REVIEW ARTICLE

Cation–π bonding and amino–aromatic interactions in the biomolecular recognition of substituted ammonium ligands

Nigel S. SCRUTTON*‡ and Andrew R.C. RAINE†

* Department of Biochemistry, University of Leicester, Adrian Building, University Road, Leicester LE1 7RH, and †Cambridge Centre for Molecular Recognition, Department of Biochemistry, University of Cambridge, Tennis Court Road, Cambridge CB2 1QW, U.K.

Cation–π bonds and amino–aromatic interactions are known to be important contributors to protein architecture and stability, and their role in ligand–protein interactions has also been reported. Many biologically active amines contain substituted ammonium moieties, and cation–π bonding and amino–aromatic interactions often enable these molecules to associate with proteins. The role of organic cation–π bonding and amino–aromatic interactions in the recognition of small-molecule amines and peptides by proteins is an important topic for those involved in structure-based drug design, and although the number of structures determined for proteins displaying these interactions is small, general features are beginning to emerge. This review explores the role of cation–π bonding and amino–aromatic interactions in the biological molecular recognition of amine ligands. Perspectives on the design of ammonium-ligand-binding sites are also discussed.

INTRODUCTION

Amines are an extensive group of small molecules that are widespread in biology. They include the large and growing family of synthetic therapeutic and abused drugs, as well as several physiological amines (e.g. histamine, catecholamines and acetylcholine), peptides and 'protein elements' involved in a variety of signalling processes. Other physiological amines of direct clinical relevance include creatinine (a marker of kidney dysfunction) and creatine (released following myocardial infarction and in certain degenerative muscle diseases). Given the widespread occurrence and use of substituted ammonium compounds in biology and medicine, a study of their recognition by target receptor proteins is essential to our understanding of small-molecule-receptor signalling at the atomic level. A fuller understanding of the interactions of these compounds with proteins will impact directly on the discipline of structure-based drug design. By identifying generic features in the molecular recognition of ammonium groups by proteins, a more focused approach to the design of drugs and new target proteins will result.

Recent years have witnessed a steady growth in the number of determined structures of proteins and protein–ligand complexes. As data accumulate, it is becoming increasingly apparent that proteins can recognize ammonium cations and amino groups via interactions with aromatic side chains. Commensurately, there has been renewed interest in these interactions from both the chemical and biological viewpoints [1]. Two types of interaction are recognized: (i) the interaction of an organic cation (e.g. a protonated or quaternary amine bearing a formal positive charge) with aromatic side chains, termed a cation–π interaction, and (ii) the interaction of an amino group bearing a $\delta(+)$ charge with aromatic side chains, here termed an amino–aromatic interaction. Cation–π and amino–aromatic interactions are prominent in protein–small-molecule recognition, ranging from the interaction of receptors with transmitters/drugs to the recognition of substrates by enzymes, and they also contribute to the stability of the folded protein.

This review sets out to illustrate the importance of amino–aromatic/cation–π interactions in biological molecular recognition, and to heighten awareness of the need to involve aromatic residues in the design of substituted ammonium ligand-binding sites rather than relying on the more conventional ion-pair interactions seemingly favoured by protein engineers. Following brief comments on cation–π and amino–aromatic interactions in chemical systems and within protein structures, the emphasis is a structural one focusing on protein–protein, protein–peptide and protein–small-ligand associations that employ cation–π/amino–aromatic interactions. Comments on the design of ammonium ligand-binding sites are also included.

CHEMICAL SYSTEMS AND INTERACTIONS WITHIN PROTEINS

Cation–π bonding in chemical systems

Chemical studies of cation–π bonding were reported in 1986 as unconventional 'ionic' bonding between substituted ammonium ligands and π-donors, as seen in the clustering reactions of NH_4^+ and $MeNH_3^+$ with benzene derivatives [2]. Experimental results and *ab initio* calculations indicated that the interaction is predominantly electrostatic and that there is no π donation into the bond. Interaction energies were found to range from approx. 42 to 92 $kJ \cdot mol^{-1}$ (10 to 22 $kcal \cdot mol^{-1}$). For interactions of NH_4^+ with π-dimers (benzene or fluorobenzene), two of the four NH_4^+ hydrogens project towards the aromatic rings [2]. More recently, a combined quantum mechanical and molecular mechanical Monte Carlo simulation of the binding of the tetramethylammonium ion and benzene in water suggested that

Abbreviation used: SH2 domain, *src* homology 2 domain.

‡ To whom correspondence should be addressed.

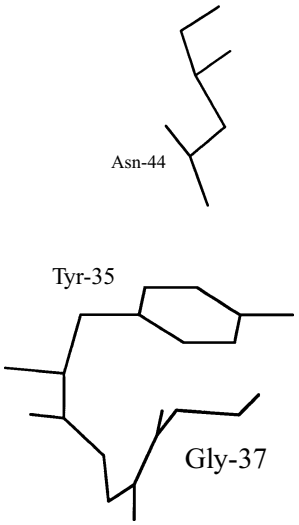

Figure 1 The amino–aromatic interaction identified by Burley and Petsko [8] in bovine pancreatic trypsin inhibitor

The backbone HN of Gly-37 and the side-chain NH_2 of Asn-44 interact with opposite faces of the aromatic ring of Tyr-35. The co-ordinates were taken from the Brookhaven Protein Databank, entry 4PTI. This and all other molecular drawings were made with the Molscript program [60].

organic-cation–π bonds are more stable than ion-pair interactions between tetra-alkylammonium ions and anionic residues in aqueous solution [3], and calculations reveal that the tetramethylammonium ion binds to benzene with an association constant of approx. 0.8 M^{-1}. In purely chemical systems, molecular recognition by cation–π interactions is now enjoying widespread study. Various synthetic host–guest systems displaying cation–π bonding [4,5], and cationic transition-state stabilization via cation–π interactions, have been reported [6,7]. The stabilization energies that cation–π interactions contribute in the biological context are expected to be less than those for the purely chemical systems; this is a necessary consequence of the fact that stabilization in proteins is more likely to be contributed by different, and fewer, aromatic residues and is also affected by geometrical constraints and the properties of the protein interior. Nonetheless, the cation–π bond is a favourable interaction in biology, and is used by proteins that interact with protonated or substituted ammonium moieties.

Interactions within proteins

For some time, cation–π/amino–aromatic interactions within protein structures have been recognized as important associations that contribute to the protein structure. In an early study, Burley and Petsko [8] performed a geometric analysis of 33 high-resolution [2 Å (0.2 nm) resolution or better] protein structures. They were able to show a statistical preference for positively charged or $\delta(+)$ amino groups of lysine, arginine, asparagine, glutamine and histidine to pack within 6 Å of the centroids of phenylalanine, tyrosine and tryptophan. In this location, the amino groups make contact with the $\delta(-)$ π-electrons of the aromatic rings in these residues. The amino groups and the aromatic side chains are preferentially separated by between 3.4 Å and 6 Å, and they avoid contact with the $\delta(+)$ edge of the aromatic ring. An early, and now seminal, example of an amino–aromatic interaction within protein molecules is seen in bovine pancreatic trypsin inhibitor, where the aromatic ring of residue Tyr-35 is sandwiched between the peptide amide proton of Gly-37 and the primary amide proton of Asn-44 (Figure 1; [9,10]). The interaction is seen both in the crystal structure of bovine pancreatic trypsin inhibitor and in those structures calculated from NMR data. In the latter, the chemical shifts of the Asn-44 primary amide proton and the peptide amide proton of Gly-37 are shifted upfield by the ring current of the Tyr-35 side chain [10]. The abnormal chemical shifts confirm that the protons interact with the π-electron cloud of Tyr-35 [10]. Amino–aromatic interactions involving amide protons of asparagine residues have also been observed in the crystal structure of haemoglobin bound to the drug bezafibrate [11]. In this case, an amide proton of Asn-108 from the β_1 subunit of haemoglobin is positioned to make an unconventional hydrogen bond with the π-electrons of the first aromatic nucleus of bezafibrate.

Burley and Petsko's early analysis of amino–aromatic interactions within protein molecules has been extended by Thornton and co-workers [12,13]. Their analysis focused on sp^2 hybridized nitrogen atoms (i.e. no formal positive charge) within protein structures. Through an analysis of 55 high-resolution structures, the results of Burley and Petsko were confirmed in that a statistical preference for the positioning of $\delta(+)$ amino groups within 6 Å of the centroids of aromatic side chains exists. Additionally, Mitchell and co-workers [13] were able to define two types of amino–aromatic interaction, i.e. those in which the sp^2 nitrogen atoms form stacking interactions with the aromatic rings and, secondly, those geometries which give rise to amino–aromatic hydrogen bonds, as proposed by Levitt and Perutz [14] and subsequently elaborated [15]. The former interaction is preferred by a factor of about 2.5:1, even though *ab initio* calculations of the gas-phase interaction energies for model systems favour the amino–aromatic hydrogen bond over stacking interactions. The reason for this disparity between gas-phase behaviour and interactions formed within protein structures is that, in stacked geometries, the nitrogen-containing groups form conventional (non-amino–aromatic) hydrogen bonds with neighbouring groups in the protein or solvent; these additional hydrogen bonds, which are absent from model gas-phase studies, are sufficient in the majority of cases to stabilize the stacked geometry within protein molecules. A similar analysis of the stacking interactions of arginine and aromatic side chains within protein structures has also been advanced by Flocco and Mowbray [16].

MOLECULAR RECOGNITION OF SMALL LIGANDS AND SUBSTRATES

Acetylcholinesterase: an archetype for cation–π bonding in biology?

Acetylcholinesterase is often considered as the foremost example of cation–π bonding in biological molecular recognition. In its interaction with acetylcholine, it serves as an excellent model for the recognition of quaternary amines by proteins. Early kinetic, spectroscopic and chemical modification studies [17] suggested that the active site of acetylcholinesterase is divided into two subsites: the 'esteratic' site (the site of bond breaking/making) and the 'anionic' (choline binding) site. The 'anionic' site is a misnomer, as this site is in fact uncharged and lipophilic. The molecular detail of acetylcholinesterase was revealed following the determination of the crystal structure of the enzyme from *Torpedo californicans* [18]. A structure for the enzyme–substrate complex is not available, but the details of substrate binding can be extrapolated from the structure of the enzyme alone [18] and those of the enzyme complexed with tacrine, edrophonium and decamethonium [19].

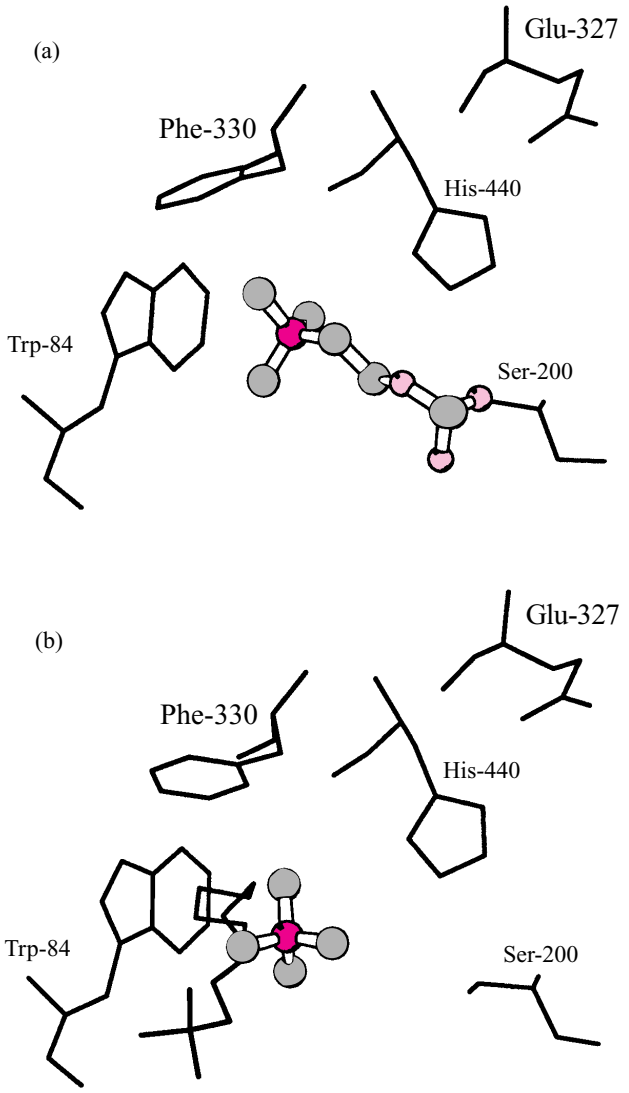

Figure 2 (a) Proposed binding of acetylcholine to acetylcholinesterase, and (b) binding of decamethonium to acetylcholinesterase

(**a**) Acetylcholine has been built into the crystal structure of unliganded *Torpedo californicans* acetylcholinesterase (modelled structure taken from Brookhaven entry 1ACE). The $N(CH_3)_3$ group is positioned to make a classic cation–π bond with the side chain of Trp-84, and is a similar distance away from the ring of Phe-330. Ser-200, His-440 and Glu-327 form a catalytic triad that is the mirror image of those seen in serine proteinases. (**b**) In the structure of acetylcholinesterase complexed with decamethonium (modelled structure taken from Brookhaven entry 1ACL), one of the terminal $N(CH_3)_3$ groups of decamethonium binds in the same position as the proposed choline-binding site in (**a**).

Access to the active site of acetylcholinesterase is via a deep and narrow gorge lined by 14 aromatic residues (making up about 40% of the surface of the gorge) and other residues. The gorge is 20 Å in length and the residues comprising the surface of the gorge are highly conserved in acetylcholinesterases from different species. Manual docking of the substrate acetylcholine at the base of the gorge reveals the esteratic and choline-binding sites. Associated with the esteratic site, a catalytic triad and putative oxyanion hole have been identified, and modelling of the choline moiety of acetylcholine suggests that it forms a cation–π bond with Trp-84 in the 'anionic' site (Figure 2). Perhaps the most remarkable feature of acetylcholinesterase is the preponderance of aromatic residues in the active-site gorge. The chemical character of the gorge leads to the question of its function in contributing to the rapid rate of substrate binding and catalysis. Sussman and colleagues suggested two mechanisms by which the on-rate for ligand binding might be increased [18]. First, the high hydrophobicity of the gorge produces a low dielectric constant in the gorge. The effect is to enhance the effective local charge contributed by the small number of acidic groups in the vicinity of the gorge, which electrostatically 'steer' substrate to the active site. In the second scenario, the aromatic lining acts as a series of low-affinity sites for the substrate (in particular, the choline moiety), and guides the trapped substrate to the active site. Because of the reduction-in-dimensionality, the rate of substrate binding is increased. Relatively weak cation–π interactions may, therefore, have a major role to play in directing the substrate towards the productive enzyme–substrate complex, whereas stronger cation–π bonding is presumably responsible for binding the choline moiety of acetylcholine in the enzyme–substrate complex. Given the wealth of cation–π interactions found in acetylcholinesterase, the enzyme no doubt will remain a principal target for investigating these interactions in biological macromolecules. Interestingly, chemical modification studies of the nicotinic acetylcholine receptor have also suggested that aromatic residues are located in the acetylcholine-binding site in this molecule [20,21].

Substituted 'small-molecule' ammonium cations

Cation–π bonding of quaternary ammonium ligands has been visualized in other structurally determined protein molecules. The McPC603 Fab, a mouse myeloma IgA (κ) that binds phosphocholine, was the first determined protein structure found to bind a quaternary ammonium moiety through the use of aromatic residues [22,23]. The phosphocholine hapten is in contact with only a small number of residues; these include side chains from the three heavy-chain hypervariable regions and one light-chain region. Glu-35 (heavy chain), Tyr-100 (light chain) and Trp-107 (heavy chain) form the choline-binding site. The first residue makes a hydrogen bond with the phenolic hydroxy group of Tyr-100L and in so doing orients the side chain of Tyr-100L to make a cation–π bond with the choline group of the hapten (Figure 3). Trp-107 also makes a good cation–π bond with the quaternary ammonium group. The quaternary ammonium-binding site is revealed directly in the crystal structure of the McPC603 protein–ligand complex, and the protein is therefore an excellent model for investigating the structure of quaternary ammonium-binding sites. But what of tertiary and secondary ammonium ligands? Many bioactive amines contain these groups, and consequently structural models of binding sites for these molecules are also required.

The simple tertiary and secondary amines tri- and di-methylamine are bound by the bacterial proteins trimethylamine dehydrogenase and dimethylamine dehydrogenase respectively. Both proteins are found in methylotrophic bacteria, where they are responsible for the ability of these bacteria to grow on trimethylamine and dimethylamine as the sole carbon source [24]. The crystal structure of trimethylamine dehydrogenase has been determined at 2.4 Å resolution [25] and recently refined at 1.7 Å (S. A. White and F. S. Mathews, unpublished work). More importantly, the structures of trimethylamine dehydrogenase in complex with the substrate inhibitor tetramethylammonium chloride or the substrate trimethylamine (as the protonated ammonium cation trimethylammonium) have demonstrated that these molecules are bound by cation–π bonding in an 'aromatic bowl' [26,27] (Figure 4). The bowl, comprising the three residues Tyr-60, Trp-264 and Trp-355, positions trimethylamine in the

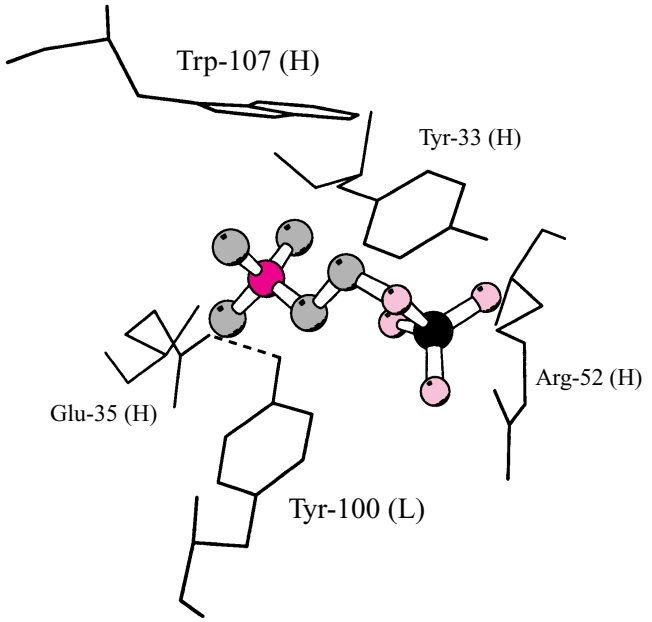

Figure 3 Phosphocholine-binding site of the McPC603 Fab

Phosphocholine is shown in ball-and-stick representation, with nitrogen dark pink and phosphorus black, while the protein residues forming the binding site are in stick representation. The choline group of the hapten is in contact with the aromatic side chains of Trp-107 (heavy chain; H) and Tyr-100 (light chain; L). Tyr-33 (H) and Arg-52 (H) interact with the phosphate. Co-ordinates were taken from Brookhaven entry 2MCP.

active site so that the substrate is located close to the N-5 atom of the enzyme-bound FMN. During catalysis, the N-5 atom is involved in the formation of a covalent enzyme–substrate intermediate [28]. In trimethylamine dehydrogenase, tertiary ammonium cations are bound in a similar fashion to that previously described for quaternary ammonium ions, and the enzyme's aromatic bowl can therefore serve as a model structure for capturing tertiary ammonium groups in proteins.

Biochemically, dimethylamine dehydrogenase is closely related to trimethylamine dehydrogenase [29], and this similarity is supported by the very high degree of sequence identity (63%) between the two proteins [30,31]. On the basis of this sequence identity, a model for the structure of dimethylamine dehydrogenase has been constructed using the crystal co-ordinates of trimethylamine dehydrogenase [27], to identify the structural changes that direct the binding of dimethylamine in dimethylamine dehydrogenase. The model of dimethylamine dehydrogenase revealed that the active sites of the two proteins are almost completely identical. Those residues involved in demethylation of substrate are totally conserved; the only change is the exchange of Tyr-60 in trimethylamine dehydrogenase (a residue of the substrate-binding aromatic bowl) for a glutamine residue in dimethylamine dehydrogenase [27,31]. On placing dimethylamine in the active site of the model, Gln-60 is ideally positioned to make a conventional hydrogen bond from the side-chain amide carbonyl to the NH hydrogen of dimethylamine (Figure 4), suggesting that this single residue change is responsible for the switch in ammonium cation specificity between the two dehydrogenases. The two remaining methyl groups of dimethylamine are positioned to make cation–π bonds with the two tryptophan residues in the same way as they do in trimethylamine dehydrogenase. In trimethylamine dehydrogenase there is no requirement for the specific orientation of substrate in the

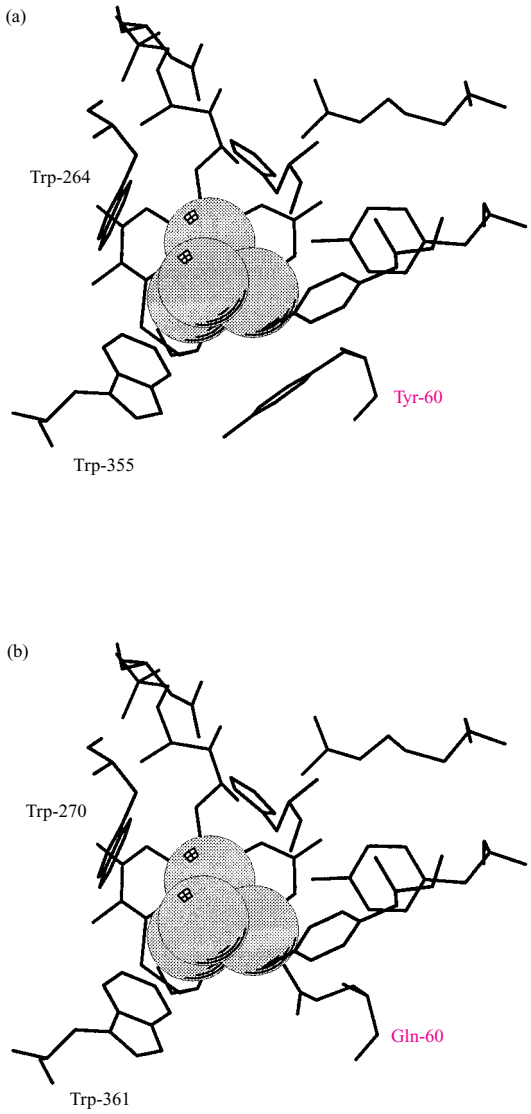

Figure 4 (a) Substrate-binding bowl in trimethylamine dehydrogenase, with tetramethylammonium (an inhibitor) bound, and (b) proposed mode of binding of the closely related dimethylamine dehydrogenase

(**a**) The bowl is formed by the aromatic side chains of Trp-264, Trp-355 and Tyr-60, which form cation–π bonds with the methyl groups of the inhibitor or, when the substrate trimethylamine is bound, with the three methyl groups of the substrate. (**b**) The only active-site difference between this enzyme and trimethylamine dehydrogenase is the replacement of Tyr-60 by a Gln. It is predicted that Trp-264 and Trp-355 will still form cation–π bonds with the two remaining methyl groups on the substrate, with Gln-60 forming a hydrogen bond with the N-H of dimethylamine.

aromatic bowl; trimethylamine has C-3 rotational symmetry, and any of the methyl groups can be oxidized during catalysis. In dimethylamine dehydrogenase the substrate is bound in a specific way, and this specificity is acquired through the provision of a conventional hydrogen bond to the substrate, ensuring that one of the methyl groups can be oxidized during catalysis.

Good models for cation–π bonding of primary ammonium cations by proteins are still lacking, but the hope is that bacterial methylamine dehydrogenase will prove to be a good example. Crystal structures exist for two methylamine dehydrogenases, one for the *Paracoccus denitrificans* enzyme [32] and the other for the *Thiobacillus versutus* enzyme [33,34]. The active-site regions

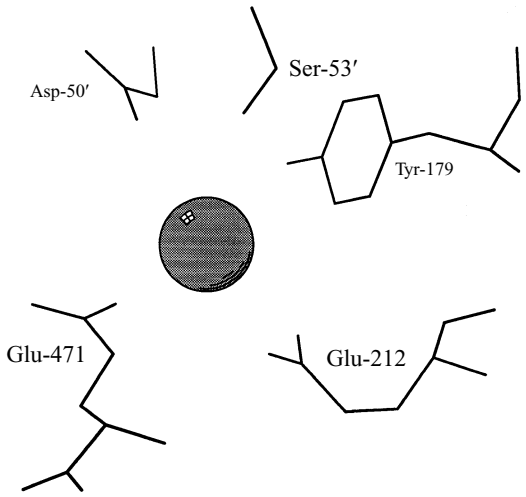

Figure 5 Ammonium-binding site in glutamine synthetase

An NH_4^+ ion has been built into the co-ordinates of Brookhaven entry 2GLS in the position at which electron density was seen for caesium and thallium ions. No energy minimization was performed after placing the ammonium ion into the ammonium-binding site of glutamine synthetase. See the text for further details.

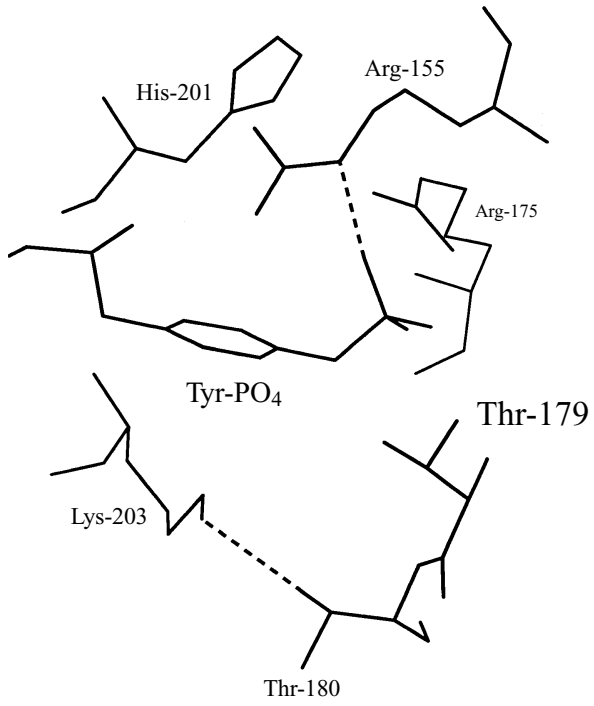

Figure 6 Detail of the binding of phosphotyrosine to the SH2 domain of the v-*src* oncogene product

One of the side-chain NH_2 groups of Arg-155 and the side-chain NH_3^+ of Lys-203 interact with the aromatic ring of the phosphotyrosine. Arg-175, Arg-155 and possibly Thr-179 interact with the phosphate group. Co-ordinates were taken from Brookhaven entry 1SHA.

of both are spatially defined, but to date only complete sequence information is available for the enzyme from *Paracoccus denitrificans* [35]. A recent refinement of this enzyme at 1.75 Å resolution indicates that the side chains of a phenylalanine (Phe-42; large subunit) and a tyrosine (Tyr-119; small subunit) residue are located in the conjectured methylamine-binding site along with an aspartate residue (F. S. Mathews, personal communication). These amino acid residues are close to the O-6 atom of the tryptophan tryptophylquinone redox cofactor of the enzyme, where the substrate forms a covalent intermediate in catalysis. A structure for methylamine dehydrogenase solved with bound substrate is lacking, but the presence of two aromatic side chains in the methylamine-binding site is highly suggestive that cation–π bonding is responsible (at least in part) for substrate recognition.

Recently, the ammonium ion-binding site has been localized in the crystal structure of glutamine synthetase [36]. Assignment of the site was made possible by performing crystal soaks of glutamine synthetase with caesium chloride and thallium acetate, the cations residing in the ammonium ion-binding site. The co-ordination shell for the ammonium ion comprises oxygen atoms donated from the γ-carboxylate of the substrate glutamate, the side-chain carboxylates of two acidic residues (Glu-212 and Asp-50′), the hydroxy group of Ser-53′ and the aromatic side chain of Tyr-179 (Figure 5). Although, in this analysis, electron densities for thallium and caesium ions were observed in the difference-density maps of glutamine synthetase, the implication is that Tyr-179 forms a cation–π interaction with the physiological substrate, the ammonium ion. Therefore, as seen for dimethylamine dehydrogenase, a mixture of cation–π bonding and more conventional bonding may well be responsible for the binding of ammonium ions in glutamine synthetase.

PROTEIN–PROTEIN AND PROTEIN–PEPTIDE RECOGNITION

Recognition of phosphotyrosine-containing peptides/proteins by the v-*src* oncogene product

Cation–π and amino–aromatic interactions are observed in protein–protein and protein–peptide recognition, and perhaps the most elegant example in this category is the recognition of tyrosine-phosphorylated peptides by the *src* homology 2 (SH2) domain of the v-*src* oncogene product. Tyrosine kinases occupy a pivotal position in signal transduction pathways [37] and, following phosphorylation of tyrosine residues, signals are transmitted by recognition of the phosphorylated tyrosines by downstream proteins containing SH2 domains. These domains reside in a variety of cytosolic non-receptor tyrosine kinases and other proteins [38]. An understanding of the molecular details of the interaction between phosphorylated-tyrosine-containing proteins and SH2 domains is, therefore, central to our understanding of signal transduction mechanisms in the cell.

Crystal structures of the v-*src* oncogene product complexed with two tyrosine-phosphorylated peptides have been solved at 1.5 Å and 2.0 Å resolution by X-ray crystallography [39]. The SH2 domain comprises an antiparallel β-sheet flanked by two α-helices, and the overall fold resembles a flattened hemisphere the face of which provides the surface for peptide binding. The phosphotyrosine residue of the bound peptide in each structure is located in a small cleft formed by residue Arg-155, the side chains of Lys-203 and His-201 and a phosphate-binding loop. Three positively charged residues (Arg-155, Arg-175 and Lys-203) are located in the phosphotyrosine-binding site (Figure 6). Arg-175 forms a conventional ion-pair interaction with two of the oxygens of phosphate. The remaining two residues (Arg-155 and Lys-203) interact with the aromatic nucleus of the phosphotyrosine. In addition to forming an optimal interaction with the aromatic ring of the phosphotyrosine, Arg-155 also forms a hydrogen bond with the phosphate of the same residue. Lys-203 is located on the opposite face of the phosphotyrosine, and is tethered by a hydrogen bond from Thr-180. In this position the

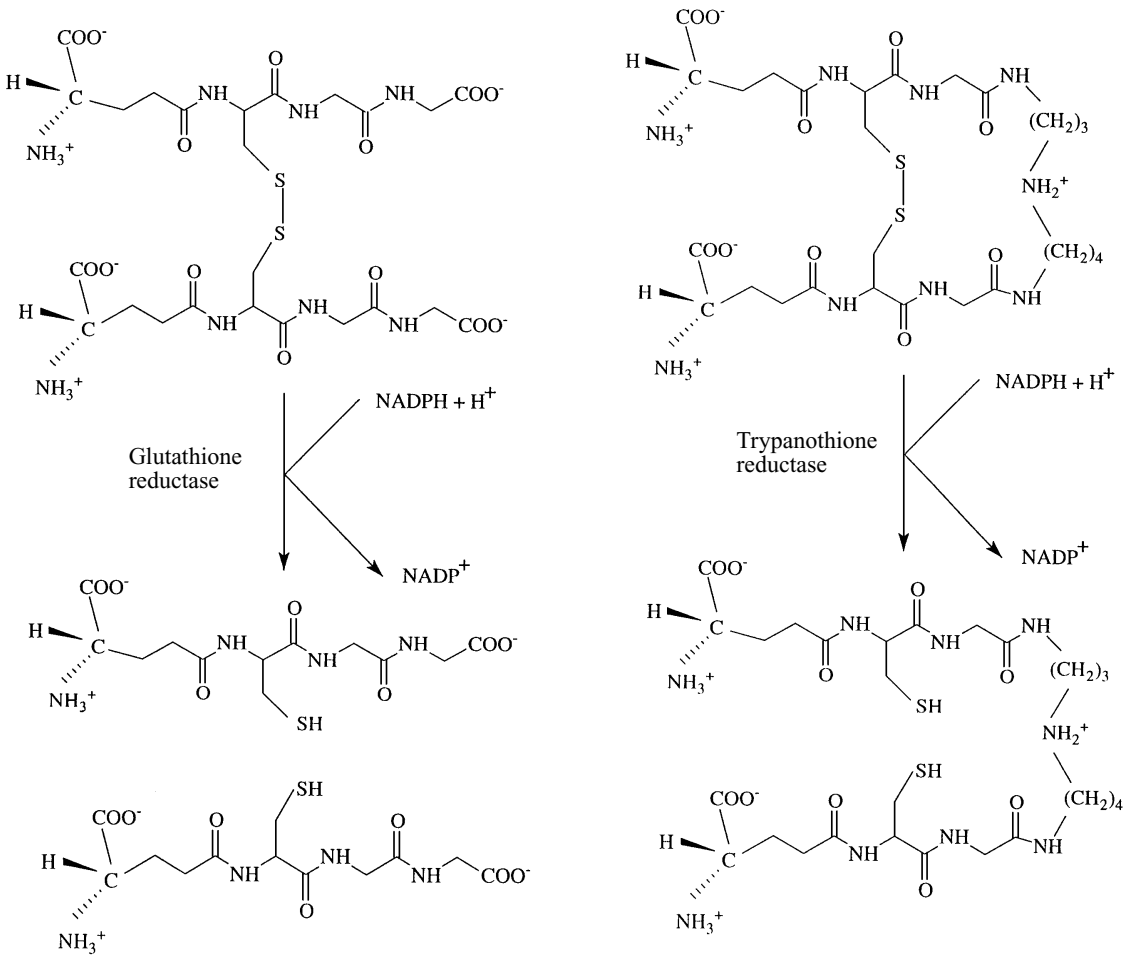

Figure 7 Reactions and substrates of glutathione reductase and trypanothione reductase

See the text for details.

terminal amino group makes no interaction with the phosphate, but forms a cation–π interaction with the aromatic side chain of the phosphorylated tyrosine residue.

The mode of binding offers a simple explanation for the selection of phosphotyrosine residues and the rejection of phosphoserine and phosphothreonine. In the latter cases, not only is the phosphate group attached to shorter side chains, which would not project sufficiently into the binding pocket to interact with those residues that form conventional hydrogen bonds with the phosphate (Arg-175, Glu-178 and Thr-179), but cation–π interactions would not form. The structural features observed in the binding of phosphotyrosine peptides are expected to hold for phosphotyrosine-containing proteins, although the contact between the SH2 domain and phosphotyrosine-containing protein is probably more extensive. Indeed, potential binding sites for residues beyond the phosphotyrosine residue (the N-terminal residue in the peptides used for the structural analysis) have been identified on the surface of the SH2 domain [39] and modelled for various SH2-domain–phosphotyrosine-peptide interactions [40]. The determined structures of other SH2-domain–phosphotyrosine-peptide complexes indicate that cation–π bonding is an important element in the recognition of the phosphorylated tyrosine residue [41–44], although this type of interaction is not a necessary hallmark of SH2-domain–phosphotyrosine complex formation [45].

Trypanothione reductase and the recognition of spermidine-linked peptides

Trypanothione reductases and glutathione reductases are related enzymes belonging to the family of disulphide oxidoreductases [46]. The enzymes are active as homodimers, using FAD as cofactor and NADPH as an electron donor. The function of trypanothione reductase is to reduce N^1,N^8-bis(glutathionyl)-spermidine (trypanothione) in trypanosomal parasites, the causative agents of African sleeping sickness, Chagas' disease and leishmaniasis. Glutathione reductase, an almost ubiquitous enzyme, is responsible for reducing disulphide-linked glutathione (Figure 7). Trypanothione reductase is unique to trypanosomal parasites and, consequently, this enzyme has become an attractive potential drug target to combat trypanosomal infections [47,48].

The structure, specificity and mechanism of human glutathione reductase have been well characterized [49–51], and the enzyme is highly specific for glutathione. In the rational design of anti-trypanosomal drugs that bind in the active site of trypanothione

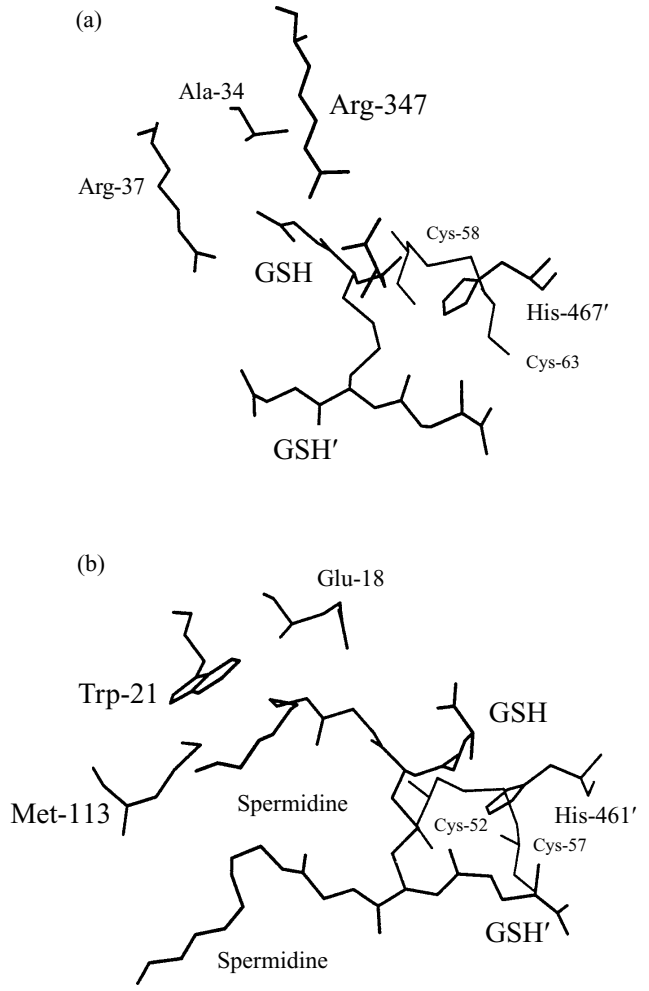

Figure 8 Detail of the binding of (a) glutathione disulphide to human glutathione reductase, and (b) N^1-glutathionyl spermidine disulphide to *Crithidia fasiculata* trypanothione reductase

(**a**) The catalytic disulphide bridge (Cys-58–Cys-63) and His-467′ (from the other subunit of the dimer) are shown, as are the residues making contact with the other end of the substrate. Co-ordinates taken from Brookhaven entry 1GRA. (**b**) Equivalent view to that of (**a**) above. The equivalent disulphide bridge (Cys-52–Cys-57) and histidine (His-461′) are shown. Met-113 and Trp-21 form a 'hydrophobic patch' where the spermidine bridge of trypanothione would be expected to bind, with the side chain of Trp-21 making a cation–π bond with the cationic ammonium group of the spermidine bridge. Co-ordinates taken from Brookhaven entry 1TYP.

reductase, any effects on human glutathione reductase must also be taken into account. An understanding of the molecular details of the trypanothione-binding site in trypanothione reductase in relation to the structure of the glutathione-binding site in glutathione reductase is, therefore, central to a programme of rational drug design. With this in mind, the crystallographic structure of trypanothione reductase from *Crithidia fasiculata* [52] and the structure of the enzyme complexed with N^1-glutathionyl spermidine disulphide [53] have been determined. The position occupied by trypanothione in the active site of trypanothione reductase has also been modelled [54]. Trypanothione carries a cationic ammonium group on the spermidine bridge linking the carboxylate groups of the glycine residues (Figure 7). Trp-21 and Met-113 in trypanothione reductase form a non-polar patch in the vicinity of the spermidine bridge of trypanothione (Figure 8); these residues are absent from the active site of human glutathione reductase. A glutamate residue (Glu-18) is also located in this region of trypanothione reductase, where it may form a favourable ion-pair interaction with the cationic ammonium group of the spermidine bridge. Trp-21 is ideally positioned to make a cation–π interaction with the ammonium group of trypanothione and is conjectured to stabilize the enzyme–substrate complex [52,53].

The observations made in the crystallographic analyses are in accord with earlier mutagenesis experiments on human [55] and *Escherichia coli* [54] glutathione reductases and on *T. congolense* trypanothione reductase [56], in which attempts were made to switch the disulphide specificity of glutathione reductase to function with trypanothione, and vice versa. In each case, the targeted residues were identified by sequence alignment of trypanothione reductase and glutathione reductase and by reference to the available high-resolution structure of human glutathione reductase. A double mutant of trypanothione reductase, in which Trp-21 was replaced by arginine and Glu-18 was replaced by alanine, acquired moderate activity as a glutathione reductase [56]. When the equivalent arginine and alanine residues were exchanged in human glutathione reductase for the naturally occurring tryptophan and glutamate residues of trypanothione reductase, the enzyme was found to discriminate against glutathione and possess significant activity with trypanothione [55], thus demonstrating the importance of these two residues in stabilizing the enzyme–substrate complex. In *E. coli* glutathione reductase, by introducing the same tryptophan and glutamate residues and an additional asparagine residue (at position 22 in the *E. coli* sequence), glutathione reductase activity is effectively abolished, but the mutant enzyme is capable of reducing trypanothione, with a selectivity coefficient (k_{cat}/K_m) 10% of that seen for wild-type trypanothione reductase [54]. Naturally, the cation–π bond and ion-pair interaction made by Trp-21 and Glu-18 respectively with the ammonium cation of the spermidine bridge of trypanothione are important discriminating interactions in substrate binding, and these residues will no doubt become a focal point in the future rational design of drugs against trypanosomal infections.

Design of ammonium-ligand-binding sites

From the available structural data, it is apparent that cation–π bonding and amino–aromatic interactions, often in combination with other more conventional interactions, can facilitate the recognition of substituted ammonium ligands, peptides and proteins by receptor protein molecules. Structural models now exist for quaternary, tertiary and secondary ammonium groups, for the ammonium ion itself and for selected peptide–protein interactions. A model for primary ammonium cations will be forthcoming in the structure of methylamine dehydrogenase. By accepting that cation–π and amino–aromatic interactions should be considered in the design of ammonium-ligand-binding sites, the question remains of how to achieve this recognition in peptide/protein design. Given the limited structural data for ammonium-ligand–protein complexes, rational design strategies seem inappropriate, especially given the generally poor record these approaches have acquired in recent years. The use of phage-display technology [57] for isolating ammonium-ligand-binding peptides seems a more attractive proposition. Ammonium-ligand-binding peptides might find use as generic clathrates for drugs and small molecules that contain ammonium groups.

The design of ammonium-ligand-binding sites in protein molecules is perhaps more challenging than the synthesis of peptide clathrates for ammonium groups. Recently, Gold and

colleagues [58] reported a modest improvement (up to 25-fold) in the specificity of *Bacillus stearothermophilus* L-lactate dehydrogenase for oxo acids containing ammonium groups. Analysis of the crystal structure of wild-type lactate dehydrogenase prompted these workers to exchange Gln-102 in the wild-type protein for acidic aspartate and glutamate residues, in the hope that the side-chain carboxylates would form ion-pair interactions with the ammonium groups on the new substrates. The mutant enzymes still show poor Michaelis constants (in the millimolar region) for the oxo acids bearing ammonium groups. From the evidence presented in this review, one might conjecture that improved catalysis would be realized through the provision of cation–π bonds with the new substrates rather than conventional ion-pair interactions. In this regard, the forced evolution of substrate specificity for *B. stearothermophilus* L-lactate dehydrogenase developed by Holbrook and co-workers might be a more appropriate route to follow. Interestingly, Holbrook has isolated a range of mutant lactate dehydrogenases that operate effectively with phenylpyruvate (a poor substrate for the wild-type enzyme). In one of these mutants, residues 101 and 102 are changed to proline and lysine respectively ([59]; J. J. Holbrook, personal communication) and the enzyme displays about a 10-fold improvement in K_m for phenylpyruvate compared with the wild-type enzyme. In this case, it is tempting to speculate that the side-chain ϵ-amino group of Lys-102 (the same residue position targeted by Gold and colleagues) might form a cation–π bond with the phenyl group of the new substrate, and the crystal structure for this mutant enzyme is therefore eagerly awaited. As the ground-breaking work on lactate dehydrogenase suggests, laboratory-based evolution may prove to be an attractive method for 'engineering' of new ammonium ligand specificities into existing protein scaffolds.

We thank Professor J.J. Holbrook and Professor F.S. Mathews for permission to discuss unpublished work from their laboratories. Some of the work discussed in this review on tertiary and secondary ammonium ligand binding is funded by the Biotechnology and Biological Sciences Research Council (N.S.S.) and the Royal Society (N.S.S.).

REFERENCES

1 Dougherty, D.A. (1996) Science **271**, 163–168
2 Deakyne, C.A. and Meot-Ner, M. (1986) J. Am. Chem. Soc. **107**, 474–479
3 Gao, J., Chou, L.W. and Auerbach, A. (1993) Biophys. J. **65**, 43–47
4 Shepodd, T.J., Petti, M.A. and Dougherty, D.A. (1986) J. Am. Chem. Soc. **108**, 6085–6087
5 Kearney, P.C., Mizoue, L.S., Kumpf, R.A., Forman, J.E., McCurdy, A. and Dougherty, D.A. (1993) J. Am. Chem. Soc. **115**, 9907–9919
6 Stauffer, D.A., Barrans, R.E. and Dougherty, D.A. (1990) Angew. Chem. Int. Ed. Engl. **29**, 915–917
7 McCurdy, A., Jimenez, L., Stauffer, D.A. and Dougherty, D.A. (1992) J. Am. Chem. Soc. **114**, 10314–10321
8 Burley, S.K. and Petsko, G.A. (1986) FEBS Lett. **203**, 139–143
9 Wlodawer, A., Walter, J., Huber, R. and Sjorn, L. (1984) J. Mol. Biol. **180**, 301–329
10 Tuchsen, E. and Woodward, C. (1987) Biochemistry **26**, 1918–1925
11 Perutz, M.F., Fermi, G., Abraham, D.J., Poyart, C. and Bursaux, E. (1986) J. Am. Chem. Soc. **108**, 1064–1078
12 Singh, J. and Thornton, J.M. (1990) J. Mol. Biol. **211**, 595–615
13 Mitchell, J.B.O., Nandi, C.L., McDonald, I.K., Thornton, J.. and Price, S.L. (1994) J. Mol. Biol. **239**, 315–331
14 Levitt, M. and Perutz, M.F. (1988) J. Mol. Biol. **201**, 751–754
15 Perutz, M.F. (1993) Philos. Trans. R. Soc. London A **345**, 105–112
16 Flocco, M.M. and Mowbray, S.L. (1994) J. Mol. Biol. **235**, 709–717
17 Nachmansohn, D. and Wilson, I.B. (1951) Adv. Enzymol. **12**, 259–339
18 Sussman, J.L., Harel, M., Frolow, F., Oefner, C., Goldman, A., Toker, L. and Silman, I. (1991) Science **253**, 872–879
19 Harel, M., Schalk, I., Ehret-Sabattier, L., Bouet, F., Goeldner, H., Hirth, C., Axelsen, P., Silman, I. and Sussman, J. (1993) Proc. Natl. Acad. Sci. U.S.A. **90**, 9031–9035
20 Dennis, M., Giraudat, J., Kotzybahibert, F., Goeldner, M., Hirth, C., Chang, J.Y., Lazure, C., Chretien, M. and Changeux, J.P. (1988) Biochemistry **27**, 2346–2357
21 Galzi, J.L., Revah, F., Black, D., Goeldner, H., Hirth, C. and Changeux, J.P. (1990) J. Biol. Chem. **265**, 10430–10437
22 Davis, D.R. and Metzger, H. (1983) Annu. Rev. Immunol. **1**, 87–117
23 Satow, Y., Cohen, G.H., Padlan, E.A. and Davis, D.R. (1986) J. Mol. Biol. **190**, 593–604
24 Steenkamp, D.J. and Mathews, F.S. (1992) in Chemistry and Biochemistry of Flavoenzymes (Muller, F., ed.), vol. 2, pp. 395–423, CRC Press, Boca Raton
25 Lim, L.W., Shamala, N., Mathews, F.S., Steenkamp, D.J., Hamlin, R. and Xuong, N. (1986) J. Biol. Chem. **261**, 15140–15146
26 Bellamy, H.D., Lim, L.W., Mathews, F.S. and Dunham, W.R. (1989) J. Biol. Chem. **264**, 11887–11892
27 Raine, A.R.C., Yang, C.-C., Packman, L.C., White, S.A., Mathews, F.S. and Scrutton, N.S. (1995) Protein Sci. **4**, 2625–2628
28 Rohlfs, R.J. and Hille, R. (1994) J. Biol. Chem. **269**, 30869–30879
29 Kasprzak, A.A., Papas, E.J. and Steenkamp, D.J. (1983) Biochem. J. **211**, 535–541
30 Boyd, G., Mathews, F.S., Packman, L.C. and Scrutton, N.S. (1992) FEBS Lett. **308**, 271–276
31 Yang, C.-C., Packman, L.C. and Scrutton, N.S. (1995) Eur. J. Biochem. **232**, 264–271
32 Chen, L., Mathews, F.S., Davidson, V.L., Huizinga, E.G., Vellieux, F.M.D. and Hol, W.G.J. (1992) Proteins Struct. Funct. Genet. **14**, 288–299
33 Vellieux, F.M.D., Huitema, F., Groendijk, H., Kalk, K.H., Frank, J., Jongejan, J.A., Duine, J.A., Petratos, K., Drenth, J. and Hol, W.G.J. (1989) EMBO J. **8**, 2171–2178
34 Vellieux, F.M.D., Kalk, K.H., Drenth, J. and Hol, W.G. (1990) Acta Crystallogr. **B46**, 806–823
35 Chisterserdov, A.Y., Boyd, J., Mathews, F.S. and Lidstrom, M.E. (1992) Biochem. Biophys. Res. Commun. **184**, 1181–1189
36 Liaw, S.-H., Kuo, I. and Eisenberg, D. (1995) Protein Sci. **4**, 2358–2365
37 Cantley, L.C., Auger, K.R., Carpenter, C., Duckworth, B., Graziani, A., Kapeller, R. and Soltoff, S. (1991) Cell **64**, 281–302
38 Koch, C.A., Anderson, D., Moran, M.F., Ellis, C. and Pawson, T. (1991) Science **252**, 668–674
39 Waksman, G., Kominos, D., Robertson, S.C., Pant, N., Baltimore, D., Birge, R., Cowburn, D., Hanafusa, H., Mayer, B.J., Overduin, M., Resh, M.D., Rios, C.B., Silverman, L. and Kuriyan, J. (1992) Nature (London) **358**, 646–653
40 Zvelebil, M.J.J.M., Panayotou, G., Linacre, J. and Waterfield, M.D. (1995) Protein Eng. **8**, 527–533
41 Eck, M.J., Shoelson, S.E. and Harrison, S.C. (1993) Nature (London) **362**, 87–91
42 Waksman, G., Shoelson, S.E., Pant, N., Cowburn, D. and Kuriyan, D. (1993) Cell **72**, 779–790
43 Narula, S.S., Yuan, R.W., Adams, S.E., Green, O.M., Green, J., Philips, T.B., Zydowsky, L.D., Botfield, M.C., Hatada, M., Laird, E.R., Zoller, M.J., Karas, J.L. and Delgarno, D.C. (1995) Structure **3**, 1061–1073
44 Zhou, M.-M., Meadows, R.R., Logan, T.M., Yoon, H.S., Wade, W.S., Ravichandran, K.S., Burakoff, S.J. and Fesik, S.W. (1995) Proc. Natl. Acad. Sci. U.S.A. **92**, 7784–7788
45 Hatada, M.H., Lu, X.D., Laird, E.R., Green, J., Morgenstern, J.P., Lou, M.Z., Marr, C.S., Phillips, T.B., Ram, M.K., Theriault, K., Zoller, M.J. and Karas, J.L. (1995) Nature (London) **377**, 32–38
46 Williams, Jr., C.H. (1992) in Chemistry and Biochemistry of Flavoenzymes (Muller, F., ed.), vol. 3, pp. 121–211, CRC Press, Boca Raton
47 Fairlamb, A.H. (1989) Parasitology **99S**, 93–112
48 Ponasik, J.A., Strickland, C., Faerman, C., Savvides, S., Karplus, P.A. and Ganem, B. (1995) Biochem. J. **311**, 371–375
49 Karplus, P.A. and Schulz, G.E. (1987) J. Mol. Biol. **195**, 701–729
50 Karplus, P.A. and Schulz, G.E. (1989) J. Mol. Biol. **210**, 163–180
51 Karplus, P.A., Pai, E.F. and Schulz, G.E. (1989) Eur. J. Biochem. **178**, 693–703
52 Hunter, W.N., Bailey, S., Habash, J., Harrop, S.J., Helliwell, J.R., Aboagye-Kwarteng, T., Smith, K. and Fairlamb, A.H. (1992) J. Mol. Biol. **227**, 322–333
53 Bailey, S., Smith, K., Fairlamb, A.H. and Hunter, W.N. (1993) Eur. J. Biochem. **213**, 67–75
54 Henderson, G.B., Murgolo, N.J., Kuriyan, J., Osapay, K., Kominos, D., Berry, A., Scrutton, N.S., Hinchliffe, N.W., Perham, R.N. and Cerami, A. (1991) Proc. Natl. Acad. Sci. U.S.A. **88**, 8769–8773
55 Bradley, M., Bucheler, U.S. and Walsh, C.T. (1991) Biochemistry **30**, 6124–6127
56 Sullivan, F.X., Sobolov, S.B., Bradley, M. and Walsh, C.T. (1991) Biochemistry **30**, 2761–2767
57 Scott, J.K. and Smith, G.P. (1990) Science **249**, 386–390
58 Hogan, J.K., Pittol, C.A., Jones, B. and Gold, M. (1995) Biochemistry **34**, 4225–4230
59 Hawrani, A.E. (1995) Ph.D. Thesis, University of Bristol, U.K.
60 Kraulis, P.J. (1991) J. Appl. Crystallogr. **24**, 946–950