BJP British Journal of Pharmacology

DOI:10.1111/bph.12182
www.brjpharmacol.org

REVIEW

Arylamine N-acetyltransferases: a structural perspective

Xiaotong Zhou[1]*, Zhiguo Ma[1]*, Dong Dong[2] and Baojian Wu[1]

[1]*Division of Pharmaceutics, College of Pharmacy, Jinan University, Guangzhou, Guangdong, China, and* [2]*Department of Pharmacological and Pharmaceutical Sciences, College of Pharmacy, University of Houston, Houston, TX, USA*

Correspondence
Baojian Wu, Division of Pharmaceutics, College of Pharmacy, Jinan University, Guangzhou, Guangdong 510632, China. E-mail: bj.wu@hotmail.com

*XZ and ZM contributed equally to this paper.

Keywords
drug metabolism; arylamine N-acetyltransferase; NAT; crystal structure; drug discovery

Received
5 December 2012
Revised
2 February 2013
Accepted
7 February 2013

Arylamine N-acetyltransferase (NAT) plays an important role in metabolism and detoxification of many compounds including drugs and environmental carcinogens through chemical modification of the amine group with an acetyl group. Recent studies have suggested that NATs are also involved in cancer cell growth and inhibition of the enzymes may be a potential target for cancer chemotherapy. Three-dimensional (3D) structures are available for NATs from both prokaryotes and eukaryotes. These structures provide valuable insights into the acetylation mechanism, features of the active site and the structural determinants that govern substrate/inhibitor-binding specificity. Such insights allow a more precise understanding of the structure–activity relationships for NAT substrates and inhibitors. Furthermore, the structural elucidation of NATs has generated powerful tools in the design of small molecule inhibitors that should alleviate cancer, based on the important role of the enzyme in cancer biology.

Abbreviations

3D, three-dimensional; 5-AS, 5-aminosalicylic acid; ANS, anisidine; CoA, Coenzyme A; HDZ, hydrazines; NAT, Arylamine N-acetyltransferase; PABA, p-aminobenzoic acid; PAS, p-aminosalicylic acid; SMZ, sulfamethazine; TZD, thiazolidinedione

Introduction

Arylamine N-acetyltransferases (NATs; EC 2.3.1.5) catalyse the acetylation reaction (generally classified as a Phase II process) by which an acetyl group from the cofactor acetyl coenzyme A (acetyl CoA) is transferred to the substrates, including drugs and carcinogens. NATs are present in most of living organisms (Butcher *et al.*, 2002; Vagena *et al.*, 2008) and NAT proteins have a characteristic motif – (Pro/Ala)-Phe-Glu-Asn-Leu [(P/A)FENL] (Payton *et al.*, 2001). Although its endogenous functions remain largely unclear, NAT is essential for synthesis of cell wall in mycobacteria (Sim *et al.*, 2008a). Human NATs include two isoenzymes, namely, NAT1 and NAT2. NAT1 is found in many tissues, particularly in the colon, whereas NAT2 is found mostly in the liver and intestine and is predominantly responsible for drug metabolism (Hickman *et al.*, 1998; Sim *et al.*, 2000). It acetylates many therapeutic agents such as isoniazid, dapsone and the sulphonamides.

In recent years, acetylation has been intensively studied due largely to its role in the carcinogenic activation of aromatic amines (Hein *et al.*, 1993). These acetylated derivatives of arylamines and N-hydroxylated arylamines can undergo further oxidation or other reactions to form highly reactive cytotoxic and carcinogenic species (Hein *et al.*, 1993; Hanna, 1994). Moreover, there is increasing evidence that NAT has an important role in folate catabolism and cancer cell biology (particularly for breast cancer) (Adam *et al.*, 2003; Butcher *et al.*, 2007; Sim *et al.*, 2008b; Butcher and Minchin, 2012). In fact, human NAT1 has emerged as a new diagnostic marker or drug target for breast cancer (Laurieri *et al.*, 2010; Russell *et al.*, 2009; Tiang *et al.*, 2010).

The first X-ray crystal structure for a NAT enzyme (from *Salmonella typhimurium*) was reported by Sinclair *et al.* (2000). To date, there are 11 crystal structures for NATs from human and five bacterial species (Table 1). This structural bank has provided valuable insights into the acetylation mechanisms and active site features. Importantly, it helps to explain why

© 2013 The Authors
British Journal of Pharmacology © 2013 The British Pharmacological Society

Table 1

List of available crystal structures for NATs (the structures can be found at http://www.pdb.org)

NAT	PDB entries	In complex with	Reference
S. typhimurium NAT	1E2T		Sinclair *et al.*, 2000
M. smegmatis NAT	1GX3		Sandy *et al.*, 2002
M. smegmatis NAT	1W6F	Isoniazid	Sandy *et al.*, 2005b
M. smegmatis NAT (C70Q)	1W5R		Sandy *et al.*, 2005a
Pseudomonas aeruginosa NAT	1W4T		Westwood *et al.*, 2005
Mesorhizobium loti NAT1	2BSZ		Holton *et al.*, 2005
M. marinum NAT	2VFB		Fullam *et al.*, 2008
M. marinum NAT	2VFC	Coenzyme A	Fullam *et al.*, 2008
Human NAT1 (F125S)	2IJA	Iodoacetamide	Wu *et al.*, 2007
Human NAT1	2PQT	2-Bromoacetanilide	Wu *et al.*, 2007
Human NAT2	2PFR	Coenzyme A	Wu *et al.*, 2007

NATs selectively catalyse aromatic amines and exhibit overlapping yet distinct substrate specificity. Both human NAT1 and NAT2 are highly polymorphic enzymes. Many of the NAT polymorphisms translate to mutations in particular protein residues. The crystal structures have suggested individual function of the residues in the protein that has greatly improved the understanding of the phenotypes associated with NAT polymorphisms. The crystal structures represent not only a tool for predicting substrate selectivity but also a powerful tool in rational design of selective inhibitors of NATs for therapeutic purposes.

NAT Structure

The secondary and tertiary structures of the NATs are highly conserved among the enzymes from prokaryotes and eukaryotes (Westwood *et al.*, 2006; Sim *et al.*, 2008a). The NAT fold is usually described in terms of three domains (Sinclair *et al.*, 2000; Wu *et al.*, 2007). For example, in human NAT, the N-terminal domain [or the helical bundle in Sinclair *et al.*, (2000)] consists of five helices (α1–α5) and one short β-strand between helices α2 and α3 (Figure 1). The second domain [or the β-barrel in Sinclair *et al.*, (2000)] consists of 10 β-strands (β2–β11) and two short helices α6 and α7 (Figure 1). The two domains connect through the α-helical interdomain (helices α8–α10) to the third domain [or the α/β lid in Sinclair *et al.*, (2000)], which has four anti-parallel β-strands (β12–β15) and helix α11 (Figure 1). The helix α11 precedes a stretch of residues that lead across the protein molecule's surface into a buried C-terminus (Figure 1).

Although a highly similar structural fold is shared by bacterial and human NATs, structural differences are evident. First, human NATs contain a 17-residue insertion (i.e. residues 167–183) between β9 and β11, which is absent in the structures of prokaryotic NATs (Wu *et al.*, 2007) (Figure 2). The insertion includes the secondary structure elements β10 and α7. Although the function of this insertion in human NATs is not fully understood, the insertion does contribute to the stability of the protein by interacting with surrounding residues (Walraven *et al.*, 2007). Second, human NATs show a significant structural divergence in the arrangement of their C-terminal residues compared with prokaryotic NATs (Figure 2). In human NAT, the C-terminal residues form a full coil and the terminus extends deep within the folded protein into close proximity to the buried catalytic triad and is capable of interacting with CoA (Wu *et al.*, 2007) (Figure 2). By contrast, the C-terminus in prokaryotic NATs is in a helical configuration, and the helix is positioned away from the protein core (Figure 2).

Active sites for cofactor and substrate binding

The crystal structures, in complex with CoA, for *Mycobacterium marinum* NAT and human NAT2 reveal that CoA binds differently to these two enzymes (Wu *et al.*, 2007; Fullam *et al.*, 2008). In human NAT, the CoA-binding site is located in a deep cleft formed between the interdomain (connecting domain II to domain III) helices α8–10 and the β-barrel (domain II) (Wu *et al.*, 2007) (Figure 3A). CoA is bound to the active site in a bent conformation, positioning the sulfhydryl group of CoA towards the protein core and the 3′-phospho-ADP group on the surface of the protein (Figure 3A). The β-mercaptoethylamine and pantothenate moieties are located deep inside the cleft, placing the cofactor's sulfhydryl group close to that of the catalytic cysteine (C68), with a sulfur–sulfur distance of ~3 Å (Wu *et al.*, 2007). Binding of CoA to the active site is mainly mediated by hydrophobic and hydrogen-bond interactions (Figure 3B). The pantotheine arm of CoA makes an extensive set of van der Waals contacts with the non-polar residues F37, F93, L209, F217 and L288 (Figure 3B and C). The adenine ring also makes contacts with the hydrophobic residues P97 and V98. In addition, the carbonyl group of the pantothenate moiety forms a hydrogen bond with the nitrogen atom of S216. The pyrophosphate group of CoA makes a series of hydrogen bonds with the residues T103, G104, Y208 and T214. The N6 of the adenine

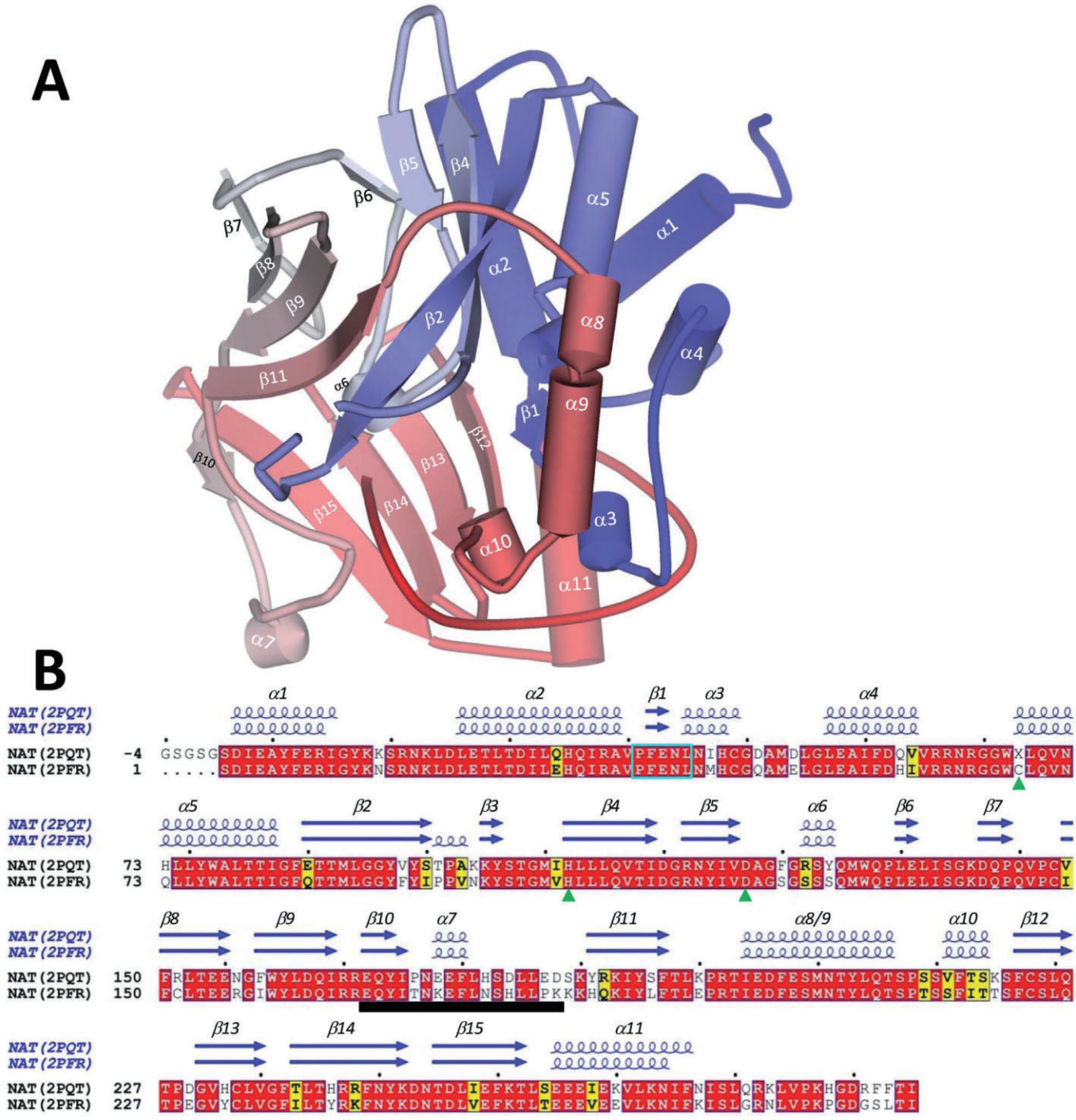

Figure 1

The tertiary and secondary structures of NAT. Panel A: The tertiary structure of NAT (using human NAT1 (PDB code: 2PQT) as an example), depicting the overall fold of a NAT structure. Panel B: Structure-based sequence alignment of human NATs (NAT1 and NAT2). The secondary structure elements are shown above the alignment. Conserved residues are highlighted in colour. The Protein Data Bank (PDB) codes for the structures are shown in parentheses. The triangles indicate the catalytic residues (please see text for details). This figure was produced using ENDscript (Gouet and Courcelle, 2002).

ring of CoA forms a hydrogen bond with the side chain of S287 (Wu *et al.*, 2007) (Figure 3B).

Unexpectedly, a different CoA-binding site is present in the prokaryotic NAT (Fullam *et al.*, 2008). Here, CoA binds to the deep cleft formed by domains II and III of the enzyme (Figure 4A). The binding site is located ~30 Å from that in human NAT (Figure 4). Recognition of the pantotheinate arm of CoA is mainly through hydrophobic interactions with

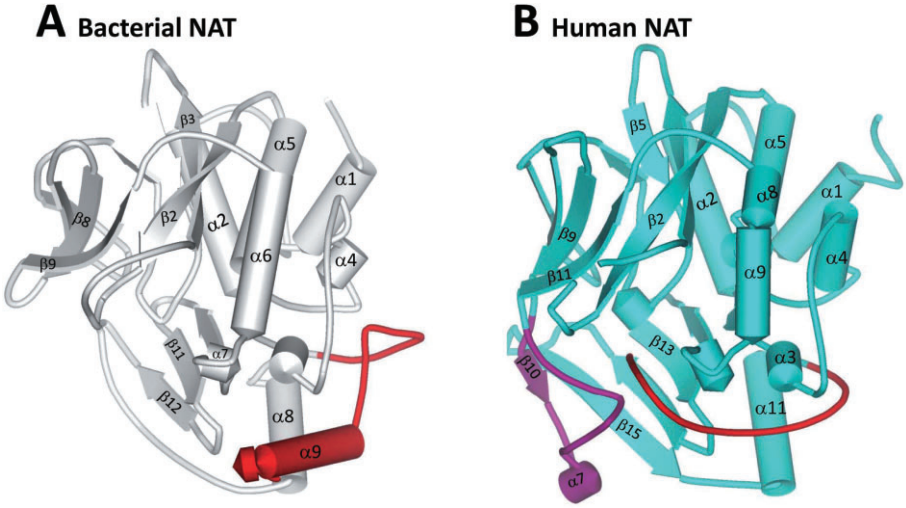

Figure 2

Structural comparison of bacterial NAT (panel A) and human NAT (panel B). The structures of NAT from *Salmonella typhimurium* (PDB code: 1E2T) and human NAT1 (PDB code: 2PQT) are shown. The insertion in the human NAT is indicated in purple. The structures of the C-terminal residues in both enzymes are shown in red.

residues F38, Y69, Y71, H110, F130 and F204. The phosphate groups of CoA are hydrogen-bonded to residues such as W97 and K236. The adenine moiety of CoA interacts with the active site residues via a mixture of hydrophobic and polar contacts. Specifically, it forms hydrophobic, hydrogen-bond and ring-stacking interactions with residues V169, E152, H229 respectively (Fullam *et al.*, 2008). The observation that human and prokaryotic NATs bind CoA in markedly different ways led Fullam *et al.* (2008) to hypothesize that the cofactor binding site may be a promising novel target for selectively inhibiting pathogenic prokaryotic NAT enzymes.

The acetyl acceptor (substrate) binding site overlaps to a great extent with the CoA-binding site and this finding is consistent with the fact that the cofactor and the substrate bind to the enzymes in a sequential manner, a feature of the Ping Pong kinetic mechanism (see discussion below). Similar to the pantotheine arm of CoA, the entire substrate molecule binds to the deep position of the cleft formed between the helical interdomain and domain II (β-barrel) (Wu *et al.*, 2007). The substrate-binding sites deduced from either prokaryotic or eukaryotic NATs consist mainly of hydrophobic residues (Figure 5A). This feature agrees with the fact that NATs have a preference for hydrophobic substrates.

Catalytic mechanism

There is strong evidence that the acetylation reaction mediated by NATs proceeds via the double displacement ('Ping Pong Bi Bi') mechanism (Andres *et al.*, 1983; 1988; Sinclair and Sim, 1997). There are two sequential steps to the reaction: firstly, the acetyl group is moved from acetyl CoA to form an acetylated enzyme intermediate, then the substrate is acetylated and CoA is released (Figure 6). Structural analysis reveals that the acetyl transfer in NATs is facilitated by a Cys-His-Asp catalytic triad (Cys^{68}-His^{107}-Asp^{122} in human NAT), which is strictly conserved in all known NATs and in identical positions in all NAT structures (Sinclair *et al.*, 2000). The triad renders a stable thiolated form of cysteine that initiates the reaction by nucleophilic attack on the carbonyl of the acetyl moiety of acetyl CoA (Figure 6) (Wang *et al.*, 2004; Sandy *et al.*, 2005a). Consistent with their critical role in catalysis, mutations of any of the triad residues abolishes the enzyme activity (Dupret and Grant, 1992; Watanabe *et al.*, 1992; Wang *et al.*, 2004). Interestingly, the latest NAT structure (from *M. marinum*) presents an extended binding interface between CoA and the protein (Fullam *et al.*, 2008). The authors propose that the nucleoside-phosphate moiety of CoA can remain associated with the protein, subsequent to acetylation of the catalytic cysteine, thereby protecting the acetylated enzyme intermediate from hydrolysis.

Wang *et al.* have proposed a model for the reaction chemistry in the NAT active site (Wang *et al.*, 2004; 2005). In this model, a thiolate–imidazolium ion pair is formed between Cys^{68} and His^{107} with a pK_a of 5.2. Asp^{122} is involved in the ionic interaction with His^{107} and is critical for optimal catalysis and structural integrity. Upon acetylation of the thiolate, half of the ion pair is lost with concomitant shift in the pK_a of His^{107} to 5.5. The process of nucleophilic attack on the thiol ester and deacetylation of the thiolate is dependent on the nucleophilic strength of the arylamine substrate. For weak nucleophiles with a pK_a <5.5, such as p-aminobenzoic acid (PABA, Figure 7A), a deprotonation step by His^{107} is required to make the substrates more nucleophilic for the attack on the thiol ester, which is followed by deacetylation of the thiolate. By contrast, those strong nucleophiles with a $pK_a \geq 5.5$ such as anisidine (ANS, Figure 7A) directly attack on the thiol ester; deacetylation of the thiolate occurs via deprotonation of a tetrahedral intermediate (Wang *et al.*, 2005) (Figure 6).

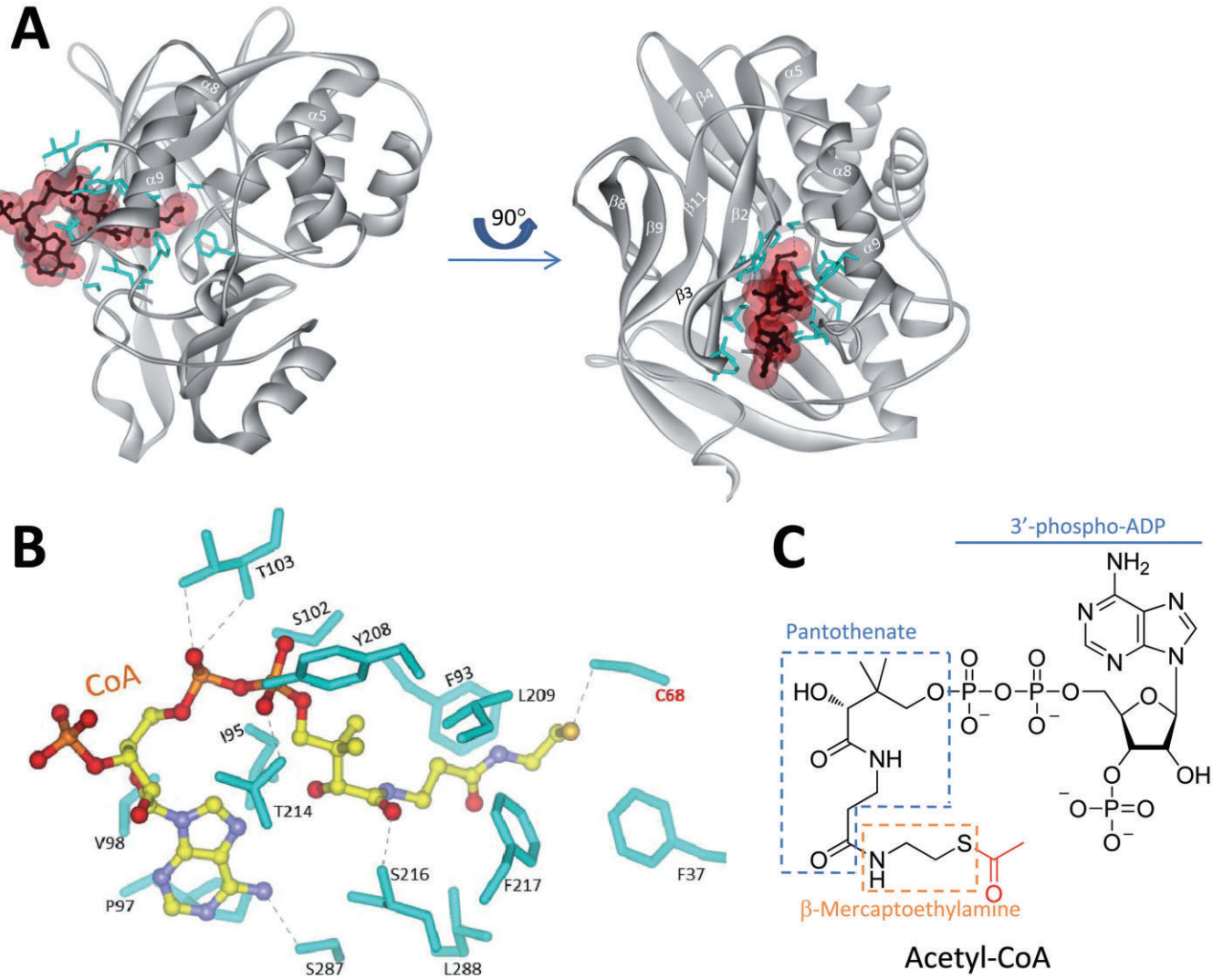

Figure 3

The CoA binding site in human NAT (PDB code: 2PFR). Panel A: A full view of CoA binding to human NAT. Panel B: Molecular interactions of CoA with the binding site residues. Panel C: Chemical structure of acetyl CoA, showing its three components.

Structure-activity relationships for NAT substrates

Human NAT1 and NAT2 exhibit an overlapping substrate specificity. Both enzymes display substrate preference for aromatic amines (Kawamura *et al.*, 2005). This is largely explained by the hydrophobic feature of the active site and the presence of aromatic residues in the active site capable of forming ring-stacking interactions (Wu *et al.*, 2007). However, distinct substrate specificity also exists between NAT1 and NAT2 (Kawamura *et al.*, 2005).

Aminobenzyl compounds such as p-aminosalicylic acid (PAS; Figure 7A) were identified as binding preferentially to NAT1. In contrast, the sulfonamide class of compounds such as sulfamethazine (SMZ) bound selectively to NAT2 (Sim *et al.*, 2008b). Wu *et al.* (2007) explored the molecular basis for this difference in substrate recognition between the two enzymes. The substrate binding pocket in NAT1 is smaller (162 Å^3) than that of NAT2 (257 Å^3) as a consequence of two key residue substitutions at positions 127 and 129, namely R127 and Y129 in NAT1, whereas in NAT2, serine residues occupy these positions (Figure 5B). The presence of two bulkier groups reduces the volume of the NAT1 pocket by ~40% compared with that of NAT2. Furthermore, V93 in NAT1 is replaced by F93 in NAT2; this substitution introduces a bump in the van der Waals surface of the pocket in NAT2, thereby significantly altering the shape of the binding pocket (Figure 5B). Molecular docking reveals that PAS and SMZ are closely fitted into the active sites of NAT1 and NAT2 respectively (Wu *et al.*, 2007). The shape and size of the substrate molecules are well matched to those of the binding pockets, which helps to explain the distinct substrate selectivity of the enzymes.

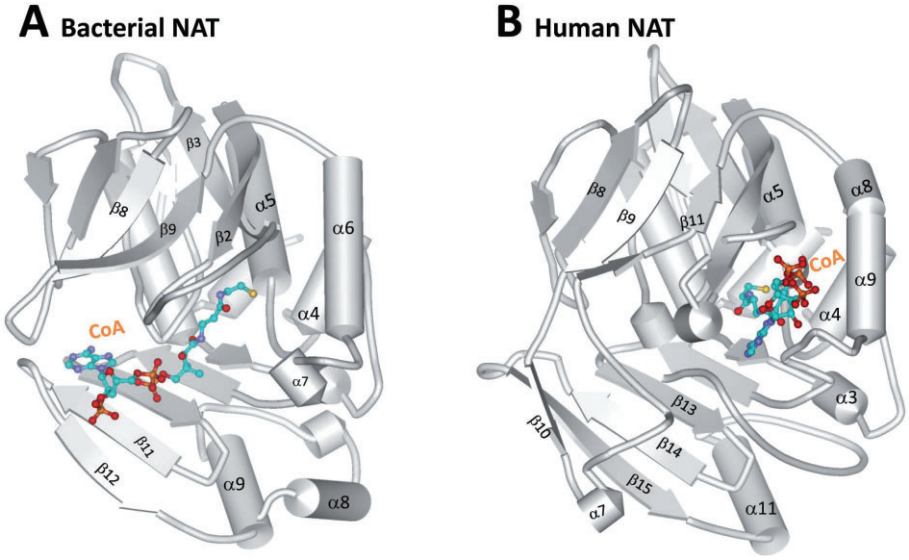

Figure 4

Comparison of the CoA-binding site in bacterial NAT from *M. marinum* (PDB code: 2VFC; panel A) with that in human NAT2 (PDB code: 2PFR; panel B).

A mutagenesis study showed that F125S substitution virtually abolished NAT1 affinity preference for PAS over SMZ and substrate selectivity of the mutant NAT1 (F125S) resembled that of wild-type NAT2 (Hein, 2002). Structural analysis and molecular docking indicate that F125 plays a key role in binding of PAS to NAT1 as the side chain of F125 forms π-π stacking interactions with the benzene ring of PAS (Wu *et al.*, 2007). Therefore, F125S mutation would reduce the enzyme's affinity for the smaller PAS due to the lack of the π-π stacking interactions (as main forces) orienting the substrate for catalysis. On the other hand, an increase in the pocket space allows accommodation of the bulkier SMZ in the active site with more favourable energy. This is consistent with the fact that the mutant has an increased affinity for SMZ compared with the wild type (Hein, 2002).

Human NAT1 acetylates PABA and 4-aminobiphenyl but not o-toluidine (Grant *et al.*, 1991; Hein *et al.*, 1993; Fretland *et al.*, 1997). The NMR structures of human NAT1 reported by Zhang *et al.* (2006) provide an explanation of such substrate selectivity. Based on the resolved structures, PABA binds favourably to the active site, consistent with the fact that it is a good NAT1 substrate (Figure 8A). However, binding of o-toluidine to the enzyme is difficult because its methyl group engages in steric clashes with F125 (Zhang *et al.*, 2006). Also o-toluidine lacks functional groups such as the carboxylic acid group of PABA that could provide favourable binding contacts, for instance, hydrogen bonding (Zhang *et al.*, 2006). The important role of F125 in determining the acetylation of o-toluidine is supported by the finding that a mutant human NAT2 in which F125 is replaced by a serine, exhibits 3- to 10-fold higher acetylation capacity for the substrate than NAT1 (Hein *et al.*, 1993; Zhang *et al.*, 2006). Based on this knowledge, Zhang *et al.* (2006) predicted that 4-amino-3-methylbenzoic acid would be a worse NAT1 substrate than PABA due to the presence of the methyl group. Interestingly, this prediction agreed well with the experimentally determined substrate activities. The authors also explained why 4-aminobiphenyl is acetylated by NAT1 by docking of the substrate with the enzyme (Zhang *et al.*, 2006), as 4-aminobiphenyl was easily accommodated in the active site (Figure 8B). Binding of 4-aminobiphenyl was largely driven by the hydrophobic interactions between residues V93, K100, I106, F125, L209, S215, V216 and F217 and the aromatic ring of the substrate molecule (Figure 8B).

Human NAT2 displays a preference for lipophilic arylamines and the acetylation rate increases as the alkyl chain length of alkoxyanilines increases (Kawamura *et al.*, 2005). On the contrary, human NAT1 displays a decrease in activity with increasing alkyl chain length (Kawamura *et al.*, 2005). It has been proposed that this is due to the smaller overall size of human NAT1 active site compared with that of human NAT2 (Kawamura *et al.*, 2005; Westwood *et al.*, 2006). Substrate profiling demonstrates that none of the arylhydrazine compounds are substrates of human NAT1, whereas these compounds can be acetylated by human NAT2 (Kawamura *et al.*, 2005). However, the reason why human NATs show this difference in substrate selection is unknown.

Brooke *et al.* (2003a) investigated the metabolism of a series of 4-alkoxyanilines by bacterial NATs (*S. typhimurium* and *M. smegmatis*) and found that there was a marked increase in the rate of acetylation with increasing alkyl chain length (Hexyl > Bu > Et > Me). This result indicated that the lipophilicity of the substrate may be a contributory factor to the rate of acetylation, which is further confirmed by a strong correlation between the enzyme activities and the calculated partition coefficients (clogP) of the substrate molecules. The preference for acetylating more lipophilic substrates by the bacterial NATs is probably due to the lipophilic nature of three conserved phenylalanine residues in the active site (Brooke *et al.*, 2003a). The dominant binding forces are con-

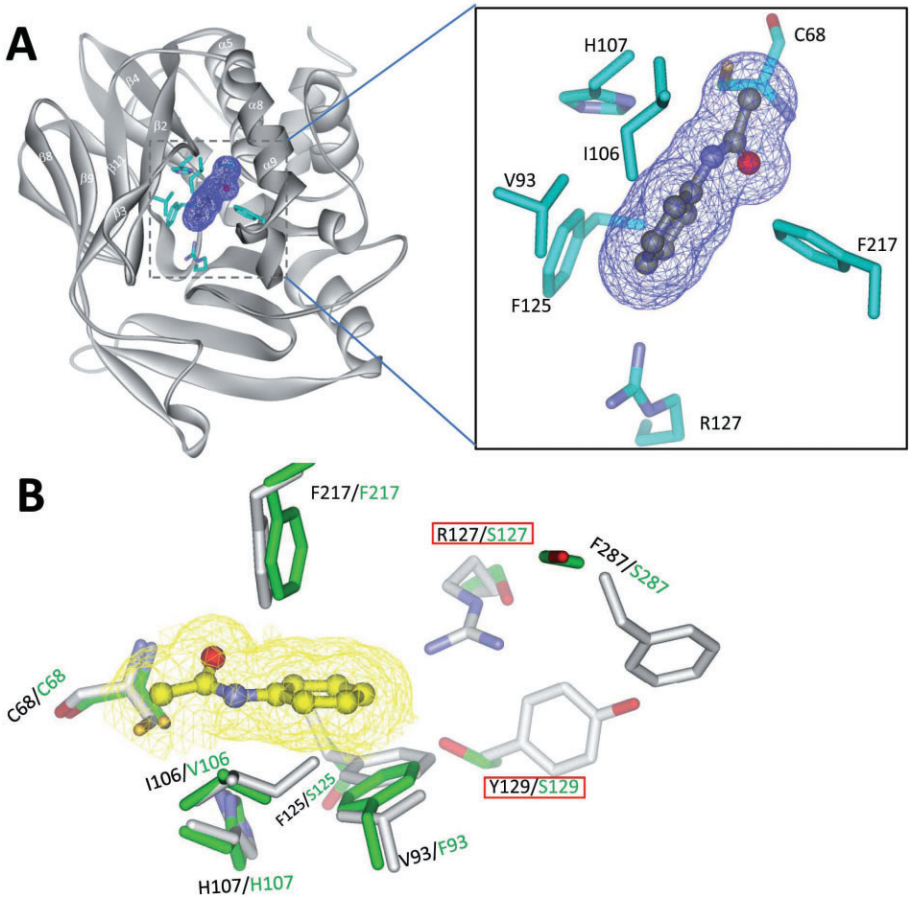

Figure 5

The substrate-binding site in human NAT. Panel A: A full view of the substrate-binding site (indicated as a blue surface) in human NAT (PDB code: 2PQT). Panel B: Comparison of the substrate-binding site residues of human NAT1 (in white; PDB code: 2PQT) with those of NAT2 (in green; PDB code: 2PFR). Residues at positions 127 and 129 play a critical role in determining the pocket size and the substrate specificity. Please see the text for details. The binding site is indicated as a yellow surface.

tributed by the lipophilic–lipophilic and π-stacking interactions between the phenylalanine residues in the enzyme and the acceptor substrate.

The crystal structure for *M. smegmatis* NAT (MSNAT)-isoniazid (INH) complex has been used to explain the enzyme substrate selectivity towards hydrazines (HDZ) and arylamines (Sandy *et al.*, 2005b). The arylhydrazines are the best substrates for MSNAT with relatively low K_m values of <5 μM for HDZ and 7.3 μM for INH, whereas the arylamines have higher K_m values – 586 μM for 5-aminosalicylic acid (5-AS), 1460 μM for ANS, 4500 μM for PAS, and 56 000 μM for PABA. This may be in part due to the formation of hydrogen bonds between the hydrazyl group and residues T109 and C70 (Figure 7B). In addition, HDZ has extra potential hydrogen bonding from the heterocyclic nitrogen to the backbone oxygen of G129 or F130 (Sandy *et al.*, 2005b). The affinity of the arylamine series depends to some extent on the capacity to form interactions with T109 and F130. 5-AS is a good arylamine substrate because it can be well oriented to form hydrogen bonds with both the side chain of T109 and the main chain carbonyl of F130. PAS has a lower apparent affinity compared with 5-AS because it is difficult for T109 to form hydrogen bonds with the substrate molecule, although hydrogen bonds to F130 can be readily achieved (Sandy *et al.*, 2005b). ANS has a lower K_m value than PAS despite offering hydrogen bonds to neither T109 nor F130. In this case, the unfavourable ligand binding in the absence of hydrogen bonding might be partly compensated by the generally lipophilic character of the methoxy group. Such groups are preferred in NAT substrates, presumably due to the formation of hydrophobic interactions in the active site. PABA has no additional chemical group on the aromatic ring facing the T109 and has an even higher K_m than that of PAS (Sandy *et al.*, 2005).

Structure–activity relationships for NAT inhibitors

There is considerable interest in developing small molecule inhibitors for NATs because (i) the inhibitors may be used for therapeutic purposes in breast cancer and tuberculosis; and (ii) the inhibitors can be used to explore the functions

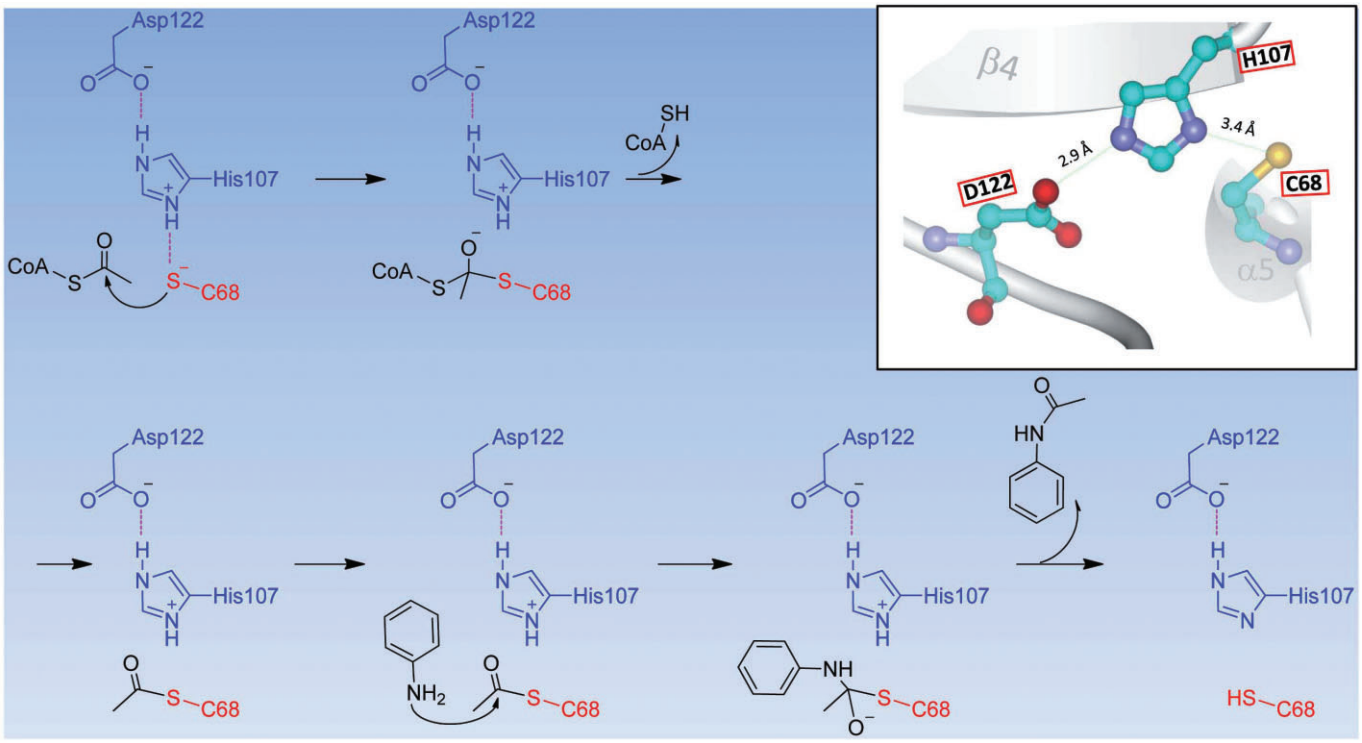

Figure 6

The proposed catalytic mechanism of NAT showing the importance of the catalytic triad Cys[68]-His[107]-Asp[122] (human NAT numbering) in initiating the acetylation reaction. The insert depicts the location of the triad residues according to the crystal structure of human NAT2 (PDB code: 2PFR).

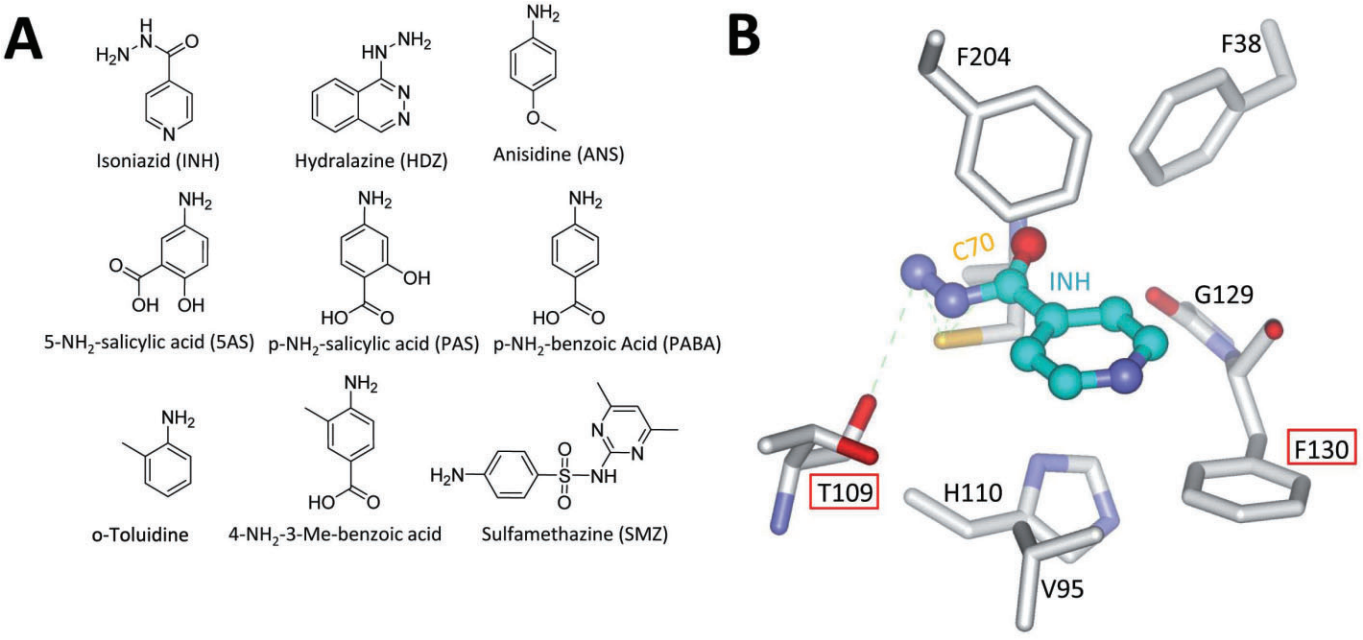

Figure 7

Substrate-binding site of NAT from *M. smegmatis* (MSNAT) and the structural determinants of its substrate preference. Panel A: Chemical structures of NAT substrates. Panel B: Diagram representation of MSNAT–isoniazid interactions (PDB code: 1W6F). Interactions with T109 and F130 are important in substrate binding.

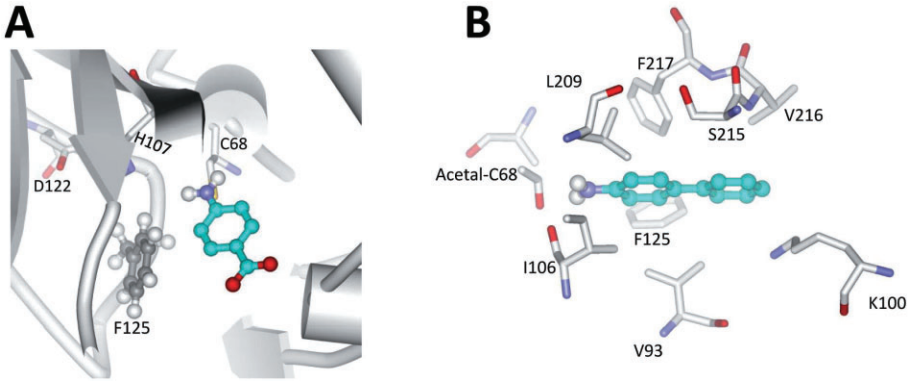

Figure 8

Predicted molecular interactions between p-aminobenzoic acid (PABA) and human NAT1 (Panel A; PDB code: 2DSS) and between 4-aminobiphenyl and human NAT1 (Panel B; PDB code: 2GWZ) based on the NMR model of human NAT1.

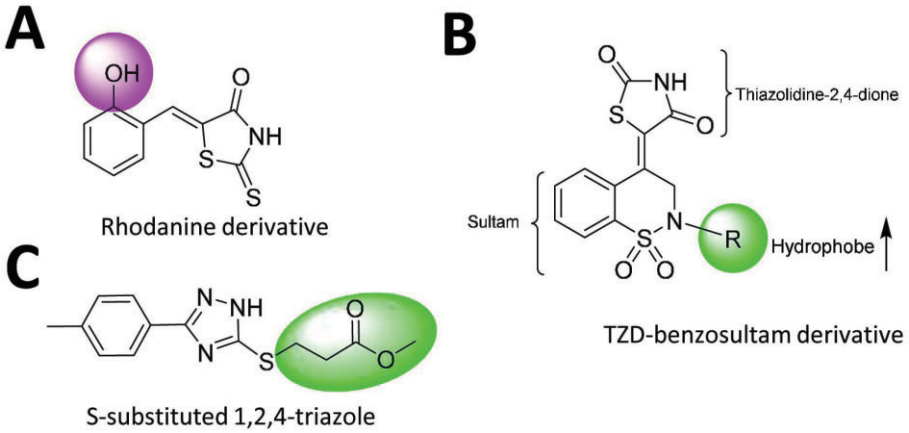

Figure 9

Chemical structures of three types of NAT inhibitors, rhodanine derivative (panel A), TZD-benzosultam derivative (panel B) and S-substituted 1,2,4-triazole (panel C). For TZD-benzosultam derivatives, the hydrophobicity of the N-substitution correlates with the binding affinity.

of NATs in various organisms (Westwood *et al.*, 2011). Russell *et al.* (2009) synthesized and evaluated a series of rhodanine (2-sulfanylidene-1,3-thiazolidin-4-one) derivatives as selective inhibitors of human NAT1. The (Z)-5-(20-hydroxybenzylidene)-2-thioxothiazolidin-4-one (Figure 9A) was one of the most potent inhibitors and this compound was docked into the crystal structure of human NAT1 to explore the inhibitor binding mode (Russell *et al.*, 2009). The aryl substituent extends into a hydrophobic pocket surrounded by residues V93, F125, V216 and F287, providing a possible explanation for the preference for lipophilic substituents. Binding of the inhibitor is also stabilized by hydrogen bonds with residues R127 and T289. The additional hydrogen bonding between the 2′-hydroxyl of the inhibitor and the enzyme may provide an explanation for the relatively high potency of 2′-hydroxyl-substituted rhodanine over its 4′-hydroxyl-substituted counterpart (Russell *et al.*, 2009).

Brooke *et al.* (2003a) identified the N-(2-naphthyl)-methyl substituted 1,1-dioxo-2,3-dihydrobenzo[1,2]-thiazine-4-ylidene thiazolidine-2,4-dione [N-(2-naphthyl)-methyl and 2,4-thiazolidinedione (TZD)-substituted benzosultam, Figure 9B] to be a weak inhibitor (IC_{50} = 78 μM) of MSNAT by high throughput screening of a proprietary small molecule library. Subsequent synthesis and screening of a series of TZD-sultam adducts revealed that the TZD moiety was essential for activity as well as the substitution on the sultam nitrogen (Brooke *et al.*, 2003b). The most effective NAT inhibitors in this series were those incorporating hydrophobic substituents at the sultam nitrogen. Docking of the TZD-sultam adducts with the crystal structure of MSNAT revealed the inhibition mechanism and the detailed molecular interactions between the inhibitor and the enzyme. The sultam moiety was located adjacent to the catalytic residue cysteine, providing a rationale for the observed competitive inhibition. The substituents at the sultam nitrogen occupied a hydrophobic groove in the protein, with a tryptophan residue, W97, appearing to hold the compound in place (Brooke *et al.*, 2003b). The imidazole moiety of H203 participated in a π-stacking interaction with the TZD moiety of the

Table 2

Mutation sites of human NAT alleles and their location in the NAT proteins as well as the functional effects of the mutations

NAT	Mutation site	*NAT* Allele	Location in the protein	Functional effect	Reference
NAT1	R117	*NAT1*5*	On the surfaces of the protein	Mutants may be subject to increased ubiquitinylation, leading to reduced protein level and reduction in the enzymic activity. Neither the V149I nor the S214A residue changes alter the structural stability of NAT1. No functional changes occur with M205V and E261K mutations.	Wu *et al.*, 2007 Hein, 2002 Liu *et al.*, 2006 Walraven *et al.*, 2008a
	V149	*NAT1*11A* *NAT1*11B* *NAT1*30*			
	R166	*NAT1*5*			
	M205	*NAT1*21*			
	S214	*NAT1*11A* *NAT1*11B* *NAT1*11C*			
	E261	*NAT1*24*			
	R64	*NAT1*17* *NAT1*19B*	On the α4-α5 loop	R64 forms H-bonds with the neighbouring residues E38 and N41. The stability of the enzyme is compromised in the absence of these interactions.	Wu *et al.*, 2007 Walraven *et al.*, 2008a
	E167	*NAT1*5*	At the beginning of β10	E167 forms H-bonds with the neighbouring residues K185 and D251. The mutant may affect protein stability.	Wu *et al.*, 2007
	R187	*NAT1*14A* *NAT1*14B*	In the 17-residue insertion	R187 forms an H-bond with E182. Substitution of R187 most likely decreases protein stability and lowers protein levels. The mutant may also alter the active site topology.	Wu *et al.*, 2007 Hughes *et al.*, 1998
	D251	*NAT1*22*	On the strand β15	D251 forms H-bonds with the neighbouring residues R242 and N245. The mutant may break these interactions and result in destabilization of the protein.	Wu *et al.*, 2007 Hein, 2002 Lin *et al.*, 1998
	I263	*NAT1*25*	In the α11	No change in protein level or catalytic activity for the I263V mutant because the hydrophobic interactions of the residue with others are preserved without introducing steric clashes.	Walraven *et al.*, 2008a
NAT2	I114	*NAT2*5* *NAT2*14C/F*	On the surfaces of the protein	Mutants may be subject to increased ubiquitinylation, leading to reduced protein level and reduction in the enzymic activity.	Wu *et al.*, 2007 Hein, 2002 Liu *et al.*, 2006
	E167	*NAT2*10*			
	R197	*NAT2*5E/J* *NAT2*6* *NAT2*14D*			
	K268	*NAT2*5* *NAT2*6C/F* *NAT2*12* *NAT2*14C/E-G/I*			
	K282	*NAT2*18*			
	G286	*NAT2*6I/J* *NAT2*7*			
	R64	*NAT2*7D* *NAT2*14* *NAT2*19*	On the α4-α5 loop	R64 forms H-bonds with the neighbouring residues E38 and N41. The stability of the enzyme is compromised in the absence of these interactions.	Wu *et al.*, 2007 Walraven *et al.*, 2008b
	D122	*NAT2*12D*	On the β5-α6 loop	D122 is a member of the catalytic triad. Mutations of D122 would adversely affect the activity of the enzyme.	Wu *et al.*, 2007 Walraven *et al.*, 2008b
	L137	*NAT2*5I*	On the β6-β7 loop	L137 makes contacts with residues L194 and W159 through hydrophobic interactions. The mutant may result in a change in secondary structure that could trigger degradation mechanisms.	Wu *et al.*, 2007 Walraven *et al.*, 2008b
	Q145	*NAT2*17*	On the β7-β8 loop	Q145 forms H-bonds with the neighbouring residues W132 and Q133. The mutant shows lower enzymic activity that may be due to reduced expression levels.	Hein, 2002 Wu *et al.*, 2007

compounds, which is consistent with the observed inactivity of the precursor ketone and ketal derivatives (Brooke *et al.*, 2003b).

A more recent study characterized the inhibitory potency of a series of 1,2,4-triazoles (with long chain aliphatic or planar aromatic substituents on sulfur; Figure 9C) against *M. tuberculosis* NAT (Westwood *et al.*, 2010). Interestingly, the potency of NAT inhibition increased with extended chain length up to approximately 10 Å, with the S-4-phenylbenzyl and S-octyl derivatives being the most potent in the series. The S-decyl derivative has an approximate chain length of 13 Å and was a less potent inhibitor of NAT than the shorter chain S-alkyl derivatives. The IC_{50} value against NAT of the S-decyl derivative was 13.1 μM, compared with an IC_{50} value of 3.8 μM for the S-octyl-1,2,4-triazole. This trend in acetylation difference was explored by molecular docking of the S-substituted 1,2,4-triazoles with NAT from *M. tuberculosis*. All except the S-decyl derivative were predicted to bind in essentially identical orientations. The triazole ring formed π-stacking interactions with F130 and formed hydrogen-bonds with the backbone carbonyls of F130 and G131. The 4′-methylphenyl group in all molecules (except the S-decyl derivative) pointed towards the active site cysteine residue. The correlation observed between the chain length and NAT inhibitory potency was closely matched by the docking scores for the same series of compounds, indicating that the crystal structures for NATs are powerful tools for structural design of inhibitors and for understanding of the affinity differences of the ligands.

NAT Polymorphisms

Acetylation was the first enzymic polymorphism to be investigated and was based on the observation that the anti-tubercular drug isoniazid (INH) caused a different level of neural toxicity across different populations (Grant *et al.*, 1997; Walker *et al.*, 2009). NAT polymorphisms have been linked to three types of acetylator phenotypes, namely, fast, intermediate and slow acetylators (Hein *et al.*, 2000; Hein, 2006). It has been suggested that slow acetylators are most at risk from INH hepatotoxicity because the toxic metabolite acetylhydrazine is retained much longer in the liver, compared with fast acetylators (Ellard *et al.*, 1978; Metushi *et al.*, 2011). In NAT1, there are 27 variant alleles that are associated with 11 amino acid alterations, whereas in NAT2, there are 65 variant alleles associated with alterations of 22 amino acids. A complete list of the NAT alleles and their associated phenotypes can be obtained online (http://louisville.edu/medschool/pharmacology). Understanding of why and how the polymorphisms alter the activity of human NATs has been significantly advanced by the availability of the crystal structures for the enzymes (Wu *et al.*, 2007). Based on the location and/or functional role of the original amino acids in the protein, the functional effects of NAT polymorphisms are well elucidated (Wu *et al.*, 2007). Table 2 summarizes the correlations of the structural data and functional implications for the NAT mutations. For more detailed discussion, one is encouraged to read the reviews by Walraven *et al.* (2008a,b).

Conclusion

NAT plays an important role in the biotransformation of many aromatic and heterocyclic amine drugs. In addition, it has been linked to cancer risk because of its roles in the metabolic activation of carcinogens and in cell growth and survival. This review has described the crystal structures for NATs and analysed the structural similarities and differences between prokaryotic and mammalian NATs. Furthermore, breakthroughs in understanding of the catalytic mechanism, substrate/inhibitor binding and polymorphisms were discussed from a structural perspective. Although the crystal structures represent a powerful tool for predicting substrate selectivity, they have also been very useful in rational design of selective inhibitors of NATs with great potential in cancer treatment.

Acknowledgement

The authors acknowledge start-up funds to BW from Jinan University.

Conflict of interest

The authors report no conflict of interest.

References

Adam PJ, Berry J, Loader JA, Tyson KL, Craggs G, Smith P *et al.* (2003). Arylamine N-acetyltransferase-1 is highly expressed in breast cancers and conveys enhanced growth and resistance to etoposide in vitro. Mol Cancer Res 1: 826–835.

Andres HH, Kolb HJ, Schreiber RJ, Weiss L (1983). Characterization of the active site, substrate specificity and kinetic properties of acetyl-CoA:arylamine N-acetyltransferase from pigeon liver. Biochim Biophys Acta 746: 193–201.

Andres HH, Klem AJ, Schopfer LM, Harrison JK, Weber WW (1988). On the active site of liver acetyl-CoA. Arylamine N-acetyltransferase from rapid acetylator rabbits (III/J). J Biol Chem 263: 7521–7527.

Brooke EW, Davies SG, Mulvaney AW, Pompeo F, Sim E, Vickers RJ (2003a). An approach to identifying novel substrates of bacterial arylamine N-acetyltransferases. Bioorg Med Chem 11: 1227–1234.

Brooke EW, Davies SG, Mulvaney AW, Okada M, Pompeo F, Sim E *et al.* (2003b). Synthesis and in vitro evaluation of novel small molecule inhibitors of bacterial arylamine N-acetyltransferases (NATs). Bioorg Med Chem Lett 13: 2527–2530.

Butcher NJ, Minchin RF (2012). Arylamine N-acetyltransferase 1: a novel drug target in cancer development. Pharmacol Rev 64: 147–165.

Butcher NJ, Boukouvala S, Sim E, Minchin RF (2002). Pharmacogenetics of the arylamine N-acetyltransferases. Pharmacogenomics J 2: 30–42.

Butcher NJ, Tetlow NL, Cheung C, Broadhurst GM, Minchin RF (2007). Induction of human arylamine N-acetyltransferase type I by androgens in human prostate cancer cells. Cancer Res 67: 85–92.

Dupret JM, Grant DM (1992). Site-directed mutagenesis of recombinant human arylamine N-acetyltransferase expressed in *Escherichia coli*. Evidence for direct involvement of Cys68 in the catalytic mechanism of polymorphic human NAT2. J Biol Chem 267: 7381–7385.

Ellard GA, Mitchison DA, Girling DJ, Nunn AJ, Fox W (1978). The hepatic toxicity of isoniazid among rapid and slow acetylators of the drug. Am Rev Respir Dis 118: 628–629.

Fretland AJ, Doll MA, Gray K, Feng Y, Hein DW (1997). Cloning, sequencing, and recombinant expression of NAT1, NAT2, and NAT3 derived from the C3H/HeJ (rapid) and A/HeJ (slow) acetylator inbred mouse: functional characterization of the activation and deactivation of aromatic amine carcinogens. Toxicol Appl Pharmacol 142: 360–366.

Fullam E, Westwood IM, Anderton MC, Lowe ED, Sim E, Noble ME (2008). Divergence of cofactor recognition across evolution: coenzyme A binding in a prokaryotic arylamine N-acetyltransferase. J Mol Biol 375: 178–191.

Gouet P, Courcelle E (2002). ENDscript: a workflow to display sequence and structure information. Bioinformatics 18: 767–768.

Grant DM, Blum M, Beer M, Meyer UA (1991). Monomorphic and polymorphic human arylamine N-acetyltransferases: a comparison of liver isozymes and expressed products of two cloned genes. Mol Pharmacol 39: 184–191.

Grant DM, Hughes NC, Janezic SA, Goodfellow GH, Chen HJ, Gaedigk A *et al*. (1997). Human acetyltransferase polymorphisms. Mutat Res 376: 61–70.

Hanna PE (1994). N-acetyltransferases, O-acetyltransferases, and N,O-acetyltransferases: enzymology and bioactivation. Adv Pharmacol 27: 401–430.

Hein DW (2002). Molecular genetics and function of NAT1 and NAT2: role in aromatic amine metabolism and carcinogenesis. Mutat Res 506–507: 65–77.

Hein DW (2006). N-acetyltransferase 2 genetic polymorphism: effects of carcinogen and haplotype on urinary bladder cancer risk. Oncogene 25: 1649–1658.

Hein DW, Doll MA, Rustan TD, Gray K, Feng Y, Ferguson RJ *et al*. (1993). Metabolic activation and deactivation of arylamine carcinogens by recombinant human NAT1 and polymorphic NAT2 acetyltransferases. Carcinogenesis 14: 1633–1638.

Hein DW, Doll MA, Fretland AJ, Leff MA, Webb SJ, Xiao GH *et al*. (2000). Molecular genetics and epidemiology of the NAT1 and NAT2 acetylation polymorphisms. Cancer Epidemiol Biomarkers Prev 9: 29–42.

Hickman D, Pope J, Patil SD, Fakis G, Smelt V, Stanley LA *et al*. (1998). Expression of arylamine N-acetyltransferase in human intestine. Gut 42: 402–409.

Holton SJ, Dairou J, Sandy J, Rodrigues-Lima F, Dupret JM, Noble ME *et al*. (2005). Structure of *Mesorhizobium loti* arylamine N-acetyltransferase 1. Acta Crystallograph Sect F Struct Biol Cryst Commun 61 (Pt 1): 14–16.

Hughes NC, Janezic SA, McQueen KL, Jewett MA, Castranio T, Bell DA *et al*. (1998). Identification and characterization of variant alleles of human acetyltransferase NAT1 with defective function using p-aminosalicylate as an in-vivo and in-vitro probe. Pharmacogenetics 8: 55–66.

Kawamura A, Graham J, Mushtaq A, Tsiftsoglou SA, Vath GM, Hanna PE *et al*. (2005). Eukaryotic arylamine N-acetyltransferase. Investigation of substrate specificity by high-throughput screening. Biochem Pharmacol 69: 347–359.

Laurieri N, Crawford MH, Kawamura A, Westwood IM, Robinson J, Fletcher AM *et al*. (2010). Small molecule colorimetric probes for specific detection of human arylamine N-acetyltransferase 1, a potential breast cancer biomarker. J Am Chem Soc 132: 3238–3239.

Lin HJ, Probst-Hensch NM, Hughes NC, Sakamoto GT, Louie AD, Kau IH *et al*. (1998). Variants of N-acetyltransferase NAT1 and a case-control study of colorectal adenomas. Pharmacogenetics 8: 269–281.

Liu F, Zhang N, Zhou X, Hanna PE, Wagner CR, Koepp DM *et al*. (2006). Arylamine N-acetyltransferase aggregation and constitutive ubiquitylation. J Mol Biol 361: 482–492.

Metushi IG, Cai P, Zhu X, Nakagawa T, Uetrecht JP (2011). A fresh look at the mechanism of isoniazid-induced hepatotoxicity. Clin Pharmacol Ther 89: 911–914.

Payton M, Mushtaq A, Yu TW, Wu LJ, Sinclair J, Sim E (2001). Eubacterial arylamine N-acetyltransferases – identification and comparison of 18 members of the protein family with conserved active site cysteine, histidine and aspartate residues. Microbiology 147 (Pt 5): 1137–1147.

Russell AJ, Westwood IM, Crawford MH, Robinson J, Kawamura A, Redfield C *et al*. (2009). Selective small molecule inhibitors of the potential breast cancer marker, human arylamine N-acetyltransferase 1, and its murine homologue, mouse arylamine N-acetyltransferase 2. Bioorg Med Chem 17: 905–918.

Sandy J, Mushtaq A, Kawamura A, Sinclair J, Sim E, Noble M (2002). The structure of arylamine N-acetyltransferase from *Mycobacterium smegmatis* – an enzyme which inactivates the anti-tubercular drug, isoniazid. J Mol Biol 318: 1071–1083.

Sandy J, Holton S, Fullam E, Sim E, Noble M (2005a). Binding of the anti-tubercular drug isoniazid to the arylamine N-acetyltransferase protein from *Mycobacterium smegmatis*. Protein Sci 14: 775–782.

Sandy J, Mushtaq A, Holton SJ, Schartau P, Noble ME, Sim E (2005b). Investigation of the catalytic triad of arylamine N-acetyltransferases: essential residues required for acetyl transfer to arylamines. Biochem J 390 (Pt 1): 115–123.

Sim E, Payton M, Noble M, Minchin R (2000). An update on genetic, structural and functional studies of arylamine N-acetyltransferases in eucaryotes and procaryotes. Hum Mol Genet 9: 2435–2441.

Sim E, Sandy J, Evangelopoulos D, Fullam E, Bhakta S, Westwood I *et al*. (2008a). Arylamine N-acetyltransferases in mycobacteria. Curr Drug Metab 9: 510–519.

Sim E, Walters K, Boukouvala S (2008b). Arylamine N-acetyltransferases: from structure to function. Drug Metab Rev 40: 479–510.

Sinclair J, Sim E (1997). A fragment consisting of the first 204 amino-terminal amino acids of human arylamine N-acetyltransferase one (NAT1) and the first transacetylation step of catalysis. Biochem Pharmacol 53: 11–16.

Sinclair JC, Sandy J, Delgoda R, Sim E, Noble ME (2000). Structure of arylamine N-acetyltransferase reveals a catalytic triad. Nat Struct Biol 7: 560–564.

Tiang JM, Butcher NJ, Minchin RF (2010). Small molecule inhibition of arylamine N-acetyltransferase Type I inhibits proliferation and invasiveness of MDA-MB-231 breast cancer cells. Biochem Biophys Res Commun 393: 95–100.

Vagena E, Fakis G, Boukouvala S (2008). Arylamine N-acetyltransferases in prokaryotic and eukaryotic genomes: a survey of public databases. Curr Drug Metab 9: 628–660.

Walker K, Ginsberg G, Hattis D, Johns DO, Guyton KZ, Sonawane B (2009). Genetic polymorphism in N-Acetyltransferase (NAT): population distribution of NAT1 and NAT2 activity. J Toxicol Environ Health B Crit Rev 12: 440–472.

Walraven JM, Trent JO, Hein DW (2007). Computational and experimental analyses of mammalian arylamine N-acetyltransferase structure and function. Drug Metab Dispos 35: 1001–1007.

Walraven JM, Trent JO, Hein DW (2008a). Structure-function analyses of single nucleotide polymorphisms in human N-acetyltransferase 1. Drug Metab Rev 40: 169–184.

Walraven JM, Zang Y, Trent JO, Hein DW (2008b). Structure/function evaluations of single nucleotide polymorphisms in human N-acetyltransferase 2. Curr Drug Metab 9: 471–486.

Wang H, Vath GM, Gleason KJ, Hanna PE, Wagner CR (2004). Probing the mechanism of hamster arylamine N-acetyltransferase 2 acetylation by active site modification, site-directed mutagenesis, and pre-steady state and steady state kinetic studies. Biochemistry 43: 8234–8246.

Wang H, Liu L, Hanna PE, Wagner CR (2005). Catalytic mechanism of hamster arylamine N-acetyltransferase 2. Biochemistry 44: 11295–11306.

Watanabe M, Sofuni T, Nohmi T (1992). Involvement of Cys69 residue in the catalytic mechanism of N-hydroxyarylamine O-acetyltransferase of *Salmonella typhimurium*. Sequence similarity at the amino acid level suggests a common catalytic mechanism of acetyltransferase for *S. typhimurium* and higher organisms. J Biol Chem 267: 8429–8436.

Westwood IM, Holton SJ, Rodrigues-Lima F, Dupret JM, Bhakta S, Noble ME *et al*. (2005). Expression, purification, characterization and structure of *Pseudomonas aeruginosa* arylamine N-acetyltransferase. Biochem J 385 (Pt 2): 605–612.

Westwood IM, Kawamura A, Fullam E, Russell AJ, Davies SG, Sim E (2006). Structure and mechanism of arylamine N-acetyltransferases. Curr Top Med Chem 6: 1641–1654.

Westwood IM, Bhakta S, Russell AJ, Fullam E, Anderton MC, Kawamura A *et al*. (2010). Identification of arylamine N-acetyltransferase inhibitors as an approach towards novel anti-tuberculars. Protein Cell 1: 82–95.

Westwood IM, Kawamura A, Russell AJ, Sandy J, Davies SG, Sim E (2011). Novel small-molecule inhibitors of arylamine N-acetyltransferases: drug discovery by high-throughput screening. Comb Chem High Throughput Screen 14: 117–124.

Wu H, Dombrovsky L, Tempel W, Martin F, Loppnau P, Goodfellow GH *et al*. (2007). Structural basis of substrate-binding specificity of human arylamine N-acetyltransferases. J Biol Chem 282: 30189–30197.

Zhang N, Liu L, Liu F, Wagner CR, Hanna PE, Walters KJ (2006). NMR-based model reveals the structural determinants of mammalian arylamine N-acetyltransferase substrate specificity. J Mol Biol 363: 188–200.