LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

UCRL-JRNL-222764

Insights into the O-Acetylation Reaction of Hydroxylated Heterocyclic Amines by Human Arylamine N-Acetyltransferases: A Computational Study

E. Y. Lau, J. S. Felton, F. C. Lightstone

July 11, 2006

Chemical Research in Toxicology

Insights into the O-Acetylation Reaction of Hydroxylated Heterocyclic Amines by Human Arylamine N-Acetyltransferases: A Computational Study.

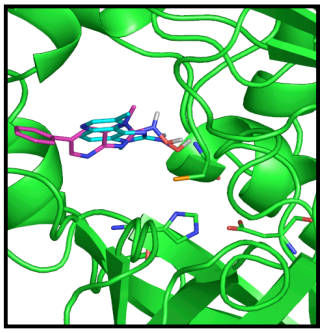
Edmond Y. Lau, James S. Felton, and Felice C. Lightstone*

Biosciences Directorate, Lawrence Livermore National Laboratory, Livermore, California, 94550

Received_____

Corresponding Author
Felice C. Lightstone
Phone: 925-423-8657
Fax: 925-424-5513
Email: felice@llnl.gov

TABLE OF CONTENTS GRAPHICS



ABSTRACT:

A computational study was performed to better understand the differences between human arylamine N-acetyltransferase (NAT) 1 and 2. Homology models were constructed from available crystal structures and comparisons of the active site residues 125, 127, and 129 for these two enzymes provide insight into observed substrate differences. The NAT2 model provided a basis for understanding how some of the common mutations may affect the structure of the protein. Molecular dynamics simulations of the human NAT models and the template structure (NAT from *Mycobacterium smegmatis*) were performed and showed the models to be stable and reasonable. Docking studies of hydroxylated heterocyclic amines in the models of NAT1 and NAT2 probed the differences exhibited by these two proteins with mutagenic agents. The hydroxylated heterocyclic amines were only able to fit into the NAT2 active site, and an alternative binding site by the P-loop was found using our models and will be discussed. Additionally, quantum mechanical calculations were performed to study the O-acetylation reaction of the hydroxylated heterocyclic amines N-OH MeIQx and N-OH PhIP. This study has given us insight into why there are substrate differences among isoenzymes and explains some of the polymorphic activity differences.

INTRODUCTION:

Arylamine N-acetyltransferase (NAT) was discovered to be the enzyme responsible for inactivation of the antitubercular drug isoniazid[1] by transfer of an acetyl group from acetyl coenzyme A (acetyl-CoA) to the amine group in isoniazid and plays an important role in the detoxification of xenobiotics.[2] These cytosolic enzymes are present in eukaryotes and prokaryotes.[3] Humans have two isoenzymes of arylamine N-acetyltransferases (NAT1 and NAT2) that are 81% homologous but differ in their tissue distribution and substrate specificity. Although NAT enzymes can detoxify molecules by N-acetylation, they can also activate other molecules through O-acetylation to form reactive acetoxy-esters that are unstable and decompose to electrophiles that bind to DNA leading to mutagenesis and carcinogenesis.[2] Polymorphisms in NAT enzymes result in differences in acetylation rate (slow and rapid) and can occur more commonly in certain ethnic populations.[4,5] The Gly191Ala mutant in NAT2 is more commonly seen in the African population, but the Gly268Glu is more commonly seen in Asians. Slow acetylation rates may make individuals more susceptible to certain diseases, and there are numerous studies that indicate a relationship between the NAT polymorphisms and certain kinds of cancers.[6,7] Slow NAT2 acetylation phenotypes have shown a relationship to urinary bladder cancer.[8] NAT polymorphisms linked to colon, breast, lung, and prostate cancers are more controversial.[9]

Crystal structures of four NATs show these enzymes are related to cysteine proteases and contain a catalytic triad of Cys-His-Asp that carries out the acetylation reaction.[10-13] The overall structures of these four NATs are very similar, and they can be easily overlaid. The catalytic triad is strictly conserved in NATs, and mutation of these

residues has shown that all three are necessary for catalysis; Asp is likely important for correct positioning of the residues rather than participates in the reaction.[14-17] The enzyme follows a ping-pong bi-bi mechanism that has the acetyl-CoA binding to NAT and transferring the acetyl group to the catalytic cysteine followed by acetylation of the substrate.[18] Although this reaction is a two-step process, substrates can bind to NAT in the absence of acetyl-CoA or an acetylated enzyme. NMR and crystallography have shown that isoniazid is able to bind to bacterial NATs in the absence of an acetylated NAT.[19,20] To better understand the differences in substrate specificity and the effect of polymorphisms in human NAT1 and NAT2, the molecular structures need to be characterized. Though the human X-ray structures for either NAT is not yet available, the recently solved crystal structure of NAT from *Mycobacterium smegmatis*[11] allowed us to build homology models of the full-length proteins of human NAT1 and NAT2. Previous homology models for human NAT1 and NAT2 were not of the complete enzymes, they consisted of only the first ~100 residues out of 290 residues for the NAT enzymes.[21,22] These models of human arylamine N-acetyltransferases provided insight into the differences in substrate specificity between the isoenzymes and the activity differences in some of the observed polymorphisms.

METHOD:

Multiple sequence alignment:

The amino acid sequence of human NAT1 (hNAT1) and NAT2 (hNAT2) is 32% and 29% identical to the NAT from *Mycobacterium smegmatis* (msNAT), respectively, a level of homology that suggests that their three-dimensional structures will be quite

similar as well. The sequence alignment between these two proteins to the template was straightforward, except for a few regions of lower homology (Figure 1). Since the correctness of the sequence-structure alignment is the major determinant of the model quality,[23] special attention was paid to the alignment in these low homology regions. To determine which regions needed further analysis PSI-BLAST-ISS was applied, an intermediate sequence search procedure, described in more detail elsewhere.[24] Briefly, a set of sequences having some level of homology to both human NAT (hNAT1 or hNAT2) and msNAT proteins were used as seeds to generate corresponding PSI-BLAST[25] profiles. Using the SEALS package[26] and in-house Perl scripts, alignments between these two sequences were extracted from the resulting individual PSI-BLAST profiles and compared. The convergence of the alignment between the human NAT's and msNAT sequences in a particular region was then used as an indicator of the alignment reliability in this region.

Model Building:

Three-dimensional models were generated automatically from the alignments described above with MODELLER.[27] There are large insertions of residues at positions 170-186 and 248-252 for both human NATs relative to msNAT. The initial models for hNAT1 and hNAT2 were created with the msNAT crystal structure (1GX3).[11] To accommodate the insertion at residues 170-186, residues 168-175 in the msNAT crystal structure were removed from the template structure. Residue side-chains for the resulting models were positioned using a backbone-dependent rotamer library implemented in SCWRL.[28] The molecular mechanics program AMBER[29] was used to further refine the

models. A force constant of 25 kcal/mole•Å^2 was placed on all atoms, and constrained energy minimization was performed using 250 steps of steepest descents and 750 steps of conjugate gradients to remove unfavorable contacts and improve bond lengths and angles. The quality of the models was evaluated using PROCHECK[30] and ProsaII.[31]

Molecular Dynamics Simulations:

Molecular dynamics (MD) was performed on each model of the human NAT and the crystal structure of msNAT (1GX3, used as a control). The simulations were performed with AMBER using the Cornell force field.[32] The catalytic cysteine and histidine in each NAT molecule was modeled as charged residues. NMR studies of the cysteine protease papain have shown the analogous residues are ionized in the ground state.[33,34] Each NAT molecule was solvated in a box of TIP3P[35] water sufficient in size to have at least 10 Å of water between the protein and the solvent interface (~81 × 81 × 81 Å^3 initial box size). The active site residues Cys68 and His107 were simulated as ionized residues.[15,33,34] To neutralize the systems 7 sodium ions were added to the msNAT simulation, 4 sodium ions to the hNAT1, and 6 sodium ions to hNAT2. The systems consisted of about 63,000 atoms (~19,000 water molecules). Each system was energy minimized using 500 steps of steepest descent and 1500 steps of conjugate gradient. Constant temperature and pressure dynamics (NPT) were performed on these minimized systems. Periodic boundary conditions were used, and electrostatic interactions were treated by particle mesh Ewald methods[36] with a 9 Å cutoff in direct space and using a 1 Å grid. Bonds containing a hydrogen were constrained using SHAKE,[37] and a time step of 2 fs was used in each simulation. The systems were

initially coupled to a heat bath at 100 K for the first 20 ps. The heat bath temperature was increased to 200 K for the next 10 ps and finally raised to 300 K for the remainder of the simulation. Coupling constants of 5 and 2 ps were used for enzyme and solvent, respectively. Each simulation was performed for a total of 1 ns, and the last 500 ps were used for analysis. The atomic fluctuations obtained from the simulations can be compared to the crystallographic results by converting the Debye-Waller (B-) factors to mean squared fluctuations using the equation $<\Delta r> = (3B/8\pi^2)^{1/2}$.

Docking Simulations:

The activated heterocyclic amines N-hydroxyl-2-amino-1-methyl-6-phenylimidazo[4,5-b]pyridine (N-OH PhIP), and N-hydroxyl-2-amino-3,8-dimethylimidazo[4,5-f]quinoxaline (N-OH MeIQx) were docked into the active sites of the NAT models from clustering the coordinate set from molecular dynamics using the program Autodock 3.05.[38] The initial geometries of the ligands were optimized at the HF/6-31G(d) level of theory using the program Gaussian98.[39] The partial atomic charges for the ligands were obtained by the RESP method.[40] Kollman-united atom charges were used for the NAT models.[41] A 60 × 60 × 60 grid was used for docking all the ligands with a spacing of 0.375 Å. The grids were centered above the catalytic histidine (His107). A Lamarckian algorithm was used to generate conformations for the ligands with the active site.[38] The parameters used for the Lamarckian genetic algorithm were the same as in Legge et. al. except the maximum number of energy evaluations was 150,000.[42] A total of 500 conformers were generated for a ligand in each NAT active site. The conformers were clustered using a 0.5 Å RMSD.

Ab initio calculations:

The acetylation reaction of N-OH PhIP and N-OH MeIQx with the thioester S-methyl thioacetate was studied by density functional theory calculations. All calculations were performed using the program Gaussian98 or Gaussian03.[39,43] The geometries were optimized using density functional theory (DFT) at the B3LYP/6-31G(d) level of theory.[44,45] Harmonic frequency calculations were performed to ensure the structures were minima and transition states. The zero point energies from the frequency calculations were scaled by 0.98. The energies of each structure from single point calculations at the B3LYP/6-311+G(d,p) level of theory using the optimized geometry (B3LYP/6-311+G(d,p)//B3LYP/6-31G(d)).

RESULTS AND DISCUSSION:

Analysis of NAT models:

Although hNAT1 and hNAT2 share 81% sequence similarity and utilize the same reaction mechanism for acetylation of substrates, there are significant differences in these two enzymes such as tissue distribution, substrate specificity, and expression levels during development. The substrate specificity of these two enzymes is different. Human NAT2 has the ability to acetylate larger molecules and appears responsible for activating hydroxylated heterocyclic aromatic molecules. Human NAT1 has lower activity than hNAT2 and is not as well studied as hNAT2. The overall fold of the arylamine N-acetyltransferases appears to be robust. So far, four bacterial NAT crystal structures have been solved for this family of enzymes, and the overall structures and the positioning of

the catalytic triads are very similar even though the homology is only ~25% between these NAT enzymes.

Both human NATs contain two significant insertions relative to the bacterial NATs (Figure 1). A long 17 residue insertion is present between α-helix 6 and β-sheet 5. A smaller insertion of 5 residues is at the end of a β-sheet. The larger insertion is distant from the active site, but the smaller insertion is adjacent to the catalytic triad, and both do not affect the positioning of the catalytic residues in either model (Figure 2). The large insertion was modeled as a loop. Structure prediction servers (PsiPred[46], Jpred[47], and SCRATCH[48]) assigned a small portion of the sequence to be β-sheet with the rest of this insertion sequence a coil. This insertion sequence was not found to be homologous with any sequence of a solved crystal structure. Recently, Josephy and coworkers have created models for the human NATs, and this region was also modeled as an extended loop.[49] The quality of our models was evaluated using ProsaII and Procheck. ProsaII scores for hNAT1 and hNAT2 were –7.99 and –7.58, respectively (a lower value indicates a more reasonable structure). The ProsaII score for the template structure (1GX3) was –10.67. Procheck G-scores for hNAT1 and hNAT2 were 0.03 and –0.07, respectively. Procheck G-scores above –0.5 imply a reasonable structure, and a higher G-score indicates a more reasonable structure. The hNAT1 had 88.5% of the residues in the most favorable regions of the Ramachandran map, and hNAT2 had a slightly lower value of 84.4%. The template structure had 88.3% of residues in the most favored regions. The sequence alignment used for these models and the one used by Josephy and coworkers are similar. Comparison of our sequence alignment with the sequence alignment by Josephy and coworkers shows that the small insertions or deletions (one to two residues) are modeled

within 3 residues of each other. The 17 residue insertion in both human NATs were modeled in the identical position. The only significant difference is a modeled insertion of 5 residues into the third domain of msNAT in our sequence alignment. This region of the enzyme was the most difficult to model and may explain the slight difference in Procheck G-scores between our models and the ones created by Josephy and coworkers (-0.21 for both human NAT models).

The active sites for the human NATs and the msNAT are similar with the catalytic triad of Cys-His-Asp positioned within an alpha and beta fold (a single α-helix with a 3 stranded β-sheet). The hNAT1 model shows that Phe125 is in close proximity to the catalytic triad, and having serine at this position in hNAT2 opens up the area (Figure 3). This change in residue has a smaller than expected effect because there is a phenylalanine at the opposite position of residue Phe125 in hNAT2 (Phe93, the analogous residue in hNAT1 is valine) that takes up some of the space vacated by a phenylalanine. Mutation experiments have shown that Phe125 in hNAT1 is an important residue in determining substrate specificity.[50]

Human NAT1 and NAT2 have different substrate specificity. Site-directed mutagenesis studies have shown that residues 125, 127, and 129 are important determinants in the specificity of human NATs.[50] In hNAT1, these residues are Phe125, Arg127, and Tyr129, and in hNAT2 all three of these residues are serines. The models of the human NATs show these residues are in close proximity to the catalytic triad, and it is readily apparent why p-aminosalicylic acid (PAS) is specific to hNAT1. The carboxylate group on PAS can form a salt bridge with the guanidinium group of Arg127, and site directed mutagenesis shows that changing this arginine to serine drastically affects the

binding constant. Mutation of Phe125 or Tyr129 has only a minor effect on the binding constant of PAS to hNAT1. Human NAT2 prefers larger molecules than hNAT1. Inspection of the molecular surfaces of these proteins provides insight into their specificity. Phe125, Arg127, and Tyr129 are positioned above the catalytic Cys68 and His107 in hNAT1 and act to restrict the length of the substrates that can bind to this enzyme (Figure 4). In hNAT2, having serine at residues 125, 127, and 129 provides more space in the active site allowing larger (longer) substrates to be accommodated such as sulphamethazine (SMZ). Individual mutations of these serines did not greatly change the binding constant of SMZ, but the double mutant of Ser127Arg and Ser129Tyr greatly increased the binding constant of SMZ. These two residues are directly above Cys68 and His107 in hNAT2, and mutating the serines to the larger phenylalanine and arginine likely interferes with SMZ fitting into the active site. Interestingly, the single mutation of Ser129 to tyrosine in hNAT2 lowered the binding constant for PAS more than the Ser127 to arginine mutation.

Polymorphisms in human NATs affect the rate of acetylation of substrates and are divided into slow and rapid phenotypes. Eleven single nucleotide polymorphisms in hNAT2 lead to six different amino acid mutations in the enzyme; Arg64Gln, Ile114Thr, Glu145Pro, Arg197Gln, Lys282Thr, and Gly286Thr. None of these mutations occurs in the active site of hNAT2 and are more likely to affect the stability of the enzyme rather than the acetylation reaction (Figure 5). As seen previously in a truncated homology model of hNAT2,[22] Arg64 forms an internal salt bridge with Glu38 providing structural stability for hNAT2. In addition to the salt bridge, the sidechain of Arg64 also forms a hydrogen bond with Asn41. Mutation of Arg64 to glutamine would remove the salt

bridge and destabilize this region of the protein. Ile114 resides in a cluster of hydrophobic residues such as Ile21, Val24, Phe84, Val112, and Tyr119, and mutation of this residue into a polar threonine could cause this residue to reorient and disrupt the hydrophobic cluster. Glu145 is located in a β-sheet; mutation of this residue to a proline would likely redirect the backbone and disrupt the secondary structure. The effect on the protein structure on mutating Arg197 to a glutamate is more difficult to determine. This residue is on the surface of the protein and distant from the active site. The effect of mutating Gly286 to a threonine is not obvious since this residue resides on the surface of the protein and close to the C-terminus. Truncation of the C-terminus of NAT from *Salmonella typhimurium* by 11 residues does impair binding of acetyl-CoA and substrate binding.[51] Mutation of the Gly286 may affect the acetylation reaction rather than affect the three-dimensional structure. A truncated version of hNAT1 (containing the first 204 amino-terminal residues) is able to form the acetylated intermediate but is unable to transfer the acetyl group to substrate.[52]

Molecular dynamics Simulations:

Molecular dynamics simulations were performed on both human NAT models and the msNAT (for comparison). All proteins were stable for the duration of the simulations. The msNAT simulation had a root mean squared deviation (RMSD) of ~1.4 Å for the backbone atoms relative to the starting structure. The models had higher RMSDs of 2.0 Å and 2.4 Å for hNAT1 and hNAT2, respectively. The major contribution to the RMSD for the NAT models were due to changes in conformation of the generated loop insertions at residues 170-186 and residues 248-252, but the changes in

conformation did not affect the overall structure of the proteins. Comparison of the root mean squared fluctuations (RMSF) of the backbone atoms of msNAT from molecular dynamics is in good agreement with RMSF estimated from the crystallographic B-factors (Figure 6). The RMSF for hNAT1 and hNAT2 are very similar to those observed for msNAT except at the residue insertions. These areas show large scale fluctuations which is not surprising since the coordinates were built-in rather than based on positions within the template. The C-terminus of hNAT2 also exhibits large scale fluctuations. This is due to the region containing two prolines causing the area to be unstructured. In hNAT1, the C-terminus was modeled with an α-helix in this region which is similar to msNAT. The catalytic triad in both human NAT models is stable through out the simulations. The sulfur of Cys68 forms a hydrogen bond with HD1 of His107, 2.01 ± 0.09 Å and 2.03 ± 0.09 Å for hNAT1 and hNAT2, respectively. The HE2 of His107 forms a hydrogen bond with one of the carboxylate oxygens of Asp122, 1.77 ± 0.10 Å and 1.93 ± 0.21 Å for hNAT1 and hNAT2, respectively. The results from the solvated molecular dynamics simulations show that the models for human NATs are stable in the nanosecond time scale and likely do not contain gross errors.

Heterocyclic Amine Docking:

The initial structures for hNAT1 and hNAT2 along with coordinates from the molecular dynamics simulations were used for docking studies on the orientational preferences of N-OH MeIQx and N-OH PhIP. Neither activated heterocyclic amine could be docked into the active site of hNAT1 for any of the coordinate sets. Both compounds were docked either on the exterior of the protein or in a pocket close to the P-

loop (residues 126, 127, 128, and 129) but distant from the active site. During dynamics, the P-loop in hNAT1 undergoes a conformational change that widens a pocket that allows binding of these heterocyclic amines. The residues Ile32, Gly126, Arg127, Ser128, Tyr129, Gln130, Ser224, Leu225, and Gln226 form this pocket. These results are consistent with mutagenesis data that shows hNAT1 is virtually unable to O-acetylate heterocyclic amines. Docking of N-OH PhIP and N-OH MeIQx to hNAT2 showed that both molecules could fit into the active site of the enzyme with similar conformations (Figure 7). The docked conformation of N-OH MeIQx is similar to the positioning of IQ in hNAT2 obtained by Josephy and coworkers.[49] These heterocyclic amine compounds interact with residues Phe37, Trp67, Phe93, Gly124, Ser125, Leu209, Ser216, and Phe217 in hNAT2. As can be seen from these residues, the binding site for hNAT2 is primarily hydrophobic. Interestingly, the crystal structure of msNAT with isoniazid has the aromatic group of the drug approximately 90° relative to both N-OH compounds. The active site of msNAT is smaller than hNAT2 because of Phe130 in msNAT occupies the analogous position of Ser125 in hNAT2, and isoniazid is able to form a π-stacking interaction with Phe130. The positioning of the aromatic groups of the N-OH compounds is similar to placement of bound *p*-bromoacetanilide inhibitor in stNAT. Although many conformers were docked into the active site, there was also another site on hNAT2 that was highly populated; and the energies were comparable to ones obtained in the active site. This alternate binding site is similar in position to one found in docking simulations by Sim and coworkers when they docked aromatic compounds to the active site of the NAT from *Salmonella typhimurium.*[51] Josephy and coworkers have shown that mutagenic activation of PhIP in *Salmonella* is not promoted by hNAT2, but MeIQx

genotoxcity is linked to hNAT2 activation.[53] These docking results are not consistent with this result because both molecules are able to bind in the active site of hNAT2. Although both molecules bind in the active site, their chemistries for O-acetylation may differ. This possibility was also investigated by calculating the reaction pathway for O-acetylation.

Ab initio Calculations of O-Acetylation Reaction Pathway:

The O-acetylation reaction of N-OH PhIP and N-OH MeIQx with S-methyl thioacetate was studied by *ab initio* calculations. Although these calculations are performed in the gas-phase, they should be relevant for these enzymes since the active site of NATs are hydrophobic and not easily accessible to solvent. To our knowledge the pK_as of N-OH PhIP and N-OH MeIQx have not been published; thus the O-acetylation reactions were calculated for N-OH MeIQx and N-OH PhIP and for both compounds deprotonated (N-O$^-$). In the case of the reaction with the N-OH group, a reactant complex is formed between S-methyl thioacetate and either N-OH PhIP or N-OH MeIQx that is structurally similar, and both complexes are approximately -4.0 kcal/mole more stable than their respective separated reactants (Figure 8). The hydrogen of the N-OH group forms a hydrogen bond with the sulfur of the thioester (2.511 Å and 2.487 Å for N-OH MeIQx and N-OH PhIP, respectively), and the oxygen of the N-OH group is ~4.0 Å from the carbonyl carbon of the thioester. The distance from the oxygen of the N-OH to the carbonyl carbon is much shorter when the transition state is formed (1.962 Å and 1.980 Å for N-OH MeIQx and N-OH PhIP, respectively). The transition state is tight with significant O-C bond formation. The hydrogen of the N-OH group is shared

between the oxygen (1.397 Å) and the sulfur of the thioester (1.528 Å) for N-OH MeIQx. Similar distances were obtained for the transition state structure of N-OH PhIP. The transition state undergoes a concerted transfer of the hydrogen from the N-OH group to the sulfur of S-methyl thioacetate and bond formation between the oxygen of the N-OH group with the carbonyl. The activation barrier for this reaction is 37.84 and 37.25 kcal/mole for N-OH MeIQx and N-OH PhIP, respectively. The transition state decays to the products, methylthiol and N-acetyl compounds. The O-acetylation reaction of N-OH MeIQx or N-OH PhIP with S-methyl thioacetate has very similar geometries and relative energies. Given the high activation barrier for this reaction, it is unlikely that NAT would utilize this reaction pathway to O-acetylate N-OH compounds.

Since the O-acetylation reaction with an N-OH group is unlikely to occur because of the high activation barrier, the O-acetylation reaction with a deprotonated N-OH group (denoted $N\text{-}O^-$) was investigated. An ion-molecule complex is first formed that is approximately –9.0 kcal/mole more stable than the separated reactants (Figure 9). The thioester and $N\text{-}O^-$ MeIQx are almost parallel to each other in the ion-molecule complex. The oxygen of the $N\text{-}O^-$ is 3.493 Å from the carbonyl carbon of the thioester. Additionally, a hydrogen bond is formed between the carbonyl oxygen of the thioester and the hydrogen on the nitrogen of the $N\text{-}O^-$ (2.524 Å). This structure proceeds to a transition state that is looser than obtained for the protonated reaction. The oxygen to carbonyl carbon distance is 2.905 Å, and the hydrogen to carbonyl oxygen distance shortens to 2.140 Å. The relative geometry of these two molecules is different from that observed for the ion-molecule complex. Whereas the ion-molecule complex had the molecules parallel, the transition state has the $N\text{-}O^-$ MeIQx almost at a 90° angle relative

to the thioester. The activation barrier for this step is only 1.14 kcal/mole. After this transition state, a tetrahedral intermediate is formed. The carbonyl oxygen-carbon bond is formed (1.576 Å), and the carbonyl carbon-sulfur bond elongates to 1.993 Å. The tetrahedral intermediate is –1.76 kcal/mole more stable than the first transition state. A second transition state is formed after the tetrahedral intermediate. There is a slight shortening of the oxygen-carbonyl carbon bond to 1.539 Å, and the carbonyl carbon-sulfur bond lengthens to 2.090 Å. The activation barrier to the second transition state is slight at 0.43 kcal/mole. Once past the second transition state, the energy steeply falls to form the products N-acetyl MeIQx and CH_3S^-. The product complex is –17.26 kcal/mole more stabile than the second transition state (Figure 10). This reaction pathway is analogous to amide hydrolysis catalyzed by cysteine proteases.[54]

The O-acetylation reaction with N-O^- PhIP and S-methyl thioacetate has similar optimized structures as N-O^- MeIQx along the reaction pathway initially, but when the zero-point energies are taken into account, the tetrahedral intermediate is higher in energy than the first transition state leading to a single energy barrier separating the reactants from the products. The transition state structure for this reaction is similar to the second transition state structure in the reaction with N-O^- MeIQx. The activation barrier for the N-O^- PhIP O-acetylation reaction is a modest 4.43 kcal/mole. Although this activation barrier is greater than the one calculated for N-O^- MeIQx, it is not high enough to significantly impede the O-acetylation reaction. These calculations imply that if an N-O^- compound is able to bind in the active site of hNAT2, then the reaction will go rapidly to completion, but if an N-OH compound is in the active site, then the O-acetylation reaction will not occur in a reasonable amount of time. A possible explanation of the

results is N-OH compounds bind to NAT but do not react until deprotonation of the hydroxyl group occurs by a general base within the active site. The histidine (His107) of the catalytic triad could be the general base for the reaction. Experimental measurements on the hamster NAT2 puts the apparent pK_a of the catalytic histidine to be around 5.5 for the acetylated enzyme but greater than 9.0 when involved in a thiolate-imidazolium pair.[15,16] An estimated pK_a for deprotonation of the hydroxyl group for N-OH PhIP (9.81) and N-OH MeIQx (9.77) was obtained using the SPARC server[55] and imply the reactivity of N-OH PhIP and N-OH MeIQx are similar. The proposed catalytic mechanism for deacetylation of hamster NAT2, similar to human NAT1, is dominated by the nucleophilicity of the substrate and assisted by a general base which is consistent with these *ab initio* calculations. It has been shown *in vitro* that N-OH PhIP and N-OH MeIQx formed DNA adducts, identified as dG-C8-PhIP and dG-C8-MeIQx, respectively, with intact prostate epithelial cells prepared from benign prostatic hypertrophy tissues known to contain hNAT2.[56] In the same study, repair synthesis of cells was observed when treated with N-OH PhIP and N-OH MeIQx but not with the parent compounds or their nitro derivatives. Our results do not explain the observed difference in mutagenic activity of PhIP relative to MeIQx in *Salmonella typhimurium* containing both human NAT enzymes. This may point to differences in reactivity or interactions of these compounds further upstream in DNA adduct formation in *Salmonella*.

CONCLUSIONS:

Three-dimensional structures for human NAT1 and NAT2 were constructed from coordinates of the NAT from *Mycobacterium smegmatis*. The modeled structures are

very similar to the template structure, although both human NATs contain a long coil insertion. The hNAT2 model is useful in rationalizing the effect of some polymorphisms on the protein structure. Molecular dynamics simulations of the models show that they are stable in solution on the nanosecond time scale. Docking of N-OH heterocyclic amines was possible in the active site of hNAT2, but the active site of hNAT1 is too small to accommodate these compounds. A second site in hNAT2 close to the P-loop and to the active site was also able to accommodate these N-OH heterocyclic amines and maybe the adenine of acetyl-CoA binding site. *Ab initio* calculations of model compounds show that the N-OH group is insufficient to cause O-acetylation with S-methyl thioacetate. A very high activation barrier separates the reactions and products. If the N-OH group is deprotonated, possibly general base catalyzed, the O-acetylation reaction is facile. These results are consistent with *in vitro* experiments of N-OH PhIP and N-OH MeIQx with prostate epithelial cells that show DNA adducts are formed with both compounds.

ACKNOWLEDGEMENTS:

This work was funded by grant CA055861 from NCI. This work was performed under the auspices of the United States Department of Energy by the University of California, Lawrence Livermore National Laboratory under contract number W-7405-ENG-48. The authors thank Wallace Bradford at SRI International for help with the SPARC calculations.

REFERENCES:

(1) Peters, J. H., Miller, K. S., Brown, P. Studies on the Metabolic Basis for the Genetically Determined Capacities for Isoniazid Inactivation in Man. *J. Pharmacol. Exp. Ther.* **1965***, 150*, 298-304.

(2) Hanna, P. E. Metabolic Activation and Detoxification of Arylamines. *Curr. Med. Chem.* **1996***, 3*, 195-210.

(3) Pompeo, F., Brooke, E., Kawamura, A., Mushtaq, A., Sim, E. The Pharmacogenetics of NAT: Structural Aspects. *Pharmacogenetics* **2002***, 3*, 19-30.

(4) Fretland, A. J., Doll, M. A., Leff, M. A., Hein, D. W. Functional Characterization of Nucleotide Polymorphisms in the Coding Region of N-acetyltransferase 1. *Pharmacogenetics* **2001***, 11*, 511-520.

(5) Fretland, A. J., Leff, M. A., Doll, M. A., Hein, D. W. Functional Characterization of Human N-acetyltransferase 2 (NAT2) Single Nucleotide Polymorphisms. *Pharmacogenetics* **2001***, 11*, 207-215.

(6) Vineis, P., Bartsch, H., Caporaso, N., Harrington, A. M., Kadlubar, F. F., Landl, M. T., Malaveille, C., Shields, P. G., Skipper, P., Talaska, G., Tannenbaum, S. R. Genetically Based N-acetyltransferase Metabolic Polymorphism and Low-Level Environmental Exposure to Carcinogens. *Nature* **1994***, 369*, 154-156.

(7) Hein, D. W. N-Acetyltransferase Genetics and Their Role in Predisposition to Aromatic and Heterocyclic Amine-Induced Carcinogenesis. *Toxicology Lett.* **2000***, 112-113*, 349-356.

(8) Garcia-Closas, M., Malats, N., Silverman, D., Dosemeci, M., Kogevinas, M., Hein, D. W., Tardon, A., Serra, C., Carrato, A., Garcia-Closas, R., Lloreta, J., Castano-Vinyals, G., Yeager, M., Welch, R., Chanock, S., Chatterjee, N., Wacholder, S., Samanic, C., Tora, M., Fernandez, F., Real, F. X., Rothman, N. NAT2 Slow Acetylation, GSTM1 Null Genotype, and Risk of Bladder Cancer: Results from the Spanish Bladder Cancer Study and Meta-Analyses. *Lancet* **2005***, 366*, 649-659.

(9) Hein, D. W., Doll, M. A., Fretland, A. J., Leff, M. A., Webb, S. J., Xiao, G. H., Devanayina, U.-S., Nangju, N. A., Feng, Y. Molecular Genetics and Epidemiology of the NAT1 and NAT2 Actylation Polymorphisms. *Cancer Epidemiol. Biomarkers Prev.* **2000***, 9*, 29-42.

(10) Sinclair, J. C., Sandy, J., Delgoda, R., Sim, E., Noble, M. E. M. Structure of Arylamine N-acetyltransferase Reveals a Cataytic Triad. *Nature Struct. Biol.* **2000***, 7*, 560-564.

(11) Sandy, J., Mushtaq, A., Kawamura, A., Sinclair, J., Sim, E., Noble, M. The Structure of Arylamine N-acetyltransferase from Mycobaterium smegmatis-An Enzyme which Inactivates the Anti-tubercular Drug, Isoniazid. *J. Mol. Biol.* **2002***, 318*, 1071-1083.

(12) Westwood, I. M., Holton, S. J., Rodrigues-Lima, F., Dupret, J.-M., Bhakta, S., Noble, M. E. M., Sim, E. Expression, Purification, Characterization and Structure of Pseudomonas aeruginosa Arylamine N-acetytransferase. *Biochem. J.* **2005***, 385*, 605-612.

(13) Holton, S. J., Dairou, J., Sandy, J., Rodrigues-Lima, F., Dupret, J.-M., Noble, M. E. M., Sim, E. Structure of Mesorhizobium loti arylamine N-acetyltransferase 1. *Acta Crystallogr., Sect. F.* **2005***, 61*, 14-16.

(14) Dupret, J.-M., Grant, D. M. Site-Directed Mutagenesis of Recombinant Human Arylamine N-Acetyltransferase Expressed in E. coli. *J. Biol. Chem.* **1992***, 267*, 7381-7385.

(15) Wang, H., Vath, G. M., Gleason, K. J., Hanna, P. E., Wagner, C. R. Probing the Mechanism of Hamster Arylamine N-acetyltrasferase 2 Acetylation by Active Site Modification, Site-Directed Mutagenesis, and Pre-Steady State and Steady State Kinetic Studies. *Biochemistry* **2004***, 43*, 8234-8246.

(16) Wang, H., Lui, L., Hanna, P. E., Wagner, C. R. Catalytic Mechanism of Hamster Arylamine N-acetyltransferase. *Biochemistry* **2005***, 44*, 11295-11306.

(17) Watanabe, M., Sofuni, T., Nohmi, T. Involvement of Cys69 Residue in the Catalytic Mechanism of N-Hydroxyarylamine O-Acetyltransferase of Salmonella typhimurium. *J. Biol. Chem.* **1992***, 267*, 8429-8436.

(18) Riddle, B., Jencks, W. P. Acetyl-Coenzyme A :Arylamine N-acetyltransferase. *J. Biol. Chem.* **1971***, 246*, 3250-3258.

(19) Delgoda, R., Lian, L.-Y., Sandy, J., Sim, E. NMR Investigation of the Catalytic Mechanism of Arylamine N-acetyltransferase from Salmonella typhimurium. *Biochim. Biophys. Acta* **2002***, 1620*, 8-14.

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

(21) Rodrigues-Lima, F., Delomenie, C., Goodfellow, G. H., Grant, D. M., Dupret, J.-M. 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.* **2001***, 356*, 327-334.

(22) Rodrigues-Lima, F., Dupret, J.-M. 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. *Biochem. Biophys. Res. Comm.* **2002***, 291*, 116-123.

(23) Tramontano, A., Morea, V., and Leplae, R. Analysis and Assessment of Comparative Modeling Predictions in CASP4. *Proteins* ***Suppl.*** **2001***, 5*, 22-38.

(24) Venclovas, C. Comparative Modeling of CASP4 Target Proteins: Combining Results of Sequence Search with Three-Dimensional Assessment. *Proteins* ***Suppl.*** **2001***, 5*, 47-54.

(25) Altschul, S. F.; Madden, T. L.; Schaffer, A. A.; Zhang, J.; Zhang, Z. et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **1997***, 25*, 3389-3402.

(26) Walker, D. R.; Koonin, E. V. SEALS: a system for easy analysis of lots of sequences. *Proc Int Conf Intell Syst Mol Biol* **1997***, 5*, 333-339.

(27) Sali, A., and Blundell, T. L. Comparative Protein Modelling by Satisfaction of Spatial Restraints. *J. Mol. Biol.* **1993***, 234*, 779-815.

(28) Bower, M. J., Cohen, F. E., and Dunbrack Jr., R. L. Prediction of Protein Side-Chain Rotamers from a Backbone-Dependent Rotamer Library: A New Homology Modeling Tool. *J. Mol. Biol.* **1997***, 267*, 1268-1282.

(29) Pearlman, D. A., Case, D. A., Caldwell, J. W., Ross, W. S., Cheatham, T. E., III, Debolt, S., Ferguson, D., Seibel, G., and Kollman, P. A. AMBER, A Package of Computer Programs for Applying Molecular Mechanics, Normal-Mode Analysis, Molecular Dynamics, and Free Energy Calculations to Simulate the Structural and Energetic Properties of Molecules. *Comput. Phys. Commun.* **1995***, 91*, 1-41.

(30) Laskowski, R. A., MacAuthur, M. W., Moss, D. S., and Thornton, J. M. PROCHECK: A Program to Check the Stereochemical Quality of Protein Structures. *J. Appl. Crystallogr.* **1993***, 26*, 283-291.

(31) Sippl, M. J. Recognition of Errors in Three-Dimensional Structures of Proteins. *Proteins* **1993***, 17*, 355-362.

(32) Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W., and Kollman, P. A. A 2nd Generation Force Field for the Simulation of Protein, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.* **1995***, 117*, 5179-5197.

(33) Lewis, S. D., Johnson, F. A., Shafer, J. A. Effect of Cysteine-25 on the Ionization of Histidine-159 in Papain As Determined by Proton Nuclear Magnetic Resonance Spectroscopy. Evidence for a His-159-Cys-25 Ion Pair and Its Possible Role in Catalysis. *Biochemistry* **1981***, 20*, 48-51.

(34) Johnson, F. A., Lewis, S. D., Shafer, J. A. Determination oof a Low pK for Histidine-159 in the S-Methylthio Derivative of Papain by Proton Nuclear Magnetic Resonance Spectroscopy. *Biochemistry* **1981***, 20*, 44-48.

(35) Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983***, 79*, 926-935.

(36) Essmann, U., Perera, L., Berkowitz, M. L., Darden, T., Lee, H., and Pedersen, L. G. A Smooth Particle Ewald Mesh Method. *J. Chem. Phys.* **1995***, 103*, 8577-8593.

(37) Ryckaert, J. P., Ciccotti, G., and Berendsen, H. J. C. Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-alkanes. *J. Comput. Phys.* **1977***, 23*, 327-341.

(38) Morris, G. M., Goodsell, D. S., Halliday, R. S., Huey, R., Hart, W. E., Belew, R. K., Olson, A. J. Automated Docking Using a Lamarckian Genetic Algorithm and Empirical Binding Free Energy Function. *J. Computat. Chem.* **1998***, 19*, 1639-1662.

(39) M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrzewski, J. A. Montgomery, Jr., R. E. Stratmann, J. C. Burant, S. Dapprich, J. M. Millam, A. D. Daniels, K. N. Kudin, M. C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G. A. Petersson, P. Y. Ayala, Q. Cui, K. Morokuma, N. Rega, P. Salvador, J. J. Dannenberg, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. Cioslowski, J. V. Ortiz, A. G. Baboul, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, J. L. Andres, C. Gonzalez, M. Head-Gordon, E. S. Replogle, J. A. Pople *Gaussian 98*; Gaussian 98, Revision A. 11.4 ed.; Gaussian, Inc.: Pittsburgh, PA.

(40) Bayly, C. I., Cieplak, P., Cornell, W., Kollman, P. A. A Well-Behaved Electrostatic Potential Based Method using Charge Restraints for Deriving Atomic Charges: The RESP Model. *J. Phys. Chem.* **1993***, 97*, 10269-10280.

(41) Weiner, S. J., Kollman, P. A., Case, D. A., Singh, U. C., Ghio, C., Alagona, G., Profeta, S., Weiner, P. A New Force Field for Molecular Mechanics Simulation of Nucleic Acids and Proteins. *J. Am. Chem. Soc.* **1984***, 106*, 765-784.

(42) Legge, G. B., Morris, G. M., Sanner, M. F., Takada, Y., Olson, A. J., Grynszpan, F. Model of the aLb2 Intergrin I-Domain/ICAM-1 DI Interface Suggests That Subtle Changes in Loop Orientation Determine Ligand Specificity. *Proteins: Struc. Func. Genetics* **2002***, 48*, 151-160.

(43) M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, J. A. Pople *Gaussian 03*; Gaussian 03, Revision C. 02 ed.; Gaussian, Inc.: Pittsburgh, PA.

(44) Becke, A. D. Density-Functional Thermochemistry .3. the Role of Exact Exchange. *J. Chem. Phys.* **1993***, 98*, 5648-5652.

(45) Lee, C., Yang, W., Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B.* **1988***, 37*, 785-789.

(46) McGuffin, L. J., Bryston, K., Jones, D. T. The PSIPRED Protein Structure Prediction Server. *Bioinformatics* **2000***, 16*, 404-440.

(47) Cuff, J. A., Barton, G. J. Evaluation and Improvement of Multiple Sequence Methods for Protein Secondary Structure Prediction. *Proteins: Struc. Func. Genetics* **1999***, 34*, 508-519.

(48) Cheng, J., Randall, A., Sweredoski, M., Baldi, P. SCRATCH: A Protein Structure and Structural Prediction Server. *Nucleic Acids Res.* **2005***, 33*, w72-w76.

(49) Savulescu, M. R., Mushtaq, A., Josephy, P. D. Screening and Characterizing Human NAT2 Variants. *Methods Enzymol.* **2005***, 400*, 192-215.

(50) Goodfellow, G. H., Dupret, J.-M., Grant, D. M. Identification of Amino Acids Imparting Acceptor Substrate Selectivity to Human Arylamine Acetyltransferases NAT1 and NAT2. *Biochem. J.* **2000***, 348*, 159-166.

(51) Mushtaq, A., Payton, M., Sim, E. The COOH Terminus of Arylamine N-acetyltransferase from Salmonella typhimurium Controls Enzymatic Activity. *J. Biol. Chem.* **2002***, 277*, 12175-12181.

(52) Sinclair, J. C., Sim, E. 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.* **1996***, 53*, 11-16.

(53) Wild, D., Feser, W., Michel, S., Lord, H. L., Josephy, P. D. Metabolic Activation of Heterocyclic Aromatic Amines Catalyzed by Human Arylamine N-acetyltransferase Isozymes (NAT1 and NAT2) Expressed in Salmonella typhimurium. *Carcinogenesis* **1995***, 16*, 643-648.

(54) Storer, A. C., Menard, R. Catalytic Mechanism in Papain Family of Cysteine Peptidases. *Methods Enzymol.* **1994***, 244*, 486-500.

(55) Hilal, S. H., Karickhoff, W., Carreira, L. A. A Rigorous Test for SPARC's Chemical Reactivity Models: Estimation of More Than 4300 Ionization pKa's. *Quant. Struct. Act. Rel.* **1995***, 14*, 348-355.

(56) Wang, C. Y., Debiec-Rychter, M., Schut, H. A. J., Morse, P., Jones, R. F., Archer, C., King, C. M., Haas, G. P. N-Acetyltransferase Expression and DNA Binding N-hydroxyheterocyclic amines in Human Prostate Epithelium. *Carcinogenesis* **1999***, 20*, 1591-1595.

Figure Legends:

Figure 1: Sequence alignment of human NAT1 (hNAT1) and NAT2 (hNAT2) with NAT from *Mycobacterium smegmatis* (msNAT).

Figure 2: Overlay of the modeled structures of hNAT1 and hNAT2. The catalytic triads of Cys-His-Asp of the two enzymes are in identical positions. Panel B is a 90º rotated view of Panel A.

Figure 3: Active site of hNAT1 showing the relative positioning of the catalytic residues relative to residues Phe125 (cyan), Arg127 (red), and Tyr129 (magenta) which are important for substrate specificity.

Figure 4: Molecular surface representation of hNAT1 (red) and hNAT2 (green). The catalytic cysteine is colored yellow. Only in hNAT2 is cysteine visible on the molecular surface of the enzyme.

Figure 5: Locations of polymorphisms in hNAT2: (A) Gly286Thr, (B) Gln145Pro, (C) Arg64Gln, and (D) Ile114Thr. Structures inside the boxes show a more detailed representation of the local environment around the mutated residue, colored green in panels (B) and (D).

Figure 6: Root mean squared fluctuations (RMSF) for msNAT (Panel A) from crystallography (red) and simulation (green). RMSF of hNAT1 (blue) and hNAT2 (violet) from simulations are shown in Panel B.

Figure 7: Docked conformations of N-OH MeIQx and N-OH PhIP in the active site of hNAT2 (Panel B). Panel C shows the two binding sites for N-HO PhIP found from the docking. The positioning of the second molecule is in close contact with the P-loop that maybe important for binding the adenine of acetyl-CoA.

Figure 8: Structures obtained from DFT calculations for the O-acetylation reaction between N-OH MeIQx and S-methyl thioacetate. Distances are in angstroms.

Figure 9: Structures obtained from DFT calculations for the O-acetylation reaction between $N\text{-}O^-$ MeIQx and S-methyl thioacetate. Distances are in angstroms.

Figure 10: Reaction pathway for the O-acetylation reaction between $N\text{-}O^-$ MeIQx and S-methyl thioacetate. Energies are in kcal/mole.

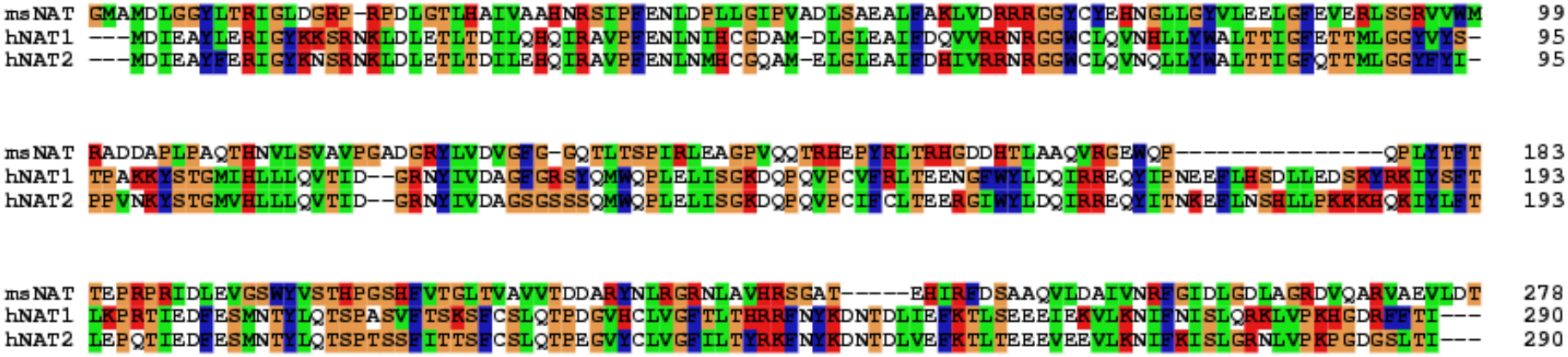

msNAT GMAMDLGGYLTRIGLDGRP-RPDLGTLHAIVAAHNRSIPFENLDPLLGIPVADLSAEALFAKLVDRRRGGYCMEHNGLLGYVLEELGFEVERLSGRVVWM 99
hNAT1 ---MDIEAYLERIGYKKSRNKLDLETLTDILQHQIRAVPFENLNIHCGDAM-DLGLEAIFDQVVRRNRGGWCLQVNHLLYWALTTIGFETTMLGGYVYS- 95
hNAT2 ---MDIEAYFERIGYKNSRNKLDLETLTDILEHQIRAVPFENLNMHCGQAM-ELGLEAIFDHIVRRNRGGWCLQVNQLLYWALTTIGFQTTMLGGYFYI- 95
msNAT RADDAPLPAQTHNVLSVAVPGADGRYLVDVGFG-GQTLTSPIRLEAGPVQQTRHEPYRLTRHGDDHTLAAQVRGEWQP---------------QPLYTFT 183
hNAT1 TPAKKYSTGMIHLLLQVTID--GRNYIVDAGFGRSYQMWQPLELISGKDQPQVPCVFRLTEENGFWYLDQIRREQYIPNEEFLHSDLLEDSKYRKIYSFT 193
hNAT2 PPVNKYSTGMVHLLLQVTID--GRNYIVDAGSGSSSQMWQPLELISGKDQPQVPCIFCLTEERGIWYLDQIRREQYITNKEFLNSHLLPKKKHQKIYLFT 193
msNAT TEPRPRIDLEVGSWYVSTHPGSHFVTGLTVAVVTDDARYNLRGRNLAVHRSGAT-----EHIREDSAAQVLDAIVNRFGIDLGDLAGRDVQARVAEVLDT 278
hNAT1 LKPRTIEDFESMNTYLQTSPASVFTSKSFCSLQTPDGVHCLVGFTLTHRRFNYKDNTDLIEFKTLSEEEIEKVLKNIFNISLQRKLVPKHGDRFFTI--- 290
hNAT2 LEPQTIEDFESMNTYLQTSPTSSFITTSFCSLQTPEGVYCLVGFILTYRKFNYKDNTDLVEFKTLTEEEVEEVLKNIFKISLGRNLVPKPGDGSLTI--- 290

Figure 1

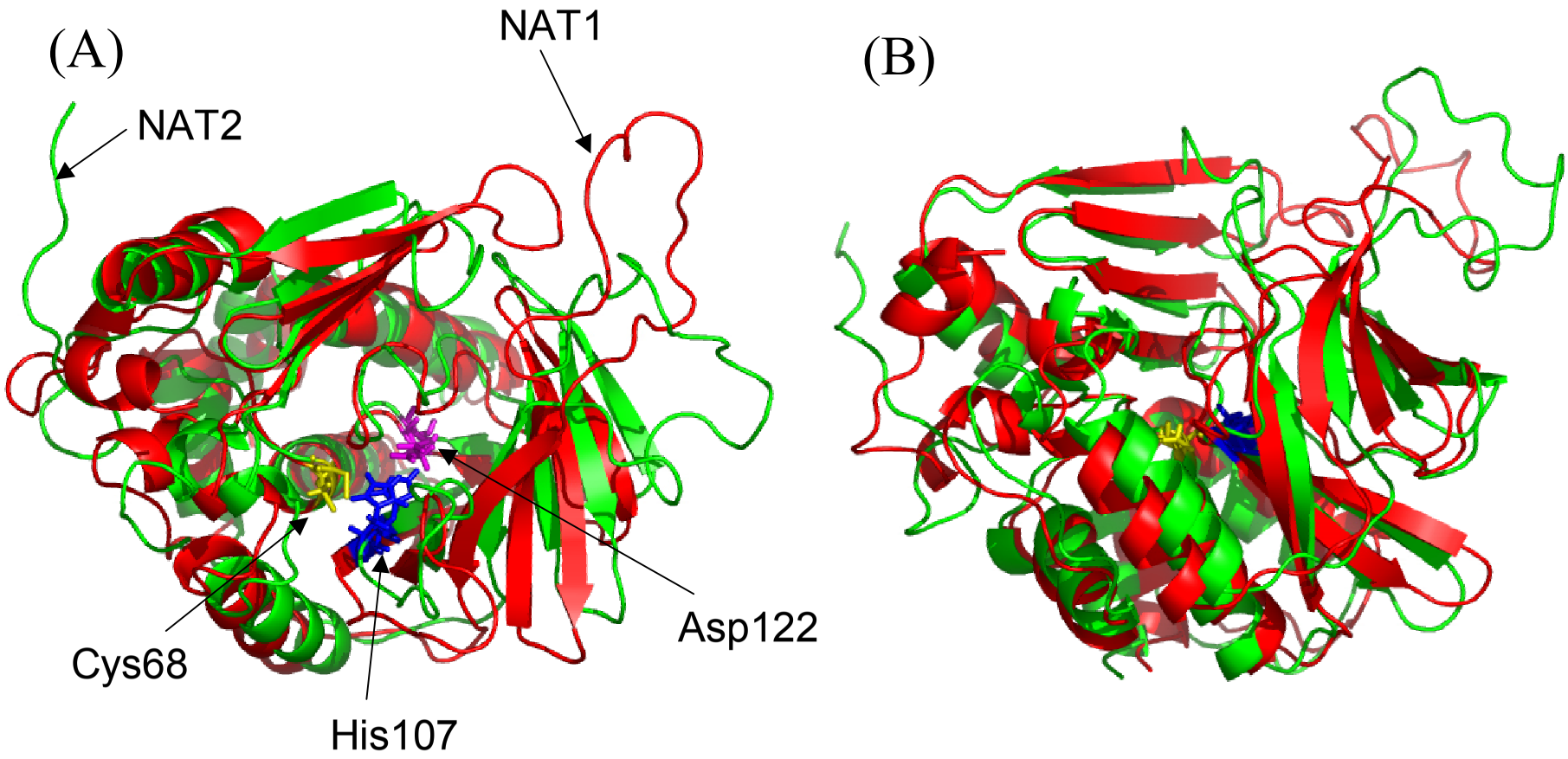
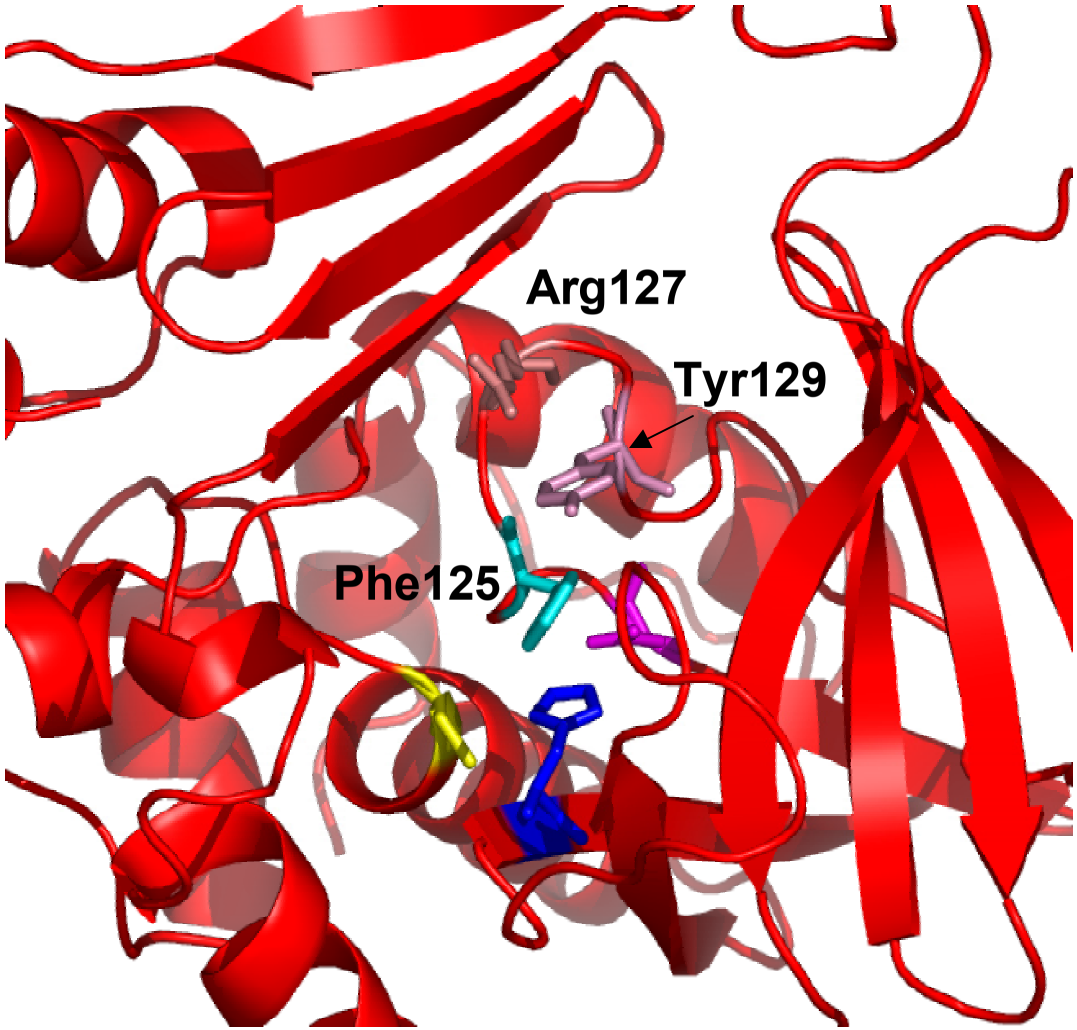
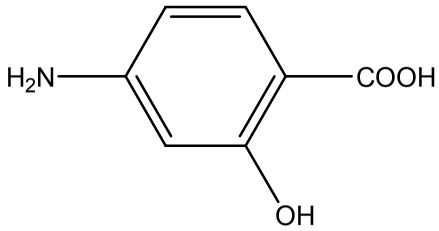


Figure 2



NAT1 substrate
P-aminosalicylic acid
PAS



NAT2 substrate
Sulphamethazine
SMZ

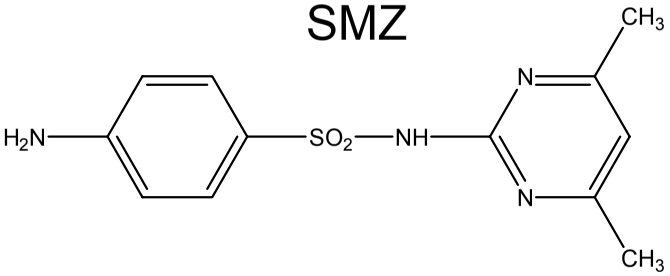


Figure 3

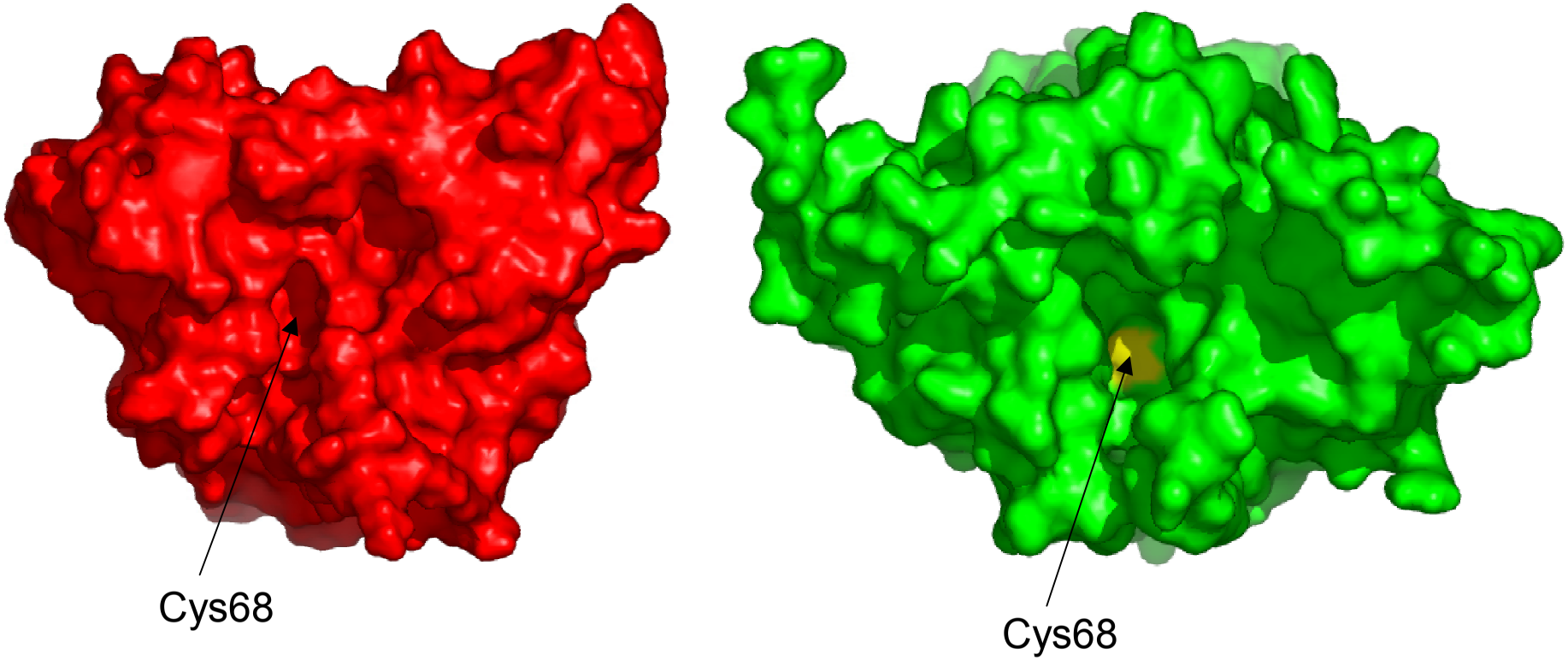


Figure 4

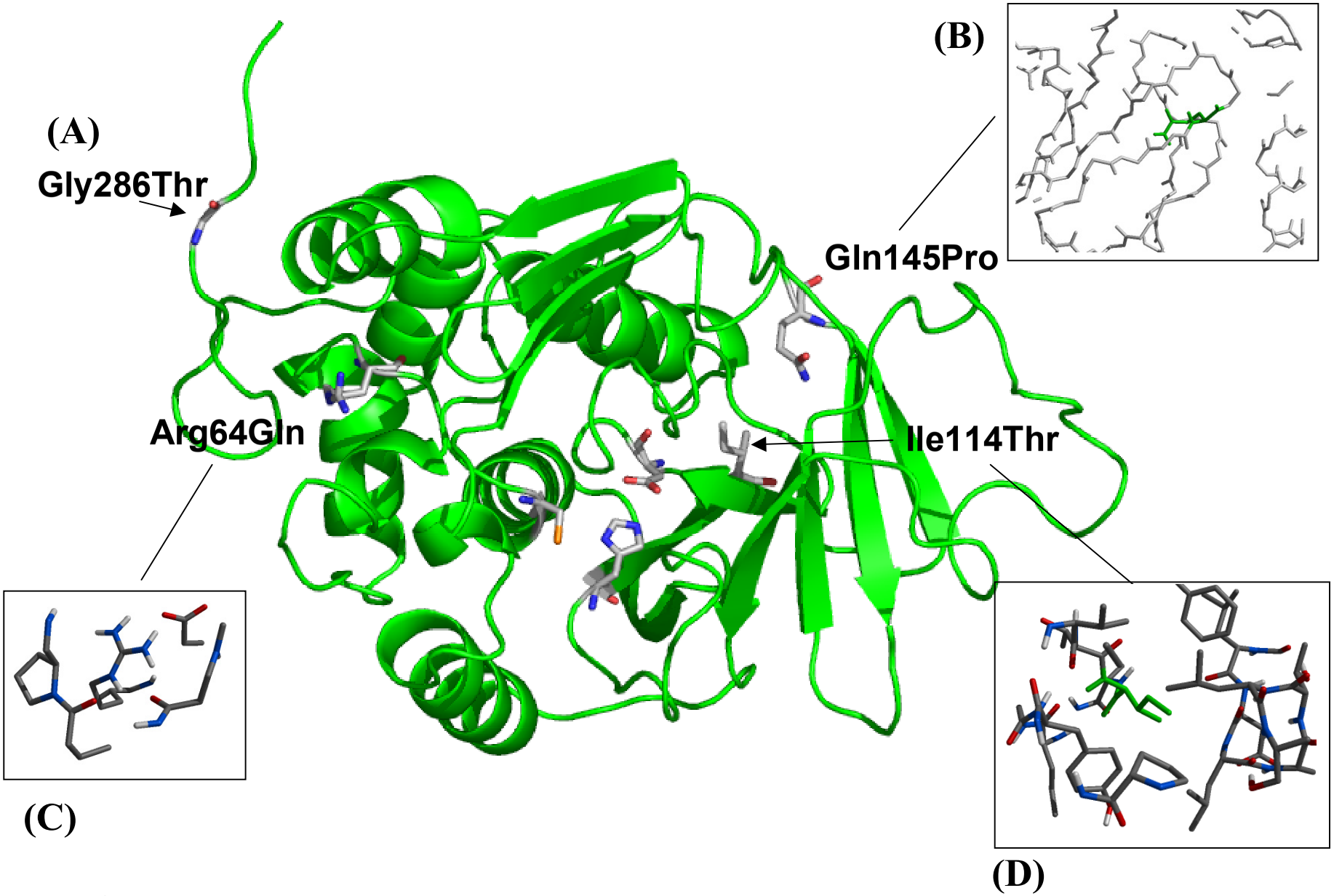


Figure 5

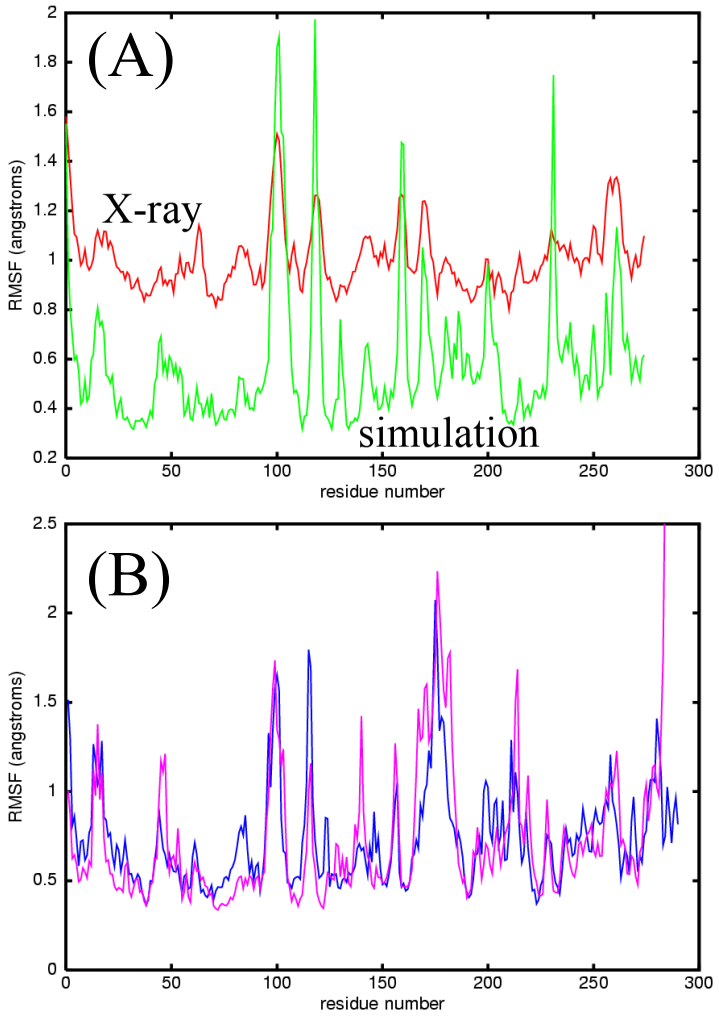
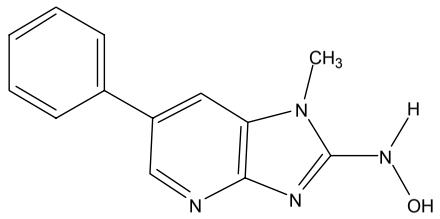
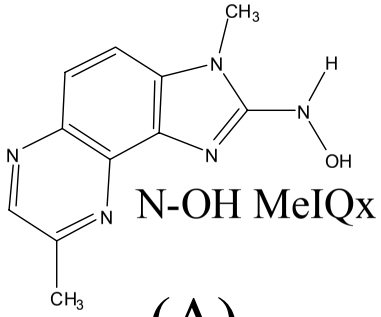

(A)
X-ray
simulation
RMSF (angstroms)
residue number
(B)
RMSF (angstroms)
residue number

Figure 6

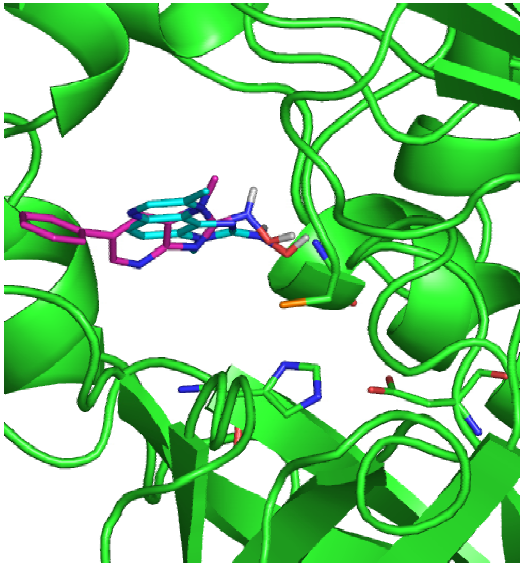


N-OH PhIP

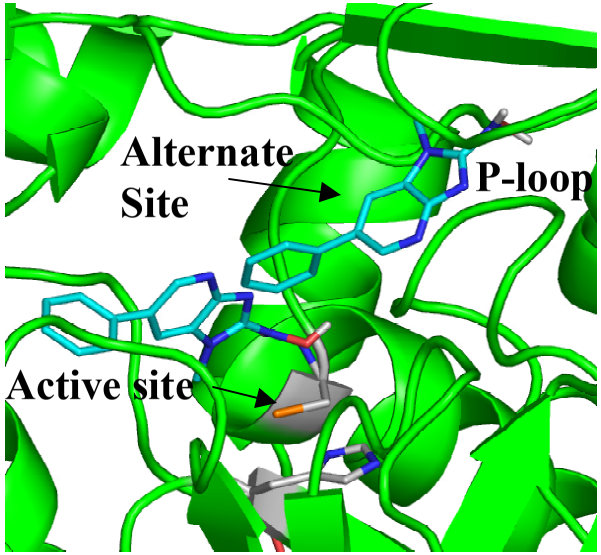


N-OH MeIQx

(A)



(B)



(C)

Figure 7

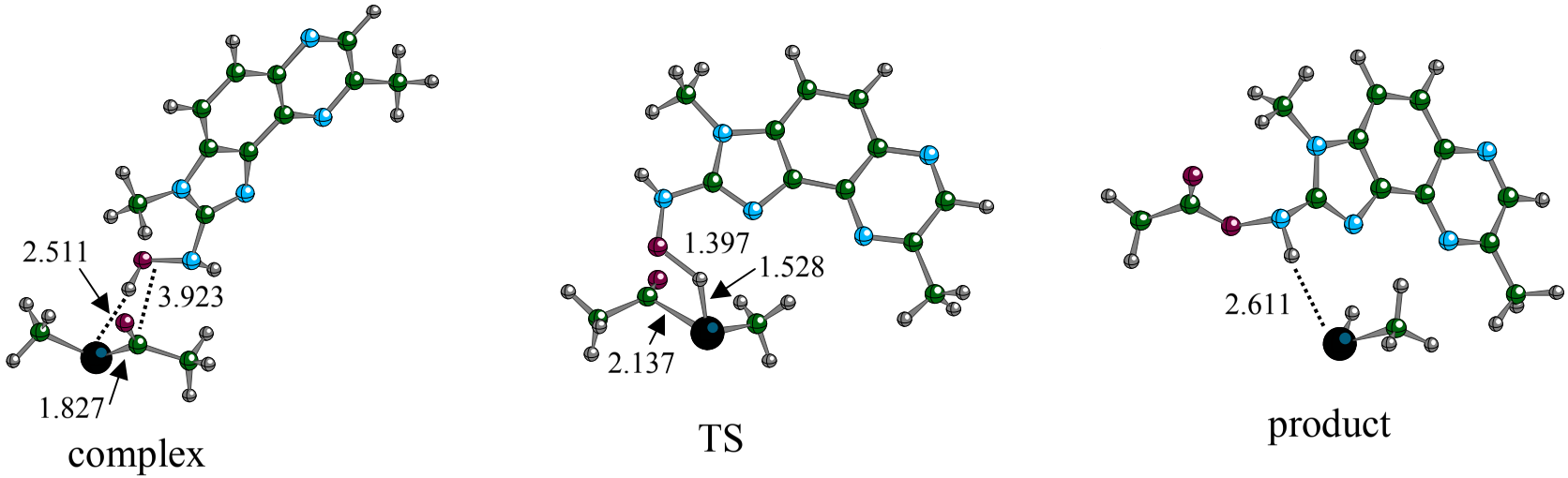
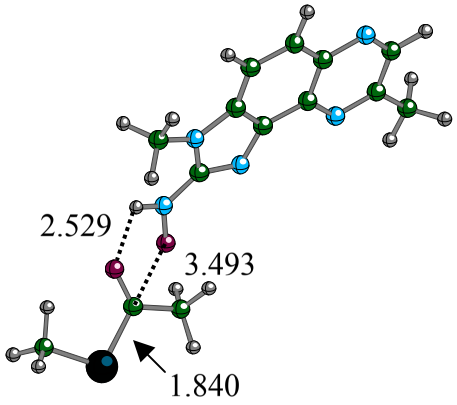
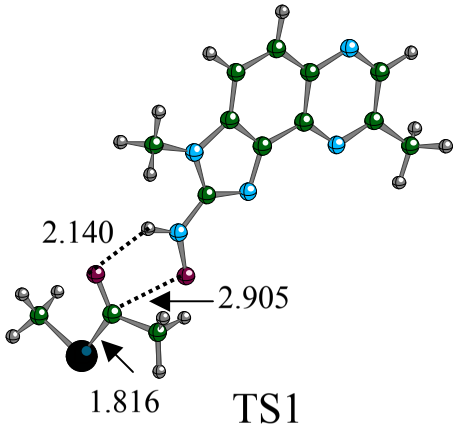


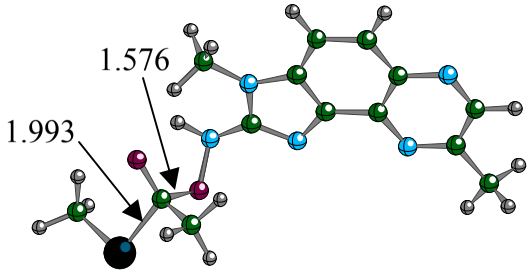
Figure 8



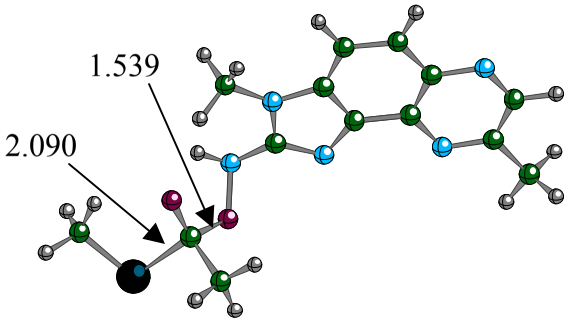
Ion-molecule complex



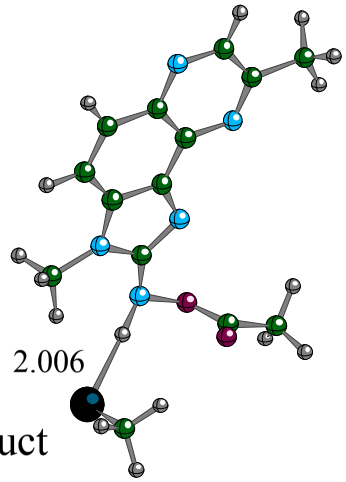
TS1



Tetrahedral intermediate



TS2



product

Figure 9

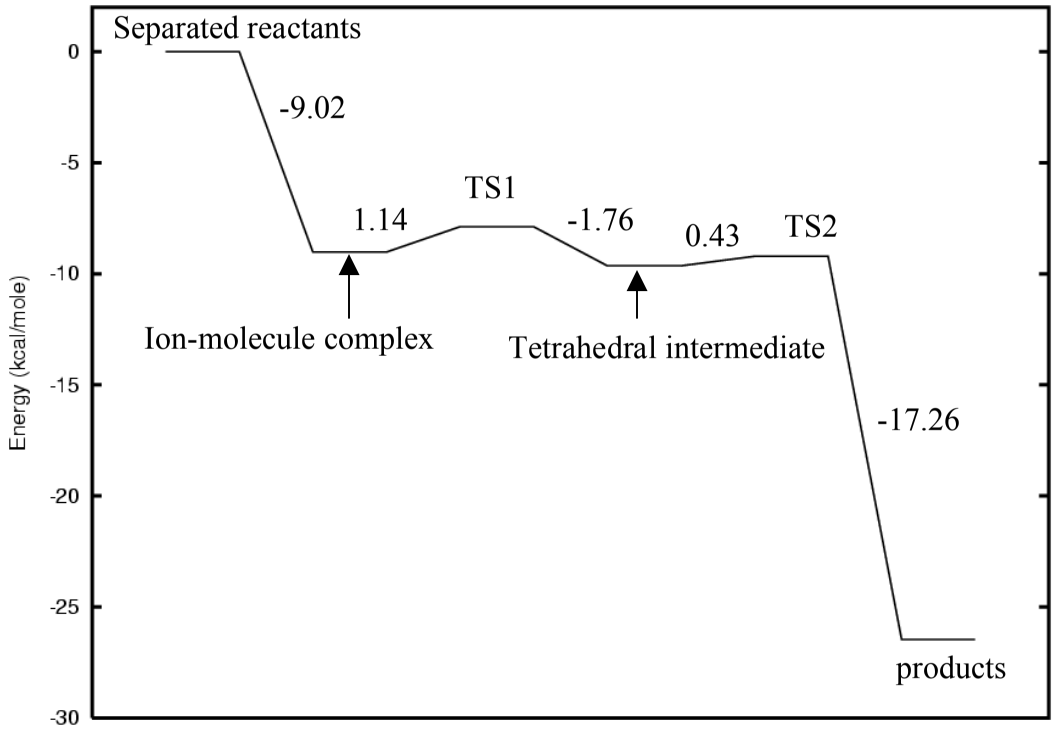

Separated reactants
0
-5
-10
-15
-20
-25
-30
Energy (kcal/mole)
-9.02
1.14
TS1
-1.76
0.43
TS2
-17.26
Ion-molecule complex
Tetrahedral intermediate
products

Figure 10