Published in final edited form as:
Biochemistry. 2011 June 7; 50(22): 4923–4935. doi:10.1021/bi101761e.

Evidence for Participation of Remote Residues in the Catalytic Activity of Co-type Nitrile Hydratase from *Pseudomonas putida*†

Heather R. Brodkin[a,‡,+], **Walter R. P. Novak**[b,§,+], **Amy C. Milne**[b], **J. Alejandro D'Aquino**[b], **N. M. Karabacak**[b], **Jeffrey N. Agar**[b], **Mark S. Payne**[c], **Gregory A. Petsko**[b], **Mary Jo Ondrechen**[a,*], and **Dagmar Ringe**[b,*]

[a]Department of Chemistry and Chemical Biology and Institute for Complex Scientific Software, Northeastern University, Boston, MA, 02115, USA

[b]Departments of Biochemistry and Chemistry and Rosenstiel Basic Medical Sciences Research Center, Brandeis University, Waltham, MA 02454-9110, USA

[c]E.I. du Pont de Nemours and Company, Wilmington, DE

Abstract

Active sites may be regarded as layers of residues, whereby the residues that interact directly with substrate also interact with residues in a second shell, and these in turn interact with residues in a third shell. These residues in the second and third layers may have distinct roles in maintaining the essential chemical properties of the first-shell catalytic residues, particularly their spatial arrangement relative to the substrate binding pocket, and their electrostatic and dynamic properties. The extent to which these remote residues participate in catalysis and precisely how they affect first-shell residues remains unexplored. In order to better understand the roles of second- and third-shell residues in catalysis, we used THEMATICS to identify residues in the second- and third-shells of the Co-type nitrile hydratase from *Pseudomonas putida* (ppNHase) that may be important for catalysis. Five of these predicted residues, plus three additional, conserved residues that were not predicted, have been conservatively mutated, and their effects studied both kinetically and structurally. All of these eight residues have no direct contact with the active site

†This research was supported by grant numbers EF-0425719, MCB-0517292, MCB-0843603, and DGE-0549344 from the National Science Foundation and an IGERT Traineeship (HRB). Use of the Advanced Photon Source (APS) is supported by the U.S. Department of Energy, Basic Energy Sciences, Office of Science, under contract No. DE-AC02-06CH11357. Use of GM/CA-CAT Sector 23 is supported in whole or in part with federal funds from the National Cancer Institute (Y1-CO-1020) and the National Institute of General Medical Science (Y1-GM-1104).

*To whom correspondence should be addressed. M.J.O.: Department of Chemistry and Chemical Biology, Northeastern University, 360 Huntington Avenue, Boston, MA 02115; telephone, (617) 373-2856; fax, (617) 373-8795; m.ondrechen@neu.edu. D.R.: Department of Chemistry, Rosenstiel Basic Medical Sciences Research Center, Brandeis University, 415 South St. Waltham, MA 02454; telephone, (781) 736-4902; fax (415) 736-2405; ringe@brandeis.edu.
‡Present address: Rosenstiel Basic Medical Sciences Research Center, Brandeis University, Waltham, MA 02454-9110, USA
§Present address: Department of Chemistry, Wabash College, Crawfordsville, IN 47933, USA
+These authors contributed equally to this work.

Supplemental Information

Supplementary Figure S1 shows a sequence alignment of four Co-type Nitrile Hydratases (NHase) and four Fe-type NHases. Known functional residues are highlighted in blue. Residues chosen for second- and third-shell mutations are highlighted in red. Residues chosen as negative controls are highlighted in green. [1] Co-type nitrile hydratases; [2] Fe-type nitrile hydratases; [3] Co-type nitrile hydratase from *Pseudomonas putida*.

Supplementary Figure S2 is provided to correct sequencing discrepancies that were published in the NCBI database. Table S2 A shows a protein sequence alignment of corrected NHase α-chain protein sequence (query) and original deposition GI # 1877504 (subject). Table S2 B shows a DNA sequence alignment of corrected NHase α-chain DNA sequence (query) and original deposition GI # 3172137:2842-3474 (subject). Table S2 C shows a protein sequence alignment of corrected NHase β-chain protein sequence (query) and original deposition GI # 1877505 (subject). Table S2 D shows a DNA sequence alignment of corrected NHase β-chain DNA sequence (query) and original deposition GI # 3172137:3485-4138 (subject).

The Supplemental materials may be accessed free of charge online at http://pubs.acs.org.

metal ion or bound substrate. These results demonstrate that three of the predicted second-shell residues, α-Asp164, β-Glu56, and β-His147, and one predicted third-shell residue β-His71, have significant effects on the catalytic efficiency of the enzyme. One of the predicted residues, α-Glu168, and the three residues not predicted, α-Arg170, α-Tyr171, and β-Tyr215, do not show any significant effects on the catalytic efficiency of the enzyme.

Enzyme active sites have evolved distinct electrostatic and chemical properties in order to facilitate catalysis and substrate recognition. Do these properties arise solely from residues immediately surrounding the reacting substrate molecule, or do the next-nearest neighbors, the “second-shell” residues, located behind the first layer, contribute also? An abundance of experimental evidence has established the importance of particular residues in the catalytic and recognition capabilities of enzymes. Typically, the residues that have been studied are, in the active form of the enzyme, in direct contact with the reacting substrate molecule or bound metal ion. These residues are considered to be in the “first coordination shell” of the bound substrate molecule or bound metal ion. However, long range interactions are known to contribute to a variety of protein processes and effects, including electron transfer (1), allostery (2), stability (3), secondary structure formation (4), and folding (5). A limited set of previous experimental data suggests that remote residues, particularly those in the second- and third-shells around the reacting substrate, also contribute to the catalytic activity of at least some enzymes.

In 1985, Leatherbarrow, Fersht, and Winter reported a startling example of participation by remote residues in catalysis for tyrosyl-tRNA synthetase (TyrRS). When His45, one residue removed from the site of reaction, is mutated to Gly, the rate of the catalyzed reaction decreases by two orders of magnitude (6). This early experiment raised the question of remote residue participation in enzyme catalysis. Evidence for the importance of residues outside the first-shell for control of enzyme specificity became evident when efforts to impart chymotrypsin specificity onto trypsin by altering only first-shell residues were unsuccessful. These two proteases have nearly identical active sites, but trypsin cleaves after positively charged residues and chymotrypsin cleaves after hydrophobic residues. Asp189 in the binding pocket of trypsin was thought to be responsible for the recognition of positively charged residues; however, mutation of Asp189 to the corresponding residue in chymotrypsin was insufficient to confer chymotrypsin specificity. Ultimately, a total of 16 mutations, including residues not in direct contact with the substrate, were required to engineer chymotrypsin specificity onto the trypsin structure (7–9).

In studies of the metal coordinating sites in metalloenzymes, distinct patterns have been observed in both the first (*i.e.* directly coordinating the metal) and second layers of residues around the metal ions, suggesting that second-shell residues are important in maintaining the chemical properties of the metal ion (10–11). For instance, mutations to both second- and third-shell residues in alkaline phosphatase (AP) have resulted in either a decrease or increase in catalytic rate depending on the mutation (12–15). Also, mutations to the second-shell residues around the zinc ion in metallo-carbonic anhydrase (CAII) result in a decrease in catalytic activity (16). Specifically, mutations at Gln92, a residue that hydrogen-bonds with the metal coordinating residue, His94, are isostructural with the wild type enzyme (17) and yet a three- to nine- fold decrease in k_{cat}/K_M is observed (16). Finally, a major change in catalytic rate was observed for a second-shell mutant of mandelate racemase (MR), an enzyme that catalyzes the interconversion of the *(R)*- and *(S)*-enantiomers of mandelate via the abstraction of a proton from the α-carbon atom (18–20). In MR, Asp270 is a second-shell residue within hydrogen bond distance to the catalytic His297 and therefore is believed to affect the orientation of His297. The single mutation Asp270Asn results in a 10,000-fold decrease in enzyme activity compared to wild type for both *(R)*- and *(S)*- mandelate

substrates (21). The mutant and wild type structures are nearly identical except that the side chain of the catalytic His is "rotated and displaced toward the binding site" in the mutant structure (21). The authors argue that Asp270 is necessary to impart the correct pK_a to the catalytic His297 (21). Taken together, these studies indicate that residues outside the active site have the ability to significantly modulate protein function.

This limited experimental evidence suggests that at least some enzyme active sites may be built in multiple layers, and that residues beyond the first-shell of the enzyme active site may be important for catalysis and/or substrate binding. However, it remains unclear to what extent recognition and catalysis are influenced by residues outside the first shell. One barrier to the investigation of such effects is the selection of residues outside the active site for study. Therefore a systematic approach for selecting first-, second- and third-shell residues is clearly indicated in order to understand the extent and mechanisms of remote residue participation in enzyme catalysis.

In this paper, first-shell residues are defined as those residues which are identified by the Catalytic Site Atlas (22) as necessary for catalysis or identified by Ligand-Protein Contacts (23) as involved in binding to substrates or cofactors. Second-shell residues are defined as those residues identified by Contacts of Structural Units (23) as contacting first-shell residues; third-shell residues are those residues identified by the same method but in contact with second-shell residues. It has been demonstrated that active site residues can be identified by abnormal theoretical titration curve shapes (Theoretical Microscopic Anomalous Titration Curve Shapes, THEMATICS) (24–26) and that active sites typically consist of clusters of residues with these perturbed titration curves. Here, we have adapted this method of using theoretical titration curves in order to identify residues in the second- and third- shells that may be important for enzyme function.

Nitrile hydratase (NHase) is a metalloenzyme that catalyzes the conversion of nitriles to amides. It has attracted much attention because it is the first example of a biocatalyst used in the large-scale industrial production of commodity chemicals such as acrylamide (27–28). In addition to the use of NHase in the production of acrylamide, it is also used in the production of nicotinamide and 5-cyanovaleramide, the latter being a starting material for the synthesis of the herbicide azafenidin (29). NHases are considered superior as catalysts compared with earlier synthetic methods due to the mild reaction conditions, high yields, absence of byproducts, and the possibility for stereoselectivity. There are two types of NHase metalloenzymes; cobalt-dependent and iron-dependent. Cobalt-dependent (Co-type) NHase possesses a noncorrin cobalt center, while iron-containing (Fe-type) NHase possesses a nonheme iron center, and, in each case, the metal center is in the +3 oxidation state (27). The metal coordination spheres for both types involve a C1-S/T-L-C2-S-C3 motif, where cysteines C2 and C3 are typically oxidized to sulfinic and sulfenic acid, respectively (30).

Both Fe- and Co-type NHases are bacterial heterodimers, made up of non-homologous α- and β-subunits, each having a molecular weight of around 23 kDa. However, both of these subunits exhibit a high degree of sequence similarity with their corresponding subunits among all known NHases (Supplementary Figure S1). Here, we focus on Co-type NHase from *Pseudomonas putida* (ppNHase, EC 4.2.1.84) and report the structures of the wild type and four mutant proteins. We also report the kinetic constants for the wild type and eight second- or third-shell mutants. Second- and third- shell residues that are most likely to be involved in catalysis were selected using the following guidelines: 1) Prediction by THEMATICS (THEMATICS-positive residues) (24–26); and 2) high sequence conservation scores, as calculated by ConSurf (31). These predicted residues include the second-shell residues α-Asp164, α-Glu168, β-Glu56 and β-His147, and the third-shell residue β-His71. Additionally, the second-shell residue α-Arg170 and the third-shell residues α-Tyr171 and

β-Tyr215 were selected for mutagenesis as negative controls. These residues also have high conservation scores, but are not identified by THEMATICS as participating in catalysis (THEMATICS-negative residues). α-Asp164, α-Glu168, β-Glu56, β-His147, α-Arg170, and β-Tyr215 are all well conserved across Co-type and Fe-type NHases, while α-Tyr171 and β-His71 are partially conserved (Supplementary Figure S1). Conservative mutations were used to probe the effects of removing the ionizable groups while retaining similar steric bulk in the residue side chain. Our results suggest that second- and third-shell residues are functionally important to varying degrees. Our results further suggest that sequence conservation alone is insufficient for determining the involvement of remote residues in catalysis, and that THEMATICS can aid in the identification of catalytically important remote residues.

Experimental Procedures

Computational Methods: Shell analysis

First-shell residues were identified using the Catalytic Site Atlas (CSA, http://www.ebi.ac.uk/thornton-srv/databases/CSA/) (22) and the Ligand Protein Contact server (LPC, http://bip.weizmann.ac.il/oca-bin/lpccsu) (23). The first shell includes literature annotated catalytic residues and those residues in contact with a bound ligand or metal ion. Second- and third-shell residues, those in contact with a given first- or second-shell residue, respectively, were determined using the Contacts of Structural Units (CSU) server (http://bip.weizmann.ac.il/oca-bin/lpccsu) (23).

Computational Methods: Selection of mutation candidates

The coordinate files in PDB format for ppNHase were analyzed by the computational method Theoretical Microscopic Anomalous Titration Curve Shapes (THEMATICS, http://pfweb.chem.neu.edu/thematics/submit.html) (24–26), first using the method of Wei (26) and then repeated using the same method but with a more inclusive cut-off. THEMATICS analyzes the shapes of the computed theoretical titration curves of the ionizable residues in a protein structure and selects the residues with the most perturbed titration curve shapes as the functionally important residues. The degree of deviation of the titration curve shape from the ideal Henderson-Hasselbalch behavior is measured by the moments of the first derivatives of the curves (24). The number of residues predicted depends on a statistical cut-off parameter; Wei *et al.* showed (26) that a cut-off parameter value of 0.99 yields the optimum predictions, taking into account both sensitivity and selectivity and using the 170 enzymes in the original, manually-curated portion of the CSA (22, 32) as the reference set. Since sensitivity and selectivity are inversely related, the optimization necessarily sacrifices some sensitivity in order to achieve good selectivity. However, since the standard cut-off of 0.99 was fixed to yield optimum performance on the CSA-annotated catalytic residues, virtually all of which are in the first shell, it was decided to repeat the THEMATICS calculation, using a more lenient cut-off value of 0.96, to seek additional residues in the second and third shells that might have some effect on catalysis. Coordinates for the cobalt were not included in the THEMATICS analysis, and active site cysteines were treated as reduced residues. Complete THEMATICS predictions for ppNHase are presented in Table 1.

Site-Directed Mutagenesis

An expression plasmid (obtained from E.I. du Pont de Nemours and Company, Wilmington, DE) containing the genes for the *Pseudomonas putida* NRRL-18668 α- and β-subunits of NHase and for the NHase activator, P14K, was used for protein expression and mutagenesis. Site-directed mutagenesis was carried out using Stratagene's Quikchange II Site Directed Mutagenesis Kit (Stratagene, La Jolla, CA). The primers are shown in Table 2. PCR

amplification was performed per the manufacturer's instructions. The mutated gene was sequenced to verify the construct. (Genewiz, South Plainfield, NJ). After confirmation of the intended mutation, the plasmid was transformed into BL21(DE3) competent cells (Stratagene, La Jolla, CA) for expression.

Protein Expression and Purification

The wild type and mutants were expressed in *E. coli* BL21(DE3) cells (Stratagene, La Jolla, CA). Cells were grown at 37 °C, in 1 L of 2XYT broth containing ampicillin (100 μg/ml). At an A_{600} of 0.8, cells were induced by the addition of 1.0 mM IPTG and 0.5 mM cobalt chloride (33). Overexpression of ppNHase continued with shaking for an additional 4–6 hours at 28 °C. All subsequent manipulations were performed at 4 °C. At all stages of purification, fractions containing ppNHase were run on an SDS gel, and those containing the purest protein were pooled for further purification and analysis. After cell harvesting by centrifugation, the pellet was resuspended in 40 mL of 50 mM Tris pH 8.0 containing 2 mM β-ME (Buffer A). Cells were sonicated for a total of 10 minutes on ice with a 10 second on pulse and 1 minute off. Supernatant was collected after centrifugation at 12, 000 × *g* for 60 minutes. The protein-containing supernatant was loaded onto a 60 mL DEAE column, and was eluted with a 700 mL linear-gradient from 80 to 200 mM NaCl in Buffer A at flow rate of 1.0 mL/min (34). Ammonium sulfate precipitation (70%) was performed on the fractions containing ppNHase. After centrifugation and reconstitution in Buffer A, the protein was further purified on a Phenyl Sepharose column (GE Healthcare, Piscataway, NJ) using a flow rate of 1.0 mL/min and a linear-gradient from 0.5 M ammonium sulfate in Buffer A to only Buffer A over 180 mL, followed by an additional 20 mL of Buffer A. The protein was then concentrated using an Amicon Ultra-15 Centrifugal Filter Unit with Ultracel-10 membrane (Millipore, Billerica, MA) with a 10 kDa nominal molecular weight limit and dialyzed 2 times (4 hours each) against Buffer A to ensure the complete removal of ammonium sulfate. For the final purification step, the protein was loaded onto a Mono-Q column (GE Healthcare, Piscataway, NJ) and eluted using a flow rate of 0.5 mL/min and a linear-gradient from 125 mM to 237.5 mM sodium chloride in Buffer A over 90 mL followed by an additional 45 mL of 237.5 mM sodium chloride in Buffer A. The protein was again concentrated using an Amicon Ultra-15 Centrifugal Filter Unit with Ultracel-10 membrane (Millipore, Billerica, MA) with 10 kDa nominal molecular weight limit and dialyzed 2 times (4 hours each) against Buffer A and stored at 4 °C. It was discovered that the α-Asp164Asn mutant protein was likely slightly destabilized by the mutation and tended to dissociate into monomers when subjected to high salt concentration, as evidenced by the disappearance of one protein band in the SDS-PAGE gel after ammonium sulfate precipitation. The purification protocol for this mutant was adjusted slightly by removing the ammonium sulfate precipitation and phenyl sepharose column steps. After the DEAE column, the α-Asp164Asn mutant protein was concentrated and subjected only to the Mono-Q column. All proteins were judged to be greater than 95% pure using SDS PAGE and Coomassie blue staining. The concentration of all proteins was determined by A_{280} measurement. The extinction coefficient used was 1.676 $mg \cdot mL^{-1} \cdot cm^{-1}$ (http://us.expasy.org/cgi-bin/protparam).

Mass Spectrometry

To perform mass spectrometry, ppNHase was desalted using Amicon Ultra centrifugal membrane filter with 10 kDa nominal molecular weight limit and diluted with 10 mM ammonium bicarbonate in HPLC grade water to a concentration of 10 μM. ppNHase was directly infused using a syringe pump into a electrospray ion source with dual ion funnel (35) (Apollo II) connected to a hybrid quadrupole Fourier transform ion cyclotron resonance mass spectrometer (Apex Qe-94, Bruker Daltonics, Billerica, MA). The resulting mass spectrum was deconvoluted using Data Analysis software (version 3.4, Bruker Daltonics)

and monoisotopic masses were calculated using the SNAP algorithm (version 2, Bruker Daltonics). The monoisotopic masses of ppNHase α-and β-subunits were calculated to be 24668.17 and 24009.10 Da using Isotope Pattern software (Bruker Daltonics), respectively. During mass calculations from the mass spectrum, the existence of three charges from the Co(III) ion were considered and three daltons were subtracted from the masses generated by the software, which assumes all charges on the ion are due to protons.

DNA and Protein Sequencing

Purified DNA samples were sent to Genewiz, Inc. (South Plainfield, NJ) for sequencing. For protein sequencing, purified protein was run on an SDS-PAGE gel, electroblotted to a PVDF membrane, stained with Coomassie, excised and submitted to the Iowa State University Protein Facility (Ames, IA) for Edman degradation sequencing.

Kinetics

NHase activity was determined by measurement of the hydration of *n*-valeronitrile in a 300 μL reaction volume, containing 100 mM HEPES pH 6.7 and 2 mM β-ME in an ice bath at 0 °C. An ice bath was used to slow the rate of reaction to a measurable rate. Non-linear regression analysis was performed in order to obtain k_{cat} and K_M values. The concentrations of *n*-valeronitrile were 0.63, 2.5, 5.0, 10, and 40 mM. The concentration of ppNHase was adjusted between 6.0 nM and 60 nM so that the concentration of product formed was within the range of the standard curve. The reaction was carried out for 40 and 60 minutes in an ice bath at 0 °C and was stopped by the addition of 0.3 N HCl. Each reaction was performed in triplicate. On average, the concentration of product formed for all experiments was between 49 μM and 2 mM.

The formation of *n*-valeramide was monitored with a Waters 2690 HPLC (Waters Corp., Milford, MA) or an Agilent 1200 HPLC (Agilent Technologies, Santa Clara, CA), using a Zorbax Aq reverse phase C_{18} column (4.6 × 150 mm) (Agilent Technologies, Santa Clara, CA) at a flow rate of 1.0 mL/min (36). Standard curves were obtained prior to each experiment and ranged between 38 μM and 2.50 mM *n*-valeramide. Running buffers were 5 mM potassium phosphate, pH 2.9 (A) and 100% acetonitrile (B), running at 1.0 ml/min. The product was eluted with a small gradient of 10% – 25% B over 7 minutes. Sample run time was 14 minutes. The absorbance of product was measured at 210 nm.

Crystallization, data collection and crystallographic refinement

Crystals of ppNHase were grown at 25 °C by vapor diffusion in 24 well hanging drop plates over 0.7 mL volume reservoirs using 1 + 1 μL drops. Wild type and mutant ppNHase crystals formed under the same conditions. Initial crystals of the β-Glu56Gln mutant were extremely small; therefore, initial crystal forms were used to streak seed drops producing larger crystals for data collection. Diffracting crystal needles were obtained using 20 mg/mL ppNHase and a reservoir containing 22% polyacrylic acid sodium salt 5100, 100 mM HEPES pH 7.5, 20 mM magnesium chloride and 4% acetone. Single crystals (needles) were dissected from clusters and transferred to a solution containing 17.6% polyacrylic acid sodium salt 5100 and 20% glycerol in 100 mM HEPES pH 7.5 and were flash-frozen in liquid nitrogen.

Data were collected at the ID-23B and the ID-23D beamlines at GM/CA-CAT (APS, Argonne, IL, USA) at 100 K using a MARMosiac 300 CCD detector and the 10 μm mini-beam. Diffraction images were indexed in space group $P2_1$, integrated and scaled using HKL2000 (37). Molecular replacement was carried out with Phaser (38) using wild type ptNHase (PDB 1IRE (30)) as a starting model for wild type ppNHase, and the resulting structure for wild type ppNHase was used as a starting model for the mutant ppNHases.

Several rounds of refinement and model building were performed using REFMAC (39) and COOT (40). Final rounds of refinement, including simulated annealing and water picking, were performed using PHENIX (41). The atomic coordinates of wild type ppNHase, as determined in this study, have been deposited in the RCSB Protein Data Bank under accession number 3QXE. The coordinates for α-E168Q have been deposited under accession number 3QZ5, for β-E56Q under 3QYG, for β-H71L under 3QYH, and for β-Y215F under 3QZ9.

Results and Discussion

THEMATICS is a computational method that predicts active site residues in enzymes using computed chemical and electrostatic properties. Specifically, perturbed ionization behavior as predicted by THEMATICS indicates that a residue is in a special environment in the protein. We have found that clusters of these residues occur in the active site of an enzyme with high frequency and with sufficient exclusivity that these computed properties are very useful for the accurate and selective identification of active sites in protein structures (26). This suggests that the electrostatic properties which cause the anomalous ionization behavior are important for catalysis. Here, using THEMATICS (24–26), we have identified a set of THEMATICS-positive residues in ppNHase outside the first-shell of the active site. Again, the anomalous ionization behavior suggests that these residues, though more distant from the active site, may have an impact on activity. Using the standard cut-off of 0.99 on the structure of ppNHase as described above, the following cluster is predicted as the active site:

[α-C112, α-C115, α-C117, β-E56, β-R52, β-H147]

This prediction includes four first-shell residues, the three active-site cysteines and the catalytic β-Arg52, plus two second-shell residues, β-Glu56 and β-His147. With the less selective cut-off of 0.96, the same six residues, plus 12 additional residues, are predicted, as shown in Table 1. This less selective prediction includes the catalytic β-Arg149 and the ligand-binding β-Tyr68, plus seven additional second-shell residues and three third-shell residues. For mutagenesis studies, we selected the two second-shell residues, β-E56 and β-H147, that were predicted by THEMATICS with the standard cut-off, plus a sample of three of the residues predicted by THEMATICS using the more lenient cut-off, α-Asp164 and α-Glu168 in the second shell and β-His71 in the third shell. We also included for study three residues not predicted by THEMATICS at all, α-Arg170 in the second shell and α-Tyr171 and β-Tyr215 in the third shell. All eight of these residues selected for mutagenesis studies are highly conserved (31) and are located in either the second or third shell of the active site. We elected to study the contributions of these residues to catalysis by a combination of mutagenic, kinetic, and structural techniques. Conservative mutations were made that preserve the relative size and shape of the residue to the extent possible, and usually also preserve hydrogen bonding capability, but remove the potential to be ionized (e.g., Glu to Gln; Asp to Asn; His to Asn, Leu, or Phe; Tyr to Phe; Arg to Gln). Thus, the causes of the decreased activities of mutants may vary depending on the interactions that the wild type and mutant residues make with surrounding residues or water molecules. In some cases the effect may be due largely to alterations in positioning of residues or water molecules, in others it may be due to electrostatic, dynamic, or quantum mechanical effects. Specifically, mutations in the second- and third-shells may exert electrostatic effects; for instance, they may alter the pK_a of one or more catalytic residues, or otherwise alter the pH range over which the catalytic residue exists in its protonated and deprotonated states. Second- and third-shell mutations may also have dynamic effects, altering the prevalence of different conformational states of the enzyme. They may also have electronic or quantum mechanical effects, altering effective charges and bond orders of the active atoms of the catalytic residues, or of the substrate molecule itself.

Kinetic Analyses

The essential roles of catalytic and ligand-binding residues have been probed by mutagenesis for both Co-type and Fe-type NHases. In particular, the three cysteine residues that are known to be coordinating ligands to the metal center, have been mutated to alanine in single, double and triple forms in rrNHase (42). All of these mutations exhibited no activity, and many could not be expressed at all according to SDS-PAGE. Additionally, mutations have been made to a tyrosine residue thought to be involved in ligand binding in ptNHase (34); the conservative β-Tyr68Phe mutation results in a 100-fold decrease in k_{cat} and a 10-fold increase in K_M with aliphatic substrates. It is believed that the hydroxyl group forms hydrogen bonds with the substrate, and that removal of this functional group affects binding and subsequently activity. Finally, the two active site arginine residues (56 and 141; equivalent to 52 and 149 in ppNHase) have both been independently mutated to Glu, Tyr and Lys in the rsNHase (43–44). These mutants exhibited sharply decreased or no enzymatic activity.

The wild type ppNHase behaves similarly to other nitrile hydratases in terms of the kinetic constants (45–46). For this study, kinetics experiments were performed at 0°C using *n*-valeronitrile as the substrate. All mutations to second- and third-shell residues in ppNHase in this study resulted in enzymes that are still active, and bind substrate with similar affinity (K_M within a factor of 4) to wild type ppNHase. However, these mutations also result in k_{cat} values that range from similar to wild type to a 100-fold decrease (Table 3).

Structure of Wild Type ppNHase

The structure of the wild type ppNHase was determined for comparison with other NHases and with the mutants described above. The data collection and refinement statistics for wild type and four mutants are shown in Table 4. During refinement, the wild type protein sequences for the α- and β-chains in the NCBI database for ppNHase (GI numbers: 1877504 and 1877505, respectively) were found to be inconsistent with the electron density in some regions. DNA sequencing determined the correct sequence for wild type ppNHase, which is in agreement with the electron density; errors in the original deposition are attributed to early DNA sequencing errors. Alignments of the original and corrected protein and DNA sequences are given in Supplementary Figure S2.

The overall structure of wild type ppNHase is similar to that of other NHase enzymes, which is not surprising given the high degree of sequence similarity among NHases (Supplementary Figure S1). Superimposing wild type ppNHase (this work) and ptNHase (PDB ID: 1IRE (30)) yields an RMSD of 0.7 Å over 177 α-carbons for the α-subunit (out of 207), and an RMSD of 0.9 Å over 183 α-carbons for the β-subunit (out of 219) when no atom pair distance is allowed to exceed 2.0 Å (Figure 1). The sequence identity between ptNHase and ppNHase is approximately 58% for the α-subunit and 43% for the β-subunit.

For the α-subunit of wild type ppNHase, no electron density was observed for the first six or seven residues, the 16 residues of the T7-tag, or the last four residues of the sequence. Backbone density was observed for the entire β-subunit. The refined model is composed of four copies of the α-subunit consisting of residues 7 (or 8) – 207, each containing one cobalt ion, four copies of the β-subunit residues 1 – 219, and 1143 water molecules. The protein exhibits good geometry as determined by PROCHECK (47); 92% of the residues are in the most favored region of the Ramachandran analysis, with 7.2% and 0.7% in the allowed and generously allowed regions, respectively. The average overall temperature factor for the structure is 23.6 $Å^2$. The average temperature factor for the cobalt ion is 13.5 $Å^2$, and for side chains and water molecules is 25.0 $Å^2$. The mutant protein models are composed of four copies of the α-subunit consisting of residues 7 (or 8) – 207, each containing one cobalt

ion and four copies of the β-subunit residues 1 – 219. The number of water molecules was different for all mutants (Table 4).

The minimal functional unit of NHases is the αβ-heterodimer, though the larger heterotetramer has been reported for the NHase from *Rhodococcus* sp. N-771 (rsNHase) (48). Under the purification conditions utilized here, ppNHase ran as the heterodimer on a size-exclusion column (data not shown). The crystal form of ppNHase, however, shows a buried surface area of approximately 1000 $Å^2$ between β-subunits as calculated by PISA (http://www.ebi.ac.uk/msd-srv/prot_int/pistart.html) (49). This region corresponds to the heterotetrameric interface reported in other NHases and is of similar size. The heterotetrameric interface is pseudo-symmetric and is composed of ten predicted hydrogen bonds. A second surface area of approximately 575 $Å^2$ between β-subunits is involved in crystal packing and is partially comprised of the α5-loop-α6 region described below. For comparison, the buried surface area between the α- and β-subunits of the obligate heterodimer is ~3800 $Å^2$.

Despite the overall similarity among NHases, comparison of wild type ppNHase with other Co-type NHase structures shows large differences in the α5-loop-α6 region in the β-subunit and in the location of the N-terminus of the α-subunit. ppNHase has an eight amino acid deletion in the β-subunit with respect to other Co-type NHases, resulting in a substantially shorter α5 helix, an alternate flexible loop location and slightly shorter α6 helix (Figure 1). In other Co-type NHases, the α5 helix interacts with much of the α1 helix of the α-subunit. In ppNHase, the shorter α5 helix makes fewer interactions with the α1 helix of the α-subunit, while the loop region between α5 and α6 caps the C-terminal end of the α1 helix of the α-subunit. The role of the α5 helix in NHases is unknown. In ppNHase, the N-terminus occludes the entrance of one opening of the large substrate tunnel. This is similar to the partial covering described in the bsNHase structure (50); however, the N-terminal interactions with the tunnel occur in trans in the bsNHase heterotetrameric complex and in cis in the ppNHase heterotetrameric complex. No blocking of this tunnel opening occurs from the N-terminal helix in ptNHase.

Crystal structures have been reported for Co-type NHases from *Pseudonocardia thermophila* (ptNHase) (PDB ID: 1IRE) (30), *Bacillus smithii* (bsNHase) (PDB ID: 1V29) (50), and *Bacillus sp. RAPc8* (bpNHase) (PDB ID: 2DPP). The active site of wild type ppNHase (Figure 2) is very similar to previously reported NHase structures where the cobalt is bound in an octahedral conformation, or pyramidal if a sixth ligand is absent, with ligands from the C1-T-L-C2-S-C3 motif: the sulfur atoms from three cysteines, α-Cys112, α-Cys115 and α-Cys117 and the backbone amide nitrogen of a cysteine (α-Cys112) and serine (α-Ser116). In the ptNHase structure, a water molecule is observed as a sixth ligand (27). In other known structures of Co-type NHases, C2 (α-Cys115) is oxidized to a sulfinic acid and C3 (α-Cys117) is oxidized to a sulfenic acid (34, 51–52). However, in the ppNHase structure reported here, two of the cysteine residues (C2 (α-Cys115) and C3 (α-Cys117)) are oxidized to sulfinic acids and held in place by interactions with two arginines, β-Arg52 and β-Arg149. The oxygen atoms of the sulfinic acids and the serine form a claw-like configuration. The sulfinic form has also recently been observed in Fe-type NHase from *Rhodococcus erythropolis* AJ270 (reNHase), but it has not been shown unequivocally that the double oxidized form of the enzyme is active (53). Kinetic analyses were performed using dissolved crystals of ppNHase and this protein was active, although five times less so than the protein in solution (data not shown).

In order to investigate the oxidation state of the active site prior to crystallization for ppNHase, mass spectrometry studies were performed. The results indicate that our purified protein samples have a monoisotopic α-subunit mass of 24668.18 Da, which is in good

agreement with the calculated monoisotopic mass of 24668.20 Da for T7-tagged protein with the N-terminal methionine cleaved, a single cobalt ion and three oxygen atoms (one sulfinic and one sulfenic acid modification) (data not shown); the different oxidation states of the three cysteine residues for NHase have been reported previously by Odaka (54). The *in vivo* cleavage of the N-terminal methionine of the ppNHase α-subunit was verified by Edman degradation sequencing. Given the mass spectrometry data, it is likely that the further oxidation of the C3 sulfenic acid to sulfinic acid witnessed in the ppNHase crystal structure occurs upon lengthy exposure to air or to trace amounts of oxidizers present in the crystallization buffer. It has been shown that anaerobic reconstitution of Fe-type NHase from *Rhododcoccus sp.* N-771 (rsNHase) resulted in an inactive protein without the C2 sulfinic modification, but that aerobic incubation induced the modification rendering the protein active (54). For thiocyanate hydrolase (SCNase), which possesses an active site similar to that of NHase, progressive oxidation states have also been reported (55), where the partially-oxidized form (C2 sulfinic acid and C3 sulfenic acid) is active and the fully-oxidized form (C2 and C3 both oxidized to sulfinic acid) is inactive.

A sixth-ligand water observed in another structure of NHase could not be clearly distinguished however a large feature of weak density is present in the active site of ppNHase; a similar feature in the equivalent position has also been reported for SCNase (55). Given the propensity of NHases to bind carboxylates and the extent of the electron density, the ligand may be polyacrylic acid, though no attempt was made to model this molecule into the density. A single molecule of glycerol is also present in the substrate tunnel almost 11 Å from the active site cobalt ion. A second glycerol molecule binds to the surface of ppNHase near β-Gly14.

Structures of Mutant ppNHases

The structures obtained for four of the mutant proteins of ppNHase show the expected differences in electron density from wild type due to changes in side chain identity, and differences that correlate with changes in electrostatic environment. The α-carbon RMSD's between wild type ppNHase and all mutant ppNHase structures were less than 0.2 Å for both the α- and β-subunits.

Mutational Analysis

The second- and third-shell residues that are the focus of this study were designed to probe the effect of these mutations on the structure of the active site and on the activity observed.

Second-shell residues

α-Asp164Asn—Asp164 on the α-subunit interacts with α-Cys112, the non-oxidized active site cysteine, through a water molecule (Figure 2). Mutation to asparagine retains the relative size, shape, and hydrogen bonding capability, while removing the ability to form a negative charge. This mutation results in a 70-fold decrease in the catalytic rate (k_{cat}), and also results in a 4-fold decrease in the Michaelis constant (K_M) compared to wild type (Table 3). Unfortunately, we were unable to determine a crystal structure for this mutant, but the circular dichroism spectrum was superimposable on that of the wild type, suggesting no gross structural changes. Thus it is likely that the most significant effect of the mutation is to alter the electrostatic environment of the active site, rather than its structure. It is possible that the loss of the ability to form a negative charge at position 164 could affect the pK_a of α-Cys112, potentially causing the destabilization of the active site. Maintenance of this electrostatic interaction is clearly important to the activity of the enzyme.

α-Glu168Gln—In the wild type structure, the side chain of α-Glu168 forms a salt bridge with the catalytic residue β-Arg52 (Figures 2 and 3A). This arginine residue in turn interacts

with the two modified cysteine residues of the active site. Again, the conservative change of glutamate to glutamine was selected in order to remove the ability to ionize but retain the hydrogen bonding capability. This mutation results in a slight decrease in K_M and a 5-fold decrease in catalytic activity, resulting in a less than four-fold reduction in the catalytic efficiency of the enzyme (Table 3). The free energy difference ($\Delta\Delta G = -RT \ln [(k_{cat}/K_M)_{mutant}/ (k_{cat}/K_M)_{wild\ type}]$) between α-Glu168Gln and the wild type enzyme is less than one kcal/mol (Table 3).

In the mutated structure (Figure 3B) the side chain of α-Gln168 is flipped away from β-Arg52, breaking the salt bridge with β-Arg52, and, instead, forming a hydrogen bond to the backbone oxygen atom of β-Val69. At the same time, the interactions between β-Arg52 and the two cysteine residues have changed slightly: the hydrogen bond distance between β-Arg52 and α-Cys117 has increased from 2.7 Å to 3.1 Å, while the hydrogen bond distance between β-Arg52 and α-Cys115 has essentially remained the same. It should be noted that the 0.4 Å difference in bond length between wild type and the α-Glu168Gln mutant is statistically significant at this resolution. This arginine residue is known to H-bond with α-Cys115 and α-Cys117 in the active site, and seems to stabilize the "claw" setting (30). This residue is completely conserved among all known NHases and it has been shown to be essential for function in Fe-type NHases (43). Thus, it is likely that β-Arg52 is also critical for function in the Co-type NHases. The removal of the salt bridge to β-Arg52 causes only a small decrease in the catalytic rate which may be due to a slight destabilization of the “claw” setting in the active site. Thus, while there may be some differences between the WT and α-Glu168Gln in the local structure and electrostatic environment of the active site, these differences have only a small effect on catalysis. Despite the large change in the side chain position, the small $\Delta\Delta G$ between α-Glu168Gln and the wild type enzyme suggests that α-Glu168 is not very important in catalysis, and should be considered a false positive prediction by THEMATICS with a reduced (non-standard) cut-off.

β-Glu56Gln—β-Glu56 is a second-shell residue that is within hydrogen bond distance to both the modified cysteine, α-Cys115, and to the functionally important arginine, β-Arg149, through a water molecule (Figure 4). β-Glu56 also forms a hydrogen bond with the side chain of β-His147. The conservative mutation of glutamate to glutamine eliminates the capability to form a negative charge, while retaining hydrogen-bonding capability and size. The β-Glu56Gln mutation shows almost no change in K_M, but displays a two order of magnitude decrease in catalytic activity compared to wild type, the largest decrease observed among all of the mutants studied (Table 3). Interestingly, however, the structures of both wild type and mutant protein are nearly identical, as are the hydrogen bond distances for the critical interactions (Figure 4). There is however a slight lengthening of the hydrogen bond distance to β-His147 in the mutant structure (0.5 Å).

Based on these results, the effect of the β-Glu56Gln mutation seems more likely to be due to its coupling to catalytic residues, rather than conformational effects. The interaction of β-Glu56 with β-Arg149, although indirect and through a water molecule, seems to have a greater effect on activity than the direct interaction that α-Glu168 has with β-Arg52 (*vide supra*). The implication is that the interaction of β-Glu56 with this water molecule in the active site plays an important role in the reaction, and the orientation and/or polarization of this water molecule may be important for catalysis. Additionally, β-Glu56 could be modulating the electrostatic properties of the modified cysteine residue at position 115, as β-Glu56 also interacts with this residue through the same water molecule.

β-His147Asn—β-His147 is located behind the active-site residues β-Arg52 and β-Arg149. The β-His147 side chain is within hydrogen bonding distance to β-Glu56, and, via a water molecule, to β-Arg52. The conservative mutation of histidine to asparagine preserves the

hydrogen bonding capability but eliminates the proton exchange capability. This mutation results in an almost 30-fold decrease in k_{cat}, while the K_M increases four-fold, corresponding to a 110-fold decrease in the catalytic efficiency (Table 3). This is similar to the result obtained via the mutagenesis of β-Glu56 to glutamine; these results suggest that the interaction between these two residues is important to the catalytic activity of the enzyme. This interaction between β-Glu56 and β-His147 may play a critical role in maintaining the pK_as of these two residues, and also of nearby residues. β-His147 may also modulate the electrostatic environment of the active site through a water molecule and β-Arg52.

Third-shell residues

β-His71Leu, β-His71Asn, β-His71Phe—The side chain of β-His71 is separated by about 7 Å from the main chain oxygen atom of α-Cys115 and the side chain of α-Ser116 via a hydrogen-bonding network through two water molecules (Figure 5). Additionally, there is a connection through a hydrogen bonding network between the side chain of β-His71 and the side chain oxygen of α-Cys115 through three water molecules. The mutation of β-His71 to phenylalanine or leucine removes the ability of the side chain to hydrogen bond and, in the β-His71Leu structure, this causes a shift of a water molecule by 0.4 Å in the active site (*, Figure 5B), disrupting the hydrogen bond network. The third-shell β-His71Leu mutation left the K_M value unaffected, but reduced k_{cat} 16-fold compared to wild type. This decrease in k_{cat} is similar to the decrease seen in the β-His71Asn and β-His71Phe mutations; however both the β-His71Asn and β-His71Phe mutations also resulted in a larger increase in the K_M value. While the shift in water position is significant, this in itself is not sufficient to explain the decrease in catalytic rate; it is therefore hypothesized that the decrease in activity is also due to alterations in the electrostatic environment of the active site.

Negative controls

α-Arg170Gln—α-Arg170 is a conserved second-shell residue located 6.2 Å from the cobalt ion, and is within hydrogen bonding distance to α-Glu168 and α-Gln93. The α-Arg170Gln mutation results in a 1.5-fold decrease in k_{cat} and a 1.2-fold increase in K_M, corresponding to a 1.8-fold decrease in the catalytic efficiency. These changes in kinetic constants result in a free energy difference of less than one kcal/mol in comparison to the wild type enzyme (Table 3), and are not significant. Hence, as predicted, α-Arg170 is not important for catalysis.

α-Tyr171Phe—α-Tyr171 is a conserved third-shell residue located 13 Å from the cobalt ion. The α-Tyr171Phe mutation causes a 1.5-fold drop in k_{cat} and a 1.4-fold increase in K_M, corresponding to a two-fold loss in the catalytic efficiency. Again, these are not significant changes in the kinetic behavior, thus confirming the prediction that α-Tyr171 is not important for catalysis.

β-Tyr215Phe—In the wild type structure, β-Tyr215 is a third-shell residue that is located approximately 14 Å from the cobalt ion (Figures 2 and 6A). The hydroxyl group of this tyrosine residue is within hydrogen-bonding distance to the main chain nitrogen atom of β-Arg149 and the side chain oxygen of β-Asp172. This suggests that β-Tyr215 assists in maintaining the structural integrity of the β-Arg149 side chain, a residue which hydrogen bonds to the metal coordinating residue α-Cys115 and is known to help stabilize the active site (30). Since these interactions involve the phenolic -OH group of the side chain, the mutation to phenylalanine should disrupt these interactions. The β-Tyr215Phe mutation results in a two-fold decrease in the Michaelis constant and an eight-fold decrease in catalytic activity (Table 3). The catalytic efficiency of the enzyme is decreased three-fold

with respect to the wild type enzyme, and $\Delta\Delta G$ between β-Tyr215Phe and the wild type enzyme is only 0.6 kcal/mol (Table 3).

Although the hydrogen bonding capabilities of the side chain have been removed, the structure does not at first glance seem to have changed significantly. There are very slight (0.1 – 0.2 Å) shifts in the hydrogen bond distances between β-Arg149 and the modified cysteine residues, but they are not significant at this resolution. There are, however, less obvious changes that have taken place that involve the salt bridge between the second-shell residue α-Glu168 and the first-shell residue β-Arg52 (Figures 6C–D). Specifically, there is a shift in the side chain of α-Glu168 causing a lengthening of the salt bridge by 1.0 Å, which is statistically significant. However, maintenance of this salt bridge does not appear to be important in preserving the catalytic efficiency of the enzyme. As in the α-Glu168Gln mutant, weakening or removal of this salt bridge interaction results only in small changes in $\Delta\Delta G$ with respect to the wild type enzyme. Estimation of the pK_a values using PROPKA 2.0 (56) indicates the pK_a of β-Arg52 is lowered only slightly in the mutant (11.3 and 10.8 in the wild type and β-Tyr215Phe structures, respectively), while the pK_a of β-Arg149 is unchanged. Despite the conservation of β-Tyr215 among NHases and some observed structural changes, the kinetic data presented here show that this residue is not very important in maintaining the catalytic efficiency of the enzyme.

Summary

Kinetics studies were performed on five second- and third- shell residues that were predicted by THEMATICS to be important in catalysis of ppNHase. Two of these are second-shell residues predicted by the standard THEMATICS calculation (26); the other three were predicted by THEMATICS with a less selective cut-off (26). The two second-shell residues predicted with the standard cut-off, β-Glu56 and β-His147, showed the greatest effect on the catalytic rate constant and on the catalytic efficiency. Of the three residues predicted with the less selective cut-off, two showed a significant decrease in the catalytic effect, although the effect is smaller. Mutations to the three residues that were not predicted by THEMATICS, even with the less selective cut-off, did not show significant difference in kinetic behavior from that of the wild type.

Interestingly, on one hand, breakage of the salt bridge between the active site β-Arg52 and the second-shell residue α-Glu168 by mutation of the glutamate to glutamine has no significant effect on catalysis, although the local structure around the active site is changed. On the other hand, loss of coupling between two residues located outside the active site and 9–10 Å away from the cobalt center, β-Glu56 and β-His147, does have a significant effect on catalysis.

The alterations in the local structure and the absence of a significant kinetic effect observed for α-Glu168Gln further suggest that the active site of ppNHase has some plasticity and therefore is tolerant of some structural variation. The apparent importance of the β-Glu56 - β-His147 couple, located well outside the active site, suggests that electrostatic effects are important.

The kinetic results reported here support the prediction that some second-shell residues and, to a lesser extent, some third-shell residues play a supporting role in catalysis in ppNHase. Furthermore, while single point mutations lead to a drop in catalytic rate by one or two orders of magnitude, collectively these residues outside the first shell could amount to a much larger effect on the catalytic rate. In other words, the composition of the second and third shell may be *critical* to achieving optimal enzymatic catalysis. There are multiple ways in which remote residues might participate in catalysis and the present study offers some suggestions as to how these residues contribute in nitrile hydratase catalysis.

Possible mechanisms for remote residue participation include: 1) local rotations or side chain shifts and more subtle changes in configuration; 2) shifts in hydrogen-bonding networks; 3) changes in dynamics that affect the active conformation of the active site; 4) changes in the intrinsic electric field in the region of the active site; or 5) quantum mechanical effects, such as changes in effective atomic partial charges, bond polarities, and polarizabilities. Small, local structural changes were observed for a few of the mutants. However, the effects on kinetics demonstrated here for second- and third-shell residues surely cannot be attributed solely to structural effects, even for those mutations where small structural shifts are observed. It is probable that structural, dynamic, and electrostatic effects are all involved in catalysis, in addition to potential recognition, binding, and alignment of key water molecules in the active site.

Understanding how nature designs enzyme active sites is a fundamental question in enzymology, and one with profound implications for protein engineering. In addition to a few examples in the literature, this study clearly demonstrates that there are long-range effects, including electrostatic effects, which influence enzyme catalysis. Second- and third-shell mutations, predicted through THEMATICS and taking into account evolutionary conservation, suggest that ppNHase has a multilayered active site. The present results suggest that computational methods could help guide the identification of such functional residues in other enzymes. Further investigation is necessary to determine whether this is a general phenomenon among all enzymes.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We are grateful to Dr. Paul Vouros, Northeastern University, and Waters Corporation, Milford, MA, for unlimited instrument use.

Abbreviations and textual footnotes

The following residues are equivalent according to sequence and structural alignment: (in order: *Pseudonorcardia thermophila* 1IRE; *Rhodococcus erythropolis* 2AHJ; *Pseudomonas putida* this work; *Bacillus sp.* 2DPP) α-subunit; Cys108, Cys109, Cys112, Cys116; Cys111, Cys112, Cys115, Cys119; Ser112 Ser113, Ser116, Ser120; Cys113, Cys114, Cys117, Cys121; Asp161, Asp161, Asp164, Asp160; Glu165, Glu165, Glu168, Glu172; β-subunit; Arg52, Arg56, Arg52, Arg56; Arg157, Arg141, Arg149, Arg160.

NHase	nitrile hydratase
bsNHase	nitrile hydratase from *Bacillus smithii*
ptNHase	nitrile hydratase from *Pseudonorcardia thermophila*
ppNHase	nitrile hydratase from *Pesudomonas putida*
bpNHase	nitrile hydratase from *Bacillus sp. RAPc8*
rrNHase	nitrile hydratase from *Rhodococcus rhodochrous*
reNHase	nitrile hydratase from *Rhodococcus erythropolis* AJ270
rsNHase	nitrile hydratase from *Rhododcoccus sp.* N-771
IPTG	isopropyl-β-D-thiogalactoside

β-ME β-mercaptoethanol

References

1. Siddarth P, Marcus RA. Electron-Transfer Reactions in Proteins: An Artificial Intelligence Approach to Electronic Coupling. J Phys Chem. 1993; 97:2400–2405.
2. Tsai C, del Sol A, Nussinov R. Protein Allostery, signal transmission and dynamics: a classification scheme of allosteric mechanisms. Molecular Biosystems. 2009; 5:207–216. [PubMed: 19225609]
3. Noivirt-Brik O, Unger R, Horovitz A. Analysing the origin of longrange interactions in proteins using lattice models. BMC Structural Biology. 2009; 9
4. Kihara D. The effect of long-range interactions on the secondary structure formation of proteins. Protein Science. 2005; 14
5. Naganathan AN, Muñoz V. Insights into protein folding mechanisms from large scale analysis of mutational effects. Proc Nat Acad Sci. 2010; 107:8611–8616. [PubMed: 20418505]
6. Leatherbarrow RJ, Fersht AR, Winter G. Transition state stabilisation in the mechanism of tyrosyl-tRNA synthetase revealed by protein engineering. Proc Natl Acad Sci USA. 1985; 82:7840–7844. [PubMed: 3865201]
7. Graf L, Craik CS, Patthy A, Roczniak S, Fletterick RJ, Rutter WJ. Selective alteration of substrate specificity by replacement of aspartic acid-189 with lysine in the binding pocket of trypsin. Biochemistry. 1987; 26:2616–2623. [PubMed: 3111531]
8. Perona JJ, Hedstrom L, Rutter WJ, Fletterick RJ. Structural origins of substrate discrimination in trypsin and chymotrypsin. Biochemistry. 1995; 34:1489–1499. [PubMed: 7849008]
9. Venekei I, Szilagyi L, Graf L, Rutter WJ. Attempts to convert chymotrypsin to trypsin. FEBS Lett. 1996; 379:143–147. [PubMed: 8635580]
10. Karlin S, Zhu Z-Y, Karlin KD. The extended environment of mononuclear metal centers in protein structures. Proc Natl Acad Sci USA. 1997; 94:14225–14230. [PubMed: 9405594]
11. Karlin S, Zhu Z-Y. Classification of mononuclear zinc metal sites in protein structures. Proc Natl Acad Sci USA. 1997; 94:14231–14236. [PubMed: 9405595]
12. Xu X, Qin XQ, Kantrowitz ER. Probing the role of histidine-372 in zinc binding and the catalytic mechanism of Escherichia coli alkaline phosphatase by site-specific mutagenesis. Biochemistry. 1994; 33:2279–2284. [PubMed: 8117685]
13. Hehir MJ, Murphy JE, Kantrowitz ER. Characterization of heterodimeric alkaline phosphatases from Escherichia coli: an investigation of intragenic complementation. J Mol Biol. 2000; 304:645–656. [PubMed: 11099386]
14. Mandecki W, Shallcross MA, Sowadski J, Tomazic-Allen S. Mutagenesis of conserved residues within the active site of Escherichia coli alkaline phosphatase yields enzymes with increased kcat. Protein Eng. 1991; 4:801–804. [PubMed: 1798702]
15. Muller BH, Lamoure C, Le Du MH, Cattolico L, Lajeunesse E, Lemaitre F, Pearson A, Ducancel F, Menez A, Boulain JC. Improving Escherichia coli alkaline phosphatase efficacy by additional mutations inside and outside the catalytic pocket. Chembiochem. 2001; 2:517–523. [PubMed: 11828484]
16. Christianson DW, Fierke CA. Carbonic anhydrase: Evolution of the zinc binding site by nature and by design. Acc. Chem. Res. 1996; 29:331–339.
17. Lesburg CA, Christianson DW. X-ray Crystallographic Studies of Engineered Hydrogen Bond Networks in a Protein-Zinc Binding Site. J Am Chem Soc. 1995; 117:6838–6844.
18. Mitra B, Kallarakal AT, Kozarich JW, Gerlt JA, Clifton JG, Petsko GA, Kenyon GL. Mechanism of the reaction catalyzed by mandelate racemase: importance of electrophilic catalysis by glutamic acid 317. Biochemistry. 1995; 34:2777–2787. [PubMed: 7893689]
19. Landro JA, Kallarakal AT, Ransom SC, Gerlt JA, Kozarich JW, Neidhart DJ, Kenyon GL. Mechanism of the reaction catalyzed by mandelate racemase. 3. Asymmetry in reactions catalyzed by the H297N mutant. Biochemistry. 1991; 30:9274–9281. [PubMed: 1909893]

20. Kallarakal AT, Mitra B, Kozarich JW, Gerlt JA, Clifton JG, Petsko GA, Kenyon GL. Mechanism of the reaction catalyzed by mandelate racemase: structure and mechanistic properties of the K166R mutant. Biochemistry. 1995; 34:2788–2797. [PubMed: 7893690]
21. Schafer SL, Barrett WC, Kallarakal AT, Mitra B, Kozarich JW, Gerlt JA, Clifton JG, Petsko GA, Kenyon GL. Mechanism of the reaction catalyzed by mandelate racemase: structure and mechanistic properties of the D270N mutant. Biochemistry. 1996; 35:5662–5669. [PubMed: 8639525]
22. Porter CT, Bartlett GJ, Thornton JM. The Catalytic Site Atlas: a resource of catalytic sites and residues identified in enzymes using structural data. Nucl. Acids Res. 2004; 32:D129–D133. [PubMed: 14681376]
23. Sobolev V, Sorokine A, Prilusky J, Abola EE, Edelman M. Automated analysis of interatomic contacts in proteins. Bioinformatics. 1999; 15:327–332. [PubMed: 10320401]
24. Ko J, Murga LF, Andre P, Yang H, Ondrechen MJ, Williams RJ, Agunwamba A, Budil DE. Statistical Criteria for the Identification of Protein Active Sites Using Theoretical Microscopic Titration Curves. Proteins: Structure Function Bioinformatics. 2005; 59:183–195.
25. Ondrechen MJ, Clifton JG, Ringe D. THEMATICS: A simple computational predictor of enzyme function from structure. Proc. Natl. Acad. Sci. (USA). 2001; 98:12473–12478. [PubMed: 11606719]
26. Wei Y, Ko J, Murga LF, Ondrechen MJ. Selective Prediction of Interaction Sites in Protein Structures with THEMATICS. BMC Bioinformatics. 2007; 8:119. [PubMed: 17419878]
27. Tyler LA, Noveron JC, Olmstead MM, Mascharak PK. Oxidation of Metal-Bound Thiolato Sulfur Centers in Fe(III) and Co(III) Complexes with Carboxamido Nitrogens and Thiolato Sulfurs as Donors: Relevance to the Active Sites of Nitrile Hydratase. Inorg. Chem. 1999; 38:616–617.
28. Kobayashi M, Shimizu S. Metalloenzyme nitrile hydratase: structure, regulation, and application to biotechnology. Nat Biotechnol. 1998; 16:733–736. [PubMed: 9702770]
29. Cowan DA, Cameron RA, Tsekoa TL. Comparative biology of mesophilic and thermophilic nitrile hydratases. Adv Appl Microbiol. 2003; 52:123–158. [PubMed: 12964242]
30. Miyanaga A, Fushinobu S, Ito K, Wakagi T. Crystal structure of cobalt-containing nitrile hydratase. Biochem Biophys Res Commun. 2001; 288:1169–1174. [PubMed: 11700034]
31. Glaser F, Pupko T, Paz I, Bell RE, Bechor-Shental D, Martz E, Ben- Tal N. ConSurf: Identification of functional regions in proteins by surfacemapping of phylogenetic information. Bioinformatics. 2003; 19:163–164. [PubMed: 12499312]
32. Bartlett GJ, Porter CT, Borkakoti N, Thornton JM. Analysis of Catalytic Residues in Enzyme Active Sites. J Mol Biol. 2002; 324:105–121. [PubMed: 12421562]
33. Wu S, Fallon RD, Payne MS. Over-production of stereoselective nitrile hydratase from Pseudomonas putida 5B in Escherichia coli: activity requires a novel downstream protein. Appl Microbiol Biotechnol. 1997; 48:704–708. [PubMed: 9457799]
34. Miyanaga A, Fushinobu S, Ito K, Shoun H, Wakagi T. Mutational and structural analysis of cobalt-containing nitrile hydratase on substrate and metal binding. EurJBiochem. 2004; 271:429–438.
35. Him T, Tolmachev AV, Harkewicz R, Prior DC, Anderson G, Udseth HR, Smith RD, Bailey TH, Rakov S, Futrell JH. Design and implementation of a new electrodynamic ion funnel. Analytical Chemistry. 2000; 72:2247–2255. [PubMed: 10845370]
36. Fallon RD, Stieglitz B, Turner I Jr. A *Pseudomonas putida* capable of stereoselective hydrolysis of nitriles. Appl. Microbiol. Biotechnol. 1997; 47:156–161.
37. Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. Macromolecular Crystallography, Pt A. 1997; 276:307–326.
38. Mccoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser crystallographic software. Journal of Applied Crystallography. 2007; 40:658–674. [PubMed: 19461840]
39. Murshudov GN, Vagin AA, Dodson EJ. Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr D Biol Crystallogr. 1997; 53:240–255. [PubMed: 15299926]
40. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr. 2004; 60:2126–2132. [PubMed: 15572765]

41. Adams PD, Grosse-Kunstleve RW, Hung LW, Ioerger TR, McCoy AJ, Moriarty NW, Read RJ, Sacchettini JC, Sauter NK, Terwilliger TC. PHENIX: building new software for automated crystallographic structure determination. Acta Crystallogr D Biol Crystallogr. 2002; 58:1948–1954. [PubMed: 12393927]
42. Hashimoto Y, Sasaki S, Herai S, Oinuma K, Shimizu S, Kobayashi M. Site-directed mutagenesis for cysteine residues of cobalt-containing nitrile hydratase. J Inorg Biochem. 2002; 91:70–77. [PubMed: 12121763]
43. Piersma SR, Nojiri M, Tsujimura M, Noguchi T, Odaka M, Yohda M, Inoue Y, Endo I. Arginine 56 mutation in the beta subunit of nitrile hydratase: importance of hydrogen bonding to the non-heme iron center. J Inorg Biochem. 2000; 80:283–288. [PubMed: 11001100]
44. Endo I, Nojiri M, Tsujimura M, Nakasako M, Nagashima S, Yohda M, Odaka M. Fe-type nitrile hydratase. J Inorg Biochem. 2001; 83:247–253. [PubMed: 11293544]
45. Nagasawa T, Nanba H, Ryuno K, Takeuchi K, Yamada H. Nitrile hydratase of Pseudomonas chlororaphis B23. Purification and characterization. Eur J Biochem. 1987; 162:691–698. [PubMed: 3830164]
46. Nagasawa T, Takeuchi K, Yamada H. Characterization of a new cobalt-containing nitrile hydratase purified from urea-induced cells of Rhodococcus rhodochrous J1. Eur J Biochem. 1991; 196:581–589. [PubMed: 2013281]
47. Laskowski RA, Macarthur MW, Moss DS, Thornton JM. Procheck - a Program to Check the Stereochemical Quality of Protein Structures. Journal of Applied Crystallography. 1993; 26:283–291.
48. Nakasako M, Odaka M, Yohda M, Dohmae N, Takio K, Kamiya N, Endo I. Tertiary and quaternary structures of photoreactive Fe-type nitrile hydratase from Rhodococcus sp. N-771: roles of hydration water molecules in stabilizing the structures and the structural origin of the substrate specificity of the enzyme. Biochemistry. 1999; 38:9887–9898. [PubMed: 10433695]
49. Krissinel E, Henrick K. Inference of macromolecular assemblies from crystalline state. J Mol Biol. 2007; 372:774–797. [PubMed: 17681537]
50. Hourai S, Miki M, Takashima Y, Mitsuda S, Yanagi K. Crystal structure of nitrile hydratase from a thermophilic Bacillus smithii. Biochem Biophys Res Commun. 2003; 312:340–345. [PubMed: 14637142]
51. Nagashima S, Nakasako M, Dohmae N, Tsujimura M, Takio K, Odaka M, Yohda M, Kamiya N, Endo I. Novel non-heme iron center of nitrile hydratase with a claw setting of oxygen atoms. Nat Struct Biol. 1998; 5:347–351. [PubMed: 9586994]
52. Tsekoa TL, Tastan-Bishop AO, Cameron RA, Sewell BT, Sayed MF, Cowan DA. Crystal structure of thermostable Bacillus sp. RAPc8 nitrile hydratase. To be Published.
53. Song L, Wang M, Shi J, Xue Z, Wang MX, Qian S. High resolution X-ray molecular structure of the nitrile hydratase from Rhodococcus erythropolis AJ270 reveals posttranslational oxidation of two cysteines into sulfinic acids and a novel biocatalytic nitrile hydration mechanism. Biochem Biophys Res Commun. 2007; 362:319–324. [PubMed: 17716629]
54. Murakami T, Nojiri M, Nakayama H, Odaka M, Yohda M, Dohmae N, Takio K, Nagamune T, Endo I. Post-translational modification is essential for catalytic activity of nitrile hydratase. Protein Sci. 2000; 9:1024–1030. [PubMed: 10850812]
55. Arakawa T, Kawano Y, Katayama Y, Nakayama H, Dohmae N, Yohda M, Odaka M. Structural Basis for Catalytic Activation of Thiocyanate Hydrolase Involving Metal-Ligated Cysteine Modification. J Am Chem Soc. 2009; 131:14838–14843. [PubMed: 19785438]
56. Bas DC, Rogers DM, Jensen JH. Very Fast Prediction and Rationalization of pKa Values for Protein-Ligand Complexes. Proteins. 2008; 73:765–783. [PubMed: 18498103]

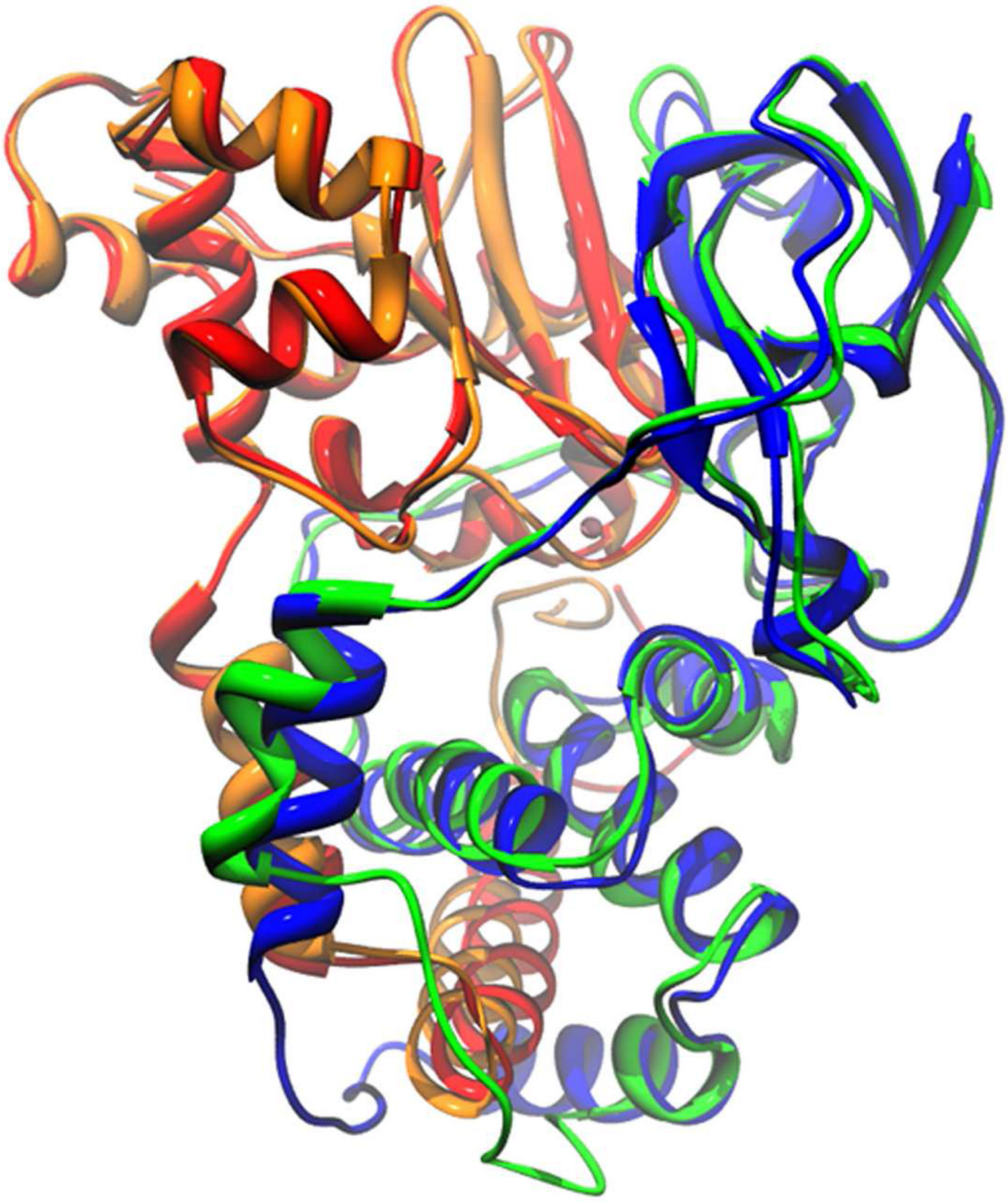

Figure 1.
Superposition of wild type ppNHase (PDB ID: 3QXE, this work) and wild type ptNHase (PDB ID: 1IRE (30)) structures. ppNHase and ptNHase α-subunits are in yellow and red and β-subunits are in green and blue, respectively. RMSD for the α subunits is 0.7 Å over 177 residues and 0.9 Å over 183 residues for the β subunits. The active site cobalt is enlarged and shown in pink. The two glycerol molecules associated with each dimer are rendered as ball and stick and shown in CPK coloring.

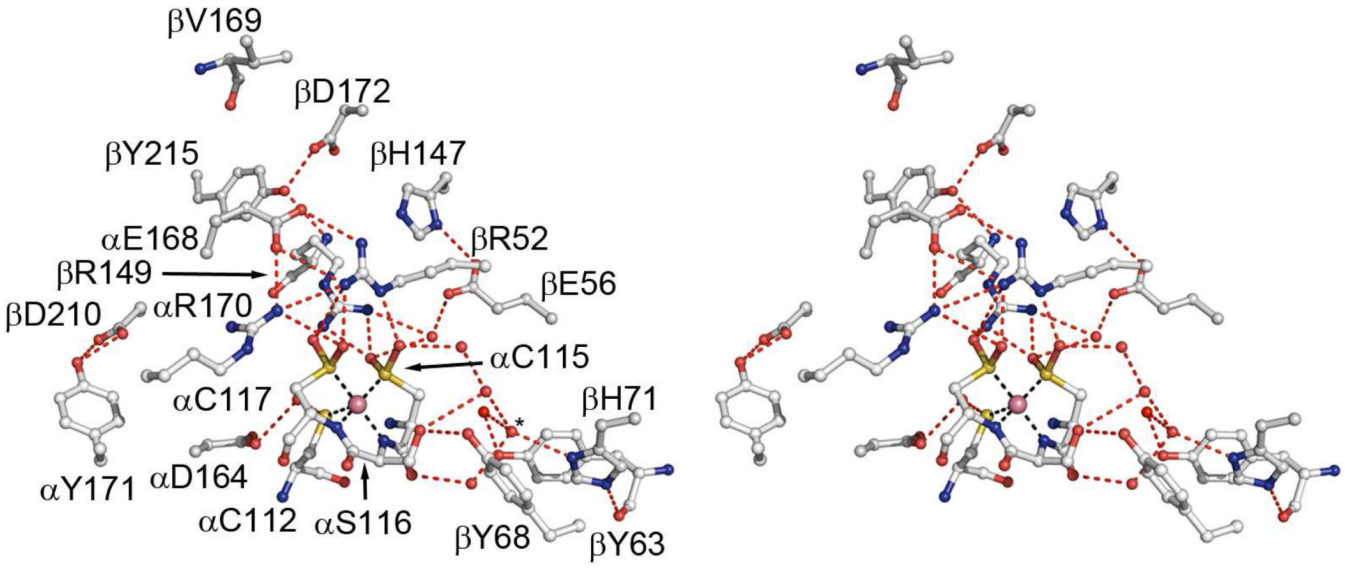

Figure 2.
Active site of Co-type nitrile hydratase from *Pseudomonas putida* shown in stereo-view. Shown are first-, second- and third-shell residues. (purple sphere = cobalt, red spheres = water)

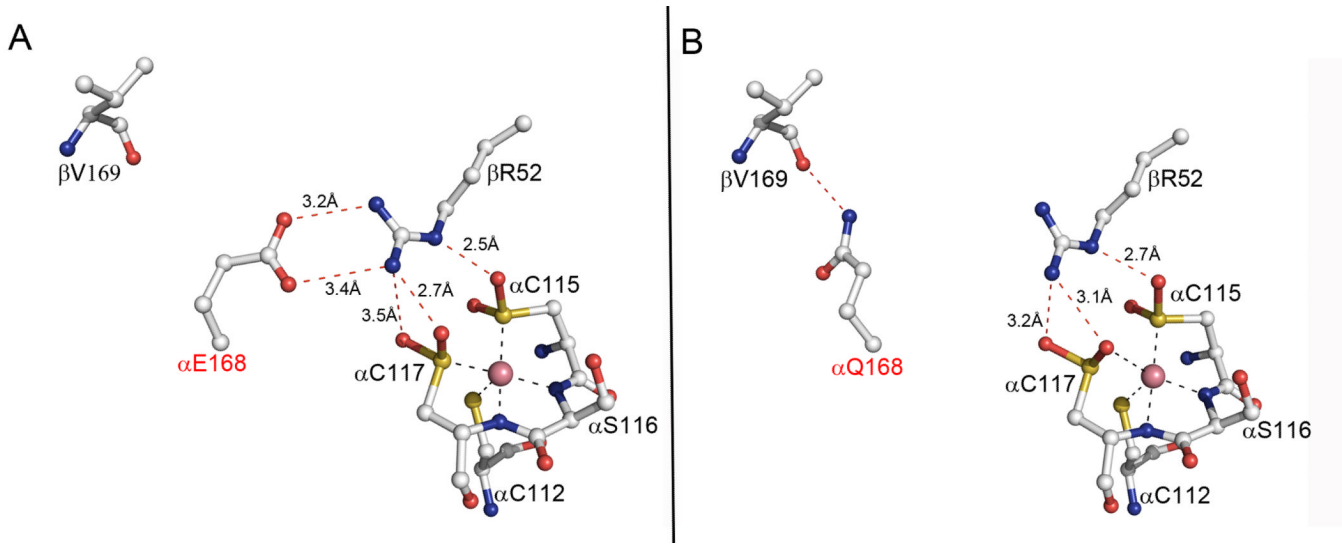

Figure 3.

Comparison of the active site of wild-type ppNHase (A) and the active site of the second-shell R-Glu168Gln mutant ppNHase (PDB entry 3QZ5, this work) (B). In the mutant structure, residue 168 has flipped out of being within salt bridge distance of β-Arg52 and forms an H-bond with the backbone oxygen atom of β-Val169 (purple sphere for cobalt).

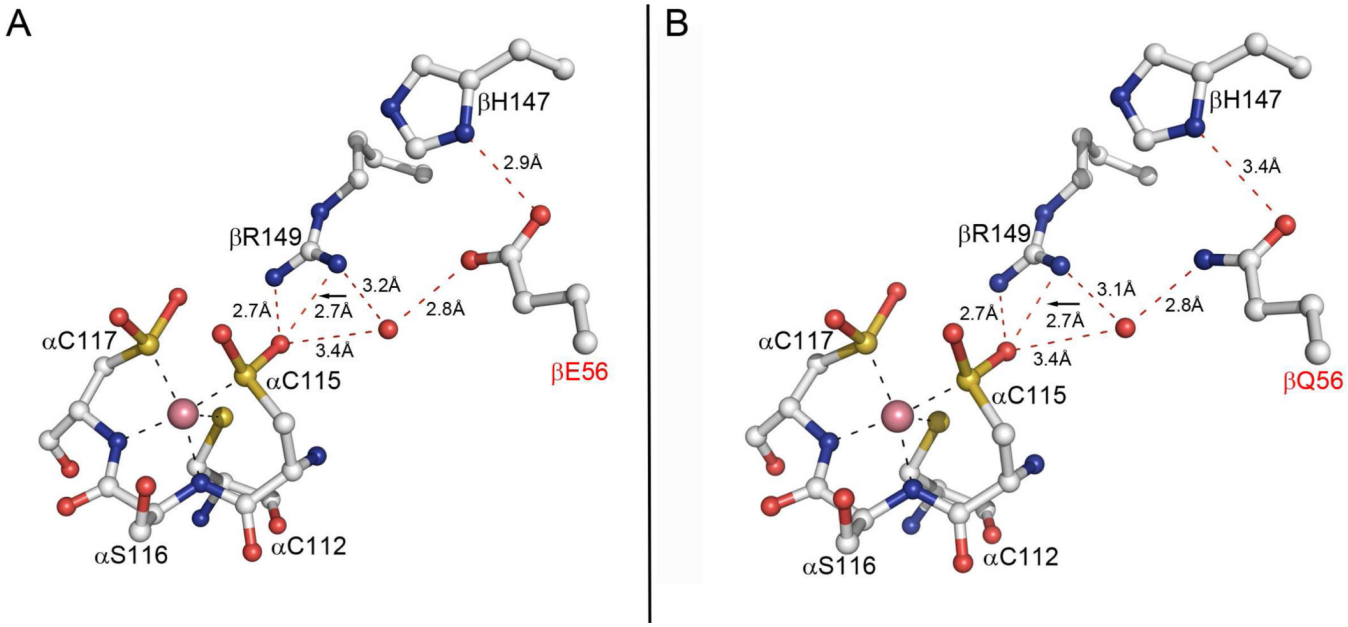

Figure 4.

Comparison of the active site of wild type ppNHase (A) and the active site of the second-shell β-Glu56Gln mutant ppNHase (PDB ID: 3QYG, this work) (B). Wild type and mutant structures are essentially the same. (purple sphere = cobalt, red sphere = water).

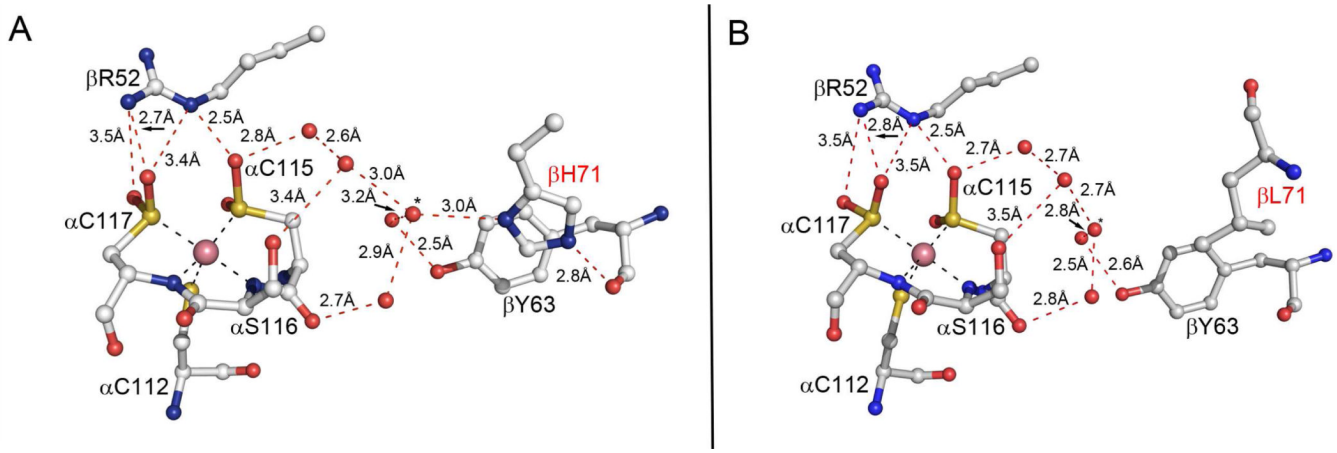

Figure 5.

Comparison of the active site of wild type ppNHase (A) and the active site of the third-shell β-His71Leu mutant ppNHase (PDB ID: 3QYH, this work) (B). Wild type and mutant structures are essentially the same, with a slight movement in one of the waters (*). (purple sphere = cobalt, red sphere = water)

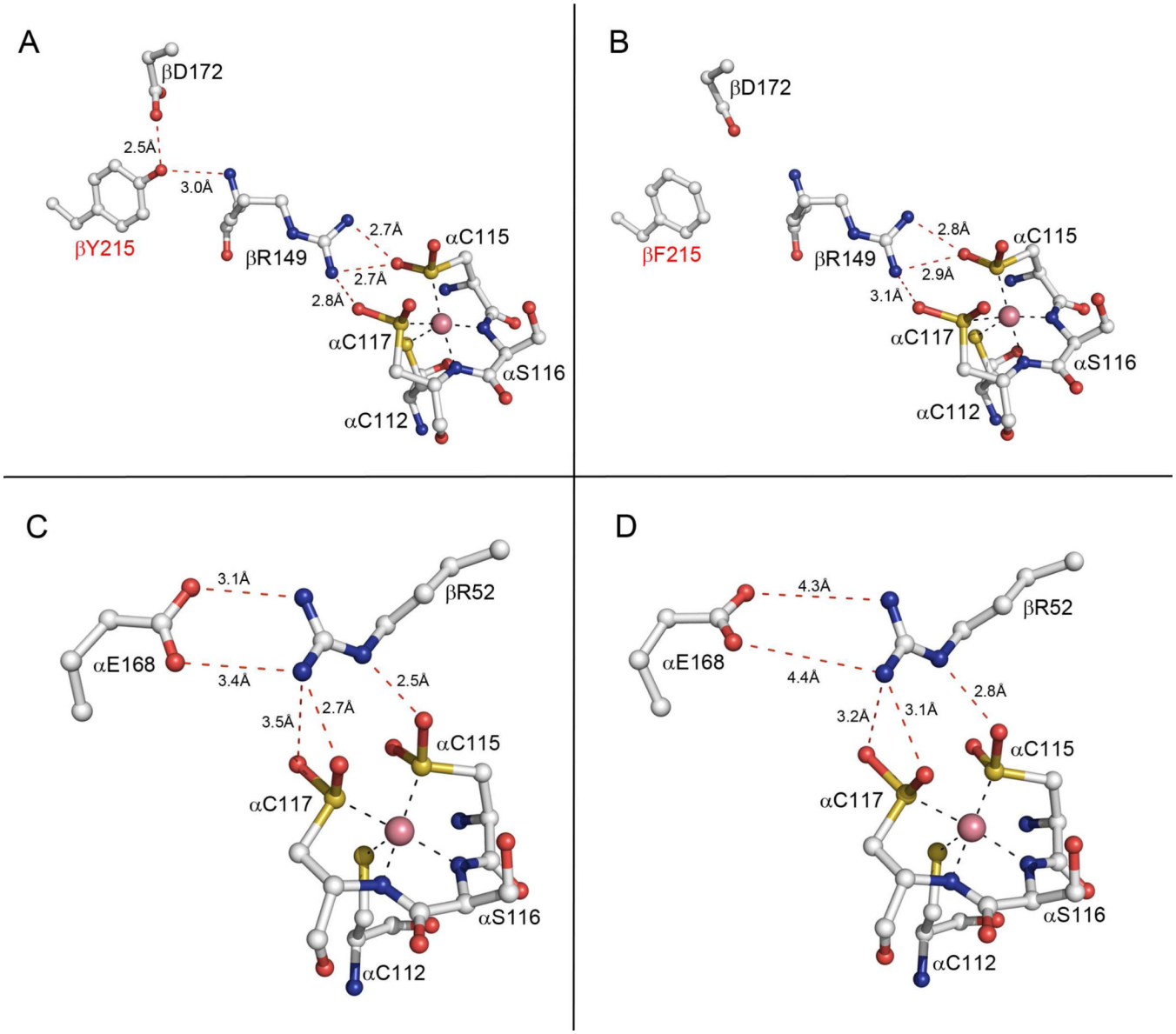

Figure 6.

Comparison of the active site of wild type ppNHase (A), (C) and the active site of the third-shell β-Tyr215Phe mutant ppNHase (PDB ID: 3QZ9, this work) (B), (D). Wild type and mutant structures are essentially the same in panels A and B. However, panels C and D show a lengthening in the salt bridge distance between α-Glu168 and β-Arg52, shown as red dotted lines. (purple sphere = cobalt).

Table 1

THEMATICS predictions of functional residues for wild type NHase from *Pseudomonas putida*. Known catalytic roles and shell assignments are listed for each residue. **Bold face** indicates residues predicted by THEMATICS that are annotated in the CSA as catalytic residues; residues predicted by THEMATICS that are found by LPC to be metal binding or ligand binding residues are shown in *italics*; those residues that are underlined have mutations reported in the literature. THEMATICS predictions are given with the standard cut-off (0.99) and a reduced cut-off (0.96). Distances from the cobalt center to the functional residue are also shown.

Assignment	Residues predicted by THEMATICS for ppNHase - 0.99 (standard) cutoff	Residues predicted by THEMATICS for ppNHase - 0.96 (reduced) cutoff	Additional mutants - NOT predicted by THEMATICS	Distance from cobalt center (Å)
Co-coordinating	*α C112**	*α C112**		2.4
Co-coordinating	**α C115***	**α C115***		2.2
Co-coordinating	**α C117***	**α C117***		2.2
Catalytic	**β R52**	**β R52**		5.6
Ligand binding		*β Y68**		6.8
Catalytic		β R149*		5.9
Second shell		β D49		10.7
Second shell		β H53		10.5
Second shell	β E56*	β E56*		8.8
Second shell		α Y118		7.7
Second shell		α Y130		6.4
Second shell		α K131*		8.3
Second shell	β H147*	β H147		10.0
Second shell		α D164*		7.6
Second shell		α E168*		8.0
Second shell			α R170	6.2
Third shell		β H5		8.8
Third shell		β Y69*		11.9
Third shell		β H71*		9.9
Third shell			α Y171	12.8
Third shell			β Y215*	13.6

Asterisk* indicates residues that were also predicted based on the NHase model from *Pseudonocardia thermophila* (PDB ID: 1IRE). Mutations at positions not predicted by THEMATICS are listed in a separate column.

Table 2

Forward primers for site-directed mutagenesis[a]

α-Asp164Asn	CCCGCCAACAAGGAAATCCGCGTCTGG**AAC**ACCACGGCCGAATTG
α-Glu168Gln	GTCTGGGACACCACGGCC**CAA**TTGCGCTACATGGTGCTG
α-Arg170Gln	GTCTGGGACACCACGGCCGAATTG**CAG**TACATGGTGCTGCCGGAACGG
β-Glu56Gln	GAATTTCGGCATTCGATC**CAG**CGAATGGGCCCGGCCCAC
β-His71Leu	GCCCACTATCTGGAGGGAACCTACTACGAA**CTC**TGGCTTCATGTCTTTGAGAACCTGCTGGTC
β-His71Asn	GAGGGAACCTACTACGAA**AAC**TGGCTTCATGTCTTTGAG
β-His71Phe	GAGGGAACCTACTACGAA**TTC**TGGCTTCATGTCTTTGAG
β-His147Asn	CTCAACAAGAACCCGGTGGGC**AAT**ACCCGCATGCCGCGC
β-Tyr215Phe	CGCGTCGACTTGTGGGATGAC**TTC**CTGGAGCCAGAGTGA

[a]Mutated codons are in bold. Primers are written 5′ - 3′.

Table 3

Catalytic rate constants for the wild type and mutant ppNHases at pH 6.7[a]

	k_{cat} (min^{-1})	K_M (mM)	k_{cat}/K_M (mM^{-1} min^{-1})	$(k_{cat}/K_M)_{WildType}$ / $(k_{cat}/K_M)_{mutant}$	$\Delta\Delta G^b$ (kcal mol^{-1})
Wild Type	19.6 (0.4)	6.6 (0.9)	3.0	-	-
α-Asp164Asn	0.27 (0.02)	1.8 (0.6)	0.15	20	1.6
α-Glu168Gln	3.7 (0.2)	4.5 (1.3)	0.82	3.7	0.7
α-Arg170Gln (−)	13.0 (0.5)	8.2 (0.9)	1.6	1.9	0.3
α-Tyr171Phe (−)	13 (1)	9.3 (2.2)	1.4	2.1	0.5
β-Glu56Gln (std)	0.21(0.03)	15.3 (3.9)	0.013	230	3.0
β-His71Leu	1.2 (0.1)	10.0 (1.9)	0.12	25	1.8
β-His71Asn	1.5 (0.2)	20.1 (3.1)	0.074	41	2.0
β-His71Phe	1.6 (0.2)	21.2 (4.8)	0.073	41	2.0
β-His147Asn (std)	0.71 (0.01)	26.4 (3.2)	0.027	110	2.6
β-Tyr215Phe (−)	2.4 (0.2)	2.7 (0.9)	0.91	3.3	0.6

[a] Estimated standard deviations are shown in parentheses.

[b] $\Delta\Delta G = -RT\ln[(k_{cat}/K_M)_{mutant}/(k_{cat}/K_M)_{wild\ type}]$

(std) Indicates a residue predicted by THEMATICS with the standard cut-off

(−) Indicates a THEMATICS negative residue (control), not predicted even with reduced cut-off

All other residues were predicted by THEMATICS with a reduced cut-off

Biochemistry. Author manuscript; available in PMC 2012 August 18.

Table 4

Data collection and refinement statistics for wild type and mutant ppNHase proteins[a]

Data collection statistics	**ppNHase wild type**	**ppNHase α-Glu168Gln**	**ppNHase β-Glu56Gln**	**ppNHase β-His71Leu**	**ppNHase β-Tyr215Phe**
Beam line	APS, GM/CA-CAT, ID-B	APS, GM/CA-CAT, ID-D	APS, GM/CA-CAT, ID-D	APS, GM/CA-CAT, ID-B	APS, GM/CA-CAT, ID-D
Wavelength	0.95 Å	0.95 Å	0.95 Å	0.95 Å	0.95 Å
Space group	$P2_1$	$P2_1$	$P2_1$	$P2_1$	$P2_1$
Cell constants	*a* = 82.2 Å *b* = 137.3 Å *c* = 85.4 Å b = 92.3°	*a* = 82.5 Å *b* = 138.0 Å *c* = 85.3 Å b = 92.0°	*a* = 81.9 Å *b* = 137.5 Å *c* = 85.4 Å b = 92.4°	*a* = 82.0 Å *b* = 137.7 Å *c* = 85.5 Å b = 92.5°	*a* = 81.9 Å *b* = 137.2 Å *c* = 86.1 Å b = 91.8°
Total reflections	385818	363594	285777	694692	434764
Unique reflections	108015	65925	80874	125221	74195
Resolution limit (Å)	2.1 (2.1 – 2.18)*	2.5 (2.5 – 2.59)*	2.3 (2.3 – 2.38)*	2.0 (2.1 – 2.00)*	2.4 (2.4 – 2.49)*
Completeness	(%) 98.6 (93.0)	99.8 (98.7)	96.4 (87.0)	98.6 (92.0)	99.1 (96.4)
Redundancy	3.6 (2.7)	5.5 (4.5)	3.6 (3.2)	5.6 (4.2)	5.9 (4.8)
$I/\sigma I$	7.7 (1.5)	10.8 (2.0)	7.8 (1.8)	13.5 (2.0)	8.9 (2.0)
R_{merge} (%)	13.1 (49.8)	20.4 (65.8)	15.1 (56.2)	11.6 (54.3)	18.0 (64.3)
Refinement statistics					
Resolution range (Å)	37.6 - 2.1	45.4 - 2.5	45.4 - 2.3	45.4 - 2.0	29.8 - 2.4
R_{free} test set size	5392 (5%)	3342 (5%)	4042 (5%)	6286 (5%)	3723 (5%)
R_{cryst} (%)	17.6	20.8	19.6	16.8	18.5
R_{free} (%)	21.7	24.9	22.4	20.2	21.5
No. Atoms					
Total	14,259	13,490	14,164	14,390	13,588
Protein	13,064	13,016	13,084	13,004	13,059
Glycerol (GOL)	48	24	24	0	24
Cobalt (Co)	4	4	4	4	4
Water	1,143	446	1,052	1,382	501
B-factors					

Data collection statistics	ppNHase wild type	ppNHase α-Glu168Gln	ppNHase β-Glu56Gln	ppNHase β-His71Leu	ppNHase β-Tyr215Phe
Overall	23.6	27.2	24.2	20.5	28.2
R.m.s. deviations					
Bond lengths (Å)	0.010	0.0030	0.0060	0.0090	0.0070
Bond angles (°)	1.2	0.75	0.94	1.2	1.0

* Highest resolution shell is shown in parentheses.