8-2023

Mapping Roles of Active Site Residues in the Acceptor Site of the PA3944 Gcn5-Related N-Acetyltransferase Enzyme

Cillian Variot
San Francisco State University

Daniel Capule
San Francisco State University

Xhulio Arolli
Loyola University Chicago

Jackson Baumgartner
San Francisco State University

Cory T. Reidl
Loyola University Chicago, reidlchem@gmail.com

Author Manuscript
See next page for additional authors
This is a pre-publication author manuscript of the final, published article.

Recommended Citation

Variot, Cillian; Capule, Daniel; Arolli, Xhulio; Baumgartner, Jackson; Reidl, Cory T.; Houseman, Charles; Ballicora, Miguel; Becker, Daniel P. Ph.D.; and Kuhn, Misty L.. Mapping Roles of Active Site Residues in the Acceptor Site of the PA3944 Gcn5-Related N-Acetyltransferase Enzyme. Protein Science, 32, 8: 1-16, 2023. Retrieved from Loyola eCommons, Chemistry: Faculty Publications and Other Works, http://dx.doi.org/10.1002/pro.4725

Mapping roles of active site residues in the acceptor site of the PA3944 Gcn5-related *N*-acetyltransferase (GNAT) enzyme

Cillian Variot[1*], Daniel Capule[1*], Xhulio Arolli[2], Jackson Baumgartner[1], Cory Reidl[2], Charles Houseman[1], Miguel A. Ballicora[2], Daniel P. Becker[2&], Misty L. Kuhn[1&]

[1]Department of Chemistry and Biochemistry, San Francisco State University, San Francisco, CA 94132, USA
[2]Department of Chemistry and Biochemistry, Loyola University Chicago, Chicago, IL 60660, USA

[&]To whom correspondence may be addressed: Department of Chemistry and Biochemistry, San Francisco State University, 1600 Holloway Ave., San Francisco, CA 94132. Tel.: 415-405-2112; E-mail: mkuhn@sfsu.edu or Department of Chemistry and Biochemistry, Loyola University Chicago, 1032 W. Sheridan Rd., Chicago, IL 60660, Tel.: 773-508-3089; E-mail: dbecke3@luc.edu.

[*]These authors contributed equally

Abstract

An increased understanding of how the acceptor site in Gcn5-related *N*-acetyltransferase (GNAT) enzymes recognizes various substrates provides important clues for GNAT functional annotation and use as chemical tools. In this study, we explored how the PA3944 GNAT enzyme from *Pseudomonas aeruginosa* recognizes three different acceptor substrates, including aspartame, NANMO, and polymyxin B, and identified acceptor site residues that are critical for substrate specificity. To achieve this, we performed a series of molecular docking experiments and tested different methods to identify acceptor substrate binding modes that were potentially catalytically relevant. We found that traditional selection of "best" docking poses by lowest S scores did not reveal acceptor substrate binding modes that were generally close enough to the donor for productive acetylation. Instead, sorting poses based on distance between the acceptor amine nitrogen atom and donor carbonyl carbon atom placed these acceptor substrates near residues that contribute to substrate specificity and catalysis. To assess whether these residues were indeed contributors to substrate specificity, we mutated seven amino acid residues to alanine and determined their kinetic parameters. We identified several residues that improved the apparent affinity and catalytic efficiency of PA3944, especially for NANMO and/or polymyxin B. Additionally, one mutant exhibited substrate inhibition toward NANMO, and we proposed different scenarios for the cause of this inhibition based on additional substrate docking studies with this mutant. Ultimately, we propose that this residue is a key gatekeeper between the acceptor and donor sites and contributes to the orientation of and restricts acceptor substrate binding.

Keywords: Gcn5-related *N*-acetyltransferase, GNAT, polymyxin B, aspartame, molecular docking, acetylation, enzyme kinetics, substrate docking, docking visualization

Statement of importance/impact: This study pairs enzyme kinetics with molecular docking experiments to deepen knowledge about how diverse acceptor substrates are acetylated by GNAT enzymes so they can be utilized as chemical tools. We also describe a method to evaluate and visualize substrate docking results to identify catalytically relevant substrate conformations. The combination of these results and approaches provides new insight as to which residues can be modified to enhance enzyme activity of GNAT enzymes.

Introduction

One important family of enzymes that utilizes a distinctive protein scaffold to catalyze diverse acylation reactions is the Gcn5-related *N*-acetyltransferase (GNAT) superfamily. These enzymes transfer an acyl group from a donor molecule to an acceptor molecule, which binds within specific sites of the protein. A wide range of acceptor substrates are acylated by GNATs, including proteins, antibiotics, RNA, polyamines, and other metabolites [1], yet many protein members within this family remain uncharacterized. Major hurdles to studying GNAT functions include inconsistent conservation of these enzymes across diverse genomes and limited genomic annotations of functions. GNATs can also exhibit low sequence identities even between proteins with similar functions, which adds to the complexity of computational annotation of genomes. Additionally, small molecule substrates can sometimes be identified *in vitro,* but they are not necessarily reflective of enzyme physiological function *in vivo*. Structurally, GNATs adopt a common fold where the donor site has remained relatively conserved; however, the acceptor site can vary dramatically, which enables these enzymes to acylate quite diverse substrates. An additional level of complexity of GNATs is that many of them are also capable of modifying more than one substrate, and the promiscuous nature of some GNATs makes it more challenging to determine their native substrates *in vitro.*

A GNAT that exhibits substrate promiscuity and has become a model enzyme for exploring GNAT substrate specificity and mechanisms in our laboratory is the PA3944 enzyme from *Pseudomonas aeruginosa*. We previously showed that this enzyme acetylates multiple substrates, including polymyxins B and E (colistin) on the 3-Dab residue [2], a synthetically designed and prepared polymyxin substrate analog (*N*-(2-aminoethyl)-*N*-methyloctanamide, NANMO) [3], and the dipeptide sweetener aspartame [2,4]. Although some GNATs provide a link between antibiotic acetylation and antibiotic resistance, PA3944's ability to acetylate polymyxin B does not enable *P. aeruginosa* to survive in its presence, nor does polymyxin B appear to be a physiological substrate for the enzyme [2]. Regardless, we were able to use these substrates to elucidate a novel chemical and kinetic mechanism for this enzyme [3], wherein we showed that S148 was critical for catalysis and acts as an acyl intermediate during the hybrid ping-pong kinetic reaction. Additionally, we revealed that a specific residue in the acceptor site (E102) is important for substrate specificity toward polymyxin B and NANMO, despite the fact that E102 does not act as a general base in the reaction [3].

Herein, we expanded our study of the role of several PA3944 acceptor site residues and assessed their significance for substrate specificity toward polymyxin B, NANMO, and aspartame. Our experiments involved *in vitro* enzyme kinetic assays, which were used to evaluate the importance of seven acceptor site residues for discriminating between three different known substrates. Additionally, we used *in silico* molecular docking studies to observe the binding modes of these substrates in the acceptor site in the presence or absence of AcCoA/CoA or an acyl enzyme intermediate (AcS148). We also used an approach to visualize docking results of complex molecules in a simplified manner, and we designed a method for selecting poses of molecules that were near specific active site residues. These tools were also useful for easily assessing whether the docked molecules were located in the donor or acceptor sites of the protein. Our results provide a much clearer picture of the acceptor site of the PA3944 enzyme and expand our knowledge of the diverse architectures of GNAT acceptor sites, which is critical in enabling directed evolution of GNAT enzymes for tailored industrial or chemical purposes.

Results

1. Kinetic characterization of the WT PA3944 enzyme toward aspartame

We previously showed that the PA3944 enzyme acetylates polymyxin B/colistin, aspartame, and a synthetic substrate analog of polymyxin B called NANMO (*N*-(2-aminoethyl)-*N*-methyloctanamide hydrochloride) [2]. Given our unexpected results that the PA3944 enzyme uses a hybrid ping-pong mechanism for catalysis, we wanted to characterize the WT enzyme toward aspartame as well. Therefore, we performed enzyme kinetics to obtain substrate saturation curves of the WT PA3944 enzyme toward aspartame and compared its activity to our previous results with polymyxin B and NANMO. Overall, the WT enzyme exhibited a similar catalytic efficiency toward all three substrates (**Table 1**); its activity toward aspartame was slightly higher than polymyxin B and NANMO, but it had a ~1.5-fold higher apparent affinity toward NANMO (**Figure 1A, 1B**, **Table 1**). Similar to our previous results with polymyxin B and NANMO substrates, the S148A mutant was inactive toward aspartame (data not shown), indicating S148 is critical for PA3944 WT activity toward all three substrates. Since the dipeptide aspartame substrate is structurally distinct from NANMO and polymyxin B, we investigated whether the enzyme would utilize a similar kinetic mechanism when aspartame was the substrate. In theory, a kinetic mechanism should not be altered for different substrates, but we did observe some small discrepancies in our previous study [3]. For example, the best kinetic model that fit the data we previously collected when polymyxin B was the acceptor substrate was hybrid ping-pong, whereas we could not discriminate between ping-pong and hybrid ping-pong models due to their similar AICc and relative likelihood values when NANMO was the acceptor substrate. To determine the kinetic model that best fit the kinetic data when aspartame is the acceptor substrate, we generated a series of kinetic curves at varying concentrations of AcCoA and aspartame and fitted the data to a set of kinetic models as described previously [3]. Similar to our earlier data with polymyxin B, the kinetic model that best fit the data when aspartame was the acceptor substrate was a hybrid ping-pong kinetic mechanism (**Table 2, Figure SF1**). Therefore, the hybrid ping-pong model explains how the PA3944 enzyme can acetylate all three structurally distinct substrates using a combination of direct transfer and ping-pong mechanistic capabilities.

2. Molecular docking studies with the PA3944 WT enzyme and three acceptor substrates

To learn more about how the PA3944 enzyme recognizes these three diverse substrates and map key residues in the acceptor site that are important for substrate recognition, we performed a series of molecular docking studies described below. We used the crystal structure of the PA3944 enzyme in complex with CoA in the donor site (PDB ID: 6EDV) and docked acceptor substrate molecules into the acceptor site; the donor and acceptor sites of the protein are located in distinct pockets (**Figure 2A**). Since we know the enzyme acetylates substrates by direct transfer or via an acyl-enzyme intermediate, we docked the three fully protonated substrates aspartame, NANMO, and polymyxin B into this structure under three different conditions: 1) in the presence of AcCoA, 2) in the presence of CoA and AcS148 (AcS148 + CoA), and 3) in the absence of CoA and in the presence of AcS148 (AcS148 - CoA). These conditions represent different possible occupancies for the donor/product AcCoA/CoA during various steps of the kinetic mechanism and how AcCoA/CoA may influence acceptor substrate binding.

In this study, we were most interested in identifying docking poses of substrates that would reveal information about how they may bind during catalysis and which residues could be important for substrate recognition and specificity. While docked molecules were generally located in the acceptor site of the protein, we found that many of them were located in regions that were not amenable to a productive acetylation reaction, e.g. on the surface, away from the active site, or were located within the donor site. Therefore, we utilized the following approach to analyze the docking data. First, we sorted all of the docked substrate molecules in the PA3944 WT protein by lowest (best/most stable) S scores (**Table ST1)** and selected the top ten poses by lowest S score for further analysis. Next, we measured the distance between the acceptor substrate amine nitrogen atom and the acetyl donor (either AcCoA or AcS148) carbonyl carbon atom (see Materials and Methods for details) for these ten poses and compared their distances and S scores across all docking conditions (**Figure 3**). We found the distances between these atoms for each condition were widely distributed and placed the acceptor amine further from the acetyl donor regardless of docking condition (**Figure 3**). On average, we found NANMO docked closer to the acetyl donor than aspartame or polymyxin but most poses were still far from the donor (**Figure 3**). These results showed that focusing on poses with lowest S scores did not reveal binding modes of acceptor substrates that would help us identify residues important for substrate specificity and instead were more reflective of possible inhibitor binding modes.

To identify docking poses that were closer to the donor, we changed strategies and sorted all of the docked substrates by distance between the acceptor substrate amine nitrogen atom and the acetyl donor carbonyl carbon atom (**Table ST1**). We selected the top ten poses by distance between these two atoms for further analysis and found that on average the three substrates were located closer to the acetyl donor molecule than when we selected poses by lowest S score (**Figure 3** and **SF2**). Our results showed aspartame docked closer to the acetyl donor when AcCoA was present compared to the two AcS148 conditions but exhibited less favorable S scores. This indicates aspartame can bind in close enough proximity to AcCoA for a sequential transfer to occur. Additionally, the presence or absence of CoA in the AcS148 docking conditions influenced how close aspartame docked to the donor; when CoA was absent, aspartame docked closer than when CoA was present (**Figure 3**). This result follows what is generally expected for a ping-pong kinetic mechanism where CoA leaves prior to acceptor substrate binding. Thus, the collective docking results with aspartame mirror our kinetic data that show the PA3944 enzyme utilizes a hybrid ping-pong kinetic mechanism with sequential and ping-pong paths for transfer. Furthermore, sorting docking poses by distance between acceptor amine and donor yields more relevant binding modes of aspartame for acetyl transfer and thus can be used to identify residues that may be important for substrate specificity.

Next, we applied this same strategy of sorting docking poses by distance rather than S score to analyze the docking results for NANMO and polymyxin B. Compared to aspartame, a different trend in the proximity of NANMO to the donors was observed. We found that all three docking conditions placed the acceptor amine atom of NANMO within ~3-4 Å from the carbonyl carbon atom of all donors (**Figure 3**). This indicates the identity of the donor and/or presence of CoA product does not significantly influence the ability of NANMO to dock close to the donors and matches our previous kinetic data that showed we could not discriminate between a ping-pong or hybrid ping-pong mechanism for this substrate. When we analyzed the polymyxin B docking results, we found the distance between the polymyxin B 3-Dab amine nitrogen atom and the acetyl donor carbonyl carbon atom across all conditions was furthest compared to aspartame and

NANMO (**Figure 3**). While some polymyxin B docking poses in the AcCoA docking condition were closer to the donor, none of the poses across all conditions were close enough for acetyl transfer. These poses were, however, closer than when we sorted the data by lowest S score. Based on the collective results across all three substrates and docking conditions, we concluded that sorting docking data by distance rather than S score provided poses that placed the substrates in a more advantageous location for acetylation to occur. Furthermore, this strategy reveals more relevant amino acids that should be queried for their possible contributions to PA3944 substrate specificity.

3. Comparative analysis of locations of docked molecules sorted by distance using x, y, z coordinates of specific substrate atoms

Due to the difficulty of analyzing numerous docking poses of complex molecules like polymyxin B within a protein structure, we designed an approach to more easily and quickly visualize trends in docking poses using x, y, z, coordinates of specific atoms within the pdb files (**Figure SF3**). For substrates like polymyxin B that have multiple primary amines, we used a different colored-coded point in the plot for each Dab nitrogen atom. We also used different colored points to track the terminal carbon of the octanoyl tail of NANMO compared to its primary amine nitrogen to assess the position of the hydrocarbon tail in 3D space. Furthermore, we added a translucent plane to the plots to symbolize the threshold between the donor and acceptor sites of the PA3944 protein and assess where docked substrates bound. While other methods like interaction fingerprints have been used in the past [[5,6]], our approach allowed us to more easily monitor trends in locations of specific substituents on substrates as well as whether acceptor substrates docked fully into the acceptor site or traversed into the donor site. Our docking results for NANMO showed that the locations of the terminal carbon atom of the octanoyl hydrocarbon chain of NANMO remained in the acceptor site except when CoA was removed (AcS148 - CoA docking condition); the octanoyl tail of some NANMO molecules docks within the donor site in this condition. In contrast, for polymyxin B we observed that the locations of different Dab residues of polymyxin B were scattered throughout the acceptor site with some docking conditions producing poses with more ordering of specific Dab residues than others (**Figure SF3**). None of the docking conditions had polymyxin B in a consistent conformation, although none of the poses were located in the donor site. It appears polymyxin B binding is variable in the PA3944 acceptor site regardless of acetyl donor and can orient itself in multiple conformations within the acceptor site. In total, these results show that tracking specific substituents of complex molecules in a simplified way enables trends across numerous docking poses to be more easily identified.

4. Enzyme kinetics with PA3944 acceptor site mutant proteins

Since the polymyxin B substrate is much larger and more complex than NANMO and aspartame, it is possible that the protein requires a significant conformational change to bind polymyxin B in a proper position for acetyl transfer. Therefore, we built a tetrahedral intermediate of AcCoA and polymyxin B, docked it into the PA3944 WT protein, and performed molecular dynamics to help identify an appropriate binding mode for this substrate (**Figure 2C**). Next, we analyzed this

model to determine which acceptor site residues interact with the polymyxin B portion of the tetrahedral intermediate (**Figure 2C-E**). Based on this analysis, we generated the following mutants that were grouped based on three main areas of the acceptor site: near the active site (E102A, T141A), toward the entrance of the acceptor site (R59A, L169A), and other residues lining the internal region of the acceptor pocket (H167A, H179A, and R106A) (**Figure 2B, 2D**). We screened all mutants for activity toward all three substrates (polymyxin B, NANMO, and aspartame) and generated substrate saturation curves to compare kinetic parameters.

E102A and T141A—The E102 and T141 residues are located near the active site of the protein where the donor and acceptor sites meet. First, we examined the E102A mutant activity toward aspartame since our previous studies indicated that the E102A mutant was important for substrate binding and specificity rather than playing a catalytic role [[3]]. In this study, we found that the E102A mutant exhibited an improved apparent affinity for aspartame compared to WT by ~5-fold, and the catalytic efficiency toward aspartame was greater than polymyxin B but less than NANMO for the E102A mutant (**Figure 1C, 1D, Table 1**). Thus, NANMO was the preferred substrate for the E102A enzyme, and the apparent affinity of the enzyme toward all substrates was improved by mutating the E102 residue to a more hydrophobic residue. Next, we tested activity of the T141A mutant toward all three substrates and found that it displayed an ~8-fold decrease in catalytic efficiency compared to WT when aspartame was the substrate (**Figure 1E, 1F, Table 1**). The catalytic efficiencies of the T141A enzyme toward both polymyxin B and NANMO were comparable (~1 and ~2-fold decrease compared to WT, respectively). We concluded that residues near the active site of the protein that do not participate directly in the chemical mechanism can alter the substrate specificity and apparent affinity of the enzyme when mutated (E102A), but not all residues at this location (e.g. T141A) exhibit dramatic effects on activity or substrate specificity.

R59A and L169A—The R59 and L169 residues are located on opposite sides of the entrance to the acceptor site and are approximately 7 Å away from each other in the PDB ID 6EDV crystal structure. R59 is on an alpha helix (α2) and L169 is on a mobile loop between β8 and β9 strands (**Figure SF5**). When we assayed these mutant enzymes for activity, we saw several changes in kinetic parameters depending on the acceptor substrate. For example, the most dramatic alterations to catalytic efficiency or apparent affinity toward the three substrates was observed with polymyxin B. Both R59A and L169A mutants exhibited an increased catalytic efficiency (~2- and ~4-fold, respectively) and an improved apparent affinity (~2-fold and ~5-fold, respectively) compared to WT when polymyxin B was the substrate (**Figure 1G, 1H, Table 1**). Relatively minor changes in catalytic efficiency and apparent affinity were observed when NANMO or aspartame was the acceptor substrate (**Figure 1I-L, Table 1**). Therefore, replacing a charged residue with a hydrophobic residue (R59A) or reducing the size of hydrophobic residue in the entrance of the acceptor site (L169A) leads to an improved apparent affinity for a large substrate like polymyxin B, but not for the smaller substrates aspartame or NANMO.

H167A and H179A—The H167 and H179 residues line the internal surface of the acceptor site and are located between the T141 and L169 residues on one side of the acceptor pocket (**Figure 2B, 2D**). H167 is the first residue on the mobile loop between β8 and β9 while H179 is the second residue on β9 (**Figure SF5**). The H179A mutation improved the catalytic efficiency of the enzyme toward both polymyxin B and NANMO by ~3-fold and ~2-fold, respectively

compared to the WT protein, but decreased catalytic efficiency of the enzyme toward aspartame by ~2-fold (**Figure 1M-R**, **Table 1**). This effect on catalytic efficiency was mainly driven by the improved apparent affinity of the H179A protein for polymyxin B and NANMO compared to WT. The greatest improvement was observed toward polymyxin B (~7-fold increase). When the adjacent H167 residue was mutated to alanine, this increase in apparent affinity for polymyxin B nearly doubled compared to the H179A mutant (~14-fold increase compared to WT). The catalytic efficiency was also nearly doubled compared to the H179A (~6-fold higher than WT) for the H167A mutant. In contrast, when NANMO was the substrate, the apparent affinities of the H167A and H179A mutants were nearly identical and the catalytic efficiencies for the WT and H167A proteins were similar (**Figure 1O, 1P**, **Table 1**). On the other hand, when aspartame is the substrate, a different effect was observed. The catalytic efficiencies of the H179A and H167A mutants decreased ~2 and ~3-fold compared to WT, respectively, and the apparent affinity of H179A for aspartame was nearly identical to WT, whereas H167A showed a 2-fold decrease compared to WT (**Figure 1Q, 1R**, **Table 1**). Therefore, mutating either H167 or H179 to the small hydrophobic alanine residue significantly improves the apparent affinity of the enzymes for both polymyxin B and NANMO, with the greatest improvement toward polymyxin B.

R106A—The R106 residue is located on the opposite side of the acceptor pocket from the H167 and H179 residues and is further inside the pocket than R59 (**Figure 2B**, **2D**). Surprisingly, the R106A mutant exhibited a ~3-fold and ~5-fold increase in catalytic efficiency and an ~16-fold and ~13-fold increase in apparent affinity for polymyxin B and NANMO, respectively, compared to WT. However, it had a ~46-fold decrease in catalytic efficiency and ~5-fold decrease in apparent affinity toward aspartame (**Figure 1M-R**, **Figure 4**, **Table 1**). In fact, this R106A mutation exhibited the most significant decrease in catalytic efficiency and apparent affinity of all the mutants tested toward aspartame. Additionally, the R106A mutation displayed the highest increased apparent affinity toward both polymyxin B and NANMO than any other mutant. Therefore, removing the positive charge in this region of the acceptor site enhances catalytic efficiency and the apparent affinity of the enzyme toward more hydrophobic and complex molecules than the small peptide aspartame; this positive charge is critical for activity toward aspartame. Additionally, the R106A mutant appears to exhibit substrate inhibition at higher concentrations of NANMO (**Figure 4**). The K_i for NANMO was calculated as 0.332 +/- 0.066 mM (**Table 1**).

5. NANMO substrate inhibition and double docking in R106A

Since our kinetic studies showed the R106A enzyme exhibited an enhanced apparent affinity for NANMO but substrate inhibition at higher concentrations, we performed docking studies with the R106A enzyme to generate hypotheses as to how substrate inhibition could occur. We found that when R106 is mutated to alanine, the interior of the acceptor pocket increases in volume and hydrophobicity (**Figure 5A**). Additionally, the acceptor substrates did not dock identically into the AcSer148 - CoA condition in the R106A model compared to WT (**Figure SF2, SF3, SF5**). Often, the acceptor substrate was found within the donor site when the top ten poses by distance were examined. We hypothesized that the observed substrate inhibition was due to NANMO binding in the donor site. Therefore, we focused on a deeper examination of the AcS148 - CoA docking results with NANMO. For docking poses sorted by distance, NANMO docked in a more

restricted conformation in the WT structure compared to the R106A model where the conformation of NANMO varies and it can explore alternative conformations in this pocket (**Figure 5B** and **C**). However, NANMO is also pulled deeper into the donor site in the R106A model than for WT (**Figure 5B**). A comparison of docked NANMO molecules sorted by S score showed that NANMO docks into two different locations in the acceptor pocket for the WT protein, whereas all of the NANMO molecules dock almost exclusively in the donor site for the R106A model (**Figure 5B** and **C**). Based on these results, we propose three different scenarios for the cause of the substrate inhibition for the R106A enzyme. The first scenario is that the decreased activity may be due to competition between NANMO and AcCoA for binding to the donor site when R106 is mutated. The second scenario is that NANMO may bind close to the alanine residue once R106 is removed and be more difficult to release due to the change in hydrophobicity. The third scenario is that at higher concentrations of NANMO, more than one substrate molecule may bind to the acceptor site.

To determine whether a second NANMO could bind to the acceptor site of the R106A protein once another molecule of NANMO was already bound, we performed a second round of docking we termed a "double dock" (**Table ST2**). In this experiment, we selected the pose for NANMO that had the acceptor amine closest to the AcS148 residue from the top ten poses sorted by S score in the R106A AcS148 - CoA docking condition. This pose happened to be present in the AcCoA donor site. We then performed a second round of NANMO docking into this model and found additional molecules of NANMO could be docked into the acceptor site. When the "double-docked" poses of NANMO are sorted by distance and S score we saw the conformations of the poses when sorted by S score were elongated across one side of the acceptor pocket (**Figure 5C**). This more elongated conformation forms significant van der Waals interactions with the alanine residue of R106A and lies in the newly formed hydrophobic pocket. Ultimately, R106 may act as a gatekeeper of the conformation of NANMO in the absence of CoA and prevent possible substrate inhibition through binding multiple substrate molecules due to increased access to van der Waals interactions when R106 is mutated to alanine.

6. Interactions of docked substrates with PA3944 protein acceptor site residues as a mechanism for understanding mutant enzyme kinetics results

We next explored the interactions of docked substrates into the WT protein with acceptor site residues that we mutated and screened for kinetic activity (**Figure 6**, **Table ST3**). The goal was to shed light on how the enzyme uses these residues for substrate recognition and specificity and whether the residue interactions change based on presence of acetyl donor and/or CoA product.

Aspartame—We found that aspartame formed considerable interactions with a variety of acceptor site residues when it was docked into the AcS148 - CoA condition compared to the other two conditions (**Figure 6**). In the AcCoA condition, we observed mostly ionic interactions between R59 and the aspartate sidechain of the dipeptide aspartame. These interactions with R59 became more prominent when the acetyl donor was AcS148. Most of the interactions observed with aspartame and L169 were van der Waals interactions with the peptide backbone of aspartame and were consistent across all three conditions. Mutating R59 to alanine does not eliminate activity toward aspartame, but rather reduces turnover by approximately half (**Table 1,**

Figure 1K). The interactions with L169 may stabilize the aspartame molecule sufficiently for acetylation to occur when R59 is mutated, but L169 is not required for effective turnover or apparent affinity since the catalytic efficiency of L169A is nearly identical to that of WT when aspartame is the substrate. The interactions with R59 and L169 in the presence of AcCoA or CoA may help to orient the acceptor amine of aspartame toward the active site of the protein, but when the donor pocket is unoccupied, these interactions are not sufficient to maintain a proper orientation for acetyl transfer.

NANMO—In contrast to aspartame, NANMO formed significant interactions with a variety of acceptor site residues in all docking conditions. Additionally, with a vacant donor site, NANMO exhibited more frequent interactions with a select set of residues on one side of the acceptor site (**Figure 6**). In the AcCoA docking results, the orientation of the octanoyl tail of NANMO varies and forms van der Waals interactions with a variety of residues we mutated and screened for kinetic activity throughout the acceptor site. These included R59, E102, R106, T141, and H179. Very few interactions occurred with H167, and no interactions with L169 were observed. All of these residues formed van der Waals interactions with NANMO in the WT AcS148 + CoA docking condition as the molecule adopted a more lengthened structure within the acceptor pocket. However, when CoA is removed, NANMO is pushed toward one side of the acceptor pocket and forms interactions with R59, E102, R106, T141. We did not observe any significant change in kinetic activity toward NANMO when R59 or L169 residues were mutated to alanine. However, when H167 and H179 residues were mutated to alanine, we saw a small increase in catalytic efficiency, which was primarily due to an improved K_m for NANMO. This improvement may be due to increasing the hydrophobicity of this region of the acceptor pocket to bind NANMO.

Discussion

GNAT acceptor site tunability as a chemical tool—One of the most limiting aspects in predicting protein function from sequence and structure for GNAT enzymes is the lack of understanding of how acceptor substrates are recognized by and bind to the active site. Moreover, key residues required for substrate specificity are not well understood, especially in enzymes with unknown or promiscuous functions and due to varying sizes and amino acid residue compositions of acceptor sites. In this study, we utilized several approaches to learn about roles of specific residues within the acceptor site of the PA3944 enzyme for recognizing and acetylating three diverse substrates aspartame, NANMO, and polymyxin B. Our combinatorial approach reveals remarkable promiscuity of the PA3944 enzyme and residues that enable often significant improvements in apparent affinities and catalytic efficiencies toward these substrates, especially toward polymyxin B and NANMO. These results provide support for the tunability of GNATs for specific functions and highlight the promise of these enzymes as acylating reagents in the chemists' toolkit with broad, basal activity that can be manipulated through readily prepared mutants.

Molecular docking and coordinate map approaches to identify catalytically relevant substrate poses—Molecular docking is a powerful tool for predictive ligand binding to proteins. While it is typically utilized to identify energy minima that mimic enzyme-inhibitor complexes, its use as a tool to explore catalytically relevant substrate binding is not as widespread. The

standard approach is to identify the "best pose" in docking studies which relies on a calculated S score that is not necessarily reflective of reactive enzyme-substrate complexes. Therefore, we compared and contrasted docking S scores representing low energy poses with calculations of distance between reacting groups, specifically between a nucleophilic amine of acceptor substrates and the electrophilic carbonyl carbon in an acyl donor in the PA3944 enzyme. Collectively, our analysis of the two sorting methods across all docking conditions showed that poses sorted based on distance between the acceptor amine nitrogen atom and donor carbonyl carbon atom were closer to the active site but generally had less favorable S scores compared to poses sorted by S score. Additionally, when the poses were sorted by S score instead of distance, the location of the acceptor amine nitrogen atoms in the various poses were more variable for all three substrates in all three docking conditions. These results show that when the substrates dock closer to the active site of the protein where catalysis occurs, the S score can often be less favorable. Thus, selecting docking poses based on distance provides potentially catalytically relevant substrate poses that are less reliant on stable occupancy typical of inhibitor binding. Our results add additional insight to prior studies that have called into question the selection criteria that are used for identifying the "best pose" in molecular docking experiments [7,8], predictive substrate identification for proteins of unknown function [9] and a greater understanding substrate specificity [10] with docking approaches. Furthermore, our coordinate mapping approach to help track locations of key moieties of complex molecules simplifies data analysis of docking data and provides a more global perspective of possible conformations of substrates in the acceptor sites of GNATs.

Relationship of docking data to substrate specificity and architecture of acceptor site—Our enzyme kinetics and molecular docking data provide new insight into key acceptor site residues that are critical for substrate specificity and binding and validate our previous kinetic mechanistic data. For example, we found that the specific docking condition greatly impacts the binding interactions between the enzyme and substrate for both aspartame and NANMO (**Figure 6**). Moreover, performing docking studies under three possible conditions (AcCoA, AcS148 + CoA, and AcS148 - CoA) show that for some substrates the presence of the donor influences the binding location of the acceptor substrate and highlights possible substrate binding modes for the sequential or ping-pong paths of the hybrid kinetic mechanism. In some cases, removal of CoA enabled the acceptor amine nitrogen to dock closer to the acetyl donor than when CoA was present, and for some poses portions of the acceptor molecule protrude into the donor site when CoA is absent. It is possible that the ping-pong path of acetyl transfer where CoA leaves prior to acceptor substrate binding may itself provide a strategy to enable greater substrate promiscuity through acceptor substrate binding that extends into the donor site. On the other hand, residues in a similar location as R106 within GNAT acceptor sites likely contribute to the orientation of and restrict acceptor substrate binding within this site to prevent substrate inhibition and enhance turnover.

Materials and Methods

Materials—Acetyl coenzyme A (AcCoA) trilithium salt (Cat#A2181), polymyxin B sulfate salt (Cat#P4932), and aspartame (*N*-(L-α-aspartyl)-L-phenylalanine methyl ester; Cat#A5139) were purchased from Millipore Sigma. *N*-(2-Aminoethyl)-*N*-methyloctanamide hydrochloride (NANMO) was synthesized as described previously [3]. Substrates were dissolved in nanopure

water to create 100 mM stocks for polymyxin B and NANMO and a 25 mM stock for aspartame for kinetic experiments.

Clones and site directed mutagenesis—The plasmid containing the *Pseudomonas aeruginosa* PAO1 *pa3944* gene (UniProt ID: Q9HX72) previously sub-cloned into the ampicillin resistant p11 vector [2] was used as a template for constructing point mutations of the PA3944 enzyme. The T141A mutant was a generous gift from Karolina Majorek and Wladek Minor at the University of Virginia Charlottesville, and the S148A and E102A mutants were the same as those previously reported [3]. The remaining mutants (R59A, R106A, N145A, H167A, L169A, and H179A) were constructed using the Phusion Site-Directed Mutagenesis Kit from Thermo Fisher Scientific and phosphorylated primers at the 5' position from Integrated DNA Technologies (IDT) following the manufacturer's instructions.

Protein expression and purification—All WT and mutant proteins were expressed and purified as described previously [2]. All proteins were purified to near homogeneity as assessed by SDS-PAGE, and protein concentrations were determined using Beer's Law based on the absorbance at $A_{280\ nm}$ and an extinction coefficient of 37,470 M^{-1} cm^{-1}. The N-terminal polyhistidine tag was retained on all proteins.

Steady-state enzyme kinetics—All enzyme kinetics were performed using a colorimetric assay as previously described [2,4]. Reactions (50 μL total) for substrate saturation curves contained varying concentrations of acceptor substrates ranging from 0-6 mM polymyxin B, 0-4 mM NANMO, and 0-10 mM aspartame at a constant concentration of AcCoA (0.5 mM) in 100 mM Tris-HCl pH 8.0 and 100 mM NaCl. The enzymatic reactions were initiated with 10 μL of enzyme, allowed to proceed for 10 min at 37°C, and then were terminated and reacted with DTNB as described [2,4]. Enzyme concentrations in each reaction were 1.3 μM of WT toward all three substrates; 0.86 μM, 1.1 μM, and 2.6 μM for R59A; 4.4 μM, 1.8 μM, and 4.4 μM for E102A; 1.5 μM, 1.9 μM, and 9.2 μM for R106A; 0.90 μM, 1.8 μM, and 4.5 μM for T141A; 0.64 μM, 1.3 μM, and 0.4 μM for H167A; 1.3 μM, 1.4 μM and 1.2 μM for L169A; and 1.3 μM, 1.4 μM and 1.9 μM for H179A toward polymyxin B, NANMO, and aspartame, respectively. All data were fitted to the Michaelis-Menten equation using non-linear regression in Origin 2017 with the exception of data for R106A toward NANMO, which was fitted to the Bi-Hill equation. Details are reported in **Table 1**. Values for k_{cat} were calculated using the protein molecular weight of 21.9 kDa.

Kinetic mechanism—A series of enzyme kinetic assays with the WT enzyme were performed as described above as well as previously [3] with the following modifications. The concentration of aspartame was varied from 0-10 mM at four different concentrations of AcCoA (0.1, 0.25, 0.5, and 1 mM) and 1.3 μM WT enzyme. Data were fitted to several models (random bisubstrate, ordered AB bisubstrate, ordered BA bisubstrate, ping-pong bisubstrate, and hybrid ping-pong bisubstrate steady state) as described before [3].

Protein and small molecule preparation for molecular docking studies—The Molecular Operating Environment (MOE) [11] and the X-ray crystal structure of the WT PA3944 protein (PDB ID: 6EDV) were used for docking studies. The structure was saturated with requisite hydrogens and protonated at pH 7.4 and 310K using a previously described protocol [12]. The

enzyme was then energy minimized in the gas phase using an Amber14:EHT force field with crystallographic water molecules set as non-rigid. A similar approach was used for the R106A mutant protein, in which the MOE Sketch feature with ChemDraw was used to manually mutate the R106 residue to alanine and then energy minimized. To create the AcCoA ligand in the energy minimized WT and R106A protein models, an acetyl group was added to the sulfur atom of CoA using the MOE Builder Utility, followed by energy minimization. A similar procedure was used to create the AcS148 version of the enzymes in the presence and absence of the CoA ligand. Compounds were protonated at pH 7.4 and 310K followed by energy minimization in the gas phase with the MMFF94x forcefield [13]. The different energy minimized, liganded WT and R106A structures (non-AcS148 with AcCoA, AcS148 with CoA, AcS148 without CoA) were used to dock the individual, fully protonated substrates into the acceptor sites using automatic and manual pocket detection strategies and employing the AMBER14:EHT forcefield. If CoA or AcCoA was present in the donor site, automatic detection of the acceptor site was used by MOE. Alternatively, if these ligands were absent, we manually defined the acceptor site as the Pocket site by populating it with inactive dummy atoms with the MOE utility Site Finder (**Figure SF2**). To dock a second molecule of NANMO into the R106A protein model acceptor site in the absence of CoA, we used the structure corresponding to the lowest S score pose from the single NANMO substrate docking result. Ligand placement for the docking protocol utilized the Alpha Triangle method [7] with Affinity dG scoring to generate 1000 data points per unique ligand. These were further refined using the Induced Fit method with GBVI/WSA [14] dG scoring to obtain docked poses, and were further analyzed. An S score is computed by MOE to indicate the GBVI/WSA free energy of the pose.

Generation of an AcCoA-polymyxin B tetrahedral intermediate—MOE was used to generate a hypothetical tetrahedral intermediate between the 3-Dab of polymyxin B and AcCoA. Starting with the tetrahedral carbon atom formed from polymyxin B and AcCoA, the remainder of the polymyxin B molecule was manually built into the acceptor site using the MOE Builder utility. All atoms in the system were optimized with a short, localized molecular minimization process in the gas phase with atoms further than 8 Å from the tetrahedral intermediate held in a fixed position. System refinement continued until an RMS Gradient of 0.1 kcal/mol/Å was obtained. The tetrahedral intermediate PA3944 model was resolvated in a simple water box at pH 7.4 and treated with NaCl counterions to balance the charge. Periodic boundary conditions were enabled, and the hydrogen bonding network of the model was optimized by automatically sampling different tautomer/protomer states using Protonate3D [15]. The model was then subjected to a local minimization in the liquid phase followed by a 1 ns MD equilibration run using an NPA algorithm with an Amber12:EHT force field. MD experiments utilized an initial heating phase from 0 to 300 K over 100 ps followed by equilibration for 100 ps at 300 K, 700 ps production run, and finally a 100 ps cooling from 300 to 0 K. Simulation results were minimized once more before final binding poses for the bisubstrate tetrahedral intermediate model were obtained for analysis.

Selection criteria for selecting and sorting docking poses for further analysis—We used the following strategy to select docked poses that were located adjacent to the acyl donor. To begin, we selected the top ten docking poses of NANMO and aspartame substrates for further analysis based on the lowest S scores. These poses were then visually inspected to determine which poses

had the free amine oriented toward the acyl donor. Since polymyxin B has multiple primary amines, we did not exclude any poses from the analysis. Next, we adapted the selection strategy to decrease bias based on appearance by measuring distances between the substrate primary amine nitrogen and the carbonyl carbon of the acyl donor for all poses. The location of these atoms was determined from the x, y, z coordinates in pdb files generated from the docking output database. We used the diaminobutyric acid primary amine nitrogens (1, 3, 5, 8, 9-Dab) of polymyxin B, the primary amine nitrogen of NANMO, and the primary amine nitrogen of aspartame. For each docking pose, the distance between the primary amine nitrogen of the acceptor substrate and the carbonyl carbon of the acetyl donor was determined using the equation: $\text{distance} = \sqrt{(x_N - x_C)^2 + (y_N - y_C)^2 + (z_N - z_C)^2}$ where (x_N, y_N, z_N) are the coordinates of the primary amine nitrogen atom and (x_C, y_C, z_C) are the coordinates of the carbonyl carbon atom of the donor in Å. These poses were then ranked by distance and the top ten poses were plotted as box plots in Origin with the following configurations: Box 25th to 75th percentile, data points centered, whiskers 5th percentile lower boundary, 95th percentile upper boundary.

Supplementary material

Five supplemental figures and three supplemental tables (Excel files with raw data) are included.
Supplemental Figures SF1-5 (pdf)
Supplemental Table ST1 (Excel sheet)
Supplemental Table ST2 (Excel sheet)
Supplemental Table ST3 (Excel sheet)

Acknowledgements

Federal funds from the National Science Foundation (CHE-1708863 to MLK and CHE-1708927 to DPB) were used to perform scientific studies presented in this manuscript. We thank Karolina Majorek and Wladek Minor at the University of Virginia Charlottesville for the PA3944 T141A mutant plasmid, David Tran, Melissa Law, and Layton Joe at San Francisco State University and Sarah Witzke at the Chemical Computing Group for their technical assistance.

References

1. Burckhardt RM, Escalante-Semerena JC (2020) Small-Molecule Acetylation by GCN5-Related N-Acetyltransferases in Bacteria. Microbiol. Mol. Biol. Rev. MMBR 84.

2. Czub MP, Zhang B, Chiarelli MP, Majorek KA, Joe L, Porebski PJ, Revilla A, Wu W, Becker DP, Minor W, et al. (2018) A Gcn5-Related N-Acetyltransferase (GNAT) Capable of Acetylating Polymyxin B and Colistin Antibiotics in Vitro. Biochemistry 57:7011–7020.

3. Baumgartner JT, Habeeb Mohammad TS, Czub MP, Majorek K, Arolli X, Variot C, Anonick M, Minor W, Ballicora MA, Becker DP, et al. (2021) Gcn5-related N-acetyltransferases (GNATs) with a catalytic serine residue can play ping-pong too. Front. Mol. Biosci. [Internet] 8. Available from: https://www.frontiersin.org/articles/10.3389/fmolb.2021.646046/abstract

4. Kuhn ML, Majorek KA, Minor W, Anderson WF (2013) Broad-substrate screen as a tool to identify substrates for bacterial Gcn5-related N-acetyltransferases with unknown substrate specificity. Protein Sci. Publ. Protein Soc. 22:222–30.

5. Deng Z, Chuaqui C, Singh J (2004) Structural interaction fingerprint (SIFt): a novel method for analyzing three-dimensional protein-ligand binding interactions. J. Med. Chem. 47:337–344.

6. Da C, Kireev D (2014) Structural Protein–Ligand Interaction Fingerprints (SPLIF) for Structure-Based Virtual Screening: Method and Benchmark Study. J. Chem. Inf. Model. 54:2555–2561.

7. Udatha DBRKG, Sugaya N, Olsson L, Panagiotou G (2012) How well do the substrates KISS the enzyme? Molecular docking program selection for feruloyl esterases. Sci. Rep. 2:323.

8. Ramírez D, Caballero J (2018) Is It Reliable to Take the Molecular Docking Top Scoring Position as the Best Solution without Considering Available Structural Data? Mol. J. Synth. Chem. Nat. Prod. Chem. 23:1038.

9. Johannes C. Hermann ‡, Eman Ghanem †, Yingchun Li †, Frank M. Raushel †, John J. Irwin *, Brian K. Shoichet* ‡ (2006) Predicting Substrates by Docking High-Energy Intermediates to Enzyme Structures. ACS Publ. [Internet]. Available from: https://pubs.acs.org/doi/pdf/10.1021/ja065860f

10. Juhl PB, Trodler P, Tyagi S, Pleiss J (2009) Modelling substrate specificity and enantioselectivity for lipases and esterases by substrate-imprinted docking. BMC Struct. Biol. 9:39.

11. Anon Molecular Operating Environment (MOE) (2013) Chemical Computing Group Inc., Montreal.

12. Heath TK, Lutz MR, Reidl CT, Guzman ER, Herbert CA, Nocek BP, Holz RC, Olsen KW, Ballicora MA, Becker DP (2018) Practical spectrophotometric assay for the dapE-encoded N-succinyl-L,L-diaminopimelic acid desuccinylase, a potential antibiotic target. PloS One 13:e0196010.

13. Halgren TA (1996) Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. J. Comput. Chem. 17:490–519.

14. Labute P (2008) The generalized Born/volume integral implicit solvent model: estimation of the free energy of hydration using London dispersion instead of atomic surface area. J. Comput. Chem. 29:1693–1698.

15. Shoichet BK (2004) Virtual screening of chemical libraries. Nature 432:862–865.

Table 1. PA3944 wild-type and mutant enzyme kinetic parameters toward polymyxin B, NANMO, and aspartame.

Enzyme	Polymyxin B			NANMO			Aspartame		
	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m ($M^{-1}s^{-1}$)	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m ($M^{-1}s^{-1}$)	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m ($M^{-1}s^{-1}$)
WT*	1.68 ± 0.07	0.51	$3.0x10^2$	1.07 ± 0.03	0.41	$3.8x10^2$	1.68 ± 0.07	0.61	$3.6x10^2$
R59A	0.708 ± 0.058	0.44	$6.2x10^2$	1.34 ± 0.09	0.39	$2.9x10^2$	2.00 ± 0.15	0.32	$1.5x10^2$
E102A*	1.05 ± 0.05	0.06	$5.7x10^1$	0.101 ± 0.013	0.15	$1.5x10^3$	0.346 ± 0.050	0.09	$2.6x10^2$
R106A[#]	0.107 ± 0.007	0.11	$1.0x10^3$	0.083 ± 0.007	0.16	$1.9x10^3$	7.74 ± 0.53	0.06	$7.8x10^0$
T141A	2.12 ± 0.18	0.53	$2.5x10^2$	1.40 ± 0.09	0.32	$2.3x10^2$	4.17 ± 0.20	0.19	$4.3x10^1$
H167A	0.122 ± 0.010	0.21	$1.7x10^3$	0.323 ± 0.027	0.14	$4.3x10^2$	3.38 ± 0.23	0.48	$1.4x10^2$
L169A	0.332 ± 0.039	0.42	$1.3x10^3$	0.871 ± 0.047	0.37	$4.3x10^2$	1.83 ± 0.12	0.65	$3.5x10^2$
H179A	0.250 ± 0.048	0.20	$8.0x10^2$	0.329 ± 0.035	0.27	$8.2x10^2$	1.64 ± 0.10	0.34	$1.9x10^2$

*Data for WT and E102A toward polymyxin B and NANMO were previously reported [3].

[#]Significant substrate inhibition of the R106A mutant toward NANMO was observed and data were fitted to the biphasic Hill (Bi-Hill) equation $y = \frac{P_m}{\left[1+\left(\frac{K_a}{x}\right)^{H_a}\right]\left[1+\left(\frac{x}{K_i}\right)^{H_i}\right]}$ where y is velocity, x is concentration of substrate, P_m is the maximal velocity, K_a is the concentration of substrate at half the maximal activity for activation (reported as K_m in the table), K_i is the concentration of substrate at half the maximal activity for inhibition, H_a is the Hill coefficient for activation, and H_i is the Hill coefficient for inhibition. K_i was calculated as 0.332 ± 0.066 mM, H_a as 2.02 ± 0.33, and H_i as 0.47 ± 0.06 from the fitting.

Table 2. Results of fitting selected kinetic models to a series of PA3944 WT enzyme kinetic curves toward aspartame. Enzyme activity data were collected by reacting WT enzyme with varying concentration of aspartame from 0-7.5 mM and different concentrations of AcCoA (0.1, 0.25, 0.5, and 1 mM). Models used for fitting were described previously [3]. The hybrid ping-pong model is bolded because it is the one with the lowest corrected Akaike Information Criterion (AICc) and ΔAICc values, and the highest relative likelihood, which was calculated using the equation $e^{-0.5(\Delta AIC_c)}$.

Acceptor Substrate	Model	AICc	ΔAICc	Relative likelihood
Aspartame	Random	-346	103	4.30×10^{-23}
	Ordered AB	-278	171	7.38×10^{-38}
	Ordered BA	-305	144	5.38×10^{-32}
	Ping-pong	-433	16	3.35×10^{-4}
	Hybrid	**-449**	**0**	**1**

Figure Legends

Figure 1. Substrate saturation curves for PA3944 WT and mutant proteins toward aspartame, NANMO, and polymyxin B. The average activity for each protein as a function of substrate concentration (polymyxin B or *N*-(2-aminoethyl)-*N*-methyloctanamide (NANMO) or aspartame) are shown in panels A, C, E, G, I, K, M, O, and Q; the standard deviation reflects at least two biological replicates in duplicate. Normalized activity is shown in panels B, D, F, H, J, L, N, P, and R. Panels A-F show all aspartame data with triangles, NANMO with circles, and polymyxin B with squares WT data are indicated with filled shapes and mutant data are with unfilled shapes. Panels G-L have WT data in black squares, R59A in red circles, and L169A in blue triangles. Panels M-R have WT data in black squares, H167A in green triangles, H179A in orange inverted triangles, and R106A in purple circles. Substrate inhibition of the R106A mutant toward NANMO is shown and data were fitted to the Bi-Hill equation using non-linear regression in Origin 2017. All other curves were fitted to the Michaelis-Menten equation in the same program. See Materials and Methods for more details about the enzyme assays. The WT and E102A data toward polymyxin B and NANMO were previously reported [3].

Figure 2. Location of donor and acceptor sites of the PA3944 WT enzyme and modeled tetrahedral AcCoA-polymyxin B intermediate. A) WT PA3944 crystal structure with CoA occupying the donor site (PDB ID: 6EDV). Donor and acceptor sites and CoA are labeled. B) Simplified diagram showing the general location of selected amino acids in acceptor site of PA3944. Amino acids are represented as circles using the color scheme of L169 in blue, H167 in green, H179 in orange, T141 in pink, E102 in yellow, S148 in black, R106 in purple, and R59 in red. C) PA3944 crystal structure with the modeled tetrahedral AcCoA-polymyxin B intermediate (cyan sticks) spanning both donor and acceptor sites. D) Key acceptor site residues surrounding the modeled AcCoA-polymyxin B tetrahedral intermediate. All residues are colored as described in panel B and the tetrahedral intermediate is colored in cyan. E) Interactions between residues of the acceptor site and the modeled AcCoA-polymyxin B tetrahedral intermediate. Atoms of the tetrahedral intermediate that form H-bonds with residues lining the acceptor site are indicated with 4-digit numbers in purple. Residues of the cyclic peptide polymyxin B are labeled in red. H-bonding interactions between the tetrahedral intermediate and acceptor site residues are indicated with black arrows. Amino acids selected for mutation and further study are shown in rainbow colors as described in panel B and also denoted with stars.

Figure 3. Box plots of substrate docking and selection of top ten poses by either closest distance or lowest S score in the PA3944 WT and R106A enzymes. Three substrates (aspartame, NANMO, and polymyxin B) were docked into the WT crystal structure (PDB ID: 6EDV) and R106A mutant generated in MOE based on the WT structure. The top ten poses by closest distance between the acceptor amine and the donor carbonyl carbon and by lowest S score were selected for further analysis. The data for polymyxin B were sorted based on the distance between the 3-Dab amine and carbonyl carbon of the acetyl donor. The distances between the acceptor amines and donor carbonyl carbon and their corresponding S scores were plotted as a function of the presence or absence of AcCoA/CoA and non-acetylated/acetylated S148 (AcS148) during docking experiments. The box plots are colored by ligand with solid fill for the WT enzyme and no fill for the R106A enzyme. S scores are shown with a diagonal pattern in the box. Data for aspartame are in pink, NANMO are in teal, and polymyxin B are in purple. Distance and S score data were automatically binned in Origin. Data points located on a

horizontal line belong to the same bin. Boxes show the span of quartile 1 (lower bound of box) to quartile 3 (upper bound of box). Medians are shown as a solid line across a box and whiskers are drawn from the top of the box to the highest point and the bottom of the box to the lowest point.

Figure 4. Normalized activity of PA3944 WT and R106A toward aspartame, NANMO, and polymyxin B. A) WT activity toward polymyxin B is shown with filled black squares and R106A in open black squares. B) WT activity toward NANMO is shown with filled purple circles and R106A in open purple circles. C) WT activity toward aspartame is shown with filled magenta triangles and R106A in open magenta triangles.

Figure 5. Comparison of WT and R106A acceptor site hydrophobicity and location of docked NANMO molecules. A) Surface representations of the hydrophobicity of the acceptor site with hydrophilic regions colored in blue and hydrophobic regions colored in orange. AcCoA is shown in white sticks to indicate the location of the donor site of the PA3944 protein. B) Comparison of top ten NANMO dockings in AcS148 - CoA condition when sorted by distance between its primary amine and the carbonyl carbon of AcS148 and S score. NANMO is shown in teal sticks. A black and white dashed circle indicates the location of the R106 position for the WT and R106A proteins. C) Comparison of docked NANMO locations for top ten docking poses by distance between NANMO primary amine and carbonyl carbon of AcS148 - CoA and S scores for PA3944 WT, R106A, and doubly docked R106A proteins (ribbon representations). A second round of NANMO docking experiments were performed using the R106A enzyme containing a single molecule of NANMO that was selected based on the lowest S score and closest distance to AcS148 in the AcS148 - CoA condition. AcCoA is shown as white sticks to indicate the location of the donor site.

Figure 6. Comparison of catalytic efficiency and K_m for WT and mutant PA3944 enzymes toward aspartame, NANMO, and polymyxin B. A) Fold increase or decrease in catalytic efficiency (CE) compared to WT for each substrate rounded to the nearest whole number. B) Fold increase or decrease in K_m compared to WT for each substrate rounded to the nearest whole number. Red numbers indicate fold decreases and black numbers are fold increases. The shapes for each substrate (triangle, circle, square) are consistent with the substrate saturation curves in **Figure 1**. C) Cartoon representation of the total number of interactions (e.g. van der Waals, H-bonding, and ionic interactions) identified between PA3944 residues and docked aspartame or NANMO using Chimera. The thickness of lines between residues and substrate were based on the total number of combined interactions identified from analyzing the top ten poses sorted by distance. Line thicknesses range between 1.5-16.5 pt where the thickness increases 1 pt for every 10 interactions identified. No line is shown if interactions were not present between the docked molecule and specific residue. Interactions with S148 are shown protruding toward the donor pocket.

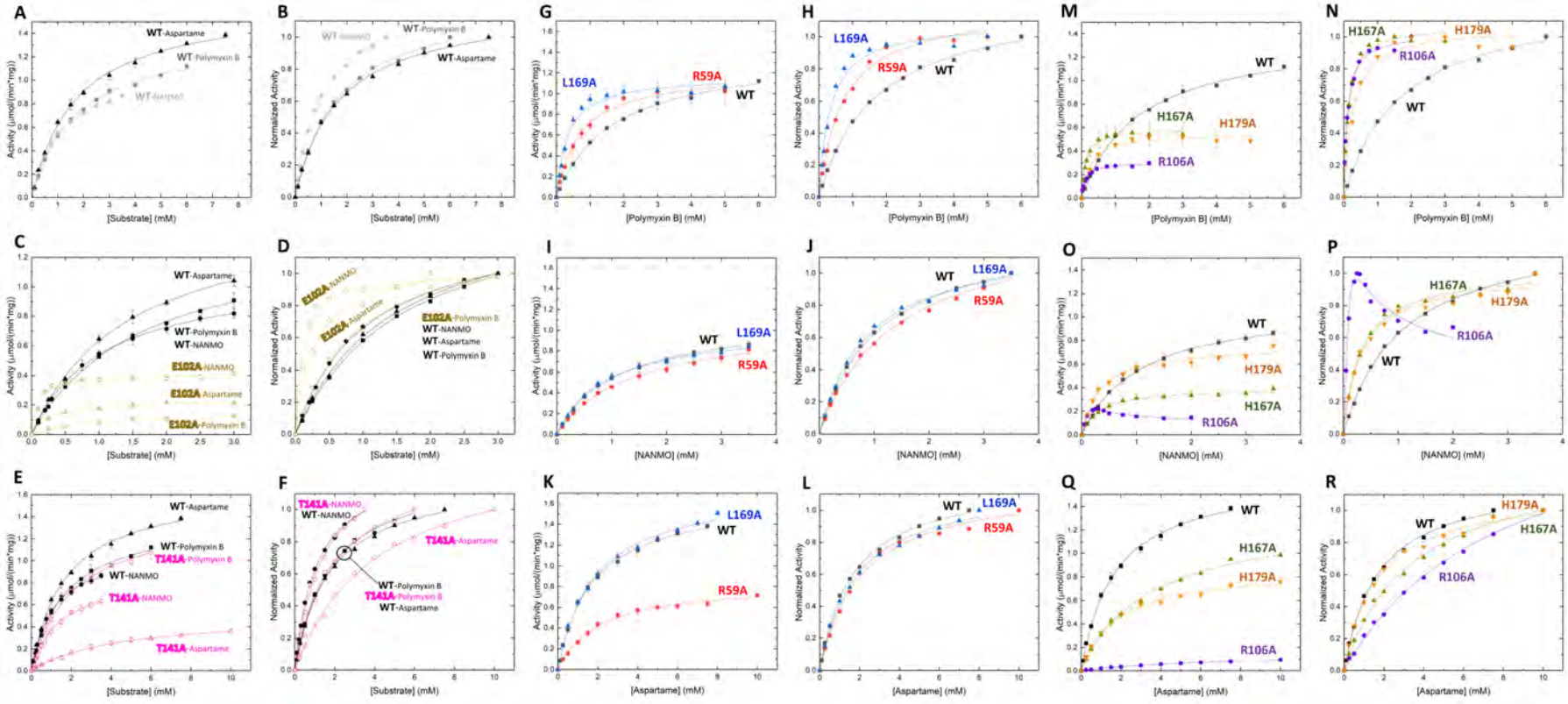
A
WT-Aspartame
WT-Polymyxin B
WT-NANMO
Activity (μmol/(min*mg))
[Substrate] (mM)
B
WT-Polymyxin B
WT-Aspartame
Normalized Activity
[Substrate] (mM)
G
L169A
R59A
WT
Activity (μmol/(min*mg))
[Polymyxin B] (mM)
H
L169A
R59A
WT
Normalized Activity
[Polymyxin B] (mM)
M
WT
H167A
H179A
R106A
Activity (μmol/(min*mg))
[Polymyxin B] (mM)
N
H167A
H179A
R106A
WT
Normalized Activity
[Polymyxin B] (mM)
C
WT-Aspartame
WT-Polymyxin B
WT-NANMO
E102A-NANMO
E102A-Aspartame
E102A-Polymyxin B
Activity (μmol/(min*mg))
[Substrate] (mM)
D
E102A-NANMO
E102A-Aspartame
E102A-Polymyxin B
WT-NANMO
WT-Aspartame
WT-Polymyxin B
Normalized Activity
[Substrate] (mM)
I
WT
L169A
R59A
Activity (μmol/(min*mg))
[NANMO] (mM)
J
WT
L169A
R59A
Normalized Activity
[NANMO] (mM)
O
WT
H179A
H167A
R106A
Activity (μmol/(min*mg))
[NANMO] (mM)
P
H167A
H179A
R106A
WT
Normalized Activity
[NANMO] (mM)
E
WT-Aspartame
WT-Polymyxin B
T141A-Polymyxin B
WT-NANMO
T141A-NANMO
T141A-Aspartame
Activity (μmol/(min*mg))
[Substrate] (mM)
F
T141A-NANMO
WT-NANMO
T141A-Aspartame
WT-Polymyxin B
T141A-Polymyxin B
WT-Aspartame
Normalized Activity
[Substrate] (mM)
K
L169A
WT
R59A
Activity (μmol/(min*mg))
[Aspartame] (mM)
L
WT
L169A
R59A
Normalized Activity
[Aspartame] (mM)
Q
WT
H167A
H179A
R106A
Activity (μmol/(min*mg))
[Aspartame] (mM)
R
H179A
WT
H167A
R106A
Normalized Activity
[Aspartame] (mM)

Figure 1

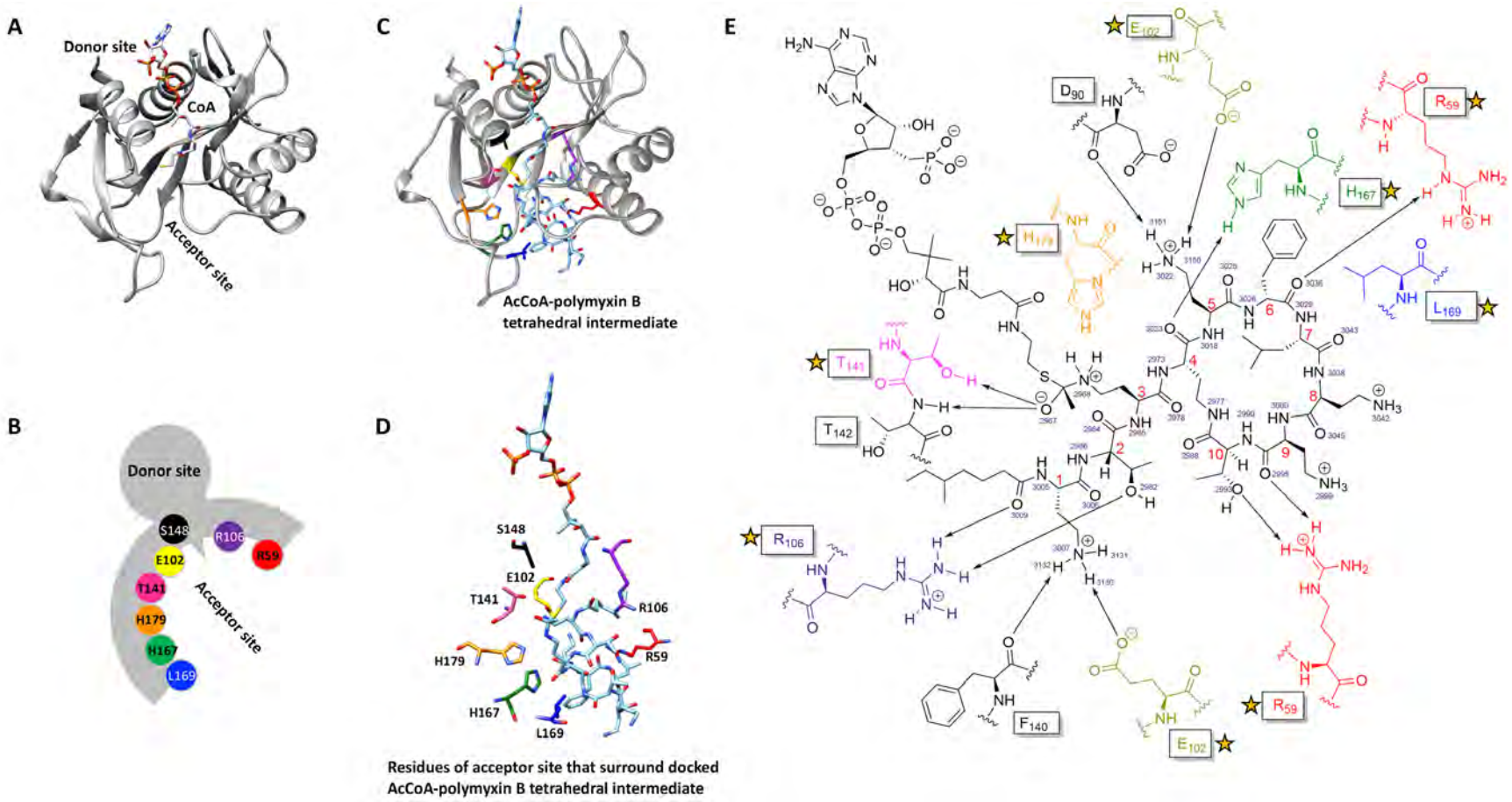
A
Donor site
CoA
Acceptor site
C
AcCoA-polymyxin B
tetrahedral intermediate
B
Donor site
S148
R106
E102
R59
T141
H179
H167
L169
Acceptor site
D
S148
E102
T141
R106
H179
R59
H167
L169
Residues of acceptor site that surround docked
AcCoA-polymyxin B tetrahedral intermediate
E
E102
D90
R59
H167
H179
L169
T141
T142
R106
F140
E102
R59

Figure 2

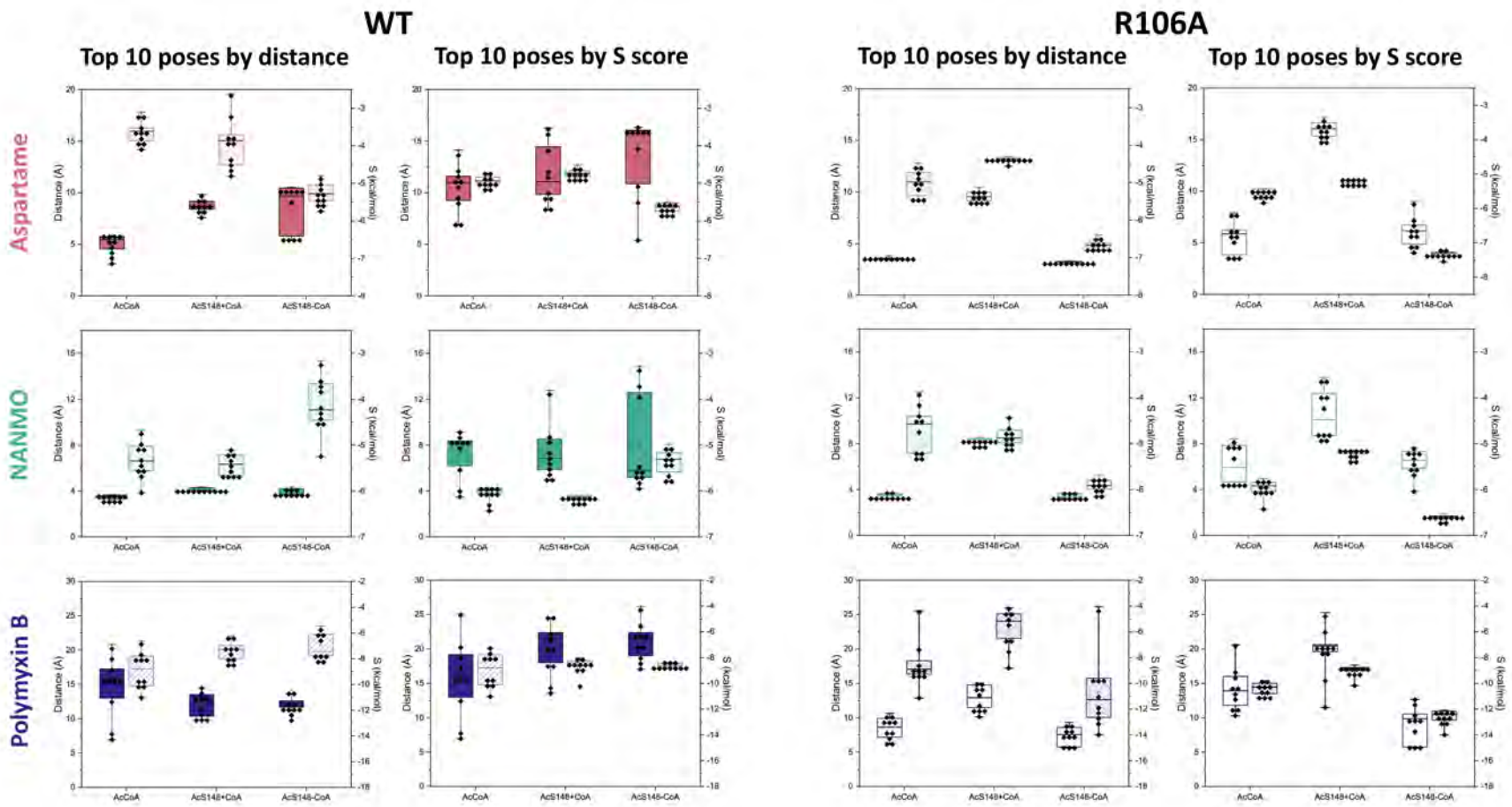
WT
R106A
Top 10 poses by distance
Top 10 poses by S score
Top 10 poses by distance
Top 10 poses by S score
Aspartame
NANMO
Polymyxin B
Distance (Å)
S (kcal/mol)
AcCoA
AcS148+CoA
AcS148-CoA

Figure 3

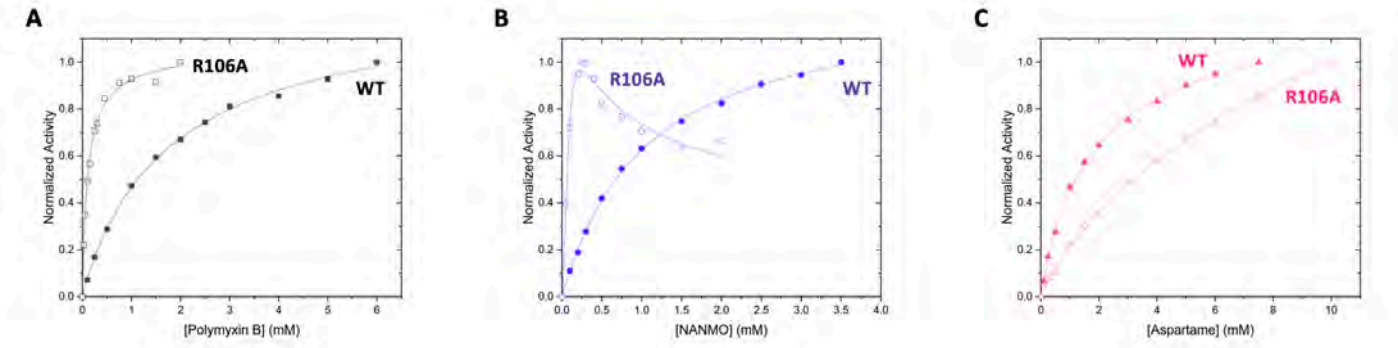
A
R106A
WT
Normalized Activity
[Polymyxin B] (mM)
B
R106A
WT
Normalized Activity
[NANMO] (mM)
C
WT
R106A
Normalized Activity
[Aspartame] (mM)

Figure 4

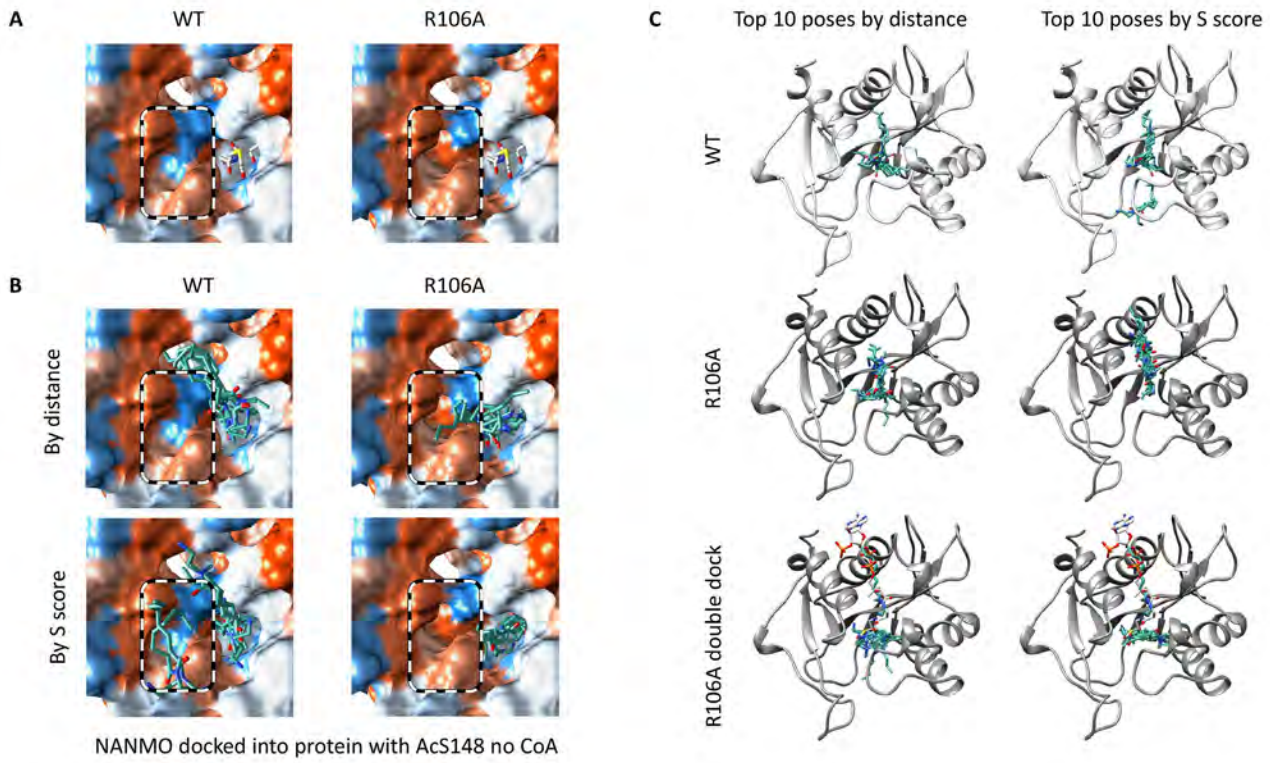
A
WT
R106A
B
WT
R106A
By distance
By S score
NANMO docked into protein with AcS148 no CoA
C
Top 10 poses by distance
Top 10 poses by S score
WT
R106A
R106A double dock

Figure 5

A

	Aspartame CE	NANMO CE	Polymyxin B CE
R59	-2	-1	+2
E102	-1	+4	-5
R106	-46	+5	+3
T141	-8	-2	-1
H167	-3	+1	+6
L169	-1	+1	+4
H179	-2	+2	+3

B

	Aspartame Km	NANMO Km	Polymyxin B Km
R59	+1	+1	-2
E102	-5	-11	-2
R106	+5	-13	-16
T141	+2	+1	+1
H167	+2	-3	-14
L169	+1	-1	-5
H179	-1	-3	-7

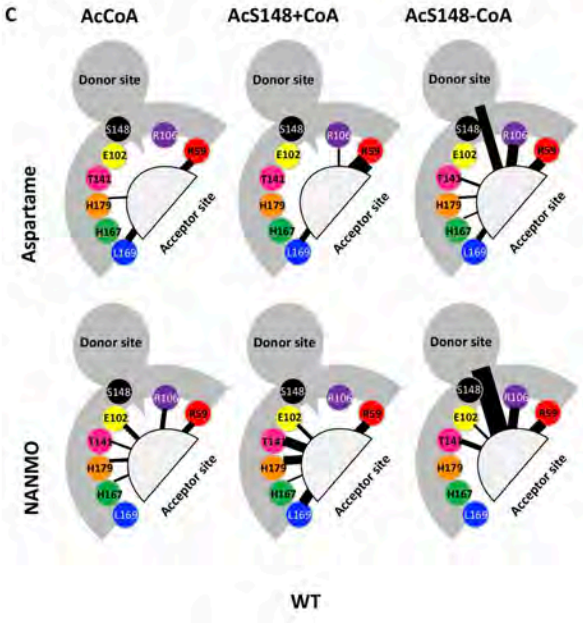

Figure 6