analysis of the chemical character of protein surfaces. We have been exploring the use of this approach to identify probable sites of protein-protein interaction.

1. Adhesion and signalling from adhesive complexes

The cell membrane is a location of bidirectional signalling. External events may activate signalling networks within the cell, while internal signals may be transduced through modulation of the properties of cell-surface proteins that mediate cellular adhesion to the extracellular matrix or to other cells. We are studying two prototypical signalling systems that involve both outside-in and inside-out signalling. These are CD44, a transmembrane protein found on the surface of many cell-types in mammals, where it acts as the major receptor for the glycosaminoglycan hyaluronan (HA), and focal adhesions, the cellular substructure that forms where integrins bind to proteinaceous components of the extracellular matrix. To date we have explored the structure of the HA binding domain of CD44, as well as various aspects of the structure and function of different parts of the focal adhesion kinase (FAK) molecule.

1.1 CD44

(With Dr. D. Jackson, IMM, Dr. A.J. Day, MRC Immunochemistry)

The binding of CD44 to HA is a key step in important biological processes, including the extrusion of immune system cells from the bloodstream. CD44-HA interactions are also implicated in several aspects of disease progression, including tumor formation and tissue invasion by streptococcal bacteria. Structural studies of CD44 have been hampered by the difficulty of growing reasonably sized crystals. Dr. Suneale Bannerji, from Dr. Jackson's laboratory, and Dr. Ed Lowe from the LMB have employed two anomalous-scattering experiments, using extremely small crystals ($5 \times 5 \times 100 \mu\text{m}^3$), that have together allowed us to determine a 2.2 Å resolution structure for the HA-binding domain of human CD44. One of these experiments, conducted at the X-ray diffraction beamline at Elettra, involved recording diffraction of 1.5 Å wavelength diffraction by CD44 crystals to allow the identification of the positions of CD44's disulphide bonds ([Figure 1a](#)). This provided crucial information for determining the non-crystallographic symmetry (NCS) transformation relating two molecules of CD44 in the asymmetric unit, which was subsequently applied to improve phases from a seleno-methionylated (SeMet) CD44 dataset collected on ID14.4 at the ESRF. The fold corresponds to that of a C-type lectin, with elaborations specific to Link domains in general, as well as major insertions at the N- and C- termini that are specific to HA-binding Link motifs found in cell surface receptors (Lyve 1 and CD44) ([Figure 1b](#)). These extensions extend the available HA interaction surface, presumably allowing for improved specificity and affinity of interactions. This extended surface would be expected to accommodate approximately six monosaccharide units, consistent with the tight affinity achieved for binding of such sugars to CD44.

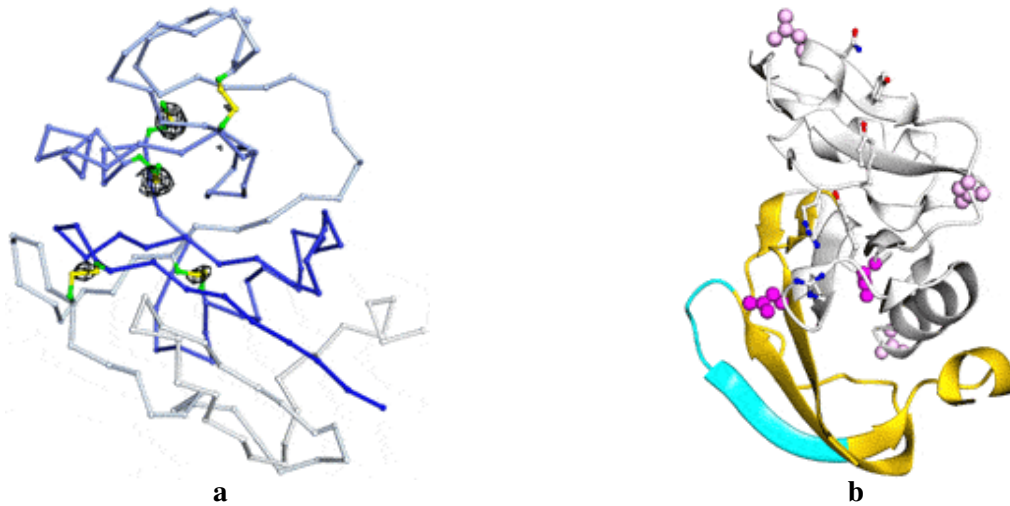


Figure 1. The crystal structure structure of the HA binding domain of CD44. [...more](#)

1.1.2 The Focal adhesion targeting domain.

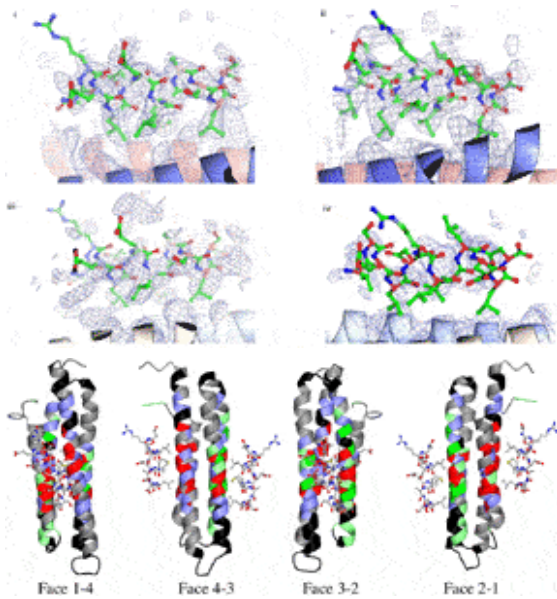


Figure 2. Experimental indications of the binding site of LD peptides on the FAT domain. [...more](#)

Maria Hoellerer(With Dr. S. Arold, CNRS Montpellier, Prof. I.Campbell, Biochemistry, Dr. J. Werner Biochemistry, Biochemistry, Dr. M. Ginsberg Scripps, San Diego)

FAK is localised to focal adhesions via its C-terminal Focal Adhesion Targeting (FAT) domain. FAT performs this function by binding to paxillin and talin, both of which are in turn associated with the cytoplasmic tails of integrins. FAT is also required for binding of FAK substrates: a phospho-tyrosine motif within FAT binds to the Src homology (SH) 2 domain of the adapter protein Grb2. A proline-rich sequence immediately upstream of the FAT domain binds to the SH3 domains of p130 CAS, Graf, and the p85 subunit of phosphatidylinositol-3 kinase.

In the period to 2001, we have solved the crystal structure of the isolated FAT domain and begun biochemical and biophysical analysis of its interactions¹. Phospho-tyrosine 925 of FAT is unable to interact with the SH2 domain of Grb2. To reconcile structural and functional data, we have proposed a switch mechanism whereby phosphorylation of Y925 triggers structural changes that lead to the exposure of protein-protein interaction motifs of FAK.

This year we have also analysed the interactions of LD-motif peptides with FAT by both crystallography and NMR. These studies have identified two peptide binding sites, both of which can be observed crystallographically ([Figure 2a](#)), and both of which are confirmed from chemical shift perturbation mapping ([Figure 2b](#)).

2. Cell cycle

We have collaborated with other groups in the LMB on various aspects of cell-cycle regulation over recent years, the results of which are discussed elsewhere in this report (see reports from [Prof. Louise Johnson](#) and [Dr. Jane Endicott](#)).

2.1 CDK inhibitor design.

David Pratt, Dr. Simon Holton, Dr. Tom Davies (With Prof. D.R. Newell, Prof. R. Griffin, Prof. B. Golding, ADDI, Newcastle, and Dr. T. Boyle, Dr. P. Jewsbury, AstraZeneca PLC).

The cyclin-dependent kinases (CDKs) are important targets for therapeutic intervention in various proliferative disease states including cancer. Although a number of small molecule CDK2 inhibitors have been identified (eg the purvalanols, olomoucine, indirubins), they frequently lack potency and/or specificity. As described in last year's report from Dr. Jane Endicott, we have developed a nanomolar inhibitor of CDK2 by employing a structure-lead approach². To describe the molecular interactions that contribute to inhibitor selectivity within the CDK family we have used CDK2 as a template to generate a "CDK4-ised CDK2" molecule. By mutating residues within the CDK2 active site we have generated a hybrid protein with the CDK4 active site sequence. In the absence of a structure for an active CDK4/6- cyclin D complex this reagent provides an opportunity to continue the design cycle to optimise inhibitor selectivity.

We have also characterised a series of pyrimidine inhibitors³. The lead compound in the series, 5-nitrosopyrimidine ([Figure 3](#)) is an inhibitor of Cdk4/cyclinD and Cdk2/cyclinA3 with IC₅₀ values of $2.9 \pm 0.1 \mu\text{M}$ and $2.2 \pm 0.6 \mu\text{M}$, respectively. This compound adopts a binding mode nearly identical to that of the 6-alkoxyguanine derivative NU2058⁴. Preliminary structure-activity studies had suggested that the presence of an intra-molecular hydrogen bond between the C5-nitroso group and the amino group at C6 was necessary to maintain the activity in the pyrimidine series. The intramolecular H-bond orientates 5-nitrosopyrimidine into a purine-like shape, enabling an optimal interaction between the 6-amino group and Glu 81 ([Figure 3a,b](#)). Further studies on the interactions made by the O6- substitution within the purine alkylguanine series have also been carried out⁵ ([Figure 3c](#)).

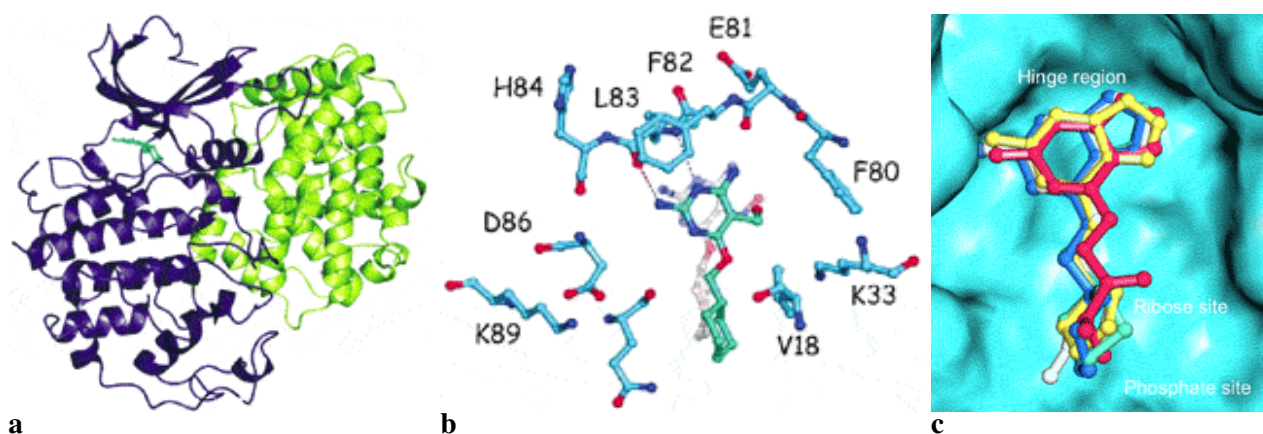


Figure 3. Crystal structures of inhibitors bound to CDK2/cyclin. [...more](#)

3. NAT enzymes

(With Prof. E. Sim, Pharmacology)

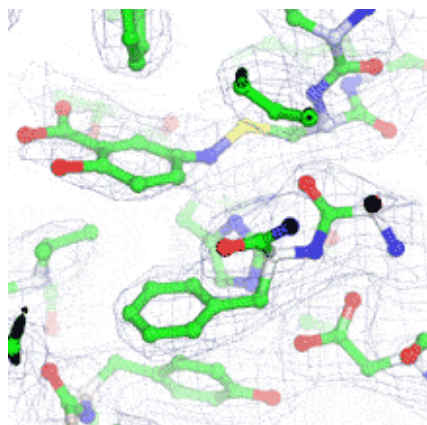


Figure 4 5-aminosalicylate bound to *M. smegmatis* NAT

[...more](#)

A growing problem in global health care is the incidence of antibiotic-resistant bacteria. Acetylation is one mechanism employed by bacteria to evade antibiotics. The family of enzymes that acetylate arylamines and hydrazines (the arylamine N-acetyltransferase or NAT family of enzymes) have a novel fold and a novel mechanism. NAT enzymes are highly active in certain *Salmonella* strains, where they have been exploited in the 'Ames' test of carcinogenicity. NAT enzymes have also been recently identified in *Mycobacterium tuberculosis*, where they may be responsible for some instances of tubercular resistance to such antibiotics as isoniazid. In order to understand their catalytic mechanism, and to design specific inhibitors of pathogenic isoforms, NAT enzymes constitute an interesting target for structural characterisation

In 1999 we published the structure of NAT from *S. typhimurium* at 2.7 Å resolution⁶, both alone and in complex with a bromoacetanilid suicide inhibitor which has confirmed the location of the active site. Subsequently, we have solved the structure of NAT from *M. smegmatis*, a close relative of the potential drug target *M. tuberculosis*^{7; 8}. *M. smegmatis* NAT crystallises in three crystal forms, one with space group P2₁2₁2₁, one with space group P4₁2₁2 and one P6₂22. These crystals diffract beyond 2.0 Å resolution, giving rise to a much more well defined structure. This year James Sandy from Professor Sim's group has worked with us to solve the structure of *M. smegmatis* NAT in complex with the substrate 5-aminosalicylate, an important compound in the treatment of Crohn's disease ([Figure 4](#)).

4. Methods

4.1 Ligand design/libraries

David Pratt, Alex Zawaira, Giles Robertson (With Dr. J. McDonnell, Dr. J. Endicott and Prof. L. Johnson).

Cyclin dependent kinases direct the progression of the eukaryotic cell cycle. Phases of this cycle are characterised by the activity of specific pairings of kinase subunits (CDKs) with cognate cyclin activators. The cyclin molecule imposes a partially active conformation upon the CDK, and further modulates its activity by recruiting substrates and inhibitor proteins. The cyclin recruitment site interacts with a degenerate "Cy- motif", which has the consensus sequence R/KXL. In previous work (see the report from Prof. Johnson), we have explored the recognition of natural and non-natural peptide ligands bound to this site, in order to learn how affinity and specificity are achieved^{9; 10}. Recently, we have begun to explore the character of interacting surfaces of CDK and cyclin molecules, by use of redundant peptide libraries. We have also begun to explore further potentially high-throughput assays for inhibitor binding to the recruitment site of cyclins, and to examine the structures of bound ligands.

4.2 Visualisation/analysis

AESOP is an interactive molecular illustration program that is able to prepare output in the form of moving as well as static images. Aesop is controlled through a graphical user interface based upon the GTK toolkit. This interface allows flexible representation of protein structures in most required forms (ball-and-stick, wire-frame, ribbon etc). Surface calculation and representation is included, as is contouring of electron density maps. Control is also offered of lighting, material properties, orientation etc. The program runs on Linux, LinuxPPC, Digital Tru64 Unix, IRIX, Windows, and MacOS X operating systems. The program is scriptable, and has its own ascii save file format for interchange and storage of data. AESOP allows the definition of views and scenes, and includes controls for animation. Animation options include morphing (linear interpolation of coordinates between two related structures), and controlled smooth trajectories between selected views. In addition to the ongoing development of AESOP, We are collaborating with Dr. Liz Potterton from the CCP4 project. CCP4 coordinates the development of a coherent suite of macromolecular crystallography programs within the British crystallographic community. The existing suite satisfies most major requirements, but lacks a tool for molecular visualization and manipulation. We have begun to address this through a Python based, object oriented molecular graphics framework called CCP4MG (working title)¹¹. In time, we hope that this program will offer all necessary tools for people analysing protein structures and developing novel algorithms in protein structure determination.

References.

1. Arold, S. T., Hoellerer, M. K. & Noble, M. E. (2002). The structural basis of localization and signaling by the focal adhesion targeting domain. *Structure (Camb)* 10, 319-27.
2. Davies, T. G., Bentley, J., Arris, C. E., Boyle, F. T., Curtin, N. J., Endicott, J. A., Gibson, A. E., Golding, B. T., Griffin, R. J., Hardcastle, I. R., Jewsbury, P., Johnson, L. N., Mesguiche, V., Newell, D. R., Noble, M. E., Tucker, J. A., Wang, L. & Whitfield, H. J. (2002). Structure-based design of a potent purine-based cyclin-dependent kinase inhibitor. *Nat Struct Biol* 9, 745-9.
3. Mesguiche, V., Parsons, R. J., Arris, C. E., Bentley, J., Boyle, F. T., Curtin, N. J., Davies, T. G., Endicott, J. A., Gibson, A. E., Golding, B. T., Griffin, R. J., Jewsbury, P., Johnson, L. N., Newell, D. R., Noble, M. E., Wang, L. Z. & Hardcastle, I. R. (2003). 4-Alkoxy-2,6-diaminopyrimidine derivatives: inhibitors of cyclin dependent kinases 1 and 2. *Bioorg Med Chem Lett* 13, 217-222.
4. Arris, C. E., Boyle, F. T., Calvert, A. H., Curtin, N. J., Endicott, J. A., Garman, E. F., Gibson, A. E., Golding, B. T., Grant, S., Griffin, R. J., Jewsbury, P., Johnson, L. N., Lawrie, A. M., Newell, D. R., Noble, M. E., Sausville, E. A., Schultz, R. & Yu, W. (2000). Identification of novel purine and pyrimidine cyclin-dependent kinase inhibitors with distinct molecular interactions and tumor cell growth inhibition profiles. *J Med Chem* 43, 2797-804.
5. Gibson, A. E., Arris, C. E., Bentley, J., Boyle, F. T., Curtin, N. J., Davies, T. G., Endicott, J. A., Golding, B. T., Grant, S., Griffin, R. J., Jewsbury, P., Johnson, L. N., Mesguiche, V., Newell, D. R., Noble, M. E., Tucker, J. A. & Whitfield, H. J. (2002). Probing the ATP Ribose-Binding Domain of Cyclin-Dependent Kinases 1 and 2 with O(6)-Substituted Guanine Derivatives. *J Med Chem* 45, 3381-3393.
6. Sinclair, J. C., Sandy, J., Delgoda, R., Sim, E. & Noble, M. E. (2000). Structure of arylamine N-acetyltransferase reveals a catalytic triad. *Nat Struct Biol* 7, 560-4.
7. Kawamura, A., Sandy, J., Upton, A., Noble, M. & Sim, E. (2003). Structural investigation of mutant *Mycobacterium smegmatis* arylamine N-acetyltransferase: a model for a naturally occurring functional polymorphism in *Mycobacterium tuberculosis* arylamine N-acetyltransferase. *Protein Expr Purif* 27, 75-84.
8. Sandy, J., Mushtaq, A., Kawamura, A., Sinclair, J., Sim, E. & Noble, M. (2002). The structure of arylamine N-acetyltransferase from *Mycobacterium smegmatis*--an enzyme which inactivates the anti-tubercular drug, isoniazid. *J Mol Biol* 318, 1071-83.
9. Lowe, E. D., Tews, I., Cheng, K. Y., Brown, N. R., Gul, S., Noble, M. E., Gamblin, S. J. & Johnson, L. N. (2002). Specificity Determinants of Recruitment Peptides Bound to Phospho-CDK2/Cyclin A(.). *Biochemistry* 41, 15625-34.
10. Leng, X., Noble, M., Adams, P. D., Qin, J. & Harper, J. W. (2002). Reversal of growth suppression by p107 via direct phosphorylation by cyclin D1/cyclin-dependent kinase 4. *Mol Cell Biol* 22, 2242-54.
11. Potterton, E., McNicholas, S., Krissinel, E., Cowtan, K. & Noble, M. (2002). The CCP4 molecular-graphics project. *Acta Crystallogr D Biol Crystallogr* 58, 1955-7.

