Biochemical and Biophysical Research Communications **291,** 116–123 (2002)
doi:10.1006/bbrc.2002.6414, available online at http://www.idealibrary.com on IDEAL®

3D Model of Human Arylamine *N*-Acetyltransferase 2: Structural Basis of the Slow Acetylator Phenotype of the R64Q Variant and Analysis of the Active-Site Loop

Fernando Rodrigues-Lima*,† and Jean-Marie Dupret*,†,[1]
*CNRS-UMR7000, Faculté de Médecine Pitié-Salpêtrière, 105 boulevard de l'Hôpital, 75013 Paris, France; and
†UFR de Biochimie, Université Denis Diderot-Paris 7, 2 Place Jussieu, Paris, France

Received January 15, 2002

The human arylamine *N*-acetyltransferase NAT2 is responsible for the biotransformation of numerous arylamine drugs and carcinogens. A common polymorphism of the *NAT2* gene has been associated with susceptibility to drug toxicity and various malignancies. In this study, we used the crystal structure of the *Salmonella typhimurium* NAT (*St*NAT) to construct a high-quality model of a catalytic N-terminal region of NAT2 (residues 34–131). We show that this region has a similar structure in *St*NAT and the human isoforms NAT1 and NAT2. Comparison of the structures of these three molecules suggests that NATs have an active-site loop with a conserved structure, which is involved in substrate recognition. Our model is consistent with previous experimental data and provides the first plausible structural basis of the effects of a common genetic polymorphism (Arg^{64}→Gln) on NAT2 activity.

Key Words: **structure–function; homology-modeling; loop; polymorphism; NAT2.**

The acetyl-CoA:arylamine *N*-acetyltransferases (NAT; EC 2.3.1.5), catalyze the transfer of an acetyl group from Ac-CoA to the nitrogen or oxygen atom of primary arylamines, hydrazines, and their N-hydroxylated metabolites. These enzymes therefore play an important role in the detoxification and potential metabolic activation of numerous xenobiotics. NAT enzymes have been identified in several vertebrate and eubacterial species (1). The human NAT1 and NAT2 enzymes are encoded by two genes located on chromosome 8 (2). Interindividual genetic variations in these genes have been identified that cause differences in NAT2 protein levels and functional activity. These variations are known to lead to the acetylation polymorphism. Susceptibility to certain types of cancer and drug responses phenotypes have been associated with the slow and rapid phenotypes (3, 4).

Population studies with NAT2-selective probe drugs and *in vitro* analysis of NAT2 isoforms have led to the identification of 19 of the 26 known variants as slow-type acetylator enzymes (5). Allelic variants (*NAT2*5, *6, *7, *14*) bearing amino acid substitutions (Arg^{64}→Gln due to G^{191}→A, Ile^{114}→Thr due to T^{341}→C, Arg^{197}→Gln due to G^{590}→A, Gly^{286}→Glu due to G^{857}→A) have been associated with slow acetylation (5). Significant ethnic differences in *NAT2** allele distribution have been observed. For example, the G^{857}→A substitution (alleles *NAT2*7A* and *NAT2*7B*) is frequent in Asians (12%) but rare in Caucasians and Africans (1–2%). The G^{191}→A substitution (present in alleles *NAT2*14A, *14B, *14C, *14D, *14E, *14F,* and **14G*), originally identified in African–Americans (6) and native Africans (7), is much more frequent among Africans (7%) than among Caucasians and Asians (1–2%). Improvements in our understanding of the relationships between polymorphic substitution at the *NAT2* locus and the activity of the variant enzymes would be of great value in the design of new drugs.

Functional analysis and site-directed mutagenesis experiments have led to the identification of a strictly conserved Cys residue (corresponding to Cys^{68} in human NATs) as essential for the catalytic function of NATs (8, 9). A highly conserved basic residue, corresponding to Arg^{64} in the human NATs, has been shown to contribute to the conformational stability of NATs (10) and regions critical for the substrate specificity of human NATs have also been mapped (11, 12).

The crystallographic structure of *Salmonella typhimurium* NAT (*St*NAT) has been resolved (13). Three critical residues have been identified, which presum-

Abbreviations used: Ac-CoA, acetyl-CoA; NAT, arylamine *N*-acetyltransferase; rmsd, root mean square deviation.

[1] To whom correspondence and reprint requests should be addressed at CNRS UMR 7000, Faculté de Médecine Pitié-Salpêtrière, 105 Bd de l'Hôpital, 75013 Paris, France. Fax: (33 1) 53 60 08 02. E-mail: jmdupret@infobiogen.fr.

0006-291X/02 $35.00

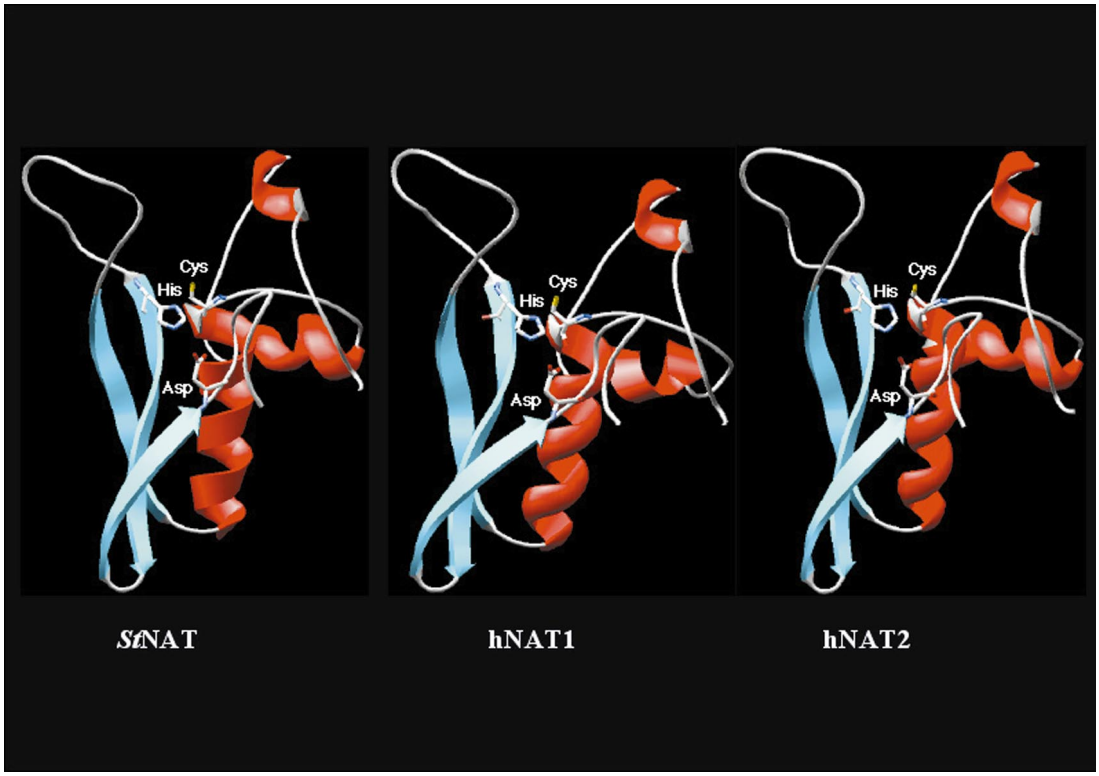

FIG. 1. Structure of the catalytic N-terminal domains of *St*NAT, human NAT1, and human NAT2. Swiss-PdbViewer (ribbon images) representation of the three-dimensional structure of *St*NAT (residues 35–131) (PDB Entry 1e2t) and of the homology models of human NAT1 (residues 34–131) (14) and NAT2 (residues 34–131). Catalytic triad residues are shown as stick models.

ably act together as a cysteine protease-like catalytic triad (corresponding to residues Cys[68], His[107], and Asp[122] in human NATs). A three-dimensional homology model of the human NAT1 N-terminal catalytic domain confirmed the existence of the catalytic triad in human NAT1 (14). Comparison of the structural elements of the active site revealed the existence in both *St*NAT and human NAT1 of a loop that forms part of the catalytic core (14). It has been suggested that some residues in this loop are involved in the binding of Ac-CoA (13) and may determine arylamine substrate selectivity (12). These residues include the amino acid at position 125 (corresponding to Phe[125] in NAT1 and Ser[125] in NAT2), which is believed to be involved in the steric control of substrate recognition (12, 14).

In a previous work, we produced a homology-based model of the N-terminal catalytic domain of human NAT1 enzyme based on the *St*NAT structure (13). Despite the high identity between the N-terminal regions of NAT1 and NAT2, we were unable to provide a model reliable enough of the NAT2 catalytic domain using the same approach (14). In this study, using the *St*NAT structure as a template and the homology-modeling software MODELLER, we report a high-quality 3D structural model of a region of the N-terminal catalytic domain of human NAT2 (residues 34–131). This model provides evidence for the structural conservation of an active-site loop in NATs and its probable involvement in substrate recognition mechanisms. This model is also the first to provide insight into the molecular basis of changes in NAT2 activity due to the polymorphic substitution Arg[64]→Gln. This model provides a rational explanation for most of the available experimental data.

MATERIALS AND METHODS

Sequence alignment and secondary structure prediction. The BLAST sequence alignment algorithm (15) was used to align the amino acid sequences of human NAT2 (Swissprot Accession No. P11245) and *St*NAT (Swissprot Accession No. PQ00267) with a gap open penality of 11 and a gap extension penality of 1 and using the BLOSUM62 matrix. PSI-Pred was used to predict secondary structure (16).

Homology modeling and structural analysis of a region of the human NAT2 N-terminal catalytic domain. A three-dimensional model of a catalytic N-terminal region (residues 34–131) of the NAT2 enzyme was constructed by comparative modeling, with the MODELLER program (17). The structure of the N-terminal domain of NAT2 considered was modeled, with the structure of *St*NAT (Protein Data Bank Entry 1e2t) used as a template. The sequence of the template (residues 35–132) was 32% identical to the sequence of NAT2 (residues 34–131) in the region considered. The alignment of the two sequences was improved by taking into account the secondary structures predicted for *St*NAT by crystallography and NAT2 by PSI-Pred. The best alignment was confirmed by comparing threading energies, using Swiss-PdbViewer version 3.7b (18), and by comparing the compatibility of the NAT2 sequence with the 3D structure

of *St*NAT, using the TITO program (19). The stereochemical geometry of the final model generated by MODELLER was evaluated with PROCHECK (20). Structures were visualized and analyzed with Swiss-PdbViewer version 3.7b (18).

RESULTS AND DISCUSSION

General Features of the Three-Dimensional Model of the Catalytic N-Terminal Region of Human NAT2

To circumvent the absence of crystallographic structures of human NATs, we carried out computational approaches to investigate the molecular structure of the catalytic region of the human NAT2. Homology modeling is now widely used for various purposes such as the comparison of macromolecules and the investigation of binding sites or catalytic pockets in proteins (21). It generates structural models that can be used to validate experimental results such as those produced in site-directed mutagenesis experiments. It can also provide starting models in X-ray crystallography and NMR spectroscopy (21).

The N-terminal catalytic domains of *St*NAT (residues 35–131) and human NAT2 (residues 34–131) have very similar sequences (14). We therefore used the N-terminal catalytic domain of *St*NAT as a structural template for homology modeling. BLAST 2 alignment of the amino acid sequences of the two N-terminal catalytic domains showed that these regions were 32% identical (data not shown). Comparison of the sequences of NATs from eukaryotes with those from prokaryotes has revealed that this N-terminal region is more highly conserved than the C-terminal region (4, 14, 22, 23), and that it is highly conserved in eukaryotes and eubacteria (22). Three strictly conserved residues (Cys^{68}, His^{107}, and Asp^{122} in human NAT2), which form a cysteine protease-like catalytic triad, are located within these conserved sequences (14). Comparison of the predicted secondary structure of the NAT2 N-terminal region with the secondary structure of *St*NAT deduced by crystallography showed that the α-helix and β-strands were similarly arranged in the two molecules (data not shown).

We used the MODELLER program to construct a model of a region (residues 34–131) of the N-terminal catalytic domain of human NAT2 encompassing the catalytic core of the enzyme (14), based on the crystal structure of *St*NAT. The best model obtained for the N-terminal region of NAT2 is shown as a Swiss-PdbViewer representation along with the equivalent *St*NAT structure (13) and the equivalent human NAT1 model (14). The stereochemical quality of the model was evaluated with PROCHECK: 85% of the residues were present in the most favored regions of the Ramachandran diagram. No residue was present in the disallowed regions. The overall quality G-factor was −0.17 (G-factor values of dihedral angles, main-chain bond lengths and main-chain bond angles were −0.12, −0.26, and −0.24, respectively), indicating that this model is of high quality [acceptable values of G-factors in PROCHECK are between 0 and −0.5 (20)].

Structural analysis of this NAT2 model showed that its backbone was extremely similar to that of the crystal structure of *St*NAT (residues 35–131) and of the human NAT1 model (Fig. 1). Indeed, structural alignments showed that α-carbon coordinates differed by only 0.7 Å between NAT2 and *St*NAT. The same rmsd value (0.7 Å) was obtained if the α-carbon atoms of the NAT2 and NAT1 models were compared (data not shown and Fig. 2). The location of the catalytic triad was found to be conserved in the structures of *St*NAT, NAT1 and NAT2 [(14); Fig. 1 and Fig. 2] with rmsd values for the three Cα atoms of only 0.2 Å (*St*NAT/NAT1) and 0.3 Å (*St*NAT/NAT2 and NAT1/NAT2). These results clearly indicate that the catalytic N-terminal domain of NAT enzymes is highly conserved. Analyses of the structure of *St*NAT and NAT1 have shown that these enzymes share a common catalytic core fold with the human transglutaminase Factor XIII and with members of the cysteine protease family such as cathepsin X (13, 14). This suggests that vertebrate and eubacterial NATs have adapted a mechanism commonly found in cysteine proteases for use in acetyl-transfer reactions (14).

Structural Analysis of the Active-Site Loop Involved in Ac-CoA Binding and Arylamine Substrate Selectivity

The structural alignment in Fig. 2 reveals the presence within the active site of human NATs of a loop of conserved structure spanning the residues Asp^{122} (catalytic triad residue) to Met^{131}. We evaluated the stereochemical geometry of this active-site loop in NAT1 and NAT2 and showed it to be correct (data not shown). Alignment of the NAT2 active-site loop with its NAT1 equivalent yielded a rmsd value for the Cα atoms of 0.3 Å (Fig. 2). The same rmsd value (0.3 Å) was obtained if the active-site loop of NAT1 or NAT2 was aligned with its *St*NAT equivalent (data not shown). These results strongly suggest that this active-site loop is structurally conserved from eubacteria to human NATs (Fig. 3A). Interestingly, the crystal structure of *Mycobacterium smegmatis* NAT supports our results by showing the extreme structural conservation of its active-site loop (24). It has been reported that this active-site loop may contain a P-loop motif (spanning residues Gly^{126} to Ser^{129} in NAT2), a nucleotide-binding structural motif involved in Ac-CoA binding (13) and arylamine substrate specificity (12, 14). The nucleotide-binding capacity of P-loops is known to depend on their three-dimensional structure rather than on their primary sequences (25). Thus, despite marked amino-acid differences at positions 127 and 129 (14), the active-site

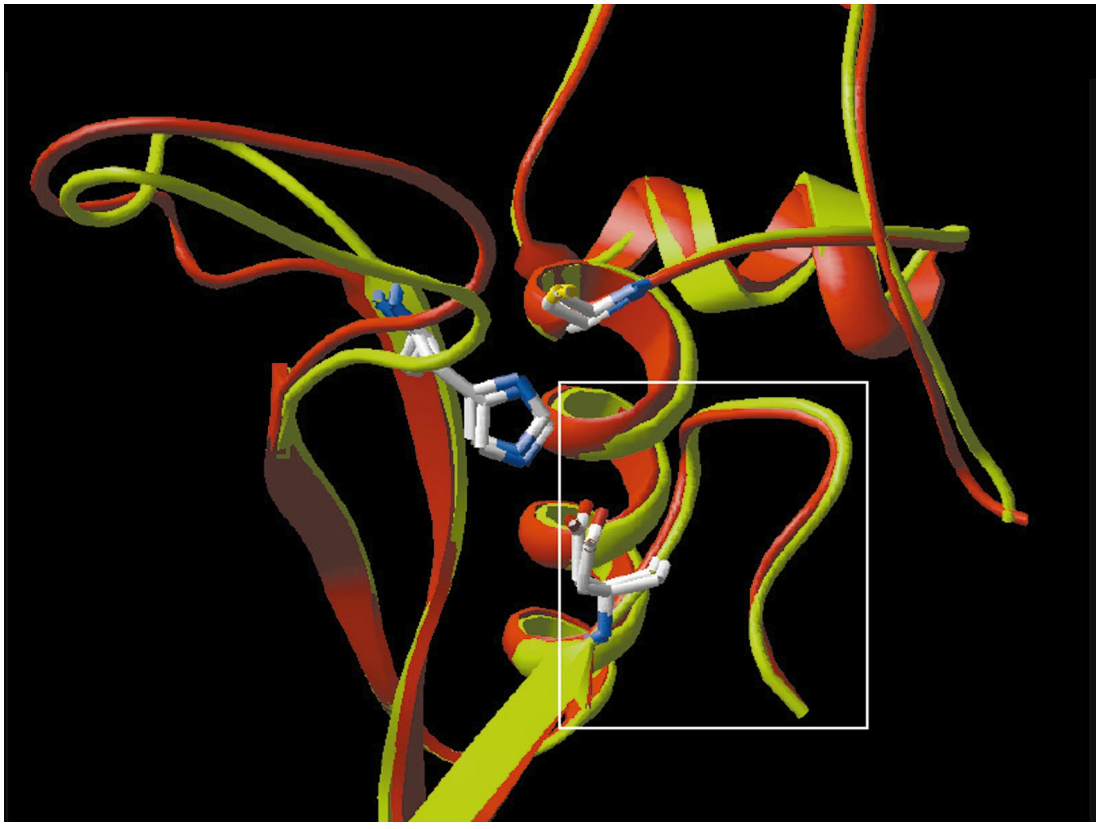

FIG. 2. Structural alignment of the catalytic cores of human NAT1 and NAT2. Swiss-PdbViewer representation of least-squares alignment of the Cα atoms of the NAT1 and NAT2 models. NAT1 is shown in green and NAT2 in red. Catalytic triad residues are shown as stick models. Active-site loops are boxed.

loops of NAT enzymes show striking structural similarity, as reported for other P-loops (26).

Further structural analysis suggested that the active-site loops of *St*NAT, NAT1 and NAT2 are stabilized by hydrogen bonds (Fig. 3A), as shown for other P-loops (26). For the three NAT active-site loops, the Asp residue (Asp^{122} in human NATs) of the strictly conserved catalytic triad is attached by hydrogen bonds to the backbone amine groups of the two adjacent C-terminal amino-acids, one of which is a highly conserved Gly residue (Gly^{124} in human NATs). These hydrogen bonds may enable the active-site loop to adopt the most favorable geometry for Ac-CoA and/or arylamine binding. Consequently, the strictly conserved catalytic Asp residue may play an additional role in the stabilization and orientation of the active-site loop. A similar hypothesis has been put forward for another active-site P-loop in a low-molecular-weight protein phosphatase (26). Another loop displaying no particular structure conservation (spanning residues Phe^{93} to Val^{106} in NAT2) is also present in the N-terminal catalytic region of NATs [(14) and Fig. 2]. However, stereochemical evaluation of the NAT2 model shows that this loop is the less reliable structural element of the model. As shown for other homology models (21), it may be difficult to model this loop due to its length, sequence and possible location at the surface of the enzyme (Fig. 2). Functional studies have suggested that this loop is unlikely to be involved in substrate selectivity (11). However the *St*NAT structure and our NAT1 and NAT2 models show that this loop is close to the active-site core [(13, 14) and Fig. 2, Fig. 3B]. Therefore, we cannot rule out a possible role of this loop in catalysis. Further functional experiments are needed to address this point.

Site-directed mutagenesis experiments have shown that the residue at position 125 in the active-site loop of human NATs is involved in arylamine substrate recognition (12). Our NAT1 (14) and NAT2 models clearly show that this active-site loop residue may be involved in arylamine substrate selectivity, and may act by a steric hindrance mechanism (Fig. 3B). The amino acid in position 125 (Phe^{125} in NAT1 and Ser^{125} in NAT2) is proximal to the catalytic triad and faces a passageway to the catalytic core [(14), Fig. 3B]. The hydroxy group of the Ser^{125} residue of NAT2 takes up less space than the bulky phenyl ring of the Phe^{125} residue in NAT1 (100 and 200 $Å^3$, respectively). The presence of a smaller group at position 125 of the NAT2 enzyme may thus facilitate the access of larger sub-

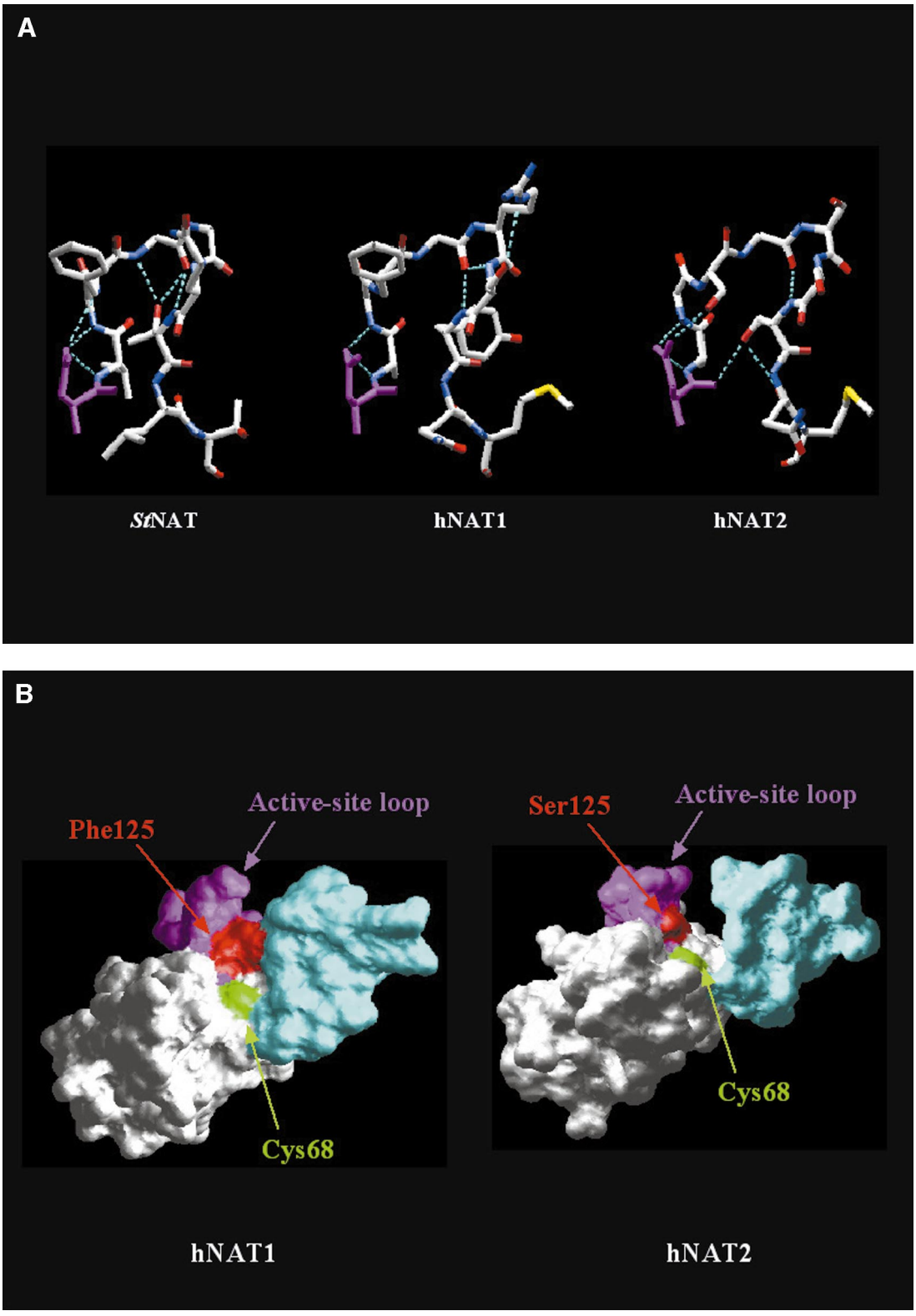

FIG. 3. Structural analysis of the active-site loop of NATs. (A) Swiss-PdbViewer (stick models) representation of the three-dimensional structure of the active-site loops of *St*NAT, human NAT1 and NAT2. The catalytic Asp residue is shown in magenta. Hydrogen bonds are shown as dashed turquoise lines. (B) Van der Waals representations of human NAT1 and NAT2 are shown under similar angles. The active-site loops of NAT1 and NAT2 are shown in magenta. The residues at position 125 (Phe^{125} for NAT1 and Ser^{125} for NAT2), which are believed to be involved in steric restriction, are shown in red. The catalytic Cys residue is shown in green. The loop spanning residues Val^{93} to Ile^{106} in NAT1 and spanning residues Phe^{93} to Val^{106} in NAT2 is shown in turquoise.

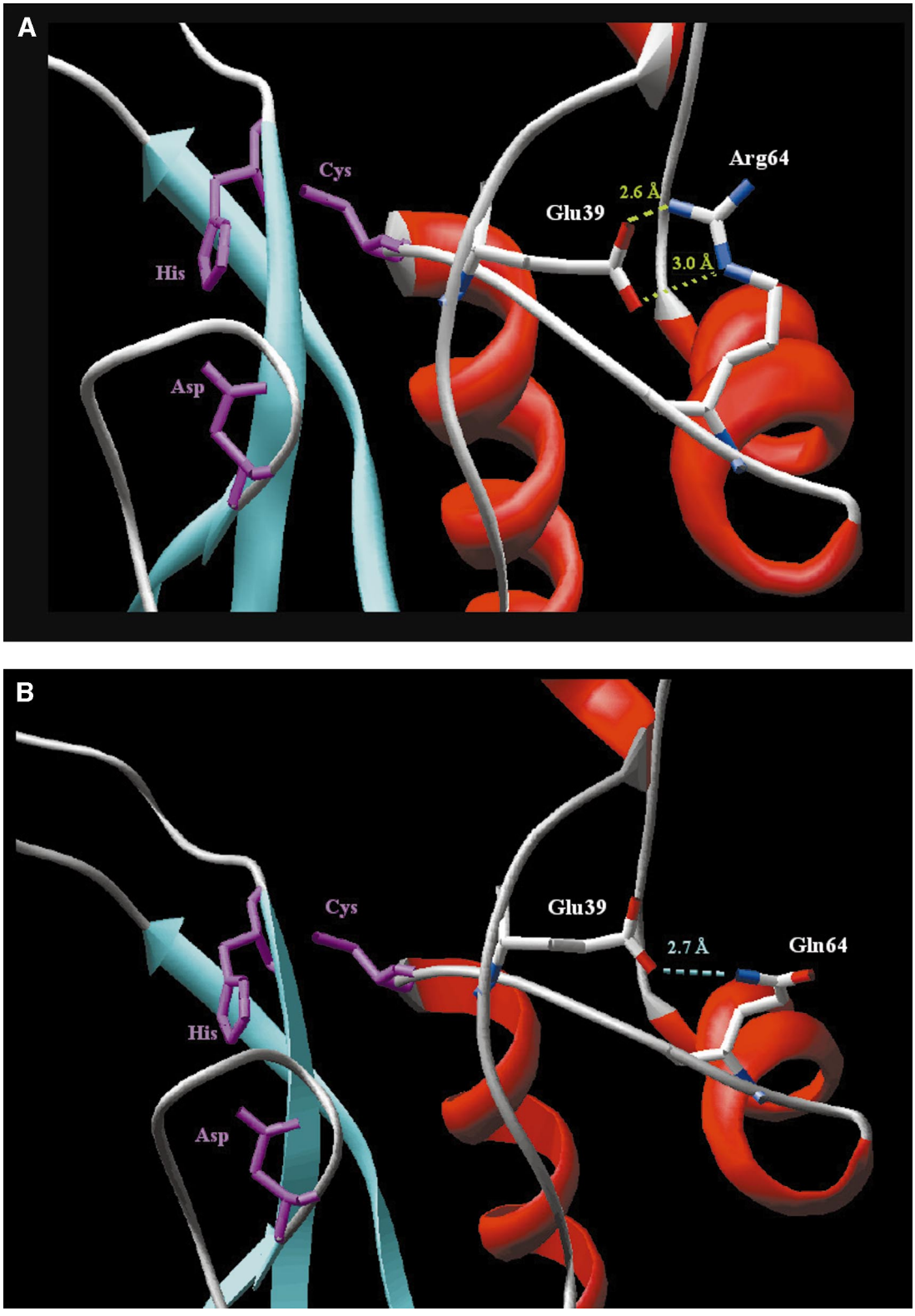

FIG. 4. Structural analysis of the polymorphic Arg^{64}→Gln substitution in the NAT2 enzyme. (A) Three-dimensional model of the catalytic core of the Arg^{64}-bearing NAT2 enzyme. Arg^{64} and Glu^{39} (shown as stick models) residues are bound by a "two-point" ionic interaction, shown as dashed green lines. Catalytic triad residues are shown as magenta stick models. (B) Three-dimensional model of the catalytic core of the Arg^{64}→Gln NAT2 variant. Gln^{64} and Glu^{39} (shown as stick models) residues are linked by a hydrogen bond, shown as a dashed turquoise line. Catalytic triad residues are shown in magenta as stick models.

strates, such as sulfamethazine (SMZ), to the active site (Fig. 3B). These structural data are consistent with functional studies and suggest that the amino acid at position 125 in NAT enzymes may be involved in arylamine substrate selectivity by restricting access to the active site (12). However, other amino acids in the C-terminal domain of NAT enzymes may also be involved in arylamine substrate recognition, as suggested by Sinclair and Sim (27).

Structural Analysis of the Arg^{64}→Gln NAT2 Variant

Several amino acid substitutions due to genetic polymorphism in *NAT1* and *NAT2* genes have been described (5). For example, the slow acetylator phenotype is associated with Arg^{64}→Gln substitution in NAT2 and the functional properties of this polymorphism have been extensively studied in prokaryote (10) and eukaryote (28) expression systems. A large reduction of the NAT2 catalytic activity (75% less active than Arg^{64}-bearing NAT2) was observed in variants with the Arg^{64}→Gln substitution (10, 28). Site-directed mutagenesis and functional experiments targeting amino acid position 64 of the NAT2 enzyme have also been performed with NAT2-specific substrates (10). We used our NAT2 model to rationalize the experimental data available and to gain insight into the structural mechanisms associated with the catalytic effects of this substitution.

Our NAT2 model suggests that in the Arg^{64}-bearing NAT2 enzyme (Fig. 4A), Arg^{64} interacts with Glu^{39}. In particular, it predicts that the carboxylate group of the conserved Glu^{39} residue is involved in a polar coplanar interaction (ionic interaction) with the positive guanidium group of Arg^{64}. Indeed, there is a "two-point" interaction between these "forked" structures (points of interaction 2.6 and 3.0 Å apart; Fig. 4A). This strong "two-point" polar interaction is probably involved in stabilization of the conformation of the NAT2 catalytic cysteine residue, thus putatively contributing to the correct spatial location and activation of this cysteine residue by the two other triad residues. Stabilization of the catalytic Cys residue of *St*NAT by ionic interactions between a Glu and an Arg residue has also been suggested (13). Given the high degree of conservation of these two residues in known NATs (23), such a mechanism probably extends to all NAT species.

The polymorphic Arg^{64}→Gln substitution in NAT2 reduces apparent maximal velocity (V_{max}) for sulfamethazine, a NAT2-specific substrate, by 75% (10, 28). Our model suggests that the Arg^{64}→Gln NAT2 variant has a single putative hydrogen bond (length 2.7 Å) between the hydrogen atom of the Glu^{39} carboxyl group and the nitrogen atom of the side chain of the Gln^{64} residue (Fig. 4B). The presence of such a hydrogen bond between Glu^{39} and Gln^{64} may provide only partial stabilization of the active-site Cys residue, thus accounting for the lower activity of the Arg^{64}→Gln NAT2 variant. Our model also predicts a lack of interaction between Glu^{39} and the amino acid at position 64, if that amino acid is an Ala or Met residue (data not shown). Interestingly, mutants with amino acid substitutions of this type (Arg^{64}→Ala and Arg^{64}→Met) have been shown to have no significant activity (10). Overall, our results suggest that a strong polar interaction between Glu^{39} and Arg^{64} is required for optimal NAT2 activity, with a weaker polar interaction resulting in low levels of activity. This notion is further supported by findings for an Arg^{64}→Lys NAT2 mutant, which has been shown to display catalytic activity similar to that of the Arg^{64}-bearing NAT2. Our model predicts that the replacement of Arg^{64} by a Lys residue should result in the maintenance of a strong polar interaction (ionic interaction) (data not shown) similar to that predicted for the Arg^{64}-bearing enzyme.

Overall, these data provide a first insight into the likely structural basis of the effects of a common single-nucleotide polymorphism in NAT2 on enzymatic activity.

ACKNOWLEDGMENTS

This work was supported by the Centre National de la Recherche Scientifique (CNRS), University Paris 6, and the Association pour la Recherche sur le Cancer (ARC).

REFERENCES

1. Upton, A., Johnson, N., Sandy, J., and Sim, E. (2001) Arylamine *N*-acetyltransferases—Of mice, men and microorganisms. *Trends Pharmacol. Sci.* **22,** 140–146.
2. Matas, N., Thygesen, P., Stacey, M., Risch, A., and Sim, E. (1997) Mapping AAC1, AAC2 and AACP, the genes for arylamine N-acetyltransferases, carcinogen metabolising enzymes on human chromosome 8p22, a region frequently deleted in tumours. *Cytogenet. Cell Genet.* **77,** 290–295.
3. Hein, D. W., Doll, M. A., Fretland, A. J., Leff, M. A., Webb, S. J., Xiao, G. H., Devanaboyina, U. S., Nangju, N. A., and Feng, Y. (2000) Molecular genetics and epidemiology of the NAT1 and NAT2 acetylation polymorphisms. *Cancer Epidemiol. Biomarkers Prev.* **9,** 29–42.
4. Sim, E., Payton, M., Noble, M., and Minchin, R. (2000) An update on genetic, structural and functional studies of arylamine N-acetyltransferases in eucaryotes and procaryotes [in process citation]. *Hum. Mol. Genet.* **9,** 2435–2441.
5. Hein, D. W., Grant, D. M., and Sim, E. (2000) Update on consensus arylamine N-acetyltransferase gene nomenclature [letter]. *Pharmacogenetics* **10,** 291–292.
6. Bell, D. A., Taylor, J. A., Butler, M. A., Stephens, E. A., Wiest, J., Brubaker, L. H., Kadlubar, F. F., and Lucier, G. W. (1993) Genotype/phenotype discordance for human arylamine N-acetyltransferase (NAT2) reveals a new slow-acetylator allele common in African–Americans. *Carcinogenesis* **14,** 1689–1692.
7. Deloménie, C., Sica, L., Grant, D. M., Krishnamoorthy, R., and Dupret, J. M. (1996) Genotyping of the polymorphic *N*-acetyltransferase (*NAT2**) gene locus in two native African populations. *Pharmacogenetics* **6,** 177–185.
8. Dupret, J. M., and Grant, D. M. (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.

9. Watanabe, M., Sofuni, T., and 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.
10. Deloménie, C., Goodfellow, G. H., Krishnamoorthy, R., Grant, D. M., and Dupret, J. M. (1997) Study of the role of the highly conserved residues Arg9 and Arg64 in the catalytic function of human *N*-acetyltransferases NAT1 and NAT2 by site-directed mutagenesis. *Biochem. J*. **323,** 207–215.
11. Dupret, J. M., Goodfellow, G. H., Janezic, S. A., and Grant, D. M. (1994) Structure–function studies of human arylamine *N*-acetyltransferases NAT1 and NAT2. Functional analysis of recombinant NAT1/NAT2 chimeras expressed in *Escherichia coli*. *J. Biol. Chem*. **269,** 26830–26835.
12. Goodfellow, G. H., Dupret, J. M., and Grant, D. M. (2000) Identification of amino acids imparting acceptor substrate selectivity to human arylamine acetyltransferases NAT1 and NAT2. *Biochem. J*. **348**(Pt. 1), 159–166.
13. Sinclair, J. C., Sandy, J., Delgoda, R., Sim, E., and Noble, M. E. (2000) Structure of arylamine *N*-acetyltransferase reveals a catalytic triad. *Nat. Struct. Biol*. **7,** 560–564.
14. Rodrigues-Lima, F., Deloménie, C., Goodfellow, G. H., Grant, D. M., and Dupret, J. M. (2001) Homology modelling and structural analysis of human arylamine N-acetyltransferase NAT1: Evidence for the conservation of a cysteine protease catalytic domain and an active-site loop. *Biochem. J*. **356,** 327–334.
15. Tatusova, T. A., and Madden, T. L. (1999) BLAST 2 sequences, a new tool for comparing protein and nucleotide sequences. *FEMS Microbiol. Lett*. **174,** 247–250.
16. Jones, D. T. (1999) Protein secondary structure prediction based on position-specific scoring matrices. *J. Mol. Biol*. **292,** 195–202.
17. Sali, A., and Blundell, T. L. (1993) Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol*. **234,** 779–815.
18. Guex, N., and Peitsch, M. C. (1997) SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. *Electrophoresis* **18,** 2714–2723.
19. Labesse, G., and Mornon, J. (1998) Incremental threading optimization (TITO) to help alignment and modelling of remote homologues. *Bioinformatics* **14,** 206–211.
20. Laskowski, R. A., Rullmannn, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) AQUA and PROCHECK-NMR: Programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR* **8,** 477–486.
21. Marti-Renom, M. A., Stuart, A. C., Fiser, A., Sanchez, R., Melo, F., and Sali, A. (2000) Comparative protein structure modeling of genes and genomes [in process citation]. *Annu. Rev. Biophys. Biomol. Struct*. **29,** 291–325.
22. Payton, M. (2000) The first 3D structure of arylamine N-acetyltransferase reveals a protease-like catalytic triad. *Trends Pharmacol. Sci*. **21,** 329–330.
23. Payton, M., Mushtaq, A., Yu, T. W., Wu, L. J., Sinclair, J., and 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,** 1137–1147.
24. Sandy, J., Kawamura, A., Noble, M., and Sim, E. (2001) The structure of *M. smegmatis* NAT at 1.7 angströms: Insight into an isoniazid metabolizing enzyme. *In* Second International NAT Workshop, Oxford, UK.
25. Via, A., Ferre, F., Brannetti, B., Valencia, A., and Helmer-Citterich, M. (2000) Three-dimensional view of the surface motif associated with the P-loop structure: Cis and trans cases of convergent evolution. *J. Mol. Biol*. **303,** 455–465.
26. Evans, B., Tishmack, P. A., Pokalsky, C., Zhang, M., and Van Etten, R. L. (1996) Site-directed mutagenesis, kinetic, and spectroscopic studies of the P-loop residues in a low molecular weight protein tyrosine phosphatase. *Biochemistry* **35,** 13609–13617.
27. Sinclair, J., and 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.
28. Fretland, A. J., Leff, M. A., Doll, M. A., and Hein, D. W. (2001) Functional characterization of human N-acetyltransferase 2 (NAT2) single nucleotide polymorphisms. *Pharmacogenetics* **11,** 207–215.