UPPSALA
UNIVERSITET

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology 382

Evasion and Attack: Structural Studies of a Bacterial Albumin-binding Protein and of a Cephalosporin Biosynthetic Enzyme

SARA LEJON

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2008

ISSN 1651-6214
ISBN 978-91-554-7062-3
urn:nbn:se:uu:diva-8399

Dissertation presented at Uppsala University to be publicly examined in B22, Biomedical Centre, Husargatan 3, Uppsala, Friday, February 1, 2008 at 13:00 for the degree of Doctor of Philosophy. The examination will be conducted in English.

Abstract
Lejon, S. 2008. Evasion and Attack: Structural Studies of a Bacterial Albumin-binding Protein and of a Cephalosporin Biosynthetic Enzyme. Acta Universitatis Upsaliensis. *Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology* 382. 65 pp. Uppsala. ISBN 978-91-554-7062-3.

This thesis describes the crystal structures of two proteins in the context of combatting bacterial infections. The GA module is a bacterial albumin-binding domain from a surface protein expressed by pathogenic strains of the human commensal bacterium *Finegoldia magna*. The structure of the GA module in complex with human serum albumin (HSA) provides insights into bacterial immune evasion, where pathogenicity is acquired by the bacterial cell through the ability to coat (and disguise) itself with serum proteins. The structure shows binding of the GA module to HSA in the presence of fatty acids, and reveals interactions responsible for the host range specificity of the invading bacterium. The complex resulting from binding of the GA module to HSA readily forms stable crystals that permit structural studies of drug binding to HSA. This was exploited to study the specific binding of the drug naproxen to the albumin molecule.

Antibiotics play a major role in controlling infections by attacking invading bacteria. The enzyme deacetylcephalosporin C acetyltransferase (DAC-AT) catalyses the last step in the biosynthesis of the beta-lactam antibiotic cephalosporin C, one of the clinically most important antibiotics in current use. The enzyme uses acetyl coenzyme A as cofactor to acetylate a biosynthetic intermediate. Structures of DAC-AT in complexes with reaction intermediates have been determined. The structures suggest that the acetyl transfer reaction proceeds through a double displacement mechanism, with acetylation of a catalytic serine by the cofactor through a suggested tetrahedral transition state, followed by acetyl transfer to the intermediate through a second suggested tetrahedral transition state. The structure of DAC-AT yields valuable information for the continued study of cephalosporin biosynthesis in the context of developing new beta-lactam compounds.

Keywords: human serum albumin, GA module, albumin-binding, Finegoldia magna, deacetylcephalosporin C acetyltransferase, cephalosporin C, antibiotic, beta-lactam, biosynthesis, Acremonium chrysogenum, X-ray crystallography

Sara Lejon, Department of Cell and Molecular Biology, Box 596, Uppsala University, SE-75124 Uppsala, Sweden

ISSN 1651-6214
ISBN 978-91-554-7062-3
urn:nbn:se:uu:diva-8399 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-8399)

Till min familj

The Road goes ever on and on
Down from the door where it began.
Now far ahead the Road has gone,
and I must follow, if I can,
Pursuing it with weary feet,
Until it joins some larger way,
Where many paths and errands meet.
And whither then? I cannot say.

(J.R.R. Tolkien, *The Lord of the Rings*)

List of publications

This thesis is based on the following papers, which will be referred to by their Roman numerals:

I **Lejon, S.**, Frick, I.M., Björck, L., Wikström, M., Svensson, S. (2004) Crystal structure and biological implications of a bacterial albumin binding module in complex with human serum albumin. *J. Biol. Chem.*, **279**, 42924-42928.

II **Lejon, S.**[*], Cramer, J.F.[*], Nordberg, P.A. Structural basis for the binding of naproxen to human serum albumin in the presence of fatty acids and the GA module. *Accepted for publication in Acta Crystallogr. F.*

III Cramer, J.F., Nordberg, P.A., Hajdu, J., **Lejon S.** (2007) Crystal structure of a bacterial albumin-binding domain at 1.4 Å resolution. *FEBS Lett.*, **581**, 3178-3182.

IV **Lejon, S.**, Ellis J., Valegård, K. The last step in cephalosporin C formation revealed: Crystal structures of deacetylcephalosporin C acetyl transferase from *Acremonium chrysogenum* in complexes with reaction intermediates. *Revision stage for J. Mol. Biol.*

[*]Sharing first authorship.

Contents

Abbreviations

aa	amino acid(s)
ABD	albumin-binding domain (from streptococcal protein G)
ACV	L-δ-(α-aminoadipoyl)-L-cysteinyl-D-valine
coA	coenzyme A
DAC	deacetylcephalosporin C
DAC-AT	deacetylcephalosporin C acetyltransferase
DAOC	deacetoxycephalosporin C
ESRF	European Synchrotron Radiation Facility
FA	fatty acid (binding site)
GA	protein G-related albumin-binding
*Hi*HTA	*Haemophilus influenzae* homoserine O-transacetylase/acetyltransferase
HSA	human serum albumin
HTA	homoserine O-transacetylase/acetyltransferase
IPN	isopenicillin N
IPNS	isopenicillin N synthase
IPTG	isopropyl-β-D-thiogalactopyranoside
K_A	association constant
K_D	dissociation constant
*Li*HTA	*Leptospira interrogans* homoserine O-transacetylase/acetyltransferase
MIR(AS)	multiple isomorphous replacement (with anomalous scattering)
NCS	non-crystallographic symmetry
NMR	nuclear magnetic resonance
NSAID	non-steroidal anti-inflammatory drug
PAB	peptostreptococcal albumin-binding
PDB	Protein Data Bank
PEG	polyethylene glycol
RMSD	root-mean-square distance
SAD	single-wavelength anomalous diffraction/dispersion
SIR(AS)	single isomorphous replacement (with anomalous scattering)
TLS	translation, libration, screw rotation

1. Introduction

In the past half-century, scientific progress and the advancement of technology have paved the way for human lifestyles vastly different to those of centuries past. A major contributor to the well-being enjoyed by people in some parts of the world today is the ready availability of medical care in the form of antibiotics and other drugs to treat infections and diseases. These advancements are, however, fraught with difficulties and limitations. The advent of antibiotic-resistant pathogenic bacterial strains was recognised already during the early days of penicillin development, and a majority of the fatal diseases that humans around the world are succumbing to today are still subjected to intense research in order to find ways of overcoming them with drug compounds and by primary and specialist care.

In the continual struggle against infections and diseases, detailed knowledge of the causative processes is of utmost importance. This thesis addresses the problem of bacterial infections from two sides: the invading pathogen's side and the side of the counter-attacking host.

1.1 Drugs and antibiotics

1.1.1 A linguistic note

The Merriam-Webster medical dictionary resource at the U.S National Library of Medicine defines the word *drug* as “a substance intended for use in the diagnosis, cure, mitigation, treatment, or prevention of disease, […] a substance other than food intended to affect the structure or function of the body” (Merriam-Webster, 2007). One of the subclasses within this definition of a drug is the *antibiotic*, which in turn is defined as “a substance produced by, or a semi-synthetic substance derived from, a microorganism and able in dilute solution to inhibit or kill another microorganism” (Merriam-Webster, 2007). Loosely, one might therefore say that a drug is a chemical substance that is administered in order to have a medical effect; in the case of an antibiotic drug, the chemical substance is specifically derived from a microorganism, and the medical effect is the killing of another (pathogenic) microorganism.

1.1.2 A short history of antibiotics

Historical evidence of the cultural use of health-improving substances (including psychoactive drugs) dates back over 5000 years. For millennia, the discovery of compounds with medically useful properties was chiefly a trial-and-error process. In the late 19th and early 20th century, serendipitous observations of the antibacterial properties of fungi belonging to the genus *Penicillium* were made by John Tyndall, Ernest Duchesne, Joseph Lister, and others (discussed by (Teiling, 1965)). The discovery of penicillin is, however, attributed to Alexander Fleming (Fleming, 1929), and it led to the isolation and large-scale production of penicillin, owing to the pioneering work by Howard Florey and Ernst Chain in the late 1930s (Chain*, et al.*, 1940). Following this successful work (reviewed by E. P. Abraham (Abraham, 1990)), numerous research efforts were mounted in the search for other natural compounds that could be used for the treatment of infectious diseases. Drug research in the 1960s also led to determination of the first quantitative relationships between chemical structure and biological activity. The new era of antibiotics was greeted with enthusiasm, and in 1969, the U.S. Surgeon General proclaimed it was time to "close the book on infectious diseases".

In sharp contrast to this statement stands the insistent call of the present day for new and more efficient antibiotics (Breithaupt, 1999). Strains of drug-resistant pathogens are evolving continually. This is troublesome both for developing countries, where diseases like tuberculosis and malaria are becoming more and more widespread and less and less susceptible to treatment (Meya & McAdam, 2007, Tilley*, et al.*, 2006), as well as for developed countries, where hospital environments and the over-use of antibiotics drive the evolution of antibiotic-resistant pathogenic strains such as MRSA (methicillin-resistant *Staphylococcus aureus*) and *Clostridium difficile* (Livermore, 2000).

It is thus clear that, in the face of the mounting threat of infectious as well as non-infectious disease, there is an ever-increasing demand for further development of efficacious medical treatments.

1.2 Aim and scope of the thesis

The global problem of infectious disease and the development of useful and sustainable treatment are immense and multi-faceted topics. The work presented here approaches the problem from two sides: the pathogenic, defense-evading side, and the defending, counter-attacking human host side.

The first part of the thesis contains a presentation of the structural properties of the protein G-related albumin-binding module (the GA module), a surface protein domain from pathogenic strains of the Gram-positive bacte-

rium *Finegoldia magna*. This protein domain confers pathogenicity by enabling the bacterium to evade the host's immune system, and the study of the details of the interactions underlying this evasion strategy could provide insights into bacterial pathogenesis.

The second part of the thesis contains a presentation of the structural basis of the last step in the biosynthesis of cephalosporin C. Cephalosporin C is a widely used and important β-lactam antibiotic that is employed for the production of semi-synthetic cephalosporin compounds, which are in turn used in the treatment of bacterial infections. Structural information of the enzymes involved in antibiotic biosynthesis could prove valuable for the development of compounds with new antimicrobial properties.

The results presented in this thesis are the outcomes of X-ray crystallography experiments, which allow the study of the three-dimensional structure of crystalline proteins and other macromolecules. The spatial arrangements of the atoms in a protein - and the way these arrangements change over time - form the basis for the function of the protein, and especially valuable is the study of proteins in complexes with interacting species such as other proteins, cofactors, substrates, products or specific ligands, which may modulate biological activity.

Chapter 2 of this thesis presents studies of the GA module and its interactions with human serum albumin (HSA). The crystal structure of the complex between HSA and the GA module is described in **Paper I**, and the structure of the HSA-GA complex with the drug molecule naproxen bound to HSA is described in **Paper II**. A crystal structure illuminating additional characteristics of the GA module in isolation is described in **Paper III**.

Chapter 3 deals with the last step in the biosynthesis of the β-lactam antibiotic cephalosporin C. This section is specifically concerned with the crystal structures of deacetylcephalosporin C acetyltransferase in complex with cofactor, substrate and product (**Paper IV**). Based on the structural data, a reaction mechanism is proposed for this enzyme.

2. The GA module and human serum albumin

2.1 Background

2.1.1 The GA module

Bacterial surface proteins serve a variety of functions that enable a bacterium to interact with a complex and changing environment. In the case of a pathogenic bacterium invading the tissues of a human host, where such interactions may aid the bacterium in its spread and growth, surface proteins play particularly important roles in ensuring the survival and propagation of the bacterial cell. Human blood plasma offers a wide range of interaction partners for an invading bacterium, and many pathogens are known to express surface receptors for binding human plasma proteins (Navarre & Schneewind, 1999). An example of such a surface protein is protein G, which is found on the cell surface of group C and G streptococci. Protein G contains several, separate domains for interacting with human serum albumin (HSA) and immunoglobulin G (IgG) (Myhre & Kronvall, 1980, Myhre & Kronvall, 1977). It has been suggested that these protein-protein interactions at the cell surface result in the bacterial cell being coated *in vivo* with layers of host serum proteins, with the result that the bacterium may circumvent detection and subsequent elimination by the host's immune system.

The anaerobic Gram-positive bacterium *Finegoldia magna* (previously known as *Peptostreptococcus magnus*) is a commensal microorganism normally found in human gastrointestinal and urogenital tracts. Certain strains of *F. magna* express albumin-binding domains as part of the cell surface protein PAB (peptostreptococcal albumin-binding). The main albumin-binding domain from PAB is homologous to the N-terminal albumin-binding domain (ABD) of streptococcal protein G, and has consequently been termed the GA (protein G-related albumin-binding) module (de Château & Björck, 1994, de Château*, et al.*, 1996).

The term "module" denotes a mobile protein domain and it has been shown that the domain is likely to have been introduced into the PAB-encoding gene via horizontal transfer from its streptococcal counterpart. This transfer is an example of "module shuffling", and it is believed that antibiotics have provided the selective pressure behind the transfer (de Château & Björck, 1996). The acquisition of the domain confers a higher growth rate and increased virulence on the bacteria when HSA is present. The domain

thereby effectively turns the bacteria into “wolves in sheep's clothing". In this manner, the bacteria are able to invade human tissues. *F. magna* has been implicated in a variety of infections, mainly vaginoses, soft tissue abscesses, infections of bone, joint and skin, and also in heart valve infections.

The GA module from *F. magna* consists of residues 213-265 of the PAB protein. The domain folds independently into a three-helix bundle structure (Figure 1B) as first shown by NMR spectroscopy (Johansson*, et al.*, 1995, Johansson*, et al.*, 1997). Initial efforts to map the HSA binding site on the GA module using NMR perturbation studies indicated that GA residues from the second α-helix and from the loop region preceding it were involved in binding HSA (Johansson*, et al.*, 2002).

The studies presented in this chapter are primarily concerned with the structural characterisation of the binding of the GA module to HSA. In addition, the crystal structure of the GA module in isolation is presented.

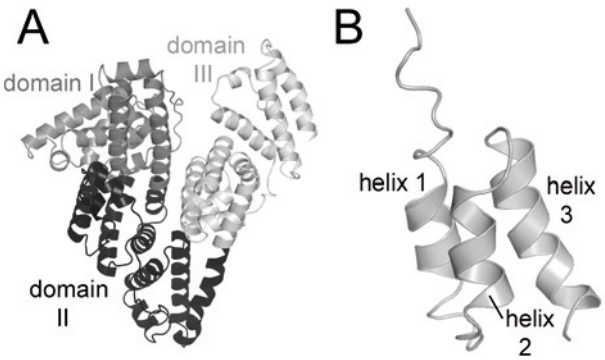

Figure 1. (A) Crystal structure of defatted human serum albumin (PDB code 1AO6) (Sugio*, et al.*, 1999). The three homologous domains are shown in different shades. (B) NMR structure of the GA module in solution (PDB code 1PRB) (Johansson*, et al.*, 1997).

2.1.2 Human serum albumin

Human serum albumin (HSA) is the predominant protein constituent of human blood plasma, where it acts as a transporter of a wide variety of compounds. Its role in transporting fatty acids to tissues for utilisation as an energy source is well-established (Fredrickson & Gordon, 1958), as is its role in the binding and transport of various steroids, hemoglobin breakdown products, hormones and other endogenous substances (Peters, 1995). In addition to these, HSA also binds an extraordinarily broad spectrum of exogenous drug compounds.

2.1.2.1 The binding properties of HSA

The propensity of HSA to bind such a large range of substances has not only earned it the reputation as the “tramp steamer” of the circulatory system, but

also notoriety for the challenges it poses in pharmaceutical drug development projects. Promising lead compounds are often reported to bind with high affinity to HSA. Although in some cases such binding might be beneficial in terms of administering drug compounds with poor solubility in water, it often leads to a severe reduction of the unbound, active concentration of the drug in plasma. As a result, high doses of a drug may be required for clinical effect, which in turn influences metabolism and elimination of the drug.

In this context, structural information on ligand binding to HSA early in drug development may be of importance in the optimisation of a lead compound through redesign (Mao*, et al.*, 2001). Following the determination of the first HSA crystal structure in 1992 (He & Carter, 1992), several X-ray crystallographic studies on HSA in complex with ligands have been reported (Bhattacharya, Curry*, et al.*, 2000, Ghuman*, et al.*, 2005, Petitpas, Bhattacharya*, et al.*, 2001, Petitpas, Grüne*, et al.*, 2001, Petitpas*, et al.*, 2003, Zunszain*, et al.*, 2003). Interpretation of structural data on albumin is, however, often problematic due to several factors; for instance, interactions between drugs binding simultaneously in neighbouring sites, allosteric effects upon binding, and the overall conformational flexibility of the albumin molecule (Kragh-Hansen*, et al.*, 2002). These factors may also play a role from a therapeutic viewpoint, by influencing the *in vivo* free concentration, distribution and half-life of a drug. The free concentration of HSA-bound ligands may also be affected by the pathophysiological status of the patient, e.g. through variations of the circulatory fatty acid level during diseased states (Cistola & Small, 1991, Kober & Sjöholm, 1980).

2.1.2.2 Binding sites in HSA

The crystal structure of HSA (He & Carter, 1992) showed a heart-shaped protein (Figure 1A), consisting of three homologous domains (I, II, III), each of which in turn may be divided into two subdomains (A and B). The 585-residue polypeptide chain folds into a mostly (67%) α-helical structure, with no β-sheets, and is stabilised by 17 disulfide bridges.

Ligand binding to HSA is usually limited to a few high-affinity sites. Many drugs bind to one or both of two well-characterised drug sites, one in subdomain IIA (drug site 1), and one in subdomain IIIA (drug site 2) (Sudlow*, et al.*, 1975, Sudlow, 1976). Drug site 1 typically accommodates bulky heterocyclic compounds such as bilirubin, indomethacin, warfarin, or β-lactam compounds such as benzylpenicillin and cephalosporins (Nerli & Pico, 1993). Drug site 2 has been implicated in the binding of smaller drug compounds such as diazepam, ibuprofen and diflunisal (Ghuman*, et al.*, 2005). Several other binding sites have also been reported.

HSA binds a range of medium-chain (C_{10}-C_{14}) and long-chain (C_{16} and longer) fatty acids in binding pockets that are asymmetrically distributed over the three homologous domains in the albumin molecule (Curry*, et al.*,

1998). Some of the fatty acid binding sites overlap with drug site 1 or 2 (Bhattacharya, Grüne*, et al.*, 2000). The binding pockets for fatty acids are predominantly apolar, allowing for burying of the hydrophobic "tails" of the fatty acid. The sites are accessible to solvent, allowing the carboxylate "head" of the fatty acid to be stabilised by hydrogen bonds or salt bridges with polar side-chains in HSA. A total of 11 distinct fatty acid binding locations have been identified in HSA (Bhattacharya, Grüne*, et al.*, 2000). Seven of these (denoted FA1 through FA7) bind both medium- and long-chain fatty acids, whereas the remaining four have only been observed to accommodate fatty acids of medium chain lengths.

Fatty acid binding to HSA is associated with a considerable change in the overall conformation of the albumin chain compared to the defatted conformation (Curry*, et al.*, 1998). Domains I and III both rotate relative to the central domain II. Although the global change in conformation is rather substantial, it is reported to be associated with only small changes in the local structures of drug sites 1 and 2 (Ghuman*, et al.*, 2005).

The HSA molecule has been shown to have one binding site for protein G, formed by residues 330-548 in domains II and III, and the binding site for the GA module was assumed to be identical (Falkenberg*, et al.*, 1992).

2.2 The HSA-GA complex **(Paper I)**

2.2.1 Background

The project on the HSA-GA complex was conducted in collaboration with the Department of Cell and Molecular Biology at Lund University and with the Structural Chemistry Laboratory at Biovitrum, Stockholm, Sweden. The Lund group provided lyophilised GA module purified from patient isolates of *F. magna* for structural studies.

2.2.2 Crystallisation, data collection and structure determination

Purified human serum albumin from plasma was obtained from Octapharma AG (Stockholm, Sweden), and was further purified with gel filtration to obtain a monodisperse sample. Purified HSA (100 mg/ml) was mixed directly with lyophilised GA module to obtain an approximate molar ratio of 1:1. The mixture was incubated at room temperature for at least 30 minutes before crystallisation trials were set up.

Crystals of the HSA-GA complex were obtained at 18°C with the hanging-drop vapour diffusion method, by mixing equal volumes of protein solution and a precipitant solution consisting of 2.2 M ammonium sulfate in 0.1 M citrate buffer, pH 6.0. Two different crystal morphologies were observed, one diamond-shaped and one rectangular prism-shaped (Figure 2). Initial

diffraction studies showed the latter to be the better-diffracting form, where the best crystals diffracted to approximately 3.5 Å resolution at room temperature. Low molecular weight PEG and various glycols were screened as cryoprotectants for data collection at 100 K, resulting in only minor improvements in diffraction quality. In the end, it was found that motor mineral oil (kindly provided by Anke Terwisscha van Scheltinga) could be used, resulting in diffraction to better than 3 Å resolution. Data collection on crystals soaked in mineral oil was performed at beamline ID14-4 at ESRF, Grenoble, with the best crystal yielding data to a resolution of 2.7 Å.

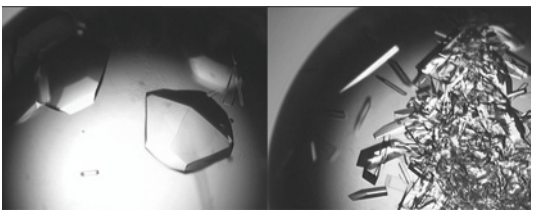

Figure 2. Crystals of the HSA-GA complex. Left: Poorly diffracting crystals. Right: Crystals used for data collection, belonging to space group $C222_1$.

After testing several HSA structures as search templates for molecular replacement, a partial solution for the complex structure was found using the apo-form of HSA as template (PDB accession code 1AO6). No satisfactory solution could be found using a fatty acid-containing HSA structure as the search template.

Clear positive difference electron density allowed manual placement of the NMR structure of the GA module (PDB accession code 1PRB) in the electron density maps. The resulting structure of the complex was subjected to simulated annealing in CNS (Brünger*, et al.*, 1998) to decrease model bias. Restrained refinement of the complex in REFMAC (Murshudov*, et al.*, 1997) was performed in combination with the use of TLS tensors to describe the anisotropic thermal motion of the albumin domains and the GA module (Winn*, et al.*, 2001, Winn*, et al.*, 2003).

2.2.3 The structure of the HSA-GA complex

The structure of the three-helix bundle of the HSA-bound GA module is highly similar to the NMR structure, apart from five N-terminal residues (Figure 3A). These residues were built as loop residues in the average NMR structure, but have been built to extend the first α-helix by one helical turn in the crystal structure. The structure of the HSA-bound GA module superposes onto the NMR structure (PDB accession code 1PRB) with an RMSD for 53 C_α atoms of 2.4 Å (calculated using least-squares superpositioning in LSQMAN (Kleywegt & Jones, 1994)). The overall similarity of the GA-bound albumin molecule to the structure of the apo-form of HSA is apparent

from superpositioning on the apo-HSA structure (PDB code 1AO6), which yields an RMSD of 0.88 Å for 556 C_α atoms. In contrast, superpositioning with a decanoate-containing HSA structure (PDB code 1E7E) yields an RMSD of 3.96 Å for 556 C_α atoms.

2.2.3.1 The interface and fatty acid binding

The GA module binds to the surface of domain II in HSA (Figure 3). The albumin part of the interface mainly involves residues 212-270 and 308-329 from the second to fourth α-helices in domain IIA and the first and second α-helix in domain IIB.

NMR studies have shown that the GA module will bind to HSA in the same single site as the albumin-binding domain from protein G (Johansson, *et al.*, 2002). Earlier binding studies localised the site to residues 330-548 on the albumin molecule; in light of the crystallographic results, this estimate requires adjustment. The crystal structure of the complex shows that binding to HSA involves residues in the GA module mainly from the second α-helix and the interhelical two loop regions, which confirms previous observations with NMR spectroscopy (Johansson, *et al.*, 2002).

Central to the binding interface is the accommodation of a phenylalanyl side-chain (Phe27) from the GA module in a hydrophobic cleft on the HSA surface, which is primarily formed by residues Phe309, Phe326 and Met329 (Figure 3B). The interaction is further stabilised by several surrounding hydrogen bonds.

After a few rounds of model rebuilding and restrained coordinate refinement, difference electron density also clearly indicated the presence of fatty acids that had copurified and were cocrystallised with the HSA-GA complex. By comparisons with previously solved structures of fatty-acid complexes of HSA, it was found that the electron density was consistent with the saturated ten-carbon fatty acid decanoate. Coordinates for the fatty acid were taken from PDB entry 1E7E, and refinement restraints were generated and inspected using the programs SKETCHER and LIBCHECK through the CCP4 interface (Collaborative Computational Project Number 4, 1994, Vagin, *et al.*, 1998). Three molecules of decanoate were subsequently modelled in fatty acid sites FA6 and FA7 (numbering according to Bhattacharya, *et al.*, 2000). Site FA6, which in this structure contains two decanoate molecules, is situated in close proximity to the HSA-GA interface. Although electron density in this site is too weak to support the modelling of a direct contact between one of the decanoate molecules and GA, there are indications that a decanoate molecule in FA6 might indirectly influence GA binding to HSA. A comparison with the apo-form of HSA suggested that the presence of a decanoate molecule in FA6 might cause the side-chain of Lys212 in HSA to rotate slightly and form a hydrogen bond with Glu47 in GA (Figure 3C).

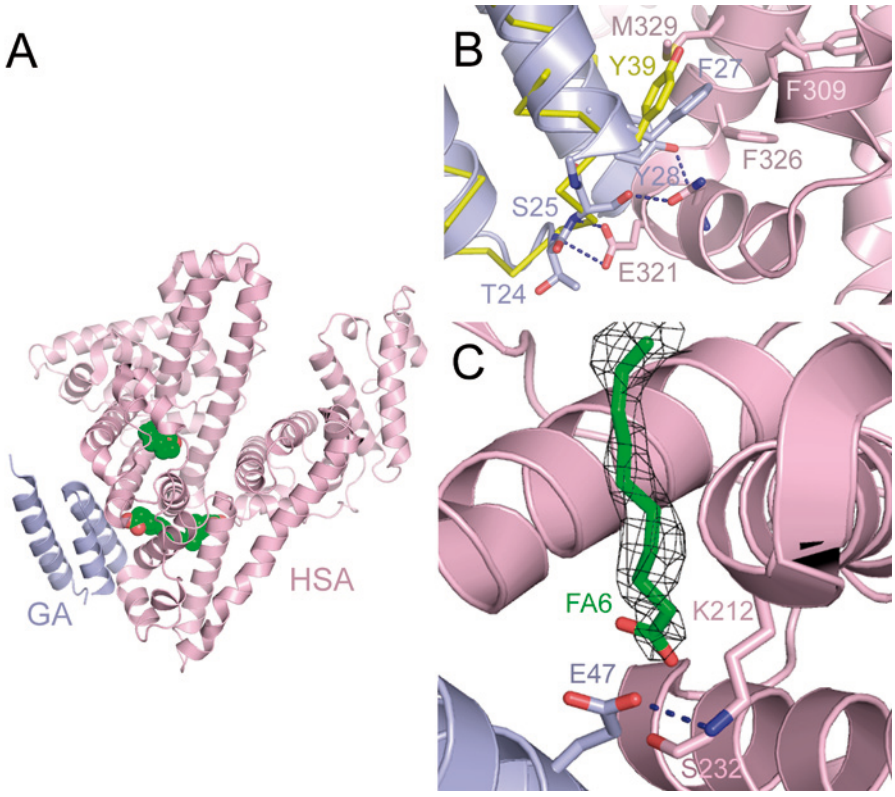

Figure 3. Structure of the HSA-GA complex. (A) Overall view, with the HSA chain in pink and the GA chain in light blue. Fatty acids are shown in green. (B) Selected residues at the HSA-GA interface. A superposed C_α trace (using least-squares superpositioning in O (Jones*, et al.*, 1991) of the albumin-binding domain from protein G (PDB code 1GJS) is shown in yellow. (C) Bound fatty acid near the interface. Electron density is shown as an omit map generated with coefficients $2mF_o$ - DF_c and phases calculated from the final model. The map is contoured at a level of 0.16 $e^-/Å^3$ (0.9σ).

2.2.3.2 Structural basis of bacterial host range

The close relation between the ABD from streptococcal protein G and the GA module from *F. magna* is reflected by an amino acid sequence identity of 59%. In addition, the two protein domains have been shown to bind to the same single site on HSA (Johansson, 2002). However, whereas group C/G streptococci have the capability of infecting a wide range of mammalian species, *F. magna* has a more restricted host range and mainly infects humans and other primates. This difference in host range displayed by the two bacteria is precisely reflected by the binding affinities of the streptococcal ABD and the GA module for albumins from different species (Johansson*, et al.*, 2002).

The hydrophobic interaction at the centre of the HSA-GA interface offers a structural basis for those observations. A sequence alignment of albumin from different mammalian species reveals variability in position 329 (Figure 4), where the primate albumins carry a residue with a non-polar side-chain (Met), whereas the mouse, rat, rabbit and bovine albumins all have polar (Ser, Thr, Lys) side-chains. Since this position is crucial for maintaining the hydrophobicity of the cleft at the HSA-GA interface, a substitution for a polar side-chain would lead to a disruption of the electrostatic complementarity of the interface.

Conversely, a superposition of the structure of the albumin-binding domain from streptococcal protein G (PDB codes 1GJS and 1GJT) reveals that it has a tyrosine residue instead of a phenylalanine in the position for interaction with the pocket in HSA domain II (Figure 3B). The polar hydroxyl group of the tyrosine is better suited for interactions with a polar side-chain in HSA position 329, as in the case of mouse, rabbit and bovine albumin, which explains the wider host range displayed by group C/G streptococci. The crystal structure of the HSA-GA complex presented here thus identifies the phenylalanine/tyrosine position to be an important determinant for the host range of bacteria expressing these albumin-binding modules. The central role of the phenylalanine/tyrosine position in the GA module has been further confirmed and characterised by other studies (He*, et al.*, 2007, He*, et al.*, 2006, Rozak*, et al.*, 2006).

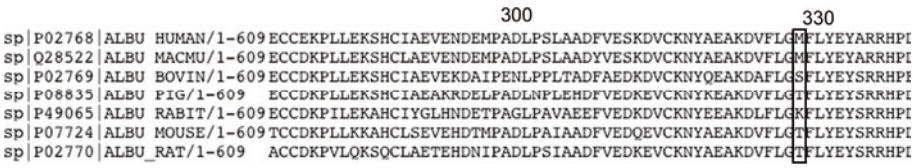

Figure 4. A section from a sequence alignment of serum albumins from various species (from top to bottom: human, macaque (93% amino acid sequence identity to HSA), bovine (75%), pig (75%), rabbit (74%), mouse (72%) and rat (73%)). Position 329 is indicated with a box.

2.3 The HSA-GA complex with naproxen **(Paper II)**

2.3.1 Background

After obtaining crystals of the HSA-GA complex as described in section 2.2.2, attempts were made to reproduce the crystals in order to continue the structural characterisation of the complex. These studies were started with a view to further investigate the influence of HSA-bound ligands (such as fatty acids and drug molecules) on binding of GA to HSA, and also to investigate

the possibility of utilising the HSA-GA complex for the study of HSA-ligand interactions.

The first crystals, described in section 2.2.2, had been obtained using freeze-dried material provided by the Lund group. This material came from a stock of GA module purified directly from patient isolates, a procedure which had probably resulted in a sample containing unknown amounts of salt and other constituents. Crystallisation of the complex using new batches of lyophilised GA from other stocks resulted in crystals of very poor quality. Dialysis of the new batches of GA protein was carried out, followed by addition of potassium or ammonium salts before setting up crystallisation trials. When these attempts did not yield better crystals, a new approach was taken. In order to gain control over crystallisation parameters, the open reading frame encoding the GA module (ALB8-GA) was instead obtained as synthetic DNA in a carrier plasmid, and subcloned into a heterologous expression vector (pET28) for subsequent overexpression in *Escherischia coli*. Cloning and refinement of the purification protocol were carried out by Peter Nordberg at the AstraZeneca Structural Chemistry Laboratory in Mölndal, Sweden. This group provided us with the plasmid for protein expression and crystallisation.

2.3.2 Crystallisation, data collection and structure determination

The plasmid encoding the GA module from *F. magna* (strain ALB8) was provided as an N-terminal histidine-tagged construct. Expression was performed in *E. coli* BL21(DE3) cells, and purification was performed using a Ni^{2+}-loaded chelating column and an imidazole gradient according to standard procedures (Nicholas*, et al.*, 1993), followed by a gel filtration step. Cleavage of the histidine tag was attempted, but due to difficulties obtaining a homogeneous cleavage product, this step was omitted for the crystallisation trials.

During subsequent crystallisation screens, it was found that reproducibility in terms of crystallising the HSA-GA complex required the inclusion of fatty acids in the preparation of the complexes. Fatty acids with methylene chain lengths of 10-14 were screened, and the best results were obtained with the ten-carbon saturated fatty acid decanoate (common name capric acid).

Crystals of the HSA-GA complex using HSA with decanoate and the new preparation of the GA module were obtained under different crystallisation conditions than those described in section 2.2.2. The new crystallisation solution contained 26-31% PEG 3350, 50 mM potassium phosphate (pH 7.5), and 0.1 M ammonium acetate. These crystals were found to belong to space group C2, and to withstand cryo-cooling when soaked in 20% glycerol. The crystals typically diffracted to a resolution of 2.5 Å or better. Stud-

ies were then undertaken to analyse the binding of fatty acids and a drug to this crystal form of the HSA-GA complex.

Naproxen (Figure 5) is a non-steroidal anti-inflammatory drug (NSAID), with high affinity for human serum albumin. The percentage of administered naproxen that binds to serum proteins is reported to be about 99.5%, which results in a low fraction of free drug in plasma (Honoré & Brodersen, 1984). High serum binding levels result in a small volume of distribution, and thereby low clearance. Consequently, secondary effects such as simultaneous binding of other ligands to HSA that cause displacement of naproxen, could have an extensive influence on the concentration of naproxen in plasma (Peters, 1995).

Studies employing equilibrium dialysis with human serum albumin immobilised in polyacrylamide microparticles have suggested that naproxen binds strongly to defatted HSA at one primary binding site, with an association constant (K_A) of 1.8 x 10^6 M^{-1} (which corresponds to a dissociation constant, K_D, of 5.6 x 10^{-7} M) (Kober & Sjöholm, 1980). The same study used drug competition assays to show that naproxen shares its primary binding site on HSA with ibuprofen and flurbiprofen. A crystal structure of defatted HSA with ibuprofen bound shows that the primary binding site for ibuprofen is drug site 2 (Ghuman*, et al.*, 2005). As yet, structural data of naproxen binding to HSA (with or without fatty acids present) have not been published.

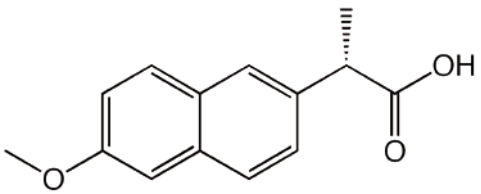

Figure 5. Structure of naproxen. Note that the stereocentre is in the S-configuration.

Complexes of HSA with the recombinantly expressed GA module crystallised readily in the presence of naproxen (crystals are shown in Figure 8B), and the crystals were found to belong be isomorphous with the HSA-GA crystals without naproxen. Exclusion of naproxen did not influence crystal formation, but decanoate was found to be required for crystallisation of the complex. A relatively high concentration of naproxen (molar ratio 6:1 with HSA) was used in order to overcome possible obstructing effects of the high (25-30% w/v) polyethylene glycol concentration. The crystals were harvested into a solution containing reservoir solution with 10 mM naproxen and 20% glycerol prior to cryo-cooling in liquid nitrogen.

Data were collected to a resolution of 2.5 Å. Molecular replacement trials using the previously determined structure of the HSA-GA complex (described in section 2.2) as a single search model did not yield a solution. Trials using the three HSA domains and the GA chain as separate search mod-

els only resulted in a satisfactory solution for the GA part. A molecular replacement solution for the entire complex was found by using a structure of HSA with bound decanoate (PDB accession code 1E7E) and the GA chain from the HSA-GA complex (PDB accession code 1TF0) as two separate search models.

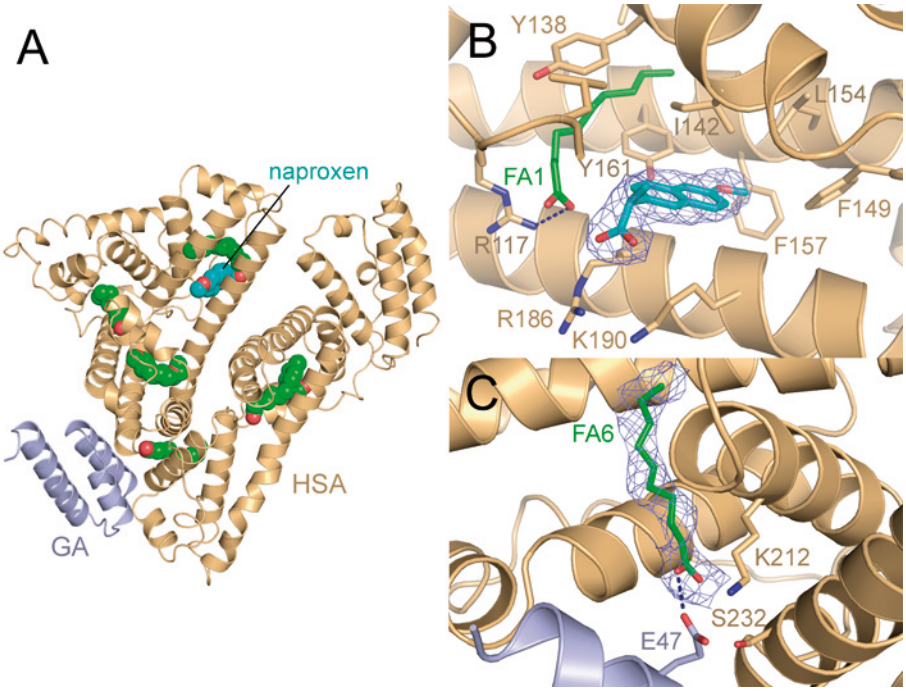

Figure 6. The HSA-GA-naproxen complex. (A) Overall view. The HSA chain is shown in light orange, and the GA module in light blue. Naproxen and fatty acids are shown as spheres (cyan and green, respectively). Electron density for ligands is shown as omit maps generated with coefficients $2mF_o - DF_c$ and phases calculated from the final model. The maps are contoured at 0.21 $e^-/Å^3$ (1.0σ). (B) The naproxen site and the adjacent fatty acid site 1. Naproxen is shown in cyan, and fatty acids in green. Residues 104-114 in HSA have been removed for clarity. (C) The decanoate molecule in FA6 near the HSA-GA interface.

2.3.3 The structure of the HSA-GA-naproxen complex

The crystal structure of the complex formed by HSA and GA cocrystallised with decanoate and naproxen (here denoted the HSA-GA-naproxen complex) is shown in Figure 6.

2.3.3.1 Binding of GA module and fatty acids

The GA module binds in the same site on HSA as reported previously (section 2.2.3). Similarly to the previously determined structure of the HSA-GA complex, decanoate is found to bind in the fatty acid binding site FA6 in domain II close to the HSA-GA interface (Figure 6C).

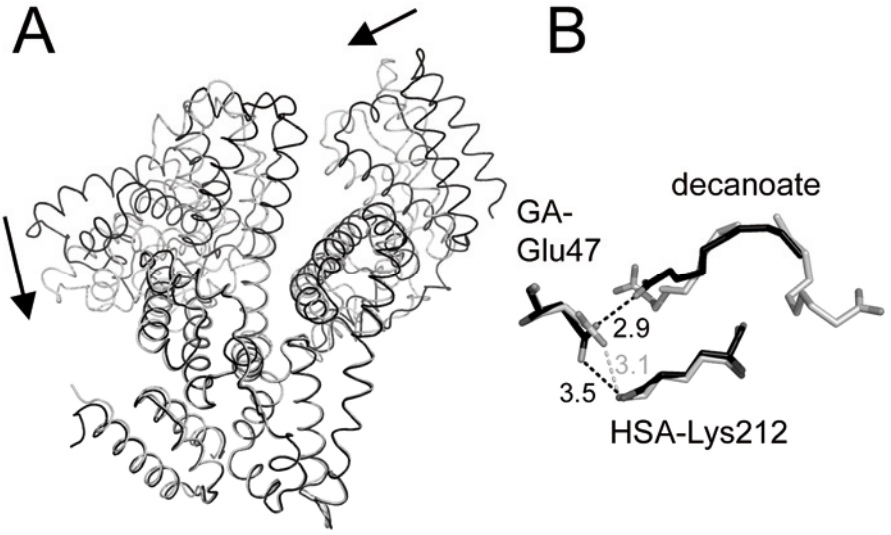

Figure 7. Superposition of the C_α traces of the two structures of the HSA-GA complex. (A) 187 C_α atoms in domain II of the albumin chain (residues 197-383) were superposed using least-squares superpositioning in O (Jones*, et al.*, 1991). The HSA-GA complex is shown in grey, and the HSA-GA-naproxen complex in black. Rotations of domains I and III are indicated (left and right arrow, respectively). (B) Close-up of the fatty acid binding site near the HSA-GA interface (FA6). Residues in the HSA-GA complex are shown in grey, and residues from the HSA-GA-naproxen complex are shown in black. Interatomic distances are shown in Å, and are indicated as dashed lines.

In contrast to the structure of the HSA-GA complex, however, electron density indicates the presence of only one molecule of decanoate in this site. The binding mode for the fatty acid in the HSA-GA-naproxen complex differs from the binding mode observed in the structure of the HSA-GA complex; it binds slightly further into the fatty acid pocket. As a consequence of the altered binding mode, interactions with neighbouring residues in HSA also change. Compared to the indirect interaction with GA observed in the structure of the HSA-GA complex, electron density for the HSA-GA-naproxen complex indicates that the interaction between the fatty acid and GA is of a more direct nature.

The closest interatomic distance between Glu47 in the GA chain and the carboxylate moiety of the fatty acid in site 6 is 2.9 Å (Figure 7B) in the HSA-GA-naproxen complex, suggesting these groups could form a hydrogen bond. Electron density for the HSA-GA-naproxen complex also shows a difference in the total number of fatty acids bound to albumin compared to the structure of the HSA-GA complex. The albumin part of the HSA-GA-naproxen complex accommodates a total of six fatty acids, as opposed to three in the structure of the HSA-GA complex. As a consequence, the HSA chain in the HSA-GA-naproxen complex shows higher structural similarity to a fatty acid-containing albumin structure than to the albumin chain in the structure of the HSA-GA complex (Figure 7A). Pairwise structural alignments by least-squares superpositioning of 555 C_α atoms yields RMSDs of 1.59 Å (comparison with HSA-decanoate complex, PDB code 1E7E) and 4.07 Å (comparison with HSA-GA complex, PDB code 1TF0), respectively. The addition of decanoate prior to crystallisation thus seems to lead to higher fatty acid occupancy in the fatty acid sites on the albumin molecule, which in turn drives a conformational change of the HSA chain, in the same manner as has been reported for previously published HSA structures (Curry*, et al.*, 1999, Curry*, et al.*, 1998). Despite the alteration in global conformation of the HSA chain, where domains I and III both rotate relative to domain II, the local structure of domain II remains comparatively unperturbed. The conserved local structure of domain II allows the GA module to form the same interactions with HSA as observed for the albumin molecule with the lower fatty acid occupancy.

2.3.3.2 Binding of naproxen

In the HSA-GA-naproxen structure, electron density consistent with one molecule of naproxen was found in a pocket in subdomain IB in the HSA molecule (Figure 6B). No evidence was found for naproxen binding in drug site 1 or drug site 2. Coordinates for the naproxen molecule were generated with the program SKETCHER, and refinement restraints were obtained with LIBCHECK (Collaborative Computational Project Number 4, 1994, Vagin*, et al.*, 1998). The binding pocket that accommodates naproxen has previously been described as the primary binding site for hemin (Zunszain*, et al.*, 2003), and as the secondary binding site for the two NSAID compounds azapropazone and indomethacin, and for the aspirin analogue triiodobenzoate (Curry*, et al.*, 1998, Ghuman*, et al.*, 2005). In the cited studies, binding of azapropazone, indomethacin and triiodobenzoate to the hemin site was only observed with one molecule of the 14-carbon fatty acid myristate simultaneously bound to HSA in a neighbouring binding compartment, denoted FA1 (fatty acid site 1).

Similarly, the HSA-GA-naproxen structure shows that naproxen binds to the hemin site with one molecule of decanoate simultaneously binding in FA1. Indeed, the hydrocarbon “tail” of the fatty acid seems to be involved in

the clustering of hydrophobic groups together with aliphatic side-chains lining the walls of the pocket. There are no hydrogen bonds anchoring the naproxen molecule in this site; the binding seems mainly governed by hydrophobic interactions.

2.3.3.3 Comparisons with previous studies of HSA-naproxen interactions

Earlier binding studies of naproxen and HSA indicated binding of naproxen to both drug sites 1 and 2 (Sjöholm*, et al.*, 1979), and also showed naproxen to bind with high affinity to the same site as ibuprofen (Kober & Sjöholm, 1980). The cited studies were mainly conducted in the absence of fatty acids.

The hemin site has not previously been implicated in the binding of naproxen to HSA. However, a published NMR study of the interactions between fatty acids and drugs binding simultaneously to HSA suggested that both site 1 and site 2 drugs may be displaced by fatty acids (Simard*, et al.*, 2006). Thus, the availability of drug sites 1 and 2 to accommodate naproxen may be affected by the presence of competing fatty acids. This suggestion agrees with our observation that the addition of an excess of naproxen to the HSA-GA-naproxen crystals before data collection did not result in the displacement of fatty acids from their respective binding sites in drug sites 1 and 2.

It has been suggested that FA6 might be a secondary binding site for ibuprofen (Ghuman*, et al.*, 2005). It is therefore possible that FA6 would also be a secondary binding site for naproxen. In the studies of the HSA-GA complex presented in the above sections, it has been observed that a fatty acid occupying FA6 might play an important role for GA binding *in vivo*, and such a situation would consequently preclude naproxen binding in this site.

In conclusion, the most plausible scenario is that several binding sites on the HSA molecule are available to naproxen, including at least drug sites 1 and 2 and the hemin site. The availability of each site will be influenced by the level of fatty acid binding, and possibly also by the binding of other components, such as the GA module.

2.4 The crystal structure of the GA module **(Paper III)**

2.4.1 Crystallisation, data collection and structure determination

During screens for improving the HSA-GA crystals (original protocol for crystallisation described in section 2.3.2), it was found that particularly well-formed crystals appeared in drops with lower (22-26%) PEG concentration (Figure 8A). These new crystals diffracted X-rays to a resolution of 1.4 Å

when cryo-cooled in liquid nitrogen and cryo-protected by 20% glycerol in reservoir solution.

Data collection was carried out at beamline ID23-2 at ESRF, Grenoble. During data processing, it became apparent that the crystals were different, in several aspects, from those previously obtained – the unit cell was drastically smaller, and the space group had changed from monoclinic C-centred C2 to monoclinic primitive $P2_1$. During structure determination, it became apparent that the crystals were of the GA module only, i.e. without HSA. Crystallisation of the GA module in isolation had occurred despite HSA being present in the drop at a concentration of at least 100 mg/ml, and despite that no other parameters for crystallisation of the HSA-GA complex had been knowingly changed except the PEG concentration.

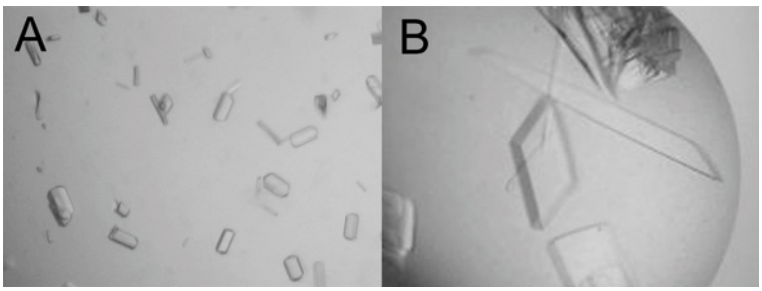

Figure 8. (A) Crystals of the GA module. (B) Crystals of HSA in complex with the recombinantly expressed GA module. These crystals were obtained with naproxen bound to HSA.

The structure of the GA module was solved by molecular replacement, with the GA chain from the HSA-GA complex structure (section 2.2; PDB accession code 1TF0, chain B) as the search template. Simulated annealing in CNS was followed by restrained refinement in REFMAC, which included refinement of individual anisotropic atomic displacement parameters.

2.4.2 The structure of the GA module

2.4.2.1 Overall structure

The structure of the GA module, which has been refined using data to a resolution of 1.4 Å, shows that the GA module crystallises as an asymmetric dimer. One monomer is positioned at an angle of approximately 25° to the other (Figure 9A), and the dimer interface is formed by residues from helix 2 and 3 of one chain (chain A) and residues from helix 1 of the other (chain B).

The overall topology of the GA module chains is highly similar to the previously described structure of the GA module in complex with human serum albumin. Least-squares superpositioning of 53 C_α atoms in the two different structures yields an RMSD of 1.00 Å. Only minor side-chain rota-

tions can be observed. The model includes, in addition to the 53 residues of the GA module, eight residues from the N-terminal histidine-tag.

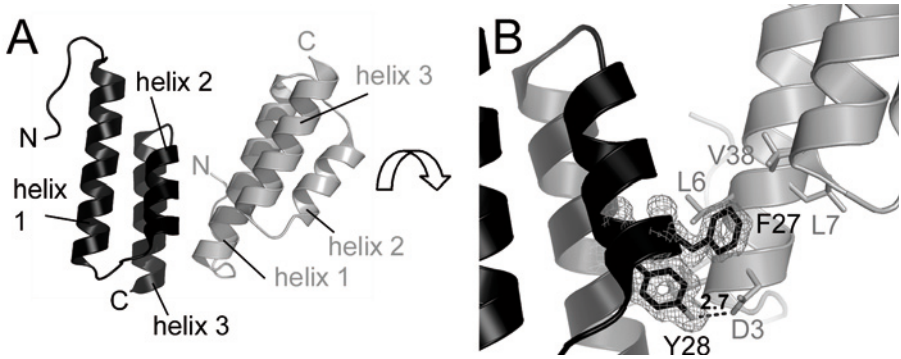

Figure 9. The crystal structure of the GA module. (A) The dimer in the asymmetric unit. Chain A in black and chain B in grey. (B) The dimer interface. Interface residues are shown as sticks. Electron density shown as an omit map generated with coefficients $2mF_o$ - DF_c and phases calculated from the final model. The map is contoured at a level of 0.41 $e^-/Å^3$ (1.0σ). A hydrogen bond is indicated as a dashed line, with the interatomic distance in Å.

2.4.2.2 Dimer interactions

The most extensive interface between the GA monomers in the crystal lattice involves interactions of Phe27 from chain A with apolar side-chains in chain B (Figure 9B). Other interactions in the lattice are different, and also less extensive. The interaction involving Phe27 is flanked by neighbouring hydrogen bonds. In addition, Glu47 from chain A participates in a salt bridge with His-tag residue Arg(-4) in chain B. The phenylalanine residue Phe27, the neighbouring hydrogen bond-forming Tyr28, and the salt bridge residue Glu47 are the same residues that were central for the interactions at interface between the GA module and albumin, described in section 2.2.3.1. In this context, it is also worth pointing out that although the interaction between Glu47 and a residue in the histidine tag of chain B (Arg(-4)) may at first seem artefactual, the fact that the native PAB polypeptide chain has a lysine residue at position –4 relative to the GA module may indicate that it is not necessarily so.

The occurrence of a GA dimer in the crystal lattice raised the question of whether the GA module would also be able to form dimeric interactions in solution. Although only circumstantial evidence for multiple binding sites in the GA module has been reported previously (Johansson*, et al.*, 2002), the crystallographic observation is underpinned by dynamic light scattering measurements indicating the presence of higher order oligomeric species in solution (see Table 2 in paper III, p. 3180). The results indicate that the dominant species in this solution is the tetrameric form of the GA module.

However, the 0.4 nm polydispersity parameter is 15% of the hydrodynamic radius, and this high value is the upper limit for which monodispersity may still be assumed. The sample may thus contain other oligomeric species (e.g. dimers and monomers) at low concentrations. An interpretation of the results from this study is that there might be transient associations between individual GA monomers in solution.

The binding surface in the dimeric GA module constituted by residues in chain B was not implicated in GA binding to HSA in the crystal structures of the HSA-GA complex (see sections 2.2.3 and 2.3.3). Nevertheless, it has been suggested in previous studies of the GA module that residues in this region may play an important role in facilitating efficient binding of the GA module to HSA (Johansson*, et al.*, 2002). In the cited study, an N-terminally truncated GA module that lacked the six N-terminal residues was observed to bind to the same site on HSA, but with 1000-fold lower efficiency, compared with the full-length GA module. This result suggests that the N-terminus may play a role in HSA-GA binding. Although this role has not yet been investigated, there is a possibility that the interactions that were observed in the crystal structure of the GA dimer may echo associations in solution that aid in the binding of the GA module to HSA.

2.5 Conclusions

2.5.1 The binding properties of the GA module

The studies of the GA module presented here have shown the domain in two binding modes; one where it binds to HSA, and one where it binds to itself. The protein-protein interactions in the two crystal forms and the presence of oligomeric states of the GA module in solution could indicate that the GA module is "sticky" and that it might be capable of forming contacts with a range of protein surfaces. Interestingly, the surface of the GA chain that is involved in binding to HSA is the same that is involved in binding to another GA chain to form a GA dimer. The hydrophobic features displayed by the other chain in the GA dimer are reminiscent of the hydrophobic core at the centre of the HSA-GA interface, although the dimer interactions are less extensive. Nevertheless, it is plausible that a transient association between GA chains could be present in solution.

In terms of bacterial pathogenesis, a tendency of the GA module from *F. magna* to engage in protein-protein interactions with a conformationally variable serum protein molecule such as HSA, and with a range of protein surfaces, could be beneficial. This tendency may allow for a certain degree of conformational freedom on the part of the protein that the GA module interacts with, which might lead to a higher likelihood that the bacterium will find a suitable interaction partner in a human host. It is also worth point-

ing out here that previous suggestions that fatty acids or other molecules carried by the albumin molecule could serve as nutrients for growing bacteria (de Château, *et al.*, 1996, Falkenberg, *et al.*, 1992) are strengthened by the observation of a fatty acid near the albumin-GA interface. Such an interaction between the bacterium and HSA could stimulate bacterial growth and thus confer additional advantages on the invading pathogen, apart from physical envelopment.

The central interactions at the interface between the GA module and HSA have been further investigated and characterised by another research group (He, *et al.*, 2007, He, *et al.*, 2006). In the course of their studies, it has emerged that the different host group specificities of various GA modules indeed may be explained by polymorphisms at positions 27 and 47; the experiments have also suggested a role for other polymorphisms at positions 26, 32 and 41. The structural characterisation of HSA-bound GA modules from other species, including group C/G streptococci, would shed more light on the role these interactions play in defining the host specificities of the respective pathogens.

2.5.2 HSA in drug development

HSA plays a central role in medicine for multiple reasons. Apart from studies exploring HSA as a possible oxygen carrier in plasma due to its ability to bind hemin (Wardell, *et al.*, 2002, Zunszain, *et al.*, 2003) or as a drug delivery system when coadministered with a drug compound (Langer, *et al.*, 2007, Stinchcombe, 2007), studies using structural data to characterise and overcome the problem of high affinity binding of promising lead compounds in drug development projects are also of scientific interest (Colmenarejo, *et al.*, 2001, Gunturi, *et al.*, 2006, Saiakhov, *et al.*, 2000). The results presented in this thesis underline previous reports of the influence of fatty acids on the conformation of HSA. In the studies presented here, two different conformations of albumin have been observed; one where three molecules of fatty acid are bound, and one where six fatty acids are bound. In the case of three bound fatty acids, fatty acid site FA2 was unoccupied, and in the case of six bound fatty acids, FA2 was occupied. The results are thereby consistent with previous observations that the accommodation of a fatty acid in FA2 seems to be the chief cause of the conformational change in the albumin chain when defatted albumin binds fatty acids (Curry, *et al.*, 1999, Curry, *et al.*, 1998).

Given that HSA contains on average 0.1-2 mol of fatty acids per mol of HSA *in vivo*, studies of ligand binding to HSA may be more relevant physiologically in the presence of fatty acids than in a completely defatted HSA molecule. (Even so, characterisation of binding sites in defatted HSA may be valuable in terms of adding information for a comprehensive mapping of the interactions between HSA and a given ligand.)

It may be argued that a situation where six molecules of fatty acid have bound to HSA is outside the average range for extracellular fluids in normal human circulation; but in fact there have been reports of chronic elevation of fatty acid levels associated with obesity and insulin-treated diabetes mellitus (Reitsma, 1967). Acute elevations of fatty acid levels could also be caused by untreated diabetes, periods of emotional stress or by Gram-negative infections, where the fatty acid:HSA molar ratio could increase to up to 6:1 (Cistola & Small, 1991).

In summary, the results presented here indicate that the readily available HSA-GA crystals may offer a new route to the crystallographic study of HSA-ligand interactions, and that such studies may also make the analysis of the effects of fatty acids on drug binding more tractable. Since the GA module is able to bind to HSA with varying levels of fatty acid bound, there is a possibility for utilising the HSA-GA complex for the study of ligand interactions with HSA and simultaneously include the effect of fatty acid binding. This will require screening for conditions that reproducibly result in a certain number of fatty acid molecules bound to HSA, with one to six probably being the relevant numbers. Additional studies to investigate the consistency of drug binding to HSA with and without the GA module, and using ligands that have previously been characterised crystallographically in terms of binding to HSA will also be valuable.

3. Cephalosporin biosynthesis

In the previous chapter, the structural basis for a bacterial immune evasion strategy was described. This strategy enables certain pathogenic bacteria to invade and infect human tissues. Antibiotics allow the counterattack against pathogenic bacteria, and in the following sections, structural studies of a cephalosporin biosynthetic enzyme are presented. Deacetylcephalosporin C acetyltransferase catalyses the last step in the biosynthesis of the β-lactam antibiotic cephalosporin C in the filamentous fungus *Acremonium chrysogenum*.

3.1 Background

3.1.1 Structure of β-lactam antibiotics

The common core of all β-lactam antibiotics is the four-membered β-lactam ring (Figure 10). The two most widely used classes of β-lactam antibiotics, the penicillins and the cephalosporins, both have a second ring fused to the β-lactam ring: a five-membered thiazolidine ring in the penicillin compounds, and a six-membered dihydrothiazine ring in the cephalosporin compounds. The β-lactam ring forms the basis for the cytolytic activity of the compounds by acting as a substrate analogue of the bacterial carboxypeptidases and transpeptidases that are involved in the biosynthesis and metabolism of the peptidoglycan cell wall layer. The biochemically inert penicillin or cephalosporin compound halts cell wall synthesis, and thereby undermines the structural integrity of the bacterium, which ultimately leads to bacterial cell lysis.

3.1.2 Development of cephalosporin antibiotics

The first cephalosporin compounds were isolated from a culture taken from a sewage outlet in Sardinia by Giuseppe Brotzu (Brotzu, 1948), who was Professor of Hygiene and Rector of the University of Cagliari, while at the same time fulfilling his roles as mayor of the city of Cagliari and president of the Sardinian government. One of the antimicrobial substances in the mixture was named cephalosporin C after the producer species, a filamentous fungus named *Cephalosporium acremonium*, now known as *Acremonium chrysogenum*.

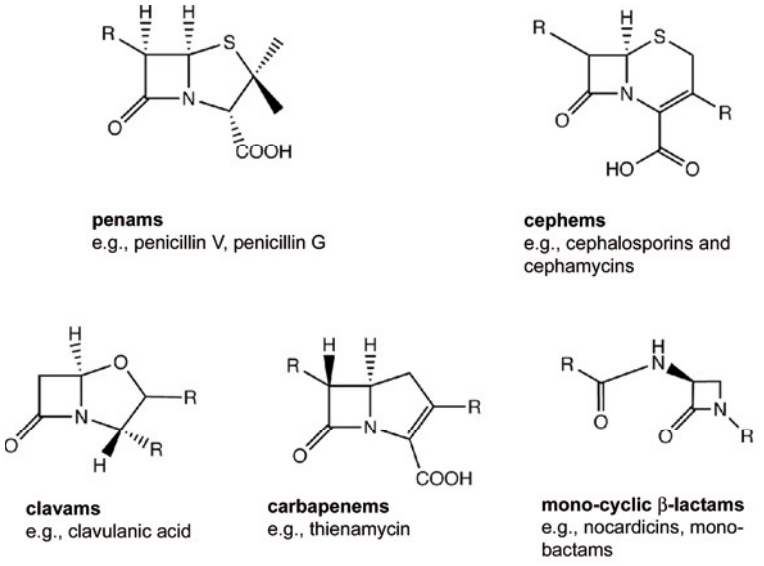

Figure 10. Some examples of structural classes of β-lactam antibiotics (see (Hamilton-Miller, 1999) and references therein). The variable chemical groups are denoted R.

The initial batches of cephalosporin C, which were isolated and purified by researchers at the University of Oxford, showed only weak antibacterial properties. However, due to its greater stability to β-lactamases and to staphylococcal penicillinases compared with penicillin compounds, and also due to its low toxicity to humans, cephalosporin C was deemed an interesting starting point for developing new antibiotic compounds.

The structure of cephalosporin C was solved in 1961 by Dorothy Crowfoot Hodgkin and Edward N. Maslen (Hodgkin & Maslen, 1961). The structure, which was determined using X-ray crystallography, showed the molecule to contain a bicyclic penicillin-like core, the so-called cephem nucleus. Cephalosporin C consists of 7-aminocephalosporanic acid (7-ACA), with two side-chains; L-γ-glutamoyl (R1) and an acetyl ester (R2) (Figure 11). By chemical modification of these side-chains, it was later possible to obtain compounds with a marked increase in clinical usefulness, and the first semi-synthetic cephalosporin agent, cephalothin, appeared on the market in 1964 (Boniece*, et al.*, 1962).

The first generation of cephalosporins, which had primary action against staphylococci and Gram-positive bacteria, was soon followed by new and improved compounds. These later generations of semi-synthetic cephalosporins have all been generated by chemical and/or enzymatic synthesis to acquire higher resistance to β-lactamases and also a wider spectrum of anti-

bacterial activity. The newer generations have significantly higher activity against Gram-negative pathogens (particularly against *Pseudomonas aeruginosa*) than the first and second generations of cephalosporins. With over 50 marketed compounds, the cephalosporins currently represent the biggest selling class of β-lactam antibiotics worldwide (Elander, 2003).

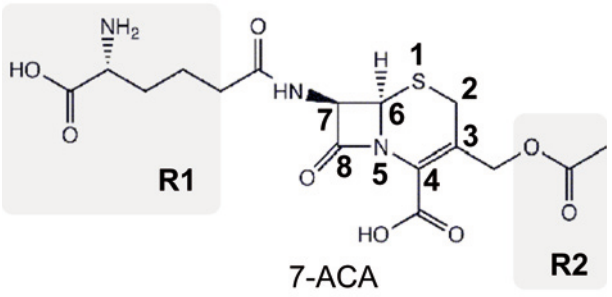

Figure 11. Structure of cephalosporin C.

3.1.3 Cephalosporin synthesis in *Acremonium chrysogenum*

In the last couple of decades, there have been several reports of detailed analyses of the biosynthetic pathways of penicillins, cephalosporins and other β-lactam antibiotics, including X-ray crystallographic studies of the enzymes in these pathways. The ultimate goal of such studies is to utilise structural and biochemical data to find new routes to synthesising new generations of β-lactam-based antibiotics, with improved antimicrobial properties (Andersson*, et al.*, 2001, Demain & Elander, 1999).

3.1.3.1 The cephalosporin C gene cluster

The ability to produce β-lactam compounds is widespread in nature – it has been identified not only in fungal species (e.g. *Aspergillus (Emericella) nidulans*, *Acremonium chrysogenum* and *Penicillium chrysogenum*), but also in some Gram-positive and Gram-negative bacteria (e.g. *Streptomyces clavuligerus*) (Brakhage*, et al.*, 2005). In all species characterised so far, the structural genes for β-lactam biosynthesis are organised as clusters. In *A. chrysogenum*, the genes for the biosynthesis of cephalosporin C are organised in two clusters. The early biosynthetic genes (*pcbAB*, *pcbC*, *cefD1* and *cefD2*), which are expressed for the first three biosynthetic steps, are located on chromosome VII or II, whereas the late genes (*cefEF* and *cefG*), which are expressed for the last two steps, are located on chromosome I or VI.

3.1.3.2 The biosynthetic pathway of cephalosporin C

All naturally occurring penicillin and cephalosporin compounds produced by bacteria or fungi are synthesised from the same three precursors, and have

the first two biosynthetic steps in common. The biosynthetic pathway of cephalosporin C in *A. chrysogenum* is depicted in Figure 12.

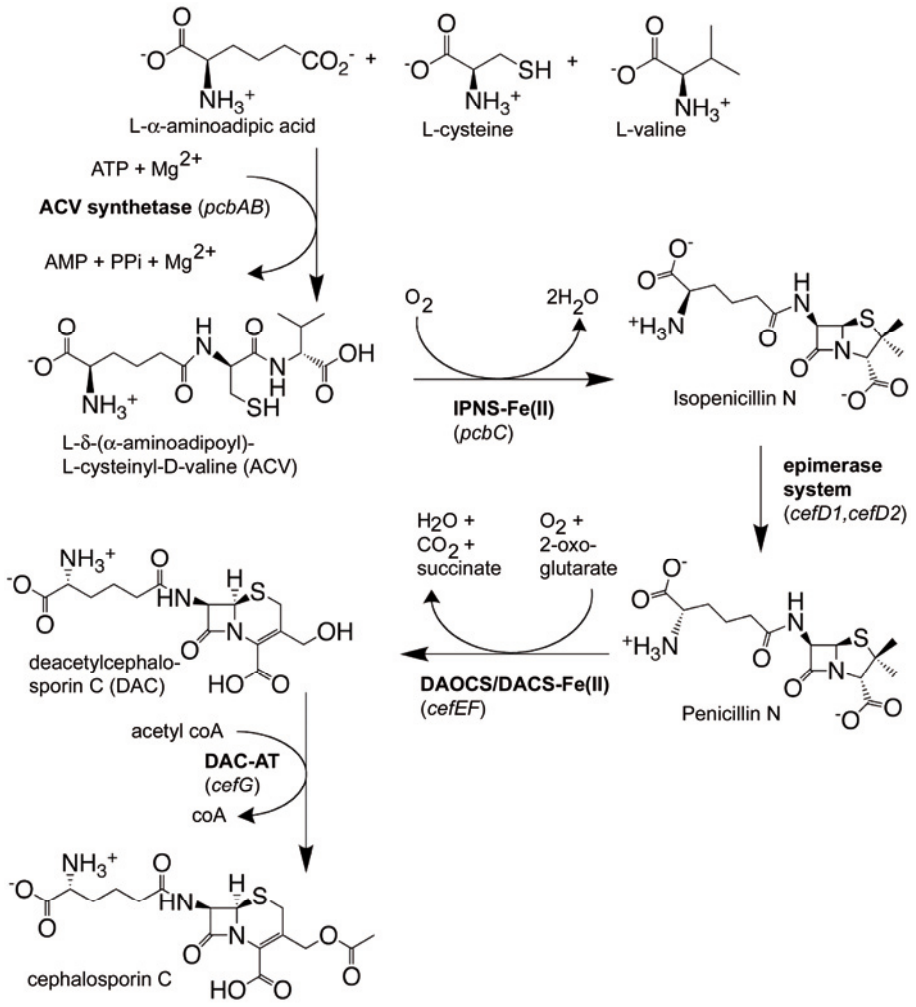

Figure 12. The biosynthetic pathway of cephalosporin C in *A. chrysogenum*. Gene names are given in italics.

The first reaction in the biosynthesis of cephalosporins and penicillins is the condensation of three amino acids; L-α-aminoadipic acid, L-cysteine and L-valine, into the tripeptide ACV. This reaction is catalysed by the enzyme ACV synthetase (Byford*, et al.*, 1997). The enzyme IPNS (Roach*, et al.*, 1997) then catalyses the oxidative ring closure of ACV to yield the five-membered thiazolidine ring of isopenicillin N (IPN). IPN constitutes the branch-point intermediate between penicillin G and cephalosporin C biosynthesis, and is also the first bioactive compound in the biosynthetic pathway.

The subsequent isomerisation (epimerisation) of the L-amino group of the aminoadipic acid moiety of IPN into the D isomer, is catalysed by the two enzymes encoded by genes *cefD1* and *cefD2* (Ullán*, et al.*, 2002). This step constitutes the committing step for cephalosporin biosynthesis. The resulting compound, penicillin N, then undergoes expansion of the thiazolidine ring into the six-membered dihydrothiazine ring of the cephem nucleus; this reaction is catalysed by the bifunctional enzyme synthase/DAC hydroxylase and gives the intermediate DAOC (deacetoxycephalosporin C). The second function of this enzyme is the oxidation of the C-3 methyl group of DAOC to yield the intermediate deacetylcephalosporin C, DAC. The final biosynthetic step, which is catalysed by deacetylcephalosporin C acetyltransferase (DAC-AT), involves the transfer of an acetyl group from the cofactor acetyl coenzyme A to the hydroxyl group of DAC to form the final product, cephalosporin C (Figure 13).

It is the last step of cephalosporin C biosynthesis that is the focus for this chapter, and a presentation of the crystal structure of DAC-AT and the implications for its catalytic mechanism will follow below.

3.2 The crystal structure of deacetylcephalosporin C acetyltransferase (**Paper IV**)

3.2.1 Background

The molecular weight, polypeptide length and subunit structure of the enzyme deacetylcephalosporin C acetyltransferase (DAC-AT) from *A. chrysogenum* have in the past been the subjects of some controversy. The *cefG* gene was cloned independently by three different groups in the years 1992 to 1993 (Gutiérrez*, et al.*, 1992, Mathison*, et al.*, 1993, Matsuda*, et al.*, 1992). All three groups used different strains of *A. chrysogenum*. Each group then proposed a different native molecular weight and polypeptide length for the resulting protein, based on varying interpretations of the three in-frame ATG start codons in the open reading frame. Gutiérrez and coworkers proposed a 444 amino acid long polypeptide, translated from the first start codon, whereas Mathison and coworkers and Matsuda and coworkers proposed shorter polypeptide chains (399 and 385 amino acids, respectively), translated from the second and third in-frame ATG codons. Matsuda and coworkers also proposed the 385 aa protein would be post-translationally processed into two subunits of 27 kDa (247 aa) and 14 kDa (126 aa), but that the active protein has a molecular weight of 55 kDa. It was later shown that all three suggested polypeptides displayed DAC-AT activity, with the second start codon resulting in the DAC-AT protein with the highest activity level (Velasco*, et al.*, 1999). The minimal polypeptide required for DAC-AT ac-

tivity is represented by the 385 amino-acid enzyme, which is expressed by *A. chrysogenum* strain IS-5 (Brakhage, 1998).

The *cefG* gene is expressed as a monocistronic 1.4 kb mRNA transcript containing two introns. The gene is expressed in the late stages of fermentation, reaching a peak after 72-96 hours (Velasco*, et al.*, 1994). It shares transcriptional regulation from the bidirectional intergenic promoter region with the *cefEF* gene, but seems to be under a different control mechanism (Gutiérrez*, et al.*, 1992). Expression of *cefG* is also set apart from the other structural genes in cephalosporin C biosynthesis in that it is not under glucose-dependent regulation, and does not respond to supplementation of methionine in the same way as the other enzymes in the pathway (Martín & Demain, 2002, Velasco*, et al.*, 1994). Studies of cephalosporin C biosynthesis in *A. chrysogenum* have unambiguously identified the acetyl transferase step to be rate-limiting (Gutiérrez*, et al.*, 1997, Mathison*, et al.*, 1993). By transforming production strains with up to five additional copies of the *cefG* gene, up to a three-fold increase in cephalosporin C titers could be obtained (Mathison*, et al.*, 1993).

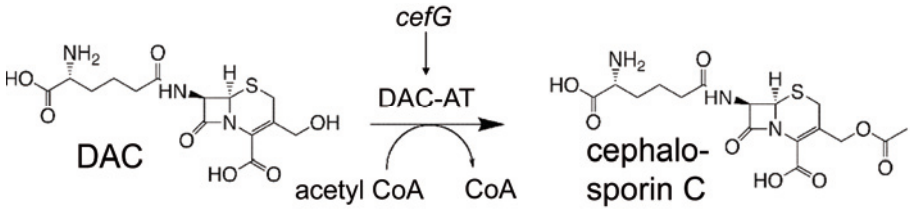

Figure 13. Reaction catalysed by DAC-AT. The acetyl group in acetyl coenzyme A is transferred to the C-3 hydroxymethyl group in the substrate DAC.

Sequence comparisons have revealed that DAC-AT shows similarity to homoserine O-transacetylases (HTAs) from bacterial and fungal species (Gutiérrez*, et al.*, 1992). The amino acid sequence identity is usually in the order of 40-50%, with the highest level of identity (55%) reported for species including *Aspergillus fumigatus*, *Neurospora crassa*, and *Schizosaccharomyces pombe*. So far, two crystal structures of the HTA family of enzymes have been reported: HTA from *Haemophilus influenzae* (PDB accession code 2B61, here denoted *Hi*HTA) and HTA from *Leptospira interrogans* (PDB accession code 2PL5, here denoted *LiHTA*). DAC-AT shares 31% and 32% amino acid sequence identity with *Hi*HTA and *Li*HTA, respectively.

Our group commenced work on DAC-AT in 1999 as a collaboration with Dr Jacqueline Ellis at the University of Leicester. The cDNA construct, which was kindly donated by Mathison and coworkers (Panlabs Inc), consists of the last 385 residues of the DAC-AT UniProt entry P39058, plus a 15

amino acid long N-terminal protease cleavage tag. This construct corresponds to the DAC-AT enzyme encoded from the third in-frame start codon.

3.2.2 Expression, purification and crystallisation of DAC-AT

Our initial efforts to obtain protein material for structural studies of DAC-AT were significantly hampered by low expression levels. Also, a propensity of DAC-AT to form inclusion bodies during IPTG-induced overexpression in complex expression media constituted a severe bottleneck in the optimisation of DAC-AT expression. At this point, only a low-level overexpression protocol would result in workable protein material. This protocol involved using leaky expression by non-induced cells transformed with the pTWIN vector (a construct with a chitin-binding domain and intein fused with a T7 promoter) at low temperatures (15-20°C). Even so, not more than 0.5-1 mg of pure protein could usually be obtained per litre of culture.

Addition of the cofactor acetyl coenzyme A to the protein solution was required for crystallisation. Small clusters of needle-like crystals were obtained in sitting drops after screening for a suitable pH with imidazole at pH 8.0. Higher ionic strength increased solubility, and after fine-tuning with sodium chloride and sodium acetate, larger needle clusters could be observed in conjunction with large granular precipitates. These crystals typically diffracted to around 2.5 Å, although data processing was usually difficult due to the relatively high mosaicity (0.8-1.2°) of the crystals.

In the course of attempting to obtain selenomethionyl protein by a metabolic inhibition protocol in a minimal expression medium (van Duyne*, et al.*, 1993) it was noticed that induction of overexpression by addition of IPTG did not produce inclusion bodies to the same extent as in the complex medium. In the same manner, when production of native protein was performed in minimal medium, and L-methionine was added instead of L-selenomethionine, the result was a 4-5-fold increase in soluble protein material compared to the leaky expression protocol in complex medium. DAC-AT protein obtained by IPTG-induced overexpression in a minimal expression medium also formed larger crystals in the shapes of flat, broad needles or swords (Figure 14). These crystals typically diffracted to better than 2.5 Å and displayed lower mosaicity (0.5-0.7°) than the initial DAC-AT crystals.

3.2.3 Data collection and structure determination

A native dataset was collected on a crystal of DAC-AT in complex with acetyl coenzyme A, to a resolution of 2.2 Å. The unit cell was found to have relatively long cell edges, which led to some uncertainty regarding the number of molecules in the asymmetric unit (10-14, resulting in solvent contents of 44-60% (Matthews, 1968)). Heavy atom derivatisation for structure determination with SIR(AS)/MIR(AS) was attempted, and although a variety

of heavy atom compounds were screened, location of heavy atom sites was troublesome due to the large and uncertain number of molecules in the asymmetric unit.

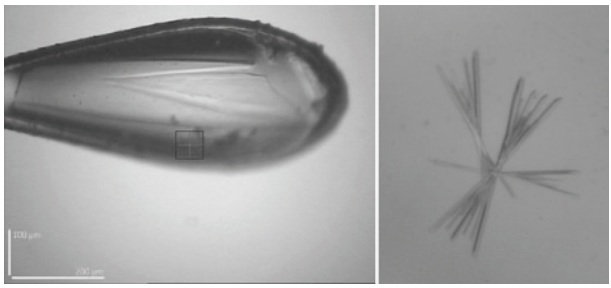

Figure 14. Left: single DAC-AT crystal in a cryoloop. Right: cluster of SeMet-derivatised DAC-AT crystals.

After optimisation of expression and crystallisation of DAC-AT, a highly redundant SAD data set could be collected on a selenomethionine derivative crystal. The data set was collected at the wavelength corresponding to the absorption energy at the selenium K-edge (E= 12.66 keV, λ= 0.979 Å), and contained diffraction data to 2.4 Å resolution. The unit cell parameters were found to be a=121.8 Å, b=109.5 Å, c=196.3 Å and β=90.05°. Using data to 2.9 Å, SHELXD (Schneider & Sheldrick, 2002) found a large number of Se sites. The 100 sites with highest occupancy were submitted for refinement in autoSHARP (Vonrhein*, et al.*, 2006), which also found additional sites. Cross-validation with MLPHARE in the CCP4 suite of programs (Collaborative Computational Project Number 4, 1994) confirmed a total of 108 selenium sites. Several rounds of SHARP refinement of the heavy atom sites (de la Fortelle & Bricogne, 1997) were carried out, followed by density modification using DM (Cowtan, 1994). An initial model for one monomer of DAC-AT was built manually, using fragments automatically built by ARP/wARP (Lamzin*, et al.*, 2001); this chain was denoted chain A. A total of 12 chains in the asymmetric unit were confirmed by using PHASER (McCoy*, et al.*, 2007) with native data to 2.2 Å. The phases thus obtained were further improved by non-crystallographic symmetry (NCS) electron density averaging using the RAVE suite of programs (Kleywegt & Jones, 1994). Firstly, the operators relating the NCS-related molecules in the asymmetric unit to chain A were calculated, and this was followed by mask calculation. A mask covering chain A was then used for 10 cycles of averaging. After rebuilding of chain A and improving the NCS operators, another 10 cycles of averaging were performed. At this point, the maps had improved significantly, and the model for the monomer could be completed to contain residues 6-241 and 268-385. In subsequent rebuilding and refinement rounds, non-averaged maps were used. NCS restraints were used for

coordinate refinement in REFMAC (Kleywegt, 1996, Murshudov, *et al.*, 1997); these restraints were gradually relaxed for certain residues, notably for some whose side-chains were involved in intermolecular contacts. For the remaining residues, tight NCS restraints were employed throughout refinement, both for main-chain and side-chain atoms.

A post-mortem analysis of the data sets collected on heavy atom derivatives that could not be used for structure determination has shown that there is some variability in the β-angle (90.0°-90.3°) for different crystals when processed in space group $P2_1$. Many of the soaked crystals displayed satisfactory statistics when data were merged in an orthorhombic space group, due to the β angle being very close to 90.0° (values of 90.00-90.03° were common).

3.2.4 The overall structure of DAC-AT

The DAC-AT monomer is composed of two domains: a 28 kDa α/β domain consisting of residues 6-177 and 304-385, capped by a 15 kDa all-α-helical domain consisting of residues 178-303 (Figure 15).

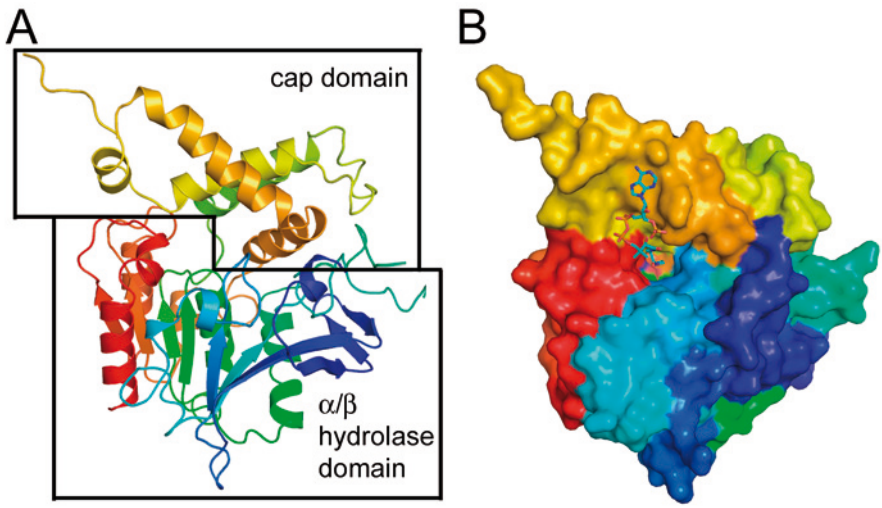

Figure 15. Tertiary structure of DAC-AT. (A) The monomer of DAC-AT in ribbon representation, colour-ramped from the N-terminus in blue to the C-terminus in red. The two-domain structure is indicated. (B) Same view of DAC-AT as in A, with the monomer shown as a molecular surface. Bound coenzyme A is shown as sticks in the binding tunnel between the cap domain and the α/β hydrolase domain.

The α/β domain adopts the canonical α/β hydrolase fold (Figure 16) (Holmquist, 2000, Ollis, *et al.*, 1992), with a central eight-stranded, mostly parallel β-sheet (β1-β8), flanked by six α-helices (αA-αF), with the first and last α-helices packing on one side of the sheet, and the remaining α-helices

on the opposite side. The β-sheet is bent to form a half-barrel with the first and final β-strands at an angle of almost 90° to each other. The cap domain, which is inserted between β6 and β7, contains a total of five α-helices. Although strictly belonging to the α/β hydrolase fold, the helix αD is actually situated in the cap domain. The remaining four α-helices of the cap domain are here denoted αD2 to αD5. No electron density could be observed for residues 242-267 of the cap domain. The juxtaposition of the two domains results in the formation of a tunnel, which harbours the active site at the deep end (see section 3.2.5).

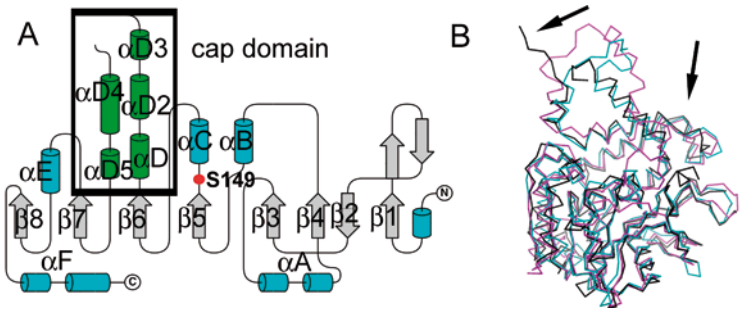

Figure 16. Topology and structural homology of DAC-AT. (A) Topological diagram, numbering of secondary structure elements of the α/β hydrolase fold (αA-αF and β1-β8) is according to Holmquist (2000). The cap domain (consisting of helices αD and αD2-αD5) is indicated in green. (B) Superposition of C_α traces of DAC-AT (in black) with *Hi*HTA (in cyan) and *Li*HTA (in magenta), using least-squares superpositioning in O. Superpositioning 307 C_α atoms from DAC-AT onto *Hi*HTA yields an RMSD of 1.51 Å, and superpositioning 288 C_α atoms from DAC-AT onto *Li*HTA yields an RMSD of 1.34 Å (with a cut-off value of 3.5 Å to identify aligned atoms). Regions of DAC-AT displaying the largest deviations from the HTA structures are indicated with arrows.

The structure of DAC-AT displays deviations from the *Hi*HTA structure (indicated with arrows in Figure 16B), and this is also reflected by the failed attempts at solving the DAC-AT structure by molecular replacement using the *Hi*HTA structure as search template. The structure of *Li*HTA was determined and published after the DAC-AT structure had been solved (Wang*, et al.*, 2007).

All DAC-AT monomers in the crystal lattice form extensive dimer contacts with another chain (Figures 17 and 18), either within or outside the asymmetric unit. No published study has previously reported that DAC-AT might form dimers, but the protein has been observed to behave as a dimer during gel filtration (J. Ellis, personal communication). Also, the size and

extent of the dimer contact area (on average 1800 Å^2) in the crystal lattice, in addition to the high structural similarity to *Hi*HTA and *Li*HTA – which have been suggested to exist physiologically as dimers (Born, *et al.*, 2000, Curry, *et al.*, 1999, Wang, *et al.*, 2007) – indicate that DAC-AT may exist as a dimer in solution.

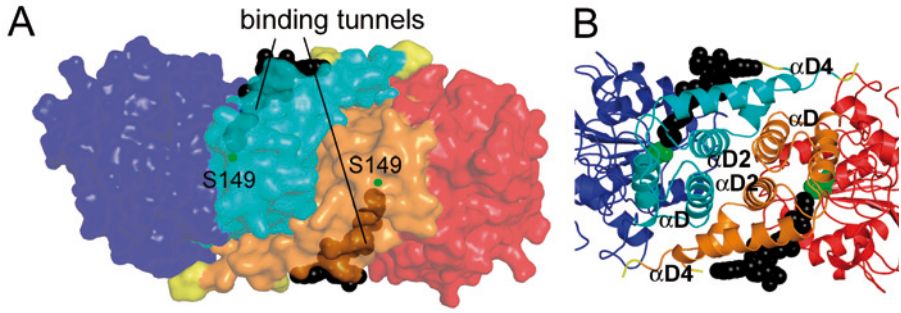

Figure 17. Quarternary structure of DAC-AT. (A) The dimer interaction viewed down the two-fold dimer axis. Monomers are shown in blue and red, cap domains are shown in lighter shades. Coenzyme A is shown as black spheres. The N- and C-terminal ends of the disordered region of the cap domain are indicated with yellow patches. (B) Helices at the dimer interface, same view as in (A).

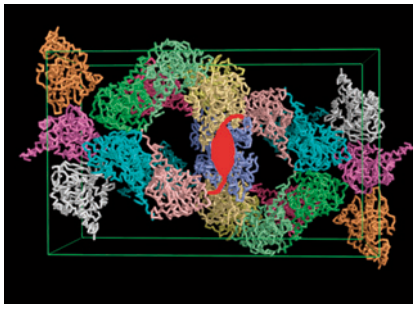

Figure 18. Packing in the crystal lattice, viewed down the unique axis *b*. The central two-fold screw axis is shown in red.

3.2.5 Complex with acetyl coenzyme A

Crystals of DAC-AT have so far only been possible to obtain with the cofactor acetyl coenzyme A present in the crystallisation solution. After a few rounds of model building in O (Jones, *et al.*, 1991), cycled with restrained refinement in REFMAC (Murshudov, *et al.*, 1997), clear difference density could be seen for the cofactor.

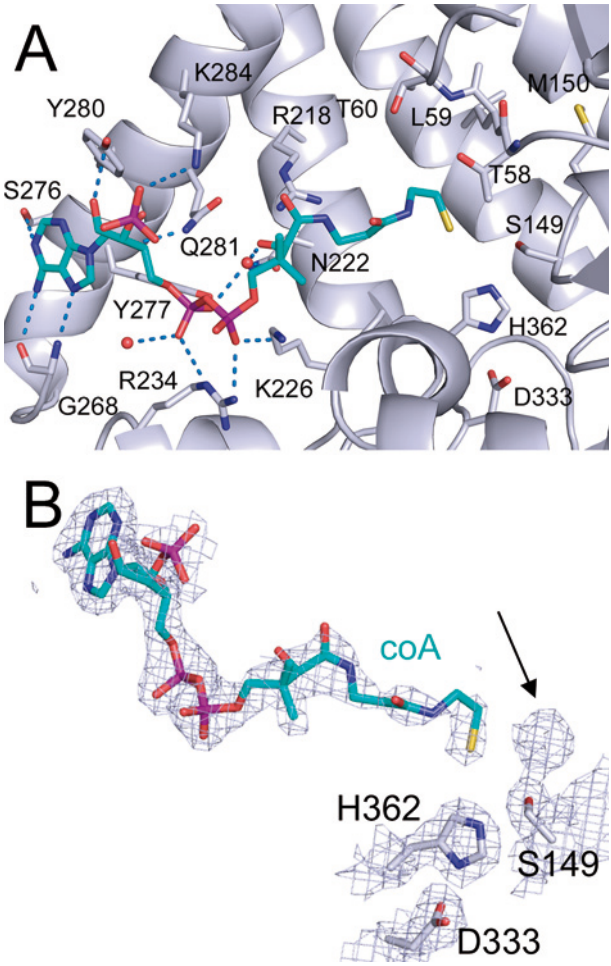

Figure 19. Binding of acetyl coenzyme A in the active site in DAC-AT. (A) Residues are shown as sticks, and coenzyme A is shown in cyan. (B) Electron density is shown as an omit map generated with coefficients $2mF_o - DF_c$ and phases calculated from the final model. The map is contoured at 0.30 $e^-/Å^3$ level (1.0σ). Probable location of the acetyl group in the acyl-enzyme complex is indicated with an arrow.

Ideal coordinates for the cofactor (and other DAC-AT ligands described below) were taken from HIC-Up (Kleywegt*, et al.*, 2003), and restraints were generated and inspected using the programs SKETCHER and LIBCHECK through the CCP4 interface (Collaborative Computational Project Number 4, 1994, Vagin*, et al.*, 1998). Acetyl coenzyme A binds in a single site in the DAC-AT monomer. The binding site is situated in the tunnel resulting from the juxtaposition of the two domains. Least-squares superpositioning of the *Hi*HTA structure on DAC-AT confirmed the active site in DAC-AT to be constituted by a serine protease-type catalytic Ser-His-Asp

triad. The active site, which involves residues Ser149, His362 and Asp333 in DAC-AT, is situated at the deep end of the binding tunnel (Figure 19).

Acetyl coenzyme A is bound in the tunnel via several hydrogen bonds and salt bridges, with the stabilising interactions primarily clustered at the tunnel entrance, where the adenine and ribose moieties of the cofactor are held in place. In contrast, the β-mercaptoethylamine thiol end of acetyl coenzyme A is more disordered, and the acetyl group itself cannot be unambiguously modelled. The final model therefore contains bound coenzyme A. Residual density near Ser149, however, indicates the presence of multiple partially occupied conformational states for the acetyl group (see Figure 19B). Electron density maps support the presence of an acyl-enzyme conformation, where the acetyl group has transferred to be covalently bound to S149, leaving coenzyme A in the tunnel.

3.2.6 Complexes with substrate and product

Crystal structures of DAC-AT with substrate and product have also been determined, and these were obtained by soaking preformed DAC-AT crystals that had cocrystallised with acetyl coenzyme A in solutions of either deacetylcephalosporin C (DAC) or cephalosporin C.

In both cases, it was found that coenzyme A is displaced from the binding tunnel, and that the same side-chains lining the binding tunnel that were observed to stabilise the cofactor are also recruited for the binding of substrate and product. The majority of the interactions stabilising DAC or cephalosporin C are clustered further into the tunnel compared with the coenzyme A complex.

3.2.6.1 Soaking with DAC

Data collected on crystals soaked in the substrate DAC indicate that the DAC-AT monomers in the crystal lattice catalyse the forward reaction – i.e. acetyl transfer from acetyl coenzyme A to the substrate DAC. The monomers in the crystal lattice consequently contain a mixture of states in their respective active sites. Electron density for the majority of active site atoms in DAC-AT side-chains and in the ligand (DAC or cephalosporin C) was well-defined, and indicated one conformation only. However, attempts to model the ligand as either DAC or cephalosporin C with full occupancy resulted in difference electron density indicating alternative conformations for the acetyl group. The position of the carbonyl oxygen of the acetyl group was clearly visible in the electron density maps, and it was found that the most plausible structure of the active site involved modelling of acetyl atoms at partial occupancy.

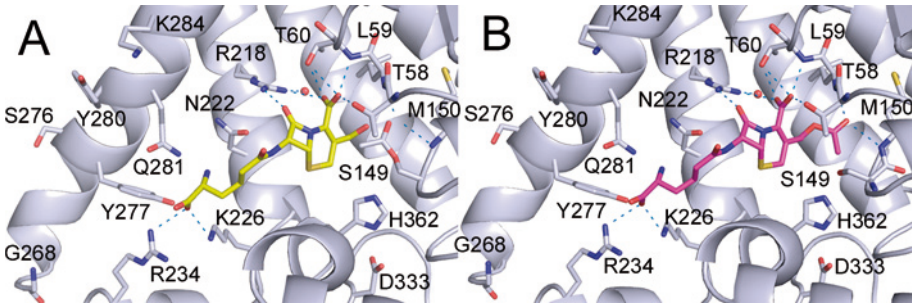

Figure 20. The active site structure of the DAC-soaked crystals. The two simultaneously modelled conformations of the active site are shown as separate images. DAC is shown in yellow, cephalosporin C in magenta. Groups within hydrogen bond distance from each other are connected by dashed lines. (A) Conformation representing the structure of the substrate DAC bound to the acyl-enzyme complex from the reaction with acetyl coenzyme A. (B) Conformation representing the structure of the product cephalosporin C resulting from the reaction with the acyl-enzyme complex.

The active site in the refined DAC-soaked structure contains two overlapping conformations of the acetyl group at partial occupancies. One conformation is represented by the acetylated Ser149 with DAC bound in the active site (i.e. the acetyl group is bound to the γ-oxygen of Ser149), and the other conformation is represented by Ser149 with cephalosporin C bound in the active site (i.e. the acetyl group is bound to DAC). It was also found that the different active sites in the asymmetric unit had reacted with the soaked substrate to different extents, and consequently some chains had to be refined with different contributions from both the substrate DAC and the resulting product cephalosporin C in their active sites.

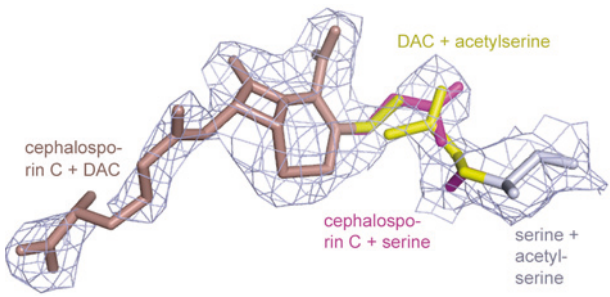

Figure 21. The overlapping conformations of DAC and cephalosporin C in the active site of the DAC-soaked crystals. Atoms belonging to the same conformation are shown in the same colour. Electron density is shown as an omit map generated with coefficients $2mF_o - DF_c$ and phase angles calculated from the final model. The map is contoured at a level of 0.27 $e^-/Å^3$ (1.0σ).

Figures 20 and 21 show an active site in the DAC-soaked structure of DAC-AT. The atoms in the acetyl moiety ($COCH_3$) are modelled with partial occupancy covalently bound to the γ-oxygen of Ser149 (Figure 20A), and with partial occupancy covalently bound to the terminal hydroxyl oxygen of the DAC molecule (Figure 20B).

3.2.6.2 Soaking with cephalosporin C

Crystals soaked with the product, cephalosporin C, were also found to react with the soaked compound. In this case, data collected on the crystals indicated that the active sites in the crystal lattice catalyse the reverse acetyl transfer reaction, i.e. the transfer of the acetyl group from cephalosporin C to Ser149 of the enzyme.

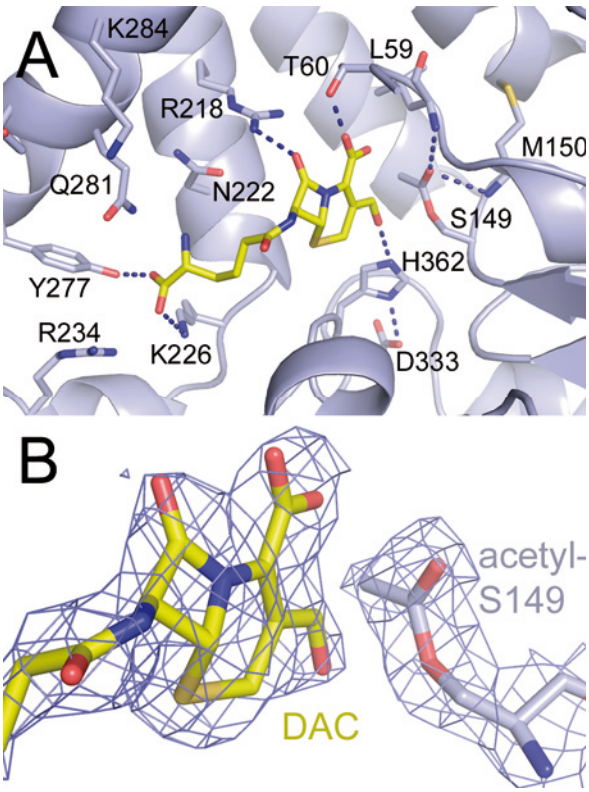

Figure 22. The active site of DAC-AT in the cephalosporin C-soaked crystals. (A) Atoms within hydrogen bond distance are connected with dashed lines. DAC is shown in yellow. (B) Close-up of the bound ligand and acetylated Ser149. Electron density is shown as an omit map generated with coefficients $2mF_o - DF_c$ and phases calculated from the final model. The map is contoured at a level of 0.26 $e^-/Å^3$ (1 σ).

Clear electron density shows a double-headed lobe on the catalytic serine, which can be satisfactorily modelled as a fully occupied acetyl group in all chains in the asymmetric unit (Figure 22). The shape of the electron density

near the catalytic serine in the cephalosporin C-soaked crystals is distinct from the electron density in the DAC-soaked structure, which instead can best be described as single-headed, as a result of the two overlapping acetyl groups (see Figure 21). Soaking of crystals in a large excess of cephalosporin C thus seems to result in a stable acyl-enzyme complex with the intermediate DAC bound.

3.2.7 Reaction mechanism

The crystallographic results presented in sections 3.2.5 and 3.2.6, which show the cofactor, substrate and product to share a single binding site in the enzyme, suggest a double displacement mechanism for the acetyl transfer reaction. The catalytic triad and charge relay system is represented in DAC-AT by the nucleophile Ser149, the catalytic base His362 and the proton acceptor Asp333. The proposed mechanism is outlined in Figure 23.

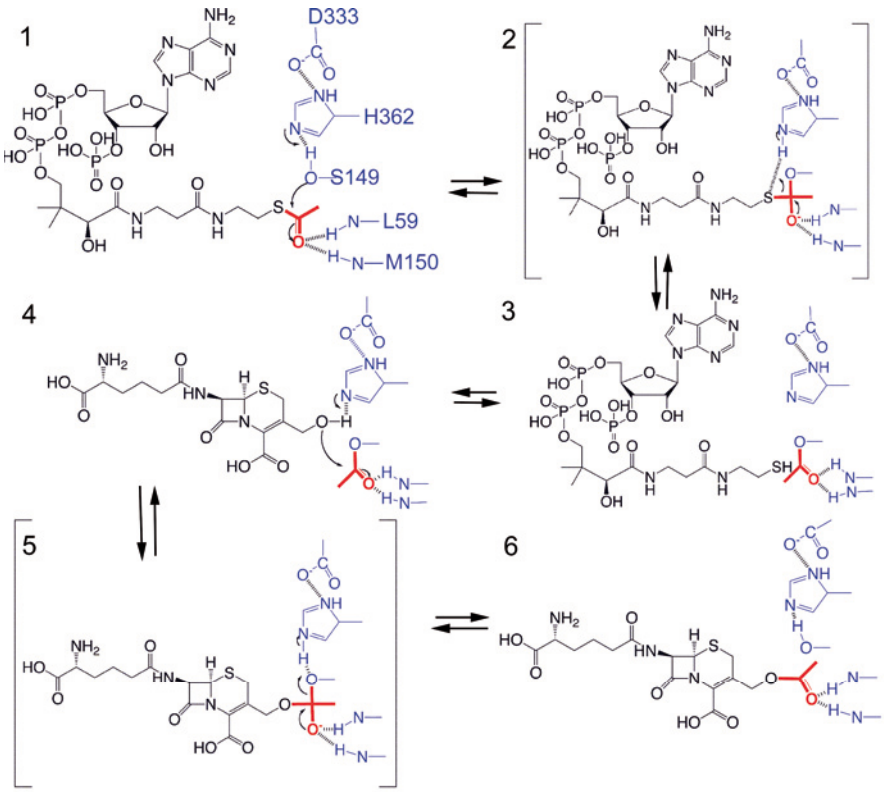

Figure 23. Proposed reaction mechanism for DAC-AT. Catalytic triad and oxyanion hole residues are shown in blue. The acetyl group is in bold red. Suggested tetrahedral intermediates are shown in square brackets.

The carbonyl oxygen of the acetyl group in the structures of the DAC-soak and of the cephalosporin C-soak are within hydrogen bonding distance from the backbone amides of Leu59 and Met150. These residues thus emerge as likely candidates for the proton-donating oxyanion hole, which is required for the stabilisation of the negatively charged tetrahedral transition state.

The double displacement reaction suggested here consists of two half-reactions, each of which carries out one acetyl transfer reaction. In the first half-reaction, the cofactor acetyl coenzyme A binds in the active site, chiefly stabilised by DAC-AT side-chains that interact with its adenine, ribose and phosphate groups (panel **1** in Figure 23). The γ-hydroxyl of Ser149, which has been rendered nucleophilic by the proton-extracting effect of the Nε group of the general base His362, then attacks the carbonyl carbon of the scissile thioester bond in acetyl coenzyme A. The nucleophilic attack results in the formation of a negatively charged tetrahedral transition state (panel **2**). The negative charge is stabilised by the backbone amide protons of Leu59 and Met150. The unstable transition state then breaks down into the more stable acyl-enzyme intermediate, with the acetyl group now covalently bound to the catalytic S149, leaving coenzyme A bound in the tunnel. The second half-reaction is initiated by the replacement of coenzyme A with the substrate DAC binding in the active site (panel **4**). The hydroxyl oxygen of DAC then acts as the nucleophile in the same manner as Ser149 in the first half-reaction, with the aid of the proton-extracting action of His362. The nucleophilic attack is followed by the formation of a second tetrahedral transition state (panel **5**), with the negative charge stabilised by the same oxyanion hole amides as in the first half-reaction. The breakdown of the transition state then results in the formation of the final product, cephalosporin C (panel **6**).

3.3 Conclusions

One of the aims of this thesis was to determine the structure of DAC-AT, the enzyme responsible for the last step in the biosynthesis of the clinically important β-lactam antibiotic cephalosporin C. Efforts were made to investigate the catalytic mechanism of the enzyme, by studying the molecular details of the interactions with substrates and products. The crystallographic results presented in this thesis pinpoint the residues that play important roles in the acetyl transfer reaction. Although the enzyme has been found to belong to the α/β hydrolase structural superfamily, which has many members of varying, mostly hydrolytic functions (e.g. cholinesterase, epoxide hydrolases and lipases), DAC-AT belongs to a relatively new subset of the α/β hydrolase fold enzymes. At the time this is put into writing, the structure of DAC-AT and the two HTA structures mentioned in the above sections are

the only acetyltransferases deposited in the PDB that have been shown to contain the α/β hydrolase fold. DAC-AT represents the first structure for which the molecular details of the acetyl transfer reaction have been elucidated.

One additional point can be mentioned here with respect to the structure and activity of DAC-AT. The two-domain architecture of DAC-AT provides an explanation for the post-translational processing of DAC-AT into two subunits that was suggested by Matsuda and coworkers. Although the crystal structure strongly suggests that the active enzyme consists of a single chain, the expected sizes of the processed subunits correlate well with the observed sizes for the two domains in the DAC-AT monomer. Thus, the observation made by Matsuda and coworkers might be explained by a presence of proteolytic enzymes in their DAC-AT preparation.

4. Future perspectives

4.1 The GA module and HSA

Within the scope of this thesis, studies of the GA module and HSA were undertaken with a view to structurally characterise the interaction between these two proteins in the context of bacterial pathogenesis. As it turns out, the structure of the complex that GA forms with HSA yielded a structural basis for the range of hosts that can be infected by the bacterium, and also offered a possibility to study the interaction of HSA with molecules present in extracellular fluids, from the point of view of the HSA molecule.

In terms of albumin science, the crystal structures presented above represent merely a small section of the interaction space available to HSA. The conformation of HSA and its binding state in terms of interaction partners are dependent on several parameters, including time, space, location in the body, age of the protein, and the nutritional and general health of the individual. To comprehensively map all possible combinations of interactions that HSA may participate in is a formidable task.

The possibility of using the GA module to obtain crystals of HSA that represent a physiologically relevant conformation of HSA, with a view to study the specific interactions between HSA and drugs sequestered by HSA has not yet been fully investigated. Although such a platform in itself will not yield a complete picture of HSA-ligand interactions, the relative ease with which crystals can be obtained and analysed is a good start for the development of such protocols. A combination of such studies with other biophysical methods, such as NMR spectroscopy and mass spectrometry, could open a powerful way of studying this physiologically and pharmacologically important protein.

4.2 DAC-AT

The growing need for controlling bacterial infections is a world-wide problem, and several routes are currently being pursued in parallel to address this issue. Some examples include hospital strategies to prevent the spread of already existing strains of antibiotic-resistant pathogens, plans to actively limit the selective pressure towards resistance by decreasing the use of anti-

biotics in farming and animal industries, the development of entirely new antibiotic substances, and the optimisation of existing antibiotics. The search for new cephalosporin antibiotics with high specific activity against MRSA is a current example that may prove fruitful in the near future (Rossolini, 2007).

Where in this perspective does DAC-AT play a role? Efforts to improve cephalosporin yields in industrial fermentation have so far mainly involved strain improvements and engineering of *A. chrysogenum* on the gene expression level (Rodríguez-Sáiz*, et al.*, 2004). Now that the structure of DAC-AT is known, are there any obvious routes to optimising cephalosporin C yields by protein engineering? The cap domain of DAC-AT could deserve some attention in this respect. It has been shown that the cap domain in other enzyme members of the α/β hydrolase superfamily plays a major role in determining the substrate specificity of the enzyme, and mutations of key residues in the cap domain have been reported to alter the catalytic properties of some of these enzymes (see (Holmquist, 2000) and references therein). In DAC-AT, there are several cap-domain residues whose side-chains are involved in interactions with the cofactor, substrate or product; specifically Arg218, Asn222, Lys226, Arg234, Gly268, Ser276, Tyr277 and Tyr280. In the case of substrate or product, these active site residues are all involved in interactions with the L-γ-glutamoyl side-chain of the cephalosporin molecule; it might be speculated that by mutating one or several of these residues, one might be able to engineer DAC-AT to recognise cephem substrates (or even other β-lactam compounds) with alternative R1 side-chains. Such efforts have been described previously for other proteins involved in β-lactam biosynthesis (Barends*, et al.*, 2004), with a view to generating semi-synthetic compounds with new antibiotic properties. With the accumulating structural knowledge of β-lactam biosynthetic enzymes, including DAC-AT, such studies will become more and more accessible.

The activity of DAC-AT and its role in the accumulation of the intermediate DAC is also worth noting when considering hydrolysis of cephalosporin C. In *A. chrysogenum,* DAC-AT and another enzyme known as cephalosporin C acetylhydrolase (Velasco*, et al.*, 2001) both catalyse the hydrolytic cleavage of the acetyl group of cephalosporin C. The biological function of such enzyme activity in *A. chrysogenum* is not yet known, and further studies of both these enzymes will be required for the understanding of the whole picture behind cephalosporin C biosynthesis. Indeed, now that preparation of active DAC-AT enzyme in the laboratory is comparatively straightforward, further studies of its activity will be more accessible than before.

5. Summary in Swedish

Kamouflage och attack: strukturstudier av ett bakteriellt albumin-bindande protein samt av ett cefalosporin-biosyntetiskt enzym

Möjligheten till behandling av bakterieinfektioner och andra sjukdomar är en av århundradets största vetenskapliga framgångar, och en av de viktigaste byggstenarna i det moderna samhället. Under de senaste decennierna har dock ett flertal problem vuxit sig allt mer insisterande, varav ett av de största ter sig vara bakteriell resistens mot antibiotika. Vissa ansträngningar görs nu därför för att motverka denna utveckling, vilken annars skulle kunna leda oss till ett pre-penicillint samhälle, det vill säga en värld där den medicinska vetenskapen står maktlös inför sjukdomsalstrande mikroorganismer. Utveckling av befintliga antibiotika samt bättre förståelse för mekanismerna bakom uppkomst och propagering av bakteriella infektioner är därför av stor vikt.

Denna avhandling behandlar två sidor av detta mångfacetterade och komplexa problem; å ena sidan hur ett visst fall av bakteriell sjukdomsalstring kan uppstå, å andra sidan hur biosyntes av ett visst antibiotikum sker. Avhandlingen bygger på studier av proteiner med hjälp av röntgenkristallografi, vilket ger en bild av ett proteins tredimensionella struktur. Detta kan i sin tur ge värdefull information om proteins funktion. Denna avhandling behandlar två proteiner; dels GA-modulen, en domän från ett ytbundet protein från patogena stammar av bakterien *Finegoldia magna*, och dels DAC-AT (deacetylcephalosporin C acetyltransferas), ett biosyntetiskt enzym från den antibiotikaproducerande svampen *Acremonium chrysogenum*.

GA- modulen, vilken utgör en 53 aminosyror lång domän från ytproteinet PAB från *F. magna*, har tidigare karakteriserats med avseende på ett flertal biokemiska egenskaper. Man fann tidigt att dess funktion var att binda till serum albumin, främst humant. Humant serum albumin (HSA) är en viktig komponent i humant blodplasma, där det verkar som transportör av ett flertal organiska molekyler, främst fettsyror. Även läkemedelsmolekyler, diverse hormoner och nedbrytningsprodukter binder till HSA. Med avseende på GA har man tidigare funnit, att patogena stammar av *F. magna* är kapabla att binda till HSA, vilket i sin tur gör att bakterien kan bilda en ”sköld” av serumproteiner. Denna sköld gör det möjligt för bakterien att lämna sin naturliga omgivning (mänsklig mag- och tarmtrakt) och invadera andra vävnader, där den kan orsaka allvarliga infektioner. Skölden av humana proteiner (bakterien kan även binda till IgG, ett annat vanligt protein i plasma) gör bakteri-

en osynlig för immunsystemet. GA-modulen hjälper alltså bakterien att förvandlas till en ”varg i fårakläder”. De röntgenkristallografiska studierna av GA-modulen som presenteras i denna avhandling har resulterat i tre proteinstrukturer; en struktur med GA-modulen bundet till HSA (här kallat HSA-GA-komplexet), en struktur med GA-modulen bundet till HSA med läkemedelsmolekylen naproxen samtidigt bunden till HSA (här kallat HSA-GA-naproxen-komplexet), samt en struktur av den isolerade GA-modulen, d.v.s. ej i komplex med HSA. Dessa tre strukturer avslöjar egenskaper inte bara hos GA-modulen, utan även hos HSA.

Strukturen av HSA-GA-komplexet är den första kristallstrukturen av humant serum albumin bundet till ett annat protein. GA-modulen binder till en yta på utsidan av HSA, och involverar ett flertal sidokedjor i vätebindningar och jonparsbindningar för stabilisering. En central interaktion mellan dessa två proteiner involverar aminosyran fenylalanin i position 27 på GA-modulen, vars opolära sidokedja begravs i en hydrofob ficka på HSA-ytan. Denna interaktion ger även en strukturell förklaring till tidigare publicerade experiment som har visat att *F. magna*-infektioner är mer begränsade med avseende på värdorganismen än infektioner orsakade av grupp C/G streptokocker. Denna skillnad är särskilt intressant eftersom GA-modulen troligtvis har överförts till icke-patogena stammar av *F. magna* via genöverföring från streptokocker. Strukturjämförelser med den albumin-bindande domänen från streptokocker visar att GA-modulen från *F. magna* interagerar med en hydrofob ficka, medan GA-modulen från streptokocker kan interagera med hjälp av polära interaktioner, med hjälp av en tyrosin i denna position. Det visar sig att albumin från andra arter än människa har varierande, främst polära, sidokedjor i den hydrofoba fickan, vilket möjliggör interaktion med en polär sidokedja i position 27, men försvårar för interaktion med en fenylalanin. Därmed gör streptokock-varianten av proteinet det möjligt för bakterien att infektera ett flertal olika däggdjursarter (t. ex. ko, kanin, mus och råtta), medan *F. magna*-varianten är mer begränsad och främst infekterar människor och andra primater.

Med utgångspunkt från HSA-GA-strukturen har även försök att vidare undersöka vilket inflytande andra serumkomponenter kan ha på denna interaktion. En sådan serumkomponent kan vara fettsyror, men även läkemedelsmolekyler. I samband med dessa studier bestämdes strukturen av HSA-GA-komplexet med läkemedelsmolekylen naproxen bundet till albumin. Denna struktur erhölls i närvaro av dekansyra (en fettsyra) bundet till albumin. Strukturen visar att GA-modulen binder till HSA på samma sätt trots att HSA har bundit en annan uppsättning av molekyler i andra bindningsställen, jämfört med den första strukturen av HSA-GA-komplexet. Ytterligare ett resultat av denna studie är karakteriseringen av naproxen-bindning till HSA i närvaro av fettsyror; HSA-GA-naproxen-komplexet visar att naproxen binder till en yta vilken tidigare inte har visats vara involverad i naproxen-bindning. De förväntade naproxen-bindningsställena är istället upptagna av

fettsyror. Sådan information om hur läkemedel interagerar med albumin är intressant ur läkemedelsutvecklingssynpunkt; med hjälp av denna och andra studier kan man karakterisera hur läkemedel binder till HSA, och på så vis förstå hur ett läkemedel beter sig när det kommer in i mänskliga vävnader och blodomlopp. Detta är i sin tur viktigt för att optimera ett läkemedels aktivitet och nerbrytningshastighet i kroppen.

En intressant observation gjordes vid bestämning av de två HSA-GA-strukturerna, nämligen att en fettsyremolekyl binder till ett bindningsställe i HSA alldeles intill bindningsytan för GA-modulen. Strukturerna visar att det eventuellt kan röra som en direkt interaktion mellan det bakteriella proteinet och fettsyran. Detta är en tydlig indikation på att fettsyror som transporteras av HSA kan ha en direkt inverkan på om en bakterie binder HSA och därmed lyckas invadera mänskliga vävnader. Tidigare studier av GA-modulen har dessutom spekulerat i att bakterien kan använda GA-modulen för att komma åt näringsämnen (t. ex. fettsyror) som är bundna till HSA. Dessa näringsämnen skulle sedan kunna användas av bakterierna för tillväxt.

Förutom dessa två strukturer av HSA-GA-komplexet bestämdes även kristallstrukturen för den isolerade GA-modulen, det vill säga utan inbindning till albumin. Denna struktur visar i sin tur att GA-modulen bildar asymmetriska dimerer (protein-par) i kristallen. Dessa protein-par bildas genom att två GA-moduler binder till varandra, och i denna kristallstruktur ser man att GA-modulen kan använda bindningsytan för HSA-interaktion även för inbindning till en annan GA-modul.

Sammanfattningsvis visar alltså dessa studier dels på GA-modulens flexibilitet vad gäller interaktioner med HSA (HSA kan anta åtminstone två olika konformationer utan att det påverkar GA-modulens förmåga att binda), och dels på att det finns indikationer på att GA-modulen själv kan binda till flera olika proteinytor. Sammantaget har alltså GA-modulen flera egenskaper vilka kan bidra till att öka infektionsförmågan hos bakterien *F. magna*. Dessutom visar dessa studier att HSA-GA-komplexet kan utgöra en plattform för vidare studier av hur läkemedel interagerar med HSA, särskilt i närvaro av olika nivåer av fettsyror i kroppens cirkulation.

Den andra delen av denna avhandling behandlar en aspekt av den andra sidan av bakteriella infektioner – det vill säga motverkan. Ett av de mest använda antibiotika i världen, vid sidan av penicillin, är penicillinets strukturella släkting cefalosporin. Cefalosporinerna härstammar från cefalosporin C, vilket produceras naturligt av svampen *A. chrysogenum*. Biosyntesen av detta s.k. β–laktam antibiotikum sker i flera steg. Det sista biosyntetiska steget är en överföring av en acetylgrupp från acetylcoenzym A till intermediatet deacetylcefalosporin C. Studierna som presenteras i denna avhandling visar de strukturella detaljerna för hur denna överföring sker. Detta har gjorts genom bestämning av kristallstrukturen för det enzym som katalyserar det sista steget i biosyntesen av cefalosporin C; deacetylcefalosporin C acetyl-

transferas (DAC-AT). Strukturbestämning av detta enzym visar att det består av två domäner; en s k α/β hydrolas-domän, samt en α-helix-innehållande domän. Dessa två domäner bildar tillsammans en tunnel, vilken agerar bindningsställe för kofaktorn acetyl coenzym A. Genom att tillföra lösningar av enzymets substrat eller den slutliga produkten till kristaller av DAC-AT, visas att kofaktorn ersätts i bindningstunneln av reaktionsintermediat - enzymet katalyserar alltså en reaktion med hjälp av dessa substanser. Med hjälp av röntgenkristallografisk strukturbestämning visas i dessa studier att katalys sker med hjälp av en aktiv yta belägen i den djupa änden av bindningstunneln. Denna aktiva yta består av en s k katalytisk triad - en trio av aminosyror - vilken är nödvändig för enzymaktivitet. Genom att först interagera med acetyl coenzym A och föra över en acetylgrupp från kofaktorn till aminosyran serin, och därefter föra över acetylgruppen från serinsidokedjan till ett biosyntetisk intermediat, katalyseras alltså en acetylöverföring. Den slutliga produkten är cephalosporin C, vilken sedan kan användas för framställning av andra, semi-syntetiska cefalosporin-antibiotika för klinisk användning mot patogena mikroorganismer.

Man har i tidigare studier sett att man kan, genom att förändra (mutera) specifika aminosyror i liknande enzymer, ändra enzymets egenskaper. Med hjälp av sådana experiment, och med ledning av strukturen som presenterats här, skulle man kunna utröna om man på så vis kan förändra DAC-AT så att enzymet kan utföra liknande eller andra reaktioner på andra substrat. Sådan strukturell information kan alltså vara värdefull för att i framtiden kunna ta fram cefalosporin-antibiotika med nya egenskaper.

Sammanfattningsvis behandlar denna avhandling några aspekter av bakterieinfektion ur ett strukturellt perspektiv. Med hjälp av dessa studier kan experiment utföras för att vidare utröna dels hur GA-modulen interagerar med HSA och/eller andra serumproteiner och dels hur HSA binder och inaktiverar läkemedelsmolekyler i närvaro av andra serumkomponenter. Vidare studier av hur biosyntes av cefalosporin C sker, och eventuellt kan moduleras, kommer också att vara värdefulla.

Acknowledgements

Thank you...
Tack till...

Janos - my heartfelt thanks for accepting me as a PhD student, for the never-ending enthusiasm & unconditional support... and for the swimming lessons

Karin - för allt! - för sällskap under sena synkrotron-nätter, för all hjälp, stöd & glada tillrop och för din osvikliga optimism (...och för match-making!)

Jackie - for all your cheerful patience and dogged persistence during the DAC-AT/cefG project

Inger - tack för all hjälp med manuskriptskrivande och för givande diskussioner

Jacob - tusind tak - du var en klippa med GA-projektet

Tomas Lundqvist & Peter Nordberg - för gott samarbete med GA-projektet

Stefan Svensson & Mats Wikström - för gott samarbete med albumin- och GA-projekten på Biovitrum

Nic - for being the best office-mate during all these years, for sunshine and monkey-feeding, and for putting up with madness and blonde-ness all in one person... also thanks to *Anca* & *Theo* for dinners, coffee, hugs, smiles and trains (... and welcome to Lucian)

Alexandra - för Cambridge-sällskap, rödvin, jagande av håriga helikopter-insekter och BodyPump... och för att du helt enkelt är snyggast!

Rosie - för att du är en så tålmodig kontorskompis och för inspirerande grrl-styrka i alla lägen

Johan - för kontorssällskap, skrattpauser, Chuck Norris & kallt kaffe

Al, Martin, Louise, Urszula, Andrea, Lotta - for all the late synchrotron night laughs, sand, rum and fellow dewar-carrying

Anke, Linda, Malin, Jenny - för att ni visade hur man gör

Dr Calle & *Dr Magnus*, *Carlos*, *Gösta*, *Eva* & *Johanna* - tack för Bonita-bröl, för punktliga Livermore-möten, för margaritas i Las Vegas & soluppgång i Grand Canyon

David & Erling - tack för avhandlingsläsning, datorhjälp & SHARP-assistans

Tex, *Emma, Alwyn* - for introducing me to xray all those years ago

Sigrid - stort tack för all hjälp och för att du alltid har tid!

Agata - for putting up with sleep-talking in Tällberg not just once, but twice!

Sara E - för tappert och tålmodigt arbete med motsträviga ORFar

Gelareh - för att du skrattar så hela BMC hör

jhgroup - past & present members - thank you all for coffee company and brilliantly nerdy discussions

alla @ xray - thanks for Tällberg, lunch room company, Lucia singing, help, support & for making BMC such a cool place to work

Sara & Björn - tack för att ni alltid finns där!

Per & *Jenny*, *Johanna*, *Patrik*, *Niklas*, *Jens*, *Nina, Ulrika* - tack för allt fika & pyssel & fest & brunch & sympatiångest & ångestsympati

Linda - för att du är världens bästa keyboard devil, för allt nattuggleri, te & kakor, The West Wing & BSG - jag saknar dej!

For company while writing the thesis: thanks to *Stargate SG-1, House MD, Blackadder* & *Red Dwarf*

Lejon-familjen - ... det här hade aldrig gått utan er... tack till mina älskade föräldrar: mamma *Kyllan* & pappa *Christer* (ni hade rätt igen, det gick det här också), och till mina underbara bröder *Tim* & *Robin* - för alla kramar, alla skratt, bästa filmsällskapen, för musiken; The Who, Led och gitarrplonkande... ni är bäst - jag älskar er!

Roger - for the last-minute kiss in Erice, for Tulum mornings, for being my sweetest thing, the love of my life, my best friend and bravest companion - without you I would never have made it... for never giving up on me, for always staying close to me, and for all your strength and love - jag älskar dig för alltid ♥

References

Abraham, E. P. (1990) Selective reminiscences of beta-lactam antibiotics: early research on penicillin and cephalosporins. *Bioessays* **12**: 601-606

Andersson, I., Terwisscha van Scheltinga, A. C. and Valegård, K. (2001) Towards new beta-lactam antibiotics. *Cell. Mol. Life Sci.* **58**: 1897-1906

Barends, T. R. M., Yoshida, H. and Dijkstra, B. W. (2004) Three-dimensional structures of enzymes useful for beta-lactam antibiotic production. *Curr. Opin. Biotechnol.* **15**: 356-363

Bhattacharya, A. A., Curry, S. and Franks, N. P. (2000) Binding of the general anesthetics propofol and halothane to human serum albumin. High resolution crystal structures. *J. Biol. Chem.* **275**: 38731-38738

Bhattacharya, A. A., Grüne, T. and Curry, S. (2000) Crystallographic analysis reveals common modes of binding of medium and long-chain fatty acids to human serum albumin. *J. Mol. Biol.* **303**: 721-732

Boniece, W. S., Wick, W. E., Holmes, D. H. and Redman, C. E. (1962) In vitro and in vivo laboratory evaluation of cephalothin, a new broad spectrum antibiotic. *J. Bacteriol.* **84**: 1292-1296

Born, T. L., Franklin, M. and Blanchard, J. S. (2000) Enzyme-catalysed acylation of homoserine: mechanistic characterisation of the Haemophilus influenzae met2-encoded homoserine transacetylase. *Biochemistry* **39**: 8556-8564

Brakhage, A. A. (1998) Molecular regulation of beta-lactam biosynthesis in filamentous fungi. *Microbiol. Mol. Biol. Rev.* **62**: 547-585

Brakhage, A. A., Al-Abdallah, Q., Tüncher, A. and Spröte, P. (2005) Evolution of beta-lactam biosynthesis genes and recruitment of trans-acting factors. *Phytochemistry* **66**: 1200-1210

Breithaupt, H. (1999) The new antibiotics. *Nat. Biotechnol.* **17**: 1165-1169

Brotzu, G. (1948) Richerche su di un nuovo antibiotico. *Lavori dell'Istituto d'Igiene di Cagliari* 1-11

Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, N., Pannu, N. J., Read, R. J., Rice, L. M., Simonson, T. and Warren, G. L. (1998) Crystallography and NMR system (CNS): A new software system for macromolecular structure determination. *Acta Crystallogr. D Biol. Crystallogr.* **54**: 905-921

Byford, M. F., Baldwin, J. E., Shiau, C. Y. and Schofield, C. J. (1997) The mechanism of ACV synthetase. *Chem. Rev.* **97**: 2631-2650

Chain, E., Florey, H. W., Gardner, D. M., Heatley, N. G., Jenning, M. A., Orr-Ewing, B. M. and Sanders, A. G. (1940) Penicillin as a chemotherapeutic agent. *Lancet* **236**: 226-228

Cistola, D. P. and Small, D. M. (1991) Fatty acid distribution in systems modeling the normal and diabetic human circulation. A 13C nuclear magnetic resonance study. *J. Clin. Invest.* **87**: 1431-1441

Collaborative Computational Project Number 4 (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **50**: 760-763

Colmenarejo, G., Alvarez-Pedraglio, A. and Lavandera, J. L. (2001) Cheminformatic models to predict binding affinities to human serum albumin. *J. Med. Chem.* **44**: 4370-4378

Cowtan, K. (1994) 'dm': An automated procedure for phase improvement by density modification. *Joint CCP4 + ESF-EAMCB Newsletter on Protein Crystallography* **31**: 34-38

Curry, S., Brick, P. and Franks, N. P. (1999) Fatty acid binding to human serum albumin: new insights from crystallographic studies. *Biochim. Biophys. Acta* **1441**: 131-140

Curry, S., Mandelkow, H., Brick, P. and Franks, N. (1998) Crystal structure of human serum albumin complexed with fatty acid reveals an asymmetric distribution of binding sites. *Nat. Struct. Biol.* **5**: 827-835

de Château, M. and Björck, L. (1996) Identification of interdomain sequences promoting the intronless evolution of a bacterial protein family. *Proc. Natl. Acad. Sci. U.S.A.* **93**: 8490-8495

de Château, M. and Björck, L. (1994) Protein PAB, a mosaic albumin-binding bacterial protein representing the first contemporary example of module shuffling. *J. Biol. Chem.* **269**: 12147-12151

de Château, M., Holst, E. and Björck, L. (1996) Protein PAB, an albumin-binding bacterial surface protein promoting growth and virulence. *J. Biol. Chem.* **271**: 26609-26615

de la Fortelle, E. and Bricogne, G. (1997) Maximum-likelihood heavy-atom parameter refinement for the multiple isomorphous replacement and multiwavelength anomalous diffraction methods. *Methods Enzymol.* **276**: 472-494

Demain, A. L. and Elander, R. P. (1999) The beta-lactam antibiotics: past, present, and future. *Antonie van Leeuwenhoek* **75**: 5-19

Elander, R. P. (2003) Industrial production of beta-lactam antibiotics. *Appl. Microbiol. Biotechnol.* **61**: 385-392

Falkenberg, C., Björck, L. and Åkerström, B. (1992) Localization of the binding site for streptococcal protein G on human serum albumin. Identification of a 5.5-kilodalton protein G binding albumin fragment. *Biochemistry* **31**: 1451-1457

Fleming, A. (1929) On the antibacterial action of Penicillium, with special reference to their use in the isolation of B. influenzae. *Brit. J. Exp. Pathol.* **10**: 226-236

Fredrickson, D. S. and Gordon, R. S. (1958) The metabolism of albumin-bound C14-labeled unesterified fatty acids in normal human subjects. *J. Clin. Invest.* **37**: 1504-1515

Ghuman, J., Zunszain, P. A., Petitpas, I., Bhattacharya, A. A., Otagiri, M. and Curry, S. (2005) Structural basis of the drug-binding specificity of human serum albumin. *J. Mol. Biol.* **353**: 38-52

Gunturi, S. B., Narayanan, R. and Khandelwal, A. (2006) In silico ADME modelling 2: computational models to predict human serum albumin binding affinity using ant colony systems. *Bioorg. Med. Chem.* **14**: 4118-4129

Gutiérrez, S., Velasco, J., Fernández, F. J. and Martín, J. F. (1992) The cefG gene of Cephalosporium acremonium is linked to the cefEF gene and encodes a deacetylcephalosporin C acetyltransferase closely related to homoserine O-acetyltransferase. *J. Bacteriol.* **174**: 3056-3064

Gutiérrez, S., Velasco, J., Marcos, A. T., Fernández, F. J., Fierro, F., Barredo, J. L., Díez, B. and Martín, J. F. (1997) Expression of the cefG gene is limiting for cephalosporin biosynthesis in Acremonium chrysogenum. *Appl. Microbiol. Biotechnol.* **48**: 606-614

Hamilton-Miller, J. M. T. (1999) Beta-lactams: variations on a chemical theme, with some surprising biological results. *J. Antimicrob. Chemother.* **44**: 729-734

He, X. M. and Carter, D. C. (1992) Atomic structure and chemistry of human serum albumin. *Nature* **358**: 209-215

He, Y., Chen, Y., Rozak, D. A., Bryan, P. N. and Orban, J. (2007) An artificially evolved albumin binding module facilitates chemical shift epitope mapping of GA domain interactions with phylogenetically diverse albumins. *Protein Sci.* **16**: 1490-1494

He, Y., Rozak, D. A., Sari, N., Chen, Y., Bryan, P. and Orban, J. (2006) Structure, dynamics and stability variation in bacterial albumin binding modules: implications for species specificity. *Biochemistry* **45**: 10102-10109

Hodgkin, D. C. and Maslen, E. N. (1961) The X-ray analysis of the structure of cephalosporin C. *Biochem. J.* **79**: 393-402

Holmquist, M. (2000) Alpha/beta-hydrolase fold enzymes: structures, functions and mechanisms. *Curr. Protein Pept. Sci.* **1**: 209-235

Honoré, B. and Brodersen, R. (1984) Albumin binding of anti-inflammatory drugs. Utility of a site-oriented versus a stoichiometric analysis. *Mol. Pharmacol.* **25**: 137-150

Johansson, M. U., de Château, M., Björck, L., Forsén, S., Drakenberg, T. and Wikström, M. (1995) The GA module, a mobile albumin-binding bacterial domain, adopts a three-helix-bundle structure. *FEBS Lett.* **374**: 257-261

Johansson, M. U., de Château, M., Wikström, M., Forsén, S., Drakenberg, T. and Björck, L. (1997) Solution structure of the albumin-binding GA module: a versatile bacterial protein domain. *J. Mol. Biol.* **266**: 859-865

Johansson, M. U., Frick, I.-M., Nilsson, H., Kraulis, P. J., Hober, S., Jonasson, P., Linhult, M., Nygren, P.-A., Uhlén, M., Björck, L., Drakenberg,

T., Forsén, S. and Wikström, M. (2002) Structure, specificity, and mode of interaction for bacterial albumin-binding modules. *J. Biol. Chem.* **277**: 8114-8120

Johansson, M. U., Nilsson, H., Evenäs, J., Forsén, S., Drakenberg, T., Björck, L., and Wikström, M. (2002) Differences in backbone dynamics of two homologous bacterial albumin-binding modules: implications for binding specificity and bacterial adaptation. *J. Mol. Biol.* **316**: 1083-1099

Jones, T. A., Zou, J.-Y., Cowan, S. W. and Kjeldgaard, M. (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* **47**: 110-119

Kleywegt, G. J. (1996) Use of non-crystallographic symmetry in protein structure refinement. *Acta Crystallogr. D Biol. Crystallogr.* **52**: 842-857

Kleywegt, G. J., Henrick, K., Dodson, E. J. and van Aalten, D. M. (2003) Pound-wise but penny-foolish: how well do micromolecules fare in macromolecular refinement? *Structure* **11**: 1051-1059

Kleywegt, G. J. and Jones, T. A. (1994) A super position. *CCP4/ESF-EACBM Newsletter on Protein Crystallography* **31**: 9-14

Kober, A. and Sjöholm, I. (1980) The binding sites on human serum albumin for some nonsteroidal antiinflammatory drugs. *Mol. Pharmacol.* **18**: 421-426

Kragh-Hansen, U., Chuang, V. T. G. and Otagiri, M. (2002) Practical aspects of the ligand-binding and enzymatic properties of human serum albumin. *Biol. Pharm. Bull.* **25**: 695-704

Langer, K., Anhorn, M. G., Steinhauser, I., Dreis, S., Celebi, D., Schrickel, N., Faust, S. and Vogel, V. (2007) Human serum albumin (HSA) nanoparticles: reproducibility of preparation process and kinetics of enzymatic degradation. *Int. J. Pharm.* **347**: 109-117

Livermore, D. M. (2000) Epidemiology of antibiotic resistance. *Intensive Care Med.* **26**: 14-21

Mao, H., Hajduk, P. J., Craig, R., Bell, R., Borre, T. and Fesik, S. W. (2001) Rational design of diflunisal analogues with reduced affinity for human serum albumin. *J. Am. Chem. Soc.* **123**: 10429-10435

Martín, J. F. and Demain, A. L. (2002) Unraveling the methionine-cephalosporin puzzle in Acremonium chrysogenum. *Trends Biotechnol.* **20**: 502-507

Mathison, L., Soliday, C., Stepan, T., Aldrich, T. and Rambosek, J. (1993) Cloning, characterisation, and use in strain improvement of the Cephalosporium acremonium gene cefG encoding acetyl transferase. *Curr. Genet.* **23**: 33-41

Matsuda, A., Sugiura, H., Matsuyama, K., Matsumoto, H., Ichikawa, S. and Komatsu, K. (1992) Molecular cloning of acetyl coenzyme A: deacetylcephalosporin C o-acetyltransferase cDNA from Acremonium chrysogenum: sequence and expression of catalytic activity in yeast. *Biochem. Biophys. Res. Commun.* **182**: 995-1001

Matthews, B. W. (1968) Solvent content of protein crystals. *J. Mol. Biol.* **33**: 491-497

McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. and Read, R. J. (2007) Phaser crystallographic software. *J. Appl. Cryst.* **40**: 658-674

Meya, D. B. and McAdam, K. P. (2007) The TB pandemic: an old problem seeking new solutions. *J. Intern. Med.* **261**: 309-329

Murshudov, G. N., Vagin, A., A. and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.* **53**: 240-255

Myhre, E. B. and Kronvall, G. (1980) Demonstration of specific binding sites for human serum albumin in group C and G streptococci. *Infect. Immun.* **27**: 6-14

Myhre, E. B. and Kronvall, G. (1977) Heterogeneity of nonimmune immunoglobulin Fc reactivitiy among gram-positive cocci: description of three major types of receptors for human immunoglobulin G. *Infect. Immun.* **17**: 475-482

Navarre, W. W. and Schneewind, O. (1999) Surface proteins of Gram-positive bacteria and mechanisms of their targeting to the cell wall envelope. *Microbiol. Mol. Biol. Rev.* **63**: 174-229

Nerli, B. and Pico, G. (1993) Identification of the cephalosporin human serum albumin binding sites. *Pharmacol. Toxicol.* **73**: 297-300

Nicholas, R. A., Lamson, D. R. and Schultz, D. E. (1993) Penicillin-binding protein 1B from Escherichia coli contains a membrane association site in addition to its transmembrane anchor. *J. Biol. Chem.* **268**: 5632-5641

Ollis, D. L., Cheah, E., Cygler, M., Dijkstra, B. W., Frolow, F., Franken, S. M., Harel, M., Remington, S. J., Silman, I., Schrag, J., Sussman, J., Verschueren, K. H. G. and Goldman, A. (1992) The alpha/beta hydrolase fold. *Protein Eng.* **5**: 197-211

Petitpas, I., Bhattacharya, A. A., Twine, S., East, M. and Curry, S. (2001) Crystal structure analysis of warfarin binding to human serum albumin: anatomy of drug site I. *J. Biol. Chem.* **276**: 22804-22809

Petitpas, I., Grüne, T., Bhattacharya, A. A. and Curry, S. (2001) Crystal structures of human serum albumin complexed with monounsaturated and polyunsaturated fatty acids. *J. Mol. Biol.* **314**: 955-960

Petitpas, I., Petersen, C. E., Ha, C. E., Bhattacharya, A. A., Zunszain, P. and Ghuman, J. (2003) Structural basis of albumin-thyroxine interactions and familial dysalbuminemic hyperthyroxinemia. *Proc. Natl. Acad. Sci. U.S.A.* **100**: 6440-6445

Reitsma, W. D. (1967) The relationship between serum free fatty acids and blood sugar in non-obese and obese diabetics. *Acta Med. Scand.* **182**: 353-361

Roach, P. L., Clifton, I. J., Hensgens, C. M., Shibata, N., Schofield, C. J., Hajdu, J. and Baldwin, J. E. (1997) Structure of isopenicillin N synthase complexed with substrate and the mechanism of penicillin formation. *Nature* **387**: 827-830

Rodríguez-Sáiz, M., Lembo, M., Bertetti, L., Muraca, R., Velasco, J., Malcangi, A., de la Fuente, J. L. and Barredo, J. L. (2004) Strain improvement for cephalosporin production by Acremonium chrysogenum us-

ing geneticin as a suitable transformation marker. *FEMS Microbiol. Lett.* **235**: 43-49

Rossolini, G. M. (2007) Redesigning beta-lactams to combat resistance: summary and conclusions. *Clin. Microbiol. Infect.* **13 (suppl 2)**: 30-33

Rozak, D. A., Alexander, P. A., He, Y., Chen, Y., Orban, J. and Bryan, P. N. (2006) Using offset recombinant polymerase chain reaction to identify functional determinants in a common family of bacterial albumin binding domains. *Biochemistry* **45**: 3263-3271

Saiakhov, R. D., Stefan, L. R. and Klopman, G. (2000) Multiple computer-automated structure evaluation model of the plasma protein binding affinity of diverse drugs. *Perspectives Drug Disc. Des.* **19**: 133-155

Schneider, T. R. and Sheldrick, G. M. (2002) Substructure solution with SHELXD. *Acta Crystallogr. D Biol. Crystallogr.* **58**: 1772-1779

Simard, J. R., Zunszain, P., Hamilton, J. A. and Curry, S. (2006) Location of high and low affinity fatty acid binding sites on human serum albumin revealed by NMR drug-competition analysis. *J. Mol. Biol.* **361**: 336-351

Sjöholm, I., Ekman, B., Kober, A., Ljungstedt-Påhlman, I., Seiving, B. and Sjödin, T. (1979) Binding of drugs to human serum albumin: XI. The specificity of three binding sites as studied with albumin immobilized in microparticles. *Mol. Pharmacol.* **16**: 767-777

Stinchcombe, T. E. (2007) Nanoparticle albumin-bound paclitaxel: a novel Cremphor-EL-free formulation of paclitaxel. *Nanomed.* **2**: 415-423

Sudlow, G., Birkett, D. J. and Wade, D. N. (1975) The characterization of two specific drug-binding sites on human serum albumin. *Mol. Pharmacol.* **11**: 824-832

Sudlow, G., Birkett, D.J., and Wade, D. N. (1976) Further characterization of specific drug binding sites on human serum albumin. *Mol. Pharmacol.* **12**: 1052-1061

Sugio, S., Kashima, A., Mochizuki, S., Noda, M. and Kobayashi, K. (1999) Crystal structure of human serum albumin at 2.5 Å resolution. *Protein Eng.* **12**: 439-446

Teiling, E. (1965) Richard Westling, discoverer of the penicillium and its miraculous action. *Sven. Farm. Tidskr.* **69**: 831-836

Tilley, L., Davis, T. M. E. and Bray, P. G. (2006) Prospects for the treatment of drug-resistant malaria parasites. *Future Microbiol.* **1**: 127-141

Ullán, R. V., Casqueiro, J., Banuelos, O., Fernández, F. J., Gutiérrez, S. and Martín, J. F. (2002) A novel epimerization system in fungal secondary metabolism involved in the conversion of isopenicillin N into penicillin N in Acremonium chrysogenum. *J. Biol. Chem.* **277**: 46216-46225

Vagin, A. A., Murshudov, G. N. and Strokopytov, B. V. (1998) BLANC: the program suite for protein crystallography. *J. Appl. Cryst.* **31**: 98-102

van Duyne, G. D., Standaert, R. F., Karplus, P. A., Schreiber, S. L. and Clardy, J. (1993) Atomic structures of the human immunophilin FKBP-12 complexes with FK506 and rapamycin. *J. Mol. Biol.* **229**: 105-124

Velasco, J., Gutiérrez, S., Campoy, S. and Martín, J. F. (1999) Molecular characterization of the Acremonium chrysogenum cefG gene product: the native deacetylcephalosporin C acetyltransferase is not processed into subunits. *Biochem. J.* **337**: 379-385

Velasco, J., Gutiérrez, S., Casqueiro, J., Fierro, F., Campoy, S. and Martín, J. F. (2001) Cloning and characterization of the gene cahB encoding a cephalosporin C acetylhydrolase from Acremonium chrysogenum. *Appl. Microbiol. Biotechnol.* **57**: 350-356

Velasco, J., Gutiérrez, S., Fernández, F. J., Marcos, A. T., Arenos, C. and Martín, J. F. (1994) Exogenous methionine increases levels of mRNAs transcribed from pcbAB, pcbC and cefEF genes, encoding enzymes of the cephalosporin biosynthetic pathway, in Acremonium chrysogenum. *J. Bacteriol.* **176**: 985-991

Vonrhein, C., Blanc, E., Roversi, P. and Bricogne, G. (2006) Automated structure solution with autoSHARP. *Methods Mol. Biol.* **364**: 215-230

Wang, M., Liu, L., Wang, Y., Wei, Z., Zhang, P., Li, Y., Jiang, X., Xu, H. and Gong, W. (2007) Crystal structure of homoserine O-acetyltransferase from Leptospira interrogans. *Biochem. Biophys. Res. Commun.* **363**: 1050-1056

Wardell, M., Wang, Z., Ho, J. X., Robert, J., Ruker, F., Ruble, J. and Carter, D. C. (2002) The atomic structure of human methemalbumin at 1.9 Å. *Biochem. Biophys. Res. Commun.* **291**: 813-819

Winn, M. D., Isupov, M. N. and Murshudov, G. N. (2001) Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Crystallogr. D Biol. Crystallogr.* **57**: 122-133

Winn, M. D., Murshudov, G. N. and Papiz, M. Z. (2003) Macromolecular TLS refinement in REFMAC at moderate resolutions. *Methods Enzymol.* **374**: 300-321

Zunszain, P., Ghuman, J., Komatsu, T., Tsuchida, E. and Curry, S. (2003) Crystal structural analysis of human serum albumin complexed with hemin and fatty acid. *BMC Struct. Biol.* **3**:

Acta Universitatis Upsaliensis

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology 382

Editor: The Dean of the Faculty of Science and Technology

A doctoral dissertation from the Faculty of Science and Technology, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology. (Prior to January, 2005, the series was published under the title "Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology".)

Distribution: publications.uu.se
urn:nbn:se:uu:diva-8399

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2008