Computational and Experimental Study of Inhibitors Design for Aldolase A

Rui Qi,[a] Brandon Walker,[a] Zhifeng Jing,[a] Maiya Yu,[b] Gabriel Stancu,[c] Ramakrishna Edupuganti,[c] Kevin N. Dalby,[c] Pengyu Ren[a*]

a. Department of Biomedical Engineering, The University of Texas at Austin, Austin, Texas 78712
b. Department of Biochemistry and Mathematics, University of Michigan, Ann Arbor, Michigan 48109
c. Division of Chemical Biology and Medicinal Chemistry, College of Pharmacy, The University of Texas at Austin, Austin, Texas 78712

Abstract

Glycolytic enzyme fructose-bisphosphate aldolase A is an emerging therapeutic target in cancer. Recently we have solved the crystal structure of murine aldolase in complex with naphthalene-2,6-diyl bisphosphate (ND1) that served as a template of the design of bisphosphate-based inhibitors. In this work, a series of ND1 analogs containing difluoromethylene (-CF_2), methylene (-CH_2) or aldehyde substitutions were designed. All designed compounds were studied using MD simulations with the AMOEBA force field. Both energetics and structural analyses have been done to understand the calculated binding free energies. The average distances between ligand and protein atoms for ND1 were very similar to the ND1 crystal structure, which indicates our MD simulation is sampling the correct conformation well. CF2 insertion lowers the binding free energy by 10~15 kcal/mol while CF2 substitution slightly increases the binding free energy, which matches the experimental measurement. In addition, we found that NDB with two CF2 insertions, the strongest binder, is entropically driven, while others including NDA with one CF2 insertion are all enthalpically driven. This work provides insights into the mechanisms underlying protein-phosphate binding and enhances the capability of applying computational and theoretical frameworks to model, predict and design diagnostic strategies targeting cancer.

Introduction

For most cells, glycolysis is critical for generating energy and supplying metabolic intermediates for cellular biomass. One of the hallmarks of cancer is the altered metabolism preferential dependence on glycolysis in an oxygen-independent manner instead of oxidative phosphorylation, known as the "Warburg effect".[1] Recently, a novel feed-forward mechanism for hypoxic cancer has been identified. While HIF-1 upregulates transcription of glycolytic enzymes, the glycolysis under inadequate oxygen supply, in turn, increases HIF1a transcriptional activity and stimulates tumor growth.[2] **(Figure1).** Tumor glycolysis has been actively studied and serves as a potential target for cancer therapy.[3-4]

Figure 1. Glycolysis acts as a feed-forward mechanism for HIF-1 action.

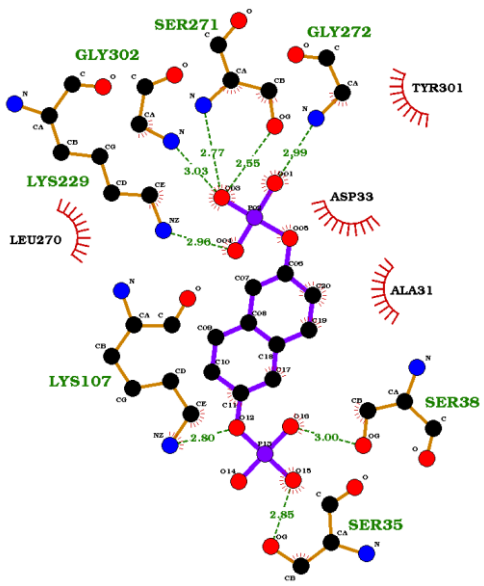


Figure 2. 2D plot of the binding pocket of ND1 in crystal structure[2] generated using LigPlot+. Left: Key residues include LYS107, SER35, SER38, SER271, GLY272, GLY302, LEU270, ALA31, ASP33, and Try301. Right: Intermolecular interactions around negatively charged phosphate groups are marked in green with distances while those hydrophobic ones involved aromatic systems are marked in red. These interactions include binding to both residues' backbone O and N, and sidechains' -OH and -NH_2.

A leading candidate for this target is the fructose-bisphosphate aldolase A (ALDOA), a central enzyme in glycolysis.[5] ALDOA is responsible for converting fructose-1,6-biphosphate (FDP) into glyceraldehyde 3-phosphate (GAP) and dihydroxyacetone phosphate (DHAP). The generally accepted catalytic mechanism for ALDOA is shown in **Figure S1**. The reaction proceeds with the formation of a Schiff base intermediate between LYS229 of the active site and the carbonyl group of the substrate FDP.[6-7] The inhibition of ALDOA has been shown to block the glycolysis, decrease HIF-1 activity and break the feed-forward loop mechanism in cells. Thus aldolase A has prospects for controlling cancer proliferation.[2, 8-10]

Aldolase inhibitors have been designed to mimic the substrate of FDP by probing the nature of the active site.[6, 11] General principles of drug design involve keeping the strong electrostatic interactions with residues in the active site while maintaining hydrophobic interactions in the linkage. Aldolase A has been co-crystallized with naphthalene-2,6-diyl bisphosphate (ND1), an active site substrate-mimetic. **Figure 2** shows the 2D structure of the ND1 and highlights the key residues in the binding pocket. **Figure S2** shows the same graphic enlarged. H-bonds have been found between the two negatively charged phosphate groups and the polar and positive charged residues including SER35, SER38, SER271, LYS229 and LYS107 as well as the neutral GLY272 and GLY302. These interactions include binding to both residues' backbone O and N, and sidechain -OH and -NH_2. Besides, hydrophobic interactions are marked in red involving LEU270, ALA31, ASP33, and TRY301. Note that negatively charged Asp33 interacts with the naphthalene ring, not the phosphate groups. Although ND1 is a potent inhibitor, with two polar phosphate groups, it is easy to be hydrolyzed and hard to deliver in vivo.

Molecular dynamics (MD) simulations are a powerful tool for understanding the driving forces underlying molecular recognition, accelerating drug discovery, and guiding molecular design.[12-18] Classical force fields such as AMBER[19], CHARMM[20], OPLS-AA[21], or GROMOS[22] are computationally efficient and sufficiently accurate for many applications.[14, 23-25] However, for highly charged species like phosphates-containing ligands, the actual charge distributions of atoms and their changes in response to the environment's electric field is complicated and challenging to model and simulate.[26-32] Recently, polarizable force fields have shown encouraging results for depicting these complicated interactions.[33-34] As a physics-grounded force field, AMOEBA depicts molecular polarizability and electrostatic potential terms by using mutual atomic dipole-dipole induction along with permanent atomic point multipoles up to quadrupole. [35-36]

In this work, we explored a series of ND1 analogs for both covalent and non-covalent inhibitors using MD simulations with AMOEBA. For the non-covalent inhibitors, the difluoromethylene (-CF_2) and methylene (-CH_2) groups have been inserted or substituted targeting the phosphate bridging oxygen. For the covalent ones, aldehyde substitutions have been done on the naphthalene rings. (**Figure 3**) For the non-covalent inhibitors, we dissected the roles of entropy and enthalpy in binding for each system based on the calculated binding free energy. We also conducted a structural analysis of the distances changing in the key interactions to further compare the ligand binding modes. For the covalent inhibitors, we conducted non-covalent binding simulations to investigate the potential sites for bond formation. This work provides insights into the mechanisms underlying protein-phosphate binding and enhance the capability of applying computational and theoretical frameworks to model, predict and design diagnostic strategies targeting caner.

Methods

Parameterization: Ab initio quantum mechanics calculations (QM) were performed using Gaussian 09[37] and PSI4 program.[38-40] All molecular mechanics (MM) force field-based calculations were performed using TINKER 8 Software.[36, 41] The parameters for ND1 and its derivatives were derived by using POLTYPE program.[42] The structures were optimized at MP2/cc-pVTZ level with Polarizable Continuum Model (PCM)[43] and the single point energy were calculated at the MP2/6-311++G(2d,2p) level. Atomic multipole moments were initially assigned from QM electron density calculated at the MP2/6-311G** level via Stone's distributed multipole analysis[44] and the optimizations were done with Tinker's POTENTIAL program to fit electrostatic potentials around molecules. The Van der Waals parameters were optimized to capture the ligand-water interaction energy at different orientations calculated at the MP2/Aug-cc-pVTZ/QZ, extrapolated to CBS level. The torsional parameters were derived to reproduce the QM conformational energy profile at MP2/6-311++G** level. For interaction energy calculations, the dimer structures were optimized using MP2/cc-pVTZ with PCM in Gaussian[4] and the interaction energies were calculated using MP2/aug-cc-pVT/QZ in Psi4[5] and extrapolated to complete basis set. All parameters for water, protein and ions were adopted from the current AMOEBA force field.[15, 45-46]

$$U = U(\lambda, \boldsymbol{R}), U(0, \boldsymbol{R}) = U_0, U(1, \boldsymbol{R}) = U_1 \quad (1)$$

$$\Delta G = \sum_{i=1}^{N-1} G(\lambda_{i+1}) - G(\lambda_i), \lambda_1 = 0, \lambda_N = 1 \quad (2)$$

$$\Delta G = \Delta H - T\Delta S \quad (3)$$

Molecular dynamics simulations: The structure of the ND1-ALDOLASE complex was taken from chain A in the crystal structure.[2] All other derivatives were aligned to ND1 by the naphthalene ring in the same protein pocket. The complexes were then solvated in periodic boxes of 82.48×82.48×82.48 $Å^3$ with NaCl added to yield 0.15 M salt concentration. All molecular dynamics simulations were run using the Tinker-OpenMM

program on GPU[47] with a RESPA integrator[48], Bussi thermostat[49], Berendsen barostat,[50] and 3.0 fs time step with hydrogen-mass repartition (heavy-hydrogen keyword). The van der Waals (vdW) iterations used a 12.0 Å cutoff, while the electrostatic interactions used a 7.0 Å cutoff. The systems were gradually heated up to 298 K and sequentially relaxed from water and ions first, and then protein-ligand complex before free energy production simulations. The non-covalent binding free energies were calculated for all the ligands by the double-decoupling method (**Figure S3**).[15, 47] A mixed potential was defined to calculate the free energy difference between the end states that were connected analytically (Eq. 1). The free energy changes from one state to the other is thus given by Eq. 2. The energy difference between adjacent states was estimated by the Bennett acceptance ratio (BAR) method.[51] For the entropy and enthalpy decomposition, the enthalpy of binding was calculated first by averaging the potential energy from each dynamic step. The entropy contribution will be obtained from the difference between free energy and enthalpy (Eq. 3). Detailed methodology and perturbation schedules can be found in previous papers.[52-53] Energies reported in this work are relative to ND1 in kcal/mol. It is noted that all of these systems do have a net charge of -4, which can cause some error in the absolute binding free energies. However, here we are only concerned with relative binding free energy changes to ND1. Error cancelation is expected to compensate for the charged system. Charge corrections schemes using a Poisson calculation with periodic boundary conditions could be used to correct this if absolute binding free energy was important[54-55].

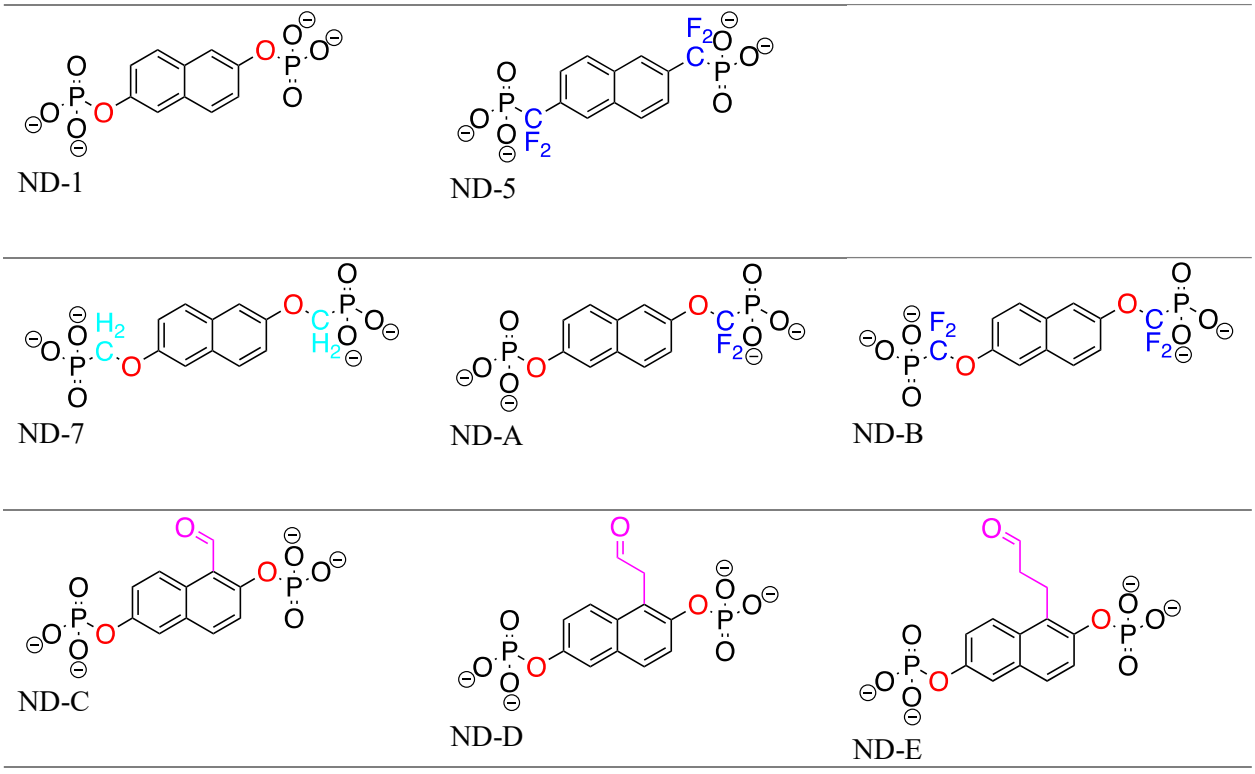


Figure 3. Structures of non-covalent and covalent inhibitors. Experimentally measured Ki values for ND1 and ND5 are 0.89 μM and 11.22 μM, respectively. Only ND1 has crystal structure[2] co-crystalized with ALDOLASE A protein. The third row contains three covalent inhibitors ND-C, ND-D, and ND-E, while all the others are noncovalent inhibitors. The bridging Os are marked in red. Substitutions and insertions of -CF_2 and -CH_2 are marked in purple and the aldehyde groups are marked in magenta.

Structure analysis: Interaction distances between protein and ligands were calculated over all frames of molecular dynamics trajectories. Then a heatmap were generated to visualize the interaction distances for all derivatives. Amino acids were chosen from being within a 4-angstrom distance of any ligand atom from the first MD trajectory frame. The heatmap has been sorted to show the interactions that are most conserved among the derivatives. For atoms that share the same force field atom type, such as O^{-}s in [PO3], Fs in [CF2], and Hs in [CH2], the heatmap only shows interactions of the atom that has the closest distance with a given protein residue. The standard deviation of distances is typically 0.1-0.4 angstroms, with few exceptions of 0.4-1 angstroms. The detailed statistics are listed in **Table S3 and S4**.

Experimental details: For the ligand ND5, the ALDOA activity was measured using standard NADH coupled enzymatic assay by measuring the absorbance at 340nm. The assay uses glycerol-3-phosphate dehydrogenase to catalyze the conversion of dihydroxyacetone phosphate (product of ALDOA substrate cleavage) to glyceraldehyde-3 phosphate in an NADH oxidation-dependent manner. The reaction was carried out in 50 mM TEA-HCl pH 7.5, 50 mM NaCl, 1 mM EDTA, 0.01% Triton-X 100, 40 u M NADH, excess of 1.7 U of glycerol-3-phosphate dehydrogenase/triosephosphate isomerase (-GDH and TPI) (Sigma-Aldrich, St. Louis, MO) and varying concentrations of fructose 1,6-bisphosphate, at 30 °C. For ND1, the chemical assay has been published previously.[9] All the detailed measurements of the Ki value can be found in **FigureS4**.

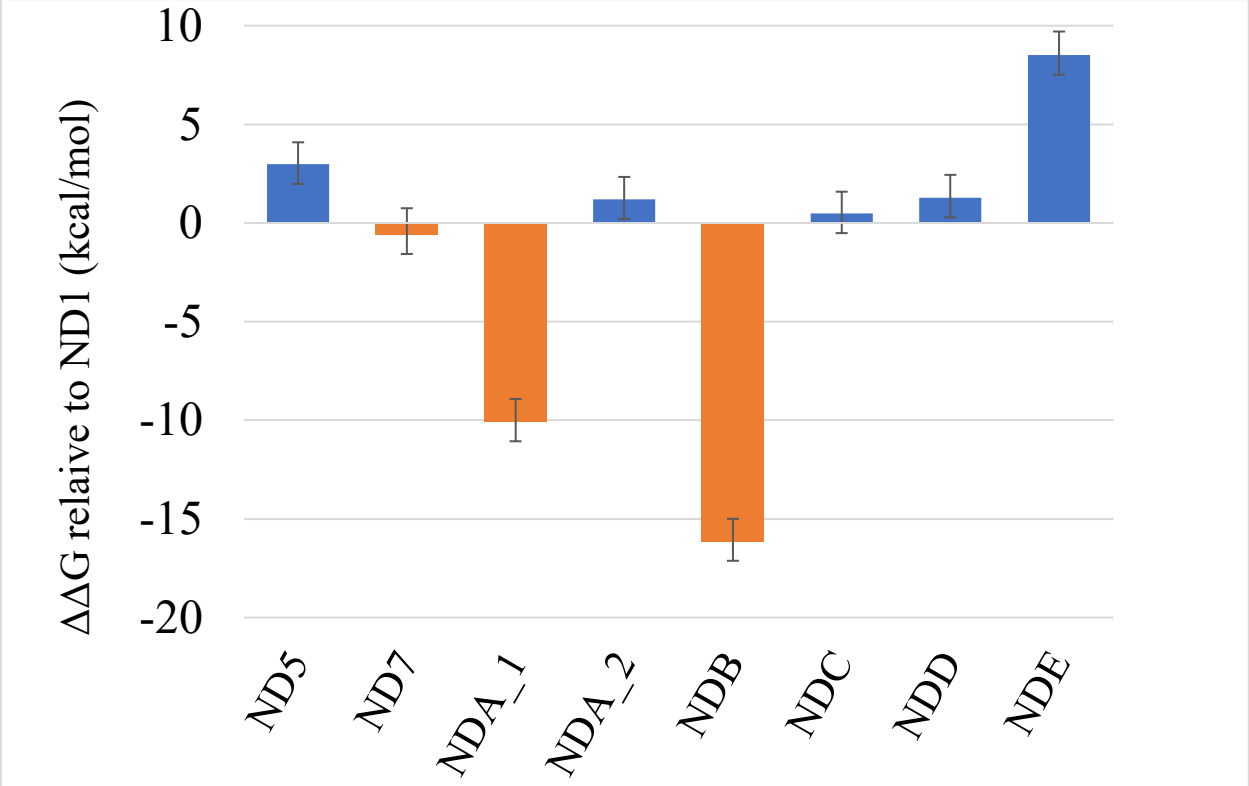


Figure 4. Relative binding free energy $\Delta\Delta G$ of all ligands to ND1 in kcal/mol. Blue ones represent that ligands that bind weaker than ND while the oranges ones represent that ligand bind tighter than ND1. For all ligands, the non-covalent binding free energies were computed. All energies are in kcal/mol and the values are listed in **Table S1**.

Results and Discussions

From chemical assays, we determined that ND1 is a strong inhibitor of ALDOA with the K_i value of 0.89 μM while ND5 is ~1.49 kcal/mol weak relative to the ND1 with the Ki value of 11.22 μM. From MD simulations using AMOEBA, the ligands ranked by calculated binding free energies from tightest to weakest are NDB, NDA1, ND7 and ND1, NDA2, ND5 for the non-covalent competitive inhibitors. All the aldehyde substituted derivatives designed to mimic the covalent inhibitors are weaker binder compared to ND1. ND5 has more positive binding free energy compare to ND1 and the relative binding free energies of ND5 to ND1 is ~2.98 kcal/mol, which matches the experimental measurement. By comparing NDA_2 and NDB, we observed that the free energy decreases when inserting another CF2 group. Neither inserting CH2 groups in ND7 nor substituting the bridge O with CF2 group in ND5 improve the binding. All the experimental details are described in the Methods section.

Table 1. Relative binding free energies ΔΔG (kcal/mol) of mutating the functional group on each side of ND5/NDB separately compared to ND1. The functional groups are on the 2, 7 positions on the naphthalene ring. In ALDOA, the protein side represents the side facing the protein binding pocket (close to LYS229 and SER271) while the water side represents the side that is close to SER38 and SER35, more exposed to water. The uncertainties are in parenthesis.

	Water side	Protein side		Both sides
ND5	0.9 (0.3)	2.9 (0.3)	ND5	3.0 (1.1)
NDA	1.2 (1.1)	-10.1 (1.1)	NDB	-16.1 (1.1)

Most of the ligands have symmetric structures with one functional group on each side. NDA was designed to study the interaction of single functional group with different side of the binding pocket. Two simulations were set up: NDA_1 with the CF2 site facing the protein pocket (close to LYS229 and SER271) and NDA_2 with the CF2 toward the water (close to SER38 and SER35). Similarly, the interactions of ND5 at two different sides were studied by mutating only the functional group at one side. The binding free energies of the single functionalization with two different orientations were compared with those of double functionalization in **Figure 4** and **Table 1**. ΔΔG are positive for ND5 with both orientations, while it is smaller when the CF2 group faces the water side. The sum of ΔΔG for two CF2 groups of ND5 is slightly higher than the total ΔΔG of ND5. For NDA, ΔΔG is close to zero when the CF2 group faces the water side (NDA_2) and much lower when CF2 faces the protein side (NDA_1). The sum of ΔΔG for NDA is significantly higher than ΔΔG of NDB, which indicates a cooperative effect of the two CF2 groups. The results of ND5 and NDA agree with each other in that modifications on ligand interacting with the protein side have a large effect.

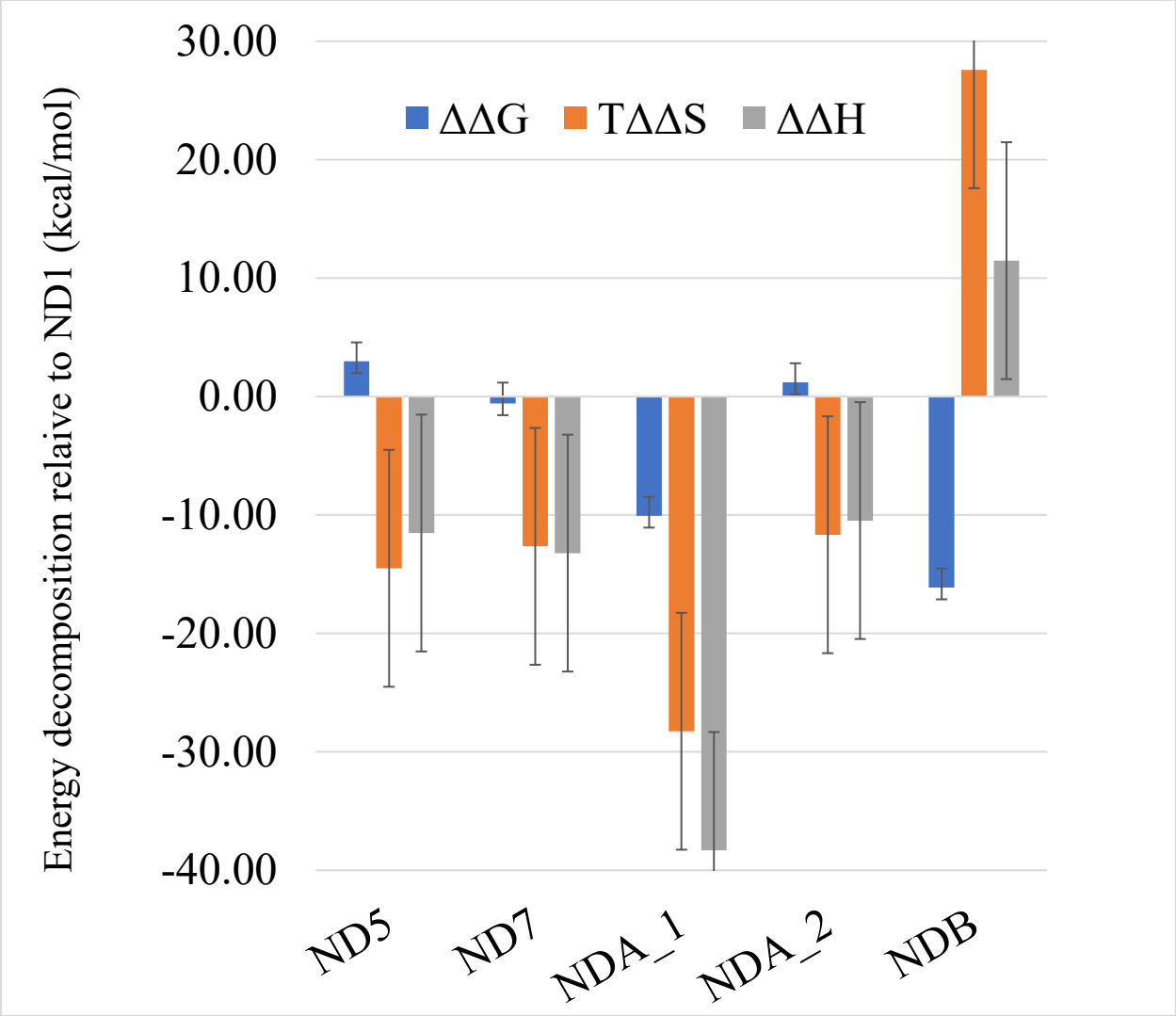


Figure 5. Relative energy decomposition results of selected non-covalent inhibitors to ND1 including the relative enthalpy (ΔΔH) in blue, relative entropy (TΔΔS) in green, and the relative binding free energy in yellow. All energies are in kcal/mol and the uncertainties are in **Table S2**.

To explore the driving force underlying ND series inhibitions, we examined the enthalpy and entropy contributions of the binding free energy. **Figure 5** lists the calculated binding enthalpy and entropy for non-covalent ligand. The binding enthalpies are vastly different. There appears to be no simple relationship between the binding thermodynamics and the size or the length of the ligands. Comparing ND5 with NDB and NDAs with NDB, we find that the binding enthalpy does not correlate with the inserting 1 or 2 atoms in the phosphate. For the two orientations of NDA, NDA_1 with the CF2 pointing to the protein side has a much lower binding enthalpy. Interestingly, the strongest binder NDB is entropically driven, while NDA_1, which also has the CF2 insertion and a negative ΔΔG, is enthapically driven. The weak binders including ND7, NDA_2 and ND5 have favorable binding enthalpies which are compensated by entropy contributions. These results suggest the importance of enthalpy-entropy compensation: either enthalpy or entropy itself cannot guide the design of potent inhibitors.

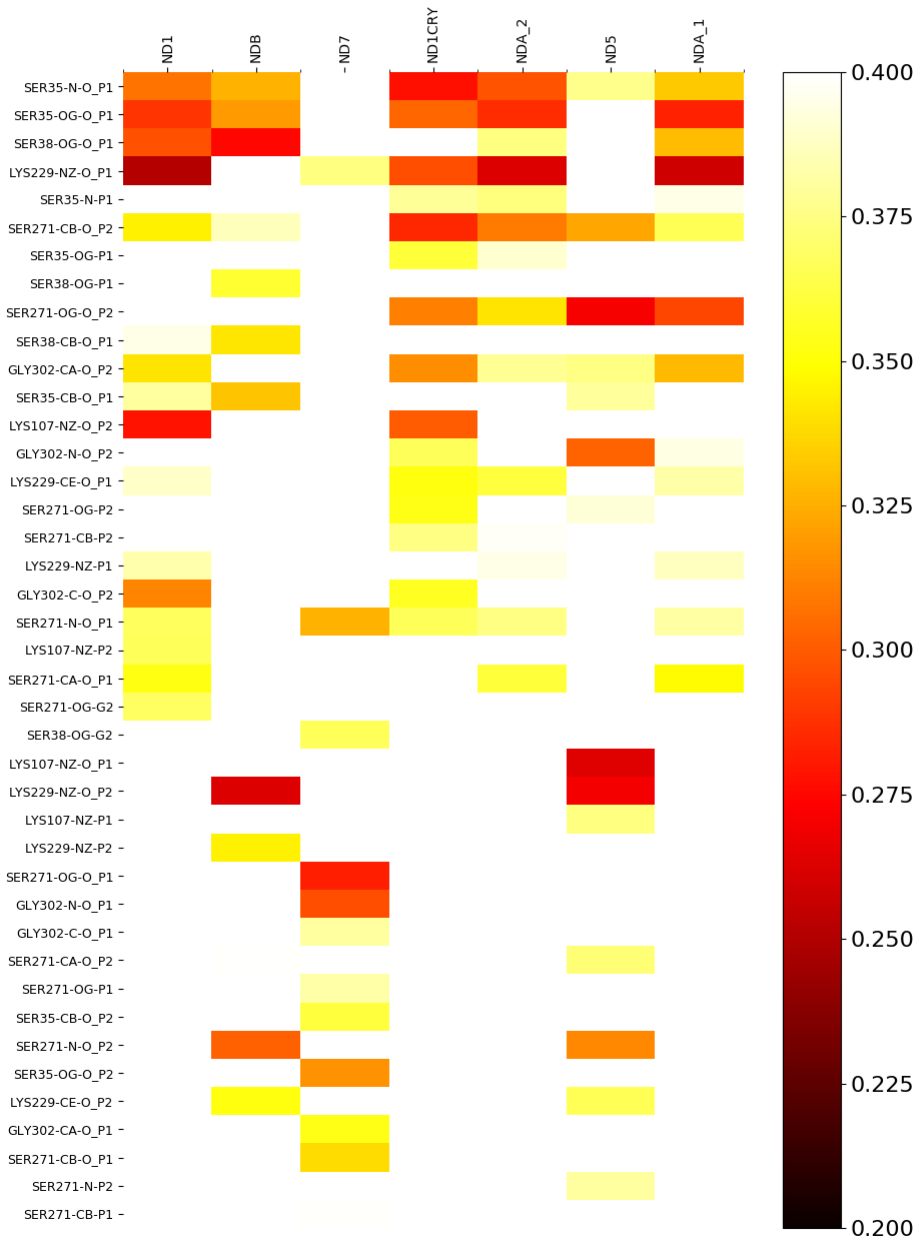
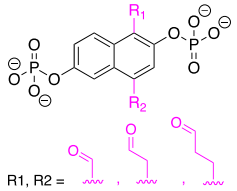
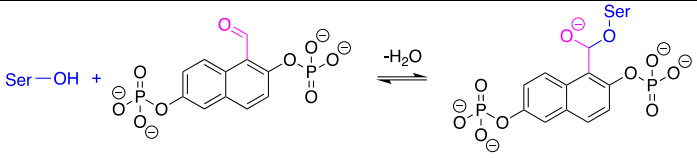
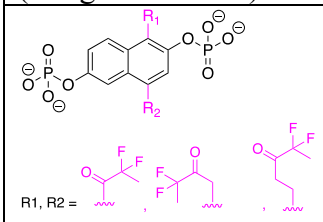
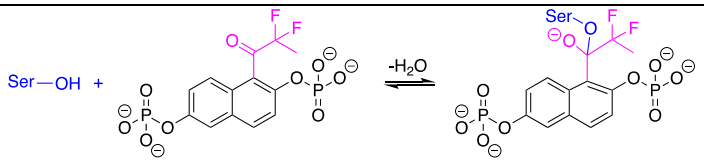
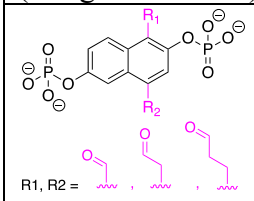
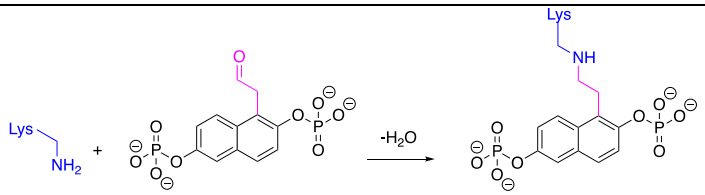
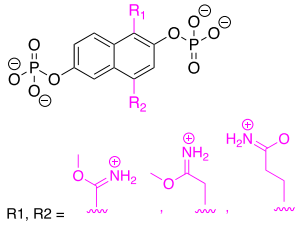
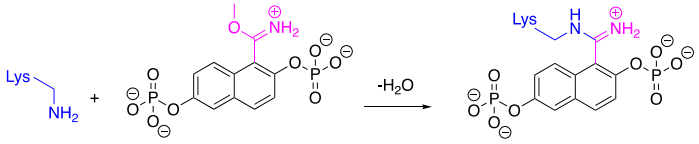


Figure 6. Heatmap of distances for conserved interactions between protein and ND1 derivatives. Interactions that are conserved on protein pocket side are defined with a 1 in their label, while atoms on the water exposed side are defined with 2 in their label. Any oxygen not on a phosphate group is defined by OB, where B refers to bridging oxygen. CF refers to the carbon attached to a Fluorine. CH refers to the carbon attached to a hydrogen. G represents the naphthalene ring. The bar on the right presents distances ranging from 0.2 nm in black to 0.4 nm in white. Blue indicates that an interaction does not exist for a given derivative. ND1CRY, refers to the crystal PDB of ND1.

Table 2. Covalent inhibitors design. 4 types of inhibitors have been designed, where two of them are categorized into covalent reversible while the other two are covalent irreversible inhibitors. Two hydrogens on the naphthalene ring were designed to be substituted by R1 or R2 highlighted in pink. The reaction site in the binding pocket has been simplified as the key residues highlighted in blue. The proposed inhibition mechanisms are illustrated by using one of the possible derivatives listed in column #2.

Designed inhibitors	Proposed Inhibition mechanism	Category
 (design #1.1~#1.3)		*covalent reversible,* aldehydes
 (design #2.1~#2.3)		*covalent reversible,* difluoroketone
 (design #3.1~#3.3)		*covalent irreversible,* "classic" affinity label
 (design #4.1~#4.3)		*covalent irreversible,*

In an attempt to understand what interactions make each derivative binding specific, we computed the average distance for interactions between protein and ND1 derivatives (**Figure 6**). With the exception of a few interactions such as ARG303-N-O_P1, ALA31-CB-O_P1, SER38-OG-G2, the crystal structure of ND1 is very close to the averaged MD trajectory of ND1. NDA_1, while keeping most of the interactions found in ND1, has an additional salt bridge between Lys107 and the PO3 group. Some of the conserved interactions across most derivatives, *i.e.* the interactions of the PO3 head inside the pocket with Ser271 and Arg303 are not observed in NDB. Instead, the CF2 group interact with these residues. The missing of the salt bridge in NDB binding may explain its unfavorable binding enthalpy.

The design of a covalent inhibitor targets the hydrophobic naphthalene on ND1 by adding functional groups to enhance the inhibitive properties of ND1. As designed, those substitution groups can form covalent bonds with polar residues in the binding pocket including Lys107, Lys229, and Ser38. The inhibition mechanisms can be categorized as covalent reversible or covalent irreversible. 12 designed covalent inhibitors in 4 categories are listed in **Table 2**. Since ND1 has a symmetric structure, two positions on the naphthalene ring (shown as R_1, R_2) have been selected. The proposed reactions are listed in the second column with residues in the binding pocket, modifications, original ND1 colored in blue, pink, and black respectively. Only one mechanism for each category is listing as an illustration while the rest of the reactions in the same row should follow the same mechanism. The second and third columns show the designed covalent reversible inhibitors (design #1.1~1.3 and design #2.1~2.3). By adding and aldehydes and difluoroketone, the new inhibitors could react with the -OH in Ser38 and form covalent bond with the binding pocket. The last two rows show the covalent irreversible designs (design #3.1~3.3 and design #4.1~4.3). Two phosphates serve as selectivity group while the aldehydes and methyl acetimidate cations are the affinity groups/warheads. The most interesting design is the last one that was inspired by the chemical reagent inhibition. Because of the intrinsic reactivity, lysine forms covalent bonds with methyl acetimidate cations.

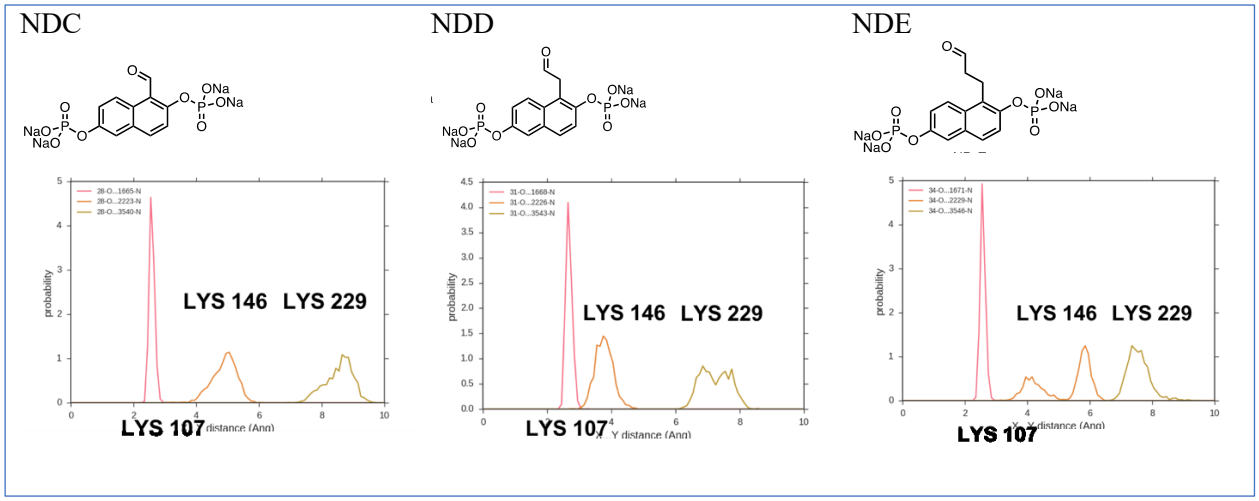


Figure 7. Distribution of distances between all covalent inhibitors with all three Lysine in the binding pocket. Close distances were found between aldehyde O and LYS N, which indicates a potential formation of the covalent bonds.

From the design of covalent inhibitors, three aldehyde derivatives, i.e. NDC, NDD and NDE were chosen for MD simulations to investigate the potential reaction site. **Figure 7** shows the distribution of distances between all covalent inhibitors with all three Lysine in the binding pocket. The close interactions with LYS107 in the binding pocket were found among all designs of the covalent inhibitors. This indicates a potential covalent bond formation while LYS146 and LYS229 are unlikely to react with the aldehyde groups. Refering back to **Figure 4,** the calculated binding free energies for design NDC, NDD and NDE are 0.3, 1.1 and 8.3 kcal/mol compared to ND1, respectively. The three positive values represent weaker binding after the modification. This might be due to the limitations of MD simulations, which can mimic the intermolecular interactions accurately without new bond formation in one continued simulation. Thus, without reactions, the addition of aldehyde makes the hydrophobic ring bulky and pushes some of the residues away from the inhibitor, and results in weak inhibitors.

Conclusion

In this work, a series of inhibitors for ALDOA have been designed in both non-covalent and covalent categories to improve the stability of ND1 and investigate the effects of CF2 group. MD simulations with the AMOEBA

force field were used to study the structures and free energies of the inhibitor binding. We observed that the average distances between ligand and protein atoms for ND1 were very similar to the ND1 crystal structure, which indicates our MD simulation is sampling the correct conformation. This gives credence to the derivatives sampling the correct conformation distribution. We found the most potent non-covalent inhibitor to be NDB with two CF2 insertions, followed by NDA with one CF2 insertion, while CF2 substitutions and CH2 insertion do not improve the binding affinity. Biochemical assays also showed that CH2 insertion weakens the inhibition effect.

Besides, it was found that CF2 substitution or insertion interacting with the protein side of the pocket has a larger contribution to free energy change than that interacting with the other side of the pocket. Notably, NDA with CF2 pointing to the protein side lowers the free energy by ~10 kcal/mol while it has little effect when CF2 points to the water side.

Thermodynamic analyses revealed nontrivial enthalpy-entropy compensation. NDB binding is entropy-driven while NDA binding is enthalpy-driven. This indicates the importance of using free energy calculations instead of energy calculation on minimized structures for lead optimization.

For the covalent inhibitors, we observed that LYS107 is a potential site for forming covalent bonds with ND1-aldehyde derivatives. This work suggests that the inhibitors containing insertions of CF2 have potential as chemotherapeutic drug leads, pending synthesis.

AUTHOR INFORMATION

Corresponding Author

*Email: pren@mail.utexas.edu Tel: 512-232-1832

Funding Sources

NIH R01GM106137 and R01GM114237; CPRIT RP160657; Welch Foundation F-1691

Notes

The authors declare no competing financial interest.

Supporting Information

The SI document contains figures and tables that elucidate more clearly, the reaction pathway for Aldolase, the most important interactions in the binding pocket, the thermodynamic cycle used to obtain binding free energy as well as tabulated values of the binding free energies relative to ND1. Additional informatin in SI includes, experimental Ki values as a function of subtrate concentration for ND1 and ND5. A table of the matrix used to generate the heatmap (Figure 6) is also included.

ACKNOWLEDGMENT

We are grateful for support by the National Institutes of Health (Nos. R01GM106137 and R01GM114237), the Robert A. Welch Foundation (No. F-1691) and CPRIT(RP160657). We thank Dr. C. Liu for careful reading of the manuscript.

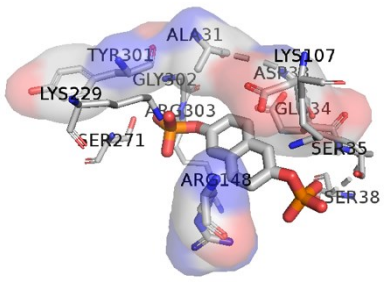
REFERENCES

1. Vander Heiden, M. G.; Cantley, L. C.; Thompson, C. B., Understanding the warburg effect: The metabolic requirements of cell proliferation. *Science* **2009,** *324* (5930), 1029-33.
2. Grandjean, G.; de Jong, P. R.; James, B.; Koh, M. Y.; Lemos, R.; Kingston, J.; Aleshin, A.; Bankston, L. A.; Miller, C. P.; Cho, E. J.*, et al.*, Definition of a novel feed-forward mechanism for glycolysis-hif1alpha signaling in hypoxic tumors highlights aldolase a as a therapeutic target. *Cancer Res* **2016,** *76* (14), 4259-4269.
3. Hwang, S. H.; Cho, A.; Yun, M.; Choi, Y. D.; Rha, S. Y.; Kang, W. J., Prognostic value of pretreatment metabolic tumor volume and total lesion glycolysis using 18f-fdg pet/ct in patients with metastatic renal cell carcinoma treated with anti-vascular endothelial growth factor-targeted agents. *Clin Nucl Med* **2017,** *42* (5), e235-e241.
4. Wang, Y.; Cai, W. S.; Chen, L.; Wang, G., Molecular dynamics simulation reveals how phosphorylation of tyrosine 26 of phosphoglycerate mutase 1 upregulates glycolysis and promotes tumor growth. *Oncotarget* **2017,** *8* (7), 12093-12107.
5. Chang, Y. C.; Yang, Y. C.; Tien, C. P.; Yang, C. J.; Hsiao, M., Roles of aldolase family genes in human cancers and diseases. *Trends Endocrinol Metab* **2018,** *29* (8), 549-559.
6. Gefflaut, T.; Blonski, C.; Perie, J.; Willson, M., Class i aldolases: Substrate specificity, mechanism, inhibitors and structural aspects. *Prog Biophys Mol Biol* **1995,** *63* (3), 301-40.
7. Garrabou, X.; Beck, T.; Hilvert, D., A promiscuous de novo retro-aldolase catalyzes asymmetric michael additions via schiff base intermediates. *Angew Chem Int Ed Engl* **2015,** *54* (19), 5609-12.
8. Chang, Y. C.; Chan, Y. C.; Chang, W. M.; Lin, Y. F.; Yang, C. J.; Su, C. Y.; Huang, M. S.; Wu, A. T. H.; Hsiao, M., Feedback regulation of aldoa activates the hif-1alpha/mmp9 axis to promote lung cancer progression. *Cancer Lett* **2017,** *403*, 28-36.
9. Cho, E. J.; Devkota, A. K.; Stancu, G.; Edupunganti, R.; Powis, G.; Dalby, K. N., A fluorescence-based high-throughput assay for the identification of anticancer reagents targeting fructose-1,6-bisphosphate aldolase. *SLAS Discov* **2018,** *23* (1), 1-10.
10. Han, X.; Zhu, X.; Hong, Z.; Wei, L.; Ren, Y.; Wan, F.; Zhu, S.; Peng, H.; Guo, L.; Rao, L.*, et al.*, Structure-based rational design of novel inhibitors against fructose-1,6-bisphosphate aldolase from candida albicans. *J Chem Inf Model* **2017,** *57* (6), 1426-1438.
11. Heron, P. W.; Abellan-Flos, M.; Salmon, L.; Sygusch, J., Bisphosphonate inhibitors of mammalian glycolytic aldolase. *J Med Chem* **2018,** *61* (23), 10558-10572.
12. Gohlke, H.; Klebe, G., Approaches to the description and prediction of the binding affinity of small-molecule ligands to macromolecular receptors. *Angew Chem Int Ed Engl* **2002,** *41* (15), 2644-76.
13. Houk, K. N.; Leach, A. G.; Kim, S. P.; Zhang, X., Binding affinities of host-guest, protein-ligand, and protein-transition-state complexes. *Angew Chem Int Ed Engl* **2003,** *42* (40), 4872-97.
14. Buch, I.; Giorgino, T.; De Fabritiis, G., Complete reconstruction of an enzyme-inhibitor binding process by molecular dynamics simulations. *Proc Natl Acad Sci U S A* **2011,** *108* (25), 10184-9.
15. Jiao, D.; Golubkov, P. A.; Darden, T. A.; Ren, P., Calculation of protein-ligand binding free energy by using a polarizable potential. *Proc Natl Acad Sci U S A* **2008,** *105* (17), 6290-5.
16. Dror, R. O.; Dirks, R. M.; Grossman, J. P.; Xu, H.; Shaw, D. E., Biomolecular simulation: A computational microscope for molecular biology. *Annu Rev Biophys* **2012,** *41* (1), 429-52.
17. Yang, W.; Gao, Y. Q.; Cui, Q.; Ma, J.; Karplus, M., The missing link between thermodynamics and structure in f1-atpase. *Proc Natl Acad Sci U S A* **2003,** *100* (3), 874-9.

18. Jorgensen, W. L., Efficient drug lead discovery and optimization. *Acc Chem Res* **2009,** *42* (6), 724-33.
19. Case, D. A.; Berryman, J. T.; Betz, R. M.; Cerutti, D. S.; T.E. Cheatham, I.; Darden, T. A.; Duke, R. E.; Giese, T. J.; Gohlke, H.; Goetz, A. W.*, et al. Amber 2015*; University of California, San Francisco: 2015.
20. Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.*, et al.*, Charmm general force field: A force field for drug-like molecules compatible with the charmm all-atom additive biological force fields. *J Comput Chem* **2010,** *31* (4), 671-90.
21. Robertson, M. J.; Tirado-Rives, J.; Jorgensen, W. L., Improved peptide and protein torsional energetics with the oplsaa force field. *J Chem Theory Comput* **2015,** *11* (7), 3499-509.
22. Reif, M. M.; Hunenberger, P. H.; Oostenbrink, C., New interaction parameters for charged amino acid side chains in the gromos force field. *J Chem Theory Comput* **2012,** *8* (10), 3705-23.
23. Modi, N.; Benz, R.; Hancock, R. E.; Kleinekathofer, U., Modeling the ion selectivity of the phosphate specific channel oprp. *J Phys Chem Lett* **2012,** *3* (23), 3639-45.
24. Noskov, S. Y.; Berneche, S.; Roux, B., Control of ion selectivity in potassium channels by electrostatic and dynamic properties of carbonyl ligands. *Nature* **2004,** *431* (7010), 830-4.
25. Modi, N.; Barcena-Uribarri, I.; Bains, M.; Benz, R.; Hancock, R. E.; Kleinekathofer, U., Tuning the affinity of anion binding sites in porin channels with negatively charged residues: Molecular details for oprp. *ACS Chem Biol* **2015,** *10* (2), 441-51.
26. Gruber, M. F.; Wood, E.; Truelsen, S.; Ostergaard, T.; Helix-Nielsen, C., Computational design of biomimetic phosphate scavengers. *Environ Sci Technol* **2015,** *49* (16), 9469-78.
27. Rao, L.; Cui, Q.; Xu, X., Electronic properties and desolvation penalties of metal ions plus protein electrostatics dictate the metal binding affinity and selectivity in the copper efflux regulator. *J Am Chem Soc* **2010,** *132* (51), 18092-102.
28. Dudev, T.; Lim, C., Competition among metal ions for protein binding sites: Determinants of metal ion selectivity in proteins. *Chem Rev* **2014,** *114* (1), 538-56.
29. Mackerell, A. D., Jr., Empirical force fields for biological macromolecules: Overview and issues. *J Comput Chem* **2004,** *25* (13), 1584-604.
30. Ponder, J. W.; Case, D. A., Force fields for protein simulations. *Adv Protein Chem* **2003,** *66*, 27-85.
31. Cieplak, P.; Dupradeau, F. Y.; Duan, Y.; Wang, J., Polarization effects in molecular mechanical force fields. *J Phys Condens Matter* **2009,** *21* (33), 333102.
32. Lopes, P. E.; Roux, B.; Mackerell, A. D., Jr., Molecular modeling and dynamics studies with explicit inclusion of electronic polarizability. Theory and applications. *Theor Chem Acc* **2009,** *124* (1-2), 11-28.
33. Jing, Z.; Liu, C.; Cheng, S. Y.; Qi, R.; Walker, B. D.; Piquemal, J. P.; Ren, P., Polarizable force fields for biomolecular simulations: Recent advances and applications. *Annu Rev Biophys* **2019,** *48*, 371-394.
34. Lemkul, J. A.; Huang, J.; Roux, B.; MacKerell, A. D., Jr., An empirical polarizable force field based on the classical drude oscillator model: Development history and recent applications. *Chem Rev* **2016,** *116* (9), 4983-5013.
35. Shi, Y.; Xia, Z.; Zhang, J.; Best, R.; Wu, C.; Ponder, J. W.; Ren, P., The polarizable atomic multipole-based amoeba force field for proteins. *J Chem Theory Comput* **2013,** *9* (9), 4046-4063.

36. Ponder, J. W.; Wu, C.; Ren, P.; Pande, V. S.; Chodera, J. D.; Schnieders, M. J.; Haque, I.; Mobley, D. L.; Lambrecht, D. S.; DiStasio, R. A., Jr.*, et al.*, Current status of the amoeba polarizable force field. *J Phys Chem B* **2010,** *114* (8), 2549-64.
37. Gomperts, R.; Frisch, M.; Scalmani, G.; Leback, B., Current status of the project to enable gaussian 09 on gpgpus. *Abstr Pap Am Chem S* **2014,** *247*.
38. Hohenstein, E. G.; Sherrill, C. D., Density fitting of intramonomer correlation effects in symmetry-adapted perturbation theory. *J Chem Phys* **2010,** *133* (1), 014101.
39. Parker, T. M.; Burns, L. A.; Parrish, R. M.; Ryno, A. G.; Sherrill, C. D., Levels of symmetry adapted perturbation theory (sapt). I. Efficiency and performance for interaction energies. *J Chem Phys* **2014,** *140* (9), 094106.
40. Turney, J. M.; Simmonett, A. C.; Parrish, R. M.; Hohenstein, E. G.; Evangelista, F. A.; Fermann, J. T.; Mintz, B. J.; Burns, L. A.; Wilke, J. J.; Abrams, M. L.*, et al.*, Psi4: An open-source ab initio electronic structure program. *Wires Comput Mol Sci* **2012,** *2* (4), 556-565.
41. Rackers, J. A.; Wang, Z.; Lu, C.; Laury, M. L.; Lagardere, L.; Schnieders, M. J.; Piquemal, J. P.; Ren, P.; Ponder, J. W., Tinker 8: Software tools for molecular design. *J Chem Theory Comput* **2018,** *14* (10), 5273-5289.
42. Wu, J. C.; Chattree, G.; Ren, P., Automation of amoeba polarizable force field parameterization for small molecules. *Theor Chem Acc* **2012,** *131* (3), 1138.
43. Mennucci, B.; Tomasi, J.; Cammi, R.; Cheeseman, J. R.; Frisch, M. J.; Devlin, F. J.; Gabriel, S.; Stephens, P. J., Polarizable continuum model (pcm) calculations of solvent effects on optical rotations of chiral molecules. *J Phys Chem A* **2002,** *106* (25), 6102-6113.
44. Stone, A. J., Distributed multipole analysis: Stability for large basis sets. *J Chem Theory Comput* **2005,** *1* (6), 1128-32.
45. Ren, P.; Ponder, J. W., Polarizable atomic multipole water model for molecular mechanics simulation. *J. Phys. Chem. B* **2003,** *107* (24), 5933-5947.
46. Jing, Z.; Liu, C.; Qi, R.; Ren, P., Many-body effect determines the selectivity for ca(2+) and mg(2+) in proteins. *Proc Natl Acad Sci U S A* **2018,** *115* (32), E7495-E7501.
47. Harger, M.; Li, D.; Wang, Z.; Dalby, K.; Lagardere, L.; Piquemal, J. P.; Ponder, J.; Ren, P., Tinker-openmm: Absolute and relative alchemical free energies using amoeba on gpus. *J Comput Chem* **2017,** *38* (23), 2047-2055.
48. Tuckerman, M.; Berne, B. J.; Martyna, G. J., Reversible multiple time scale molecular-dynamics. *J Chem Phys* **1992,** *97* (3), 1990-2001.
49. Bussi, G.; Donadio, D.; Parrinello, M., Canonical sampling through velocity rescaling. *J Chem Phys* **2007,** *126* (1), 014101.
50. Berendsen, H. J. C.; Postma, J. P. M.; Vangunsteren, W. F.; Dinola, A.; Haak, J. R., Molecular-dynamics with coupling to an external bath. *J Chem Phys* **1984,** *81* (8), 3684-3690.
51. Bennett, C. H., Efficient estimation of free-energy differences from monte-carlo data. *J Comput Phys* **1976,** *22* (2), 245-268.
52. Bell, D. R.; Qi, R.; Jing, Z.; Xiang, J. Y.; Mejias, C.; Schnieders, M. J.; Ponder, J. W.; Ren, P., Calculating binding free energies of host-guest systems using the amoeba polarizable force field. *Phys Chem Chem Phys* **2016,** *18* (44), 30261-30269.
53. Qi, R.; Jing, Z.; Liu, C.; Piquemal, J. P.; Dalby, K. N.; Ren, P., Elucidating the phosphate binding mode of phosphate-binding protein: The critical effect of buffer solution. *J Phys Chem B* **2018,** *122* (24), 6371-6376.
54. Rocklin, G. J.; Mobley, D. L.; Dill, K. A.; Hunenberger, P. H., Calculating the binding free energies of charged species based on explicit-solvent simulations employing lattice-sum methods:

An accurate correction scheme for electrostatic finite-size effects. *J Chem Phys* **2013,** *139* (18), 184103.
55. Simonson, T.; Roux, B., Concepts and protocols for electrostatic free energies. *Mol Simul* **2016,** *42* (13), 1090-1101.



TOC Graphic

TOC Graphic

Supporting information for

Computational and Experimental Study of Inhibitors Design for Aldolase A

Rui Qi,[a] Brandon Walker,[a] Zhifeng Jing,[a] Maiya Yu,[b] Gabriel Stancu,[c] Ramakrishna Edupuganti,[c] Kevin N. Dalby,[c] Pengyu Ren[a*]

a. Department of Biomedical Engineering, The University of Texas at Austin, Austin, Texas 78712
b. Department of Biochemistry and Mathematics, University of Michigan, Ann Arbor, Michigan 48109
c. Division of Chemical Biology and Medicinal Chemistry, College of Pharmacy, The University of Texas at Austin, Austin, Texas 78712

Corresponding Author Dr. Pengyu Ren. Tel: 512-232-1832, Email: pren@mail.utexas.edu

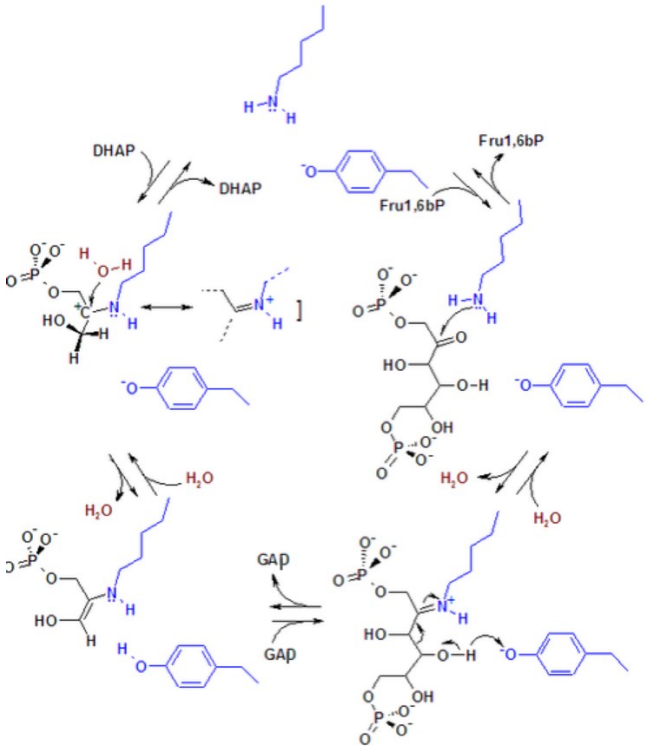


Figure S1. The catalytic mechanism for ALDOA. Two residues in the binding pocket, Lys229 and Tyr363, are highlighted in blue. Lys forms a Schiff base with substrate FDP while Tyr is involved in the protonation or deprotonation of the enamine intermediate.[3]

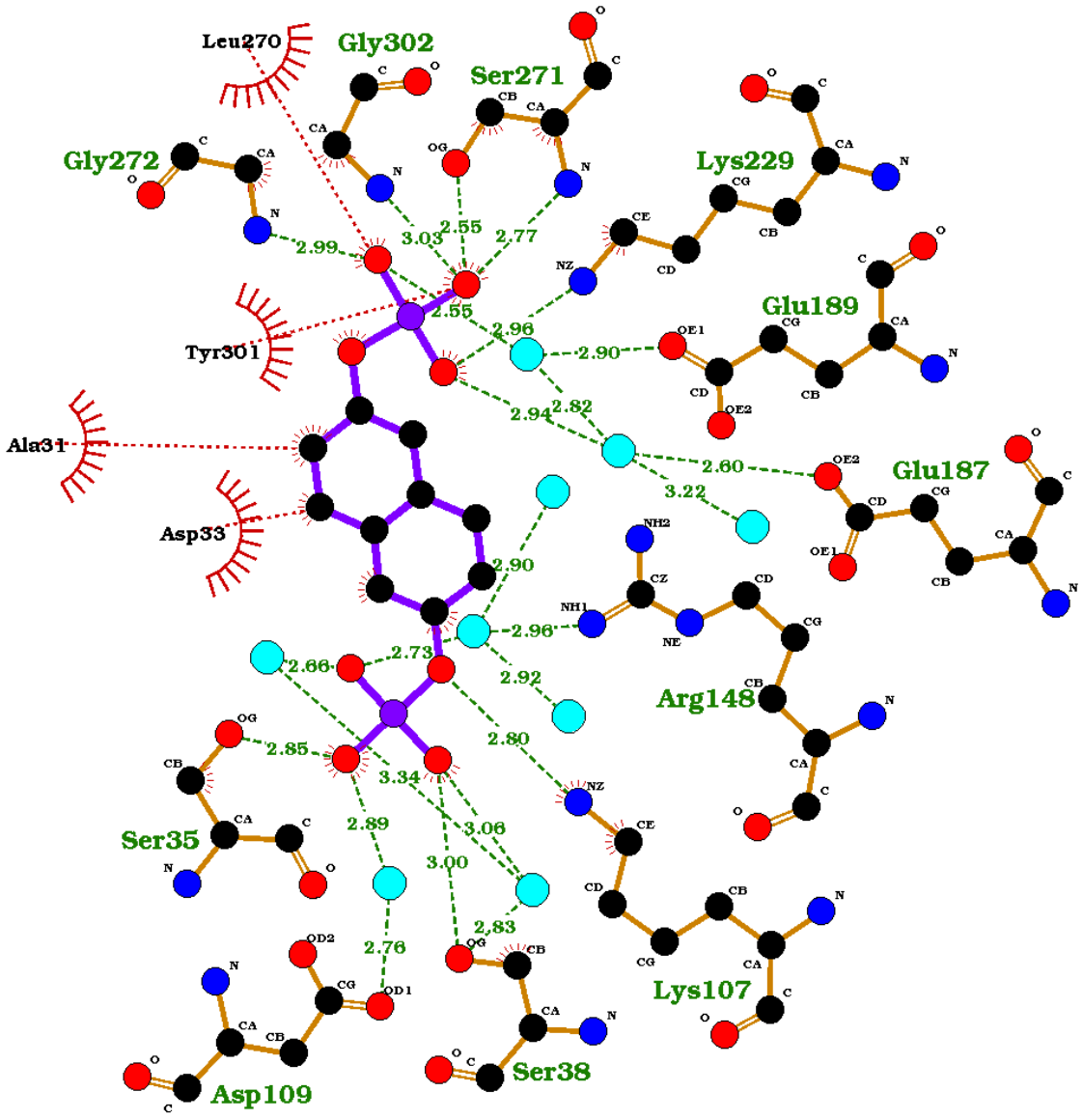


Figure S2. The 2D plot of the binding pocket of ND1 in Aldolase. Intermolecular interactions around negatively charged phosphate groups are marked in green with distances while those hydrophobic ones involved aromatic systems are marked in red. Water molecules in the crystal structure are in cyan.

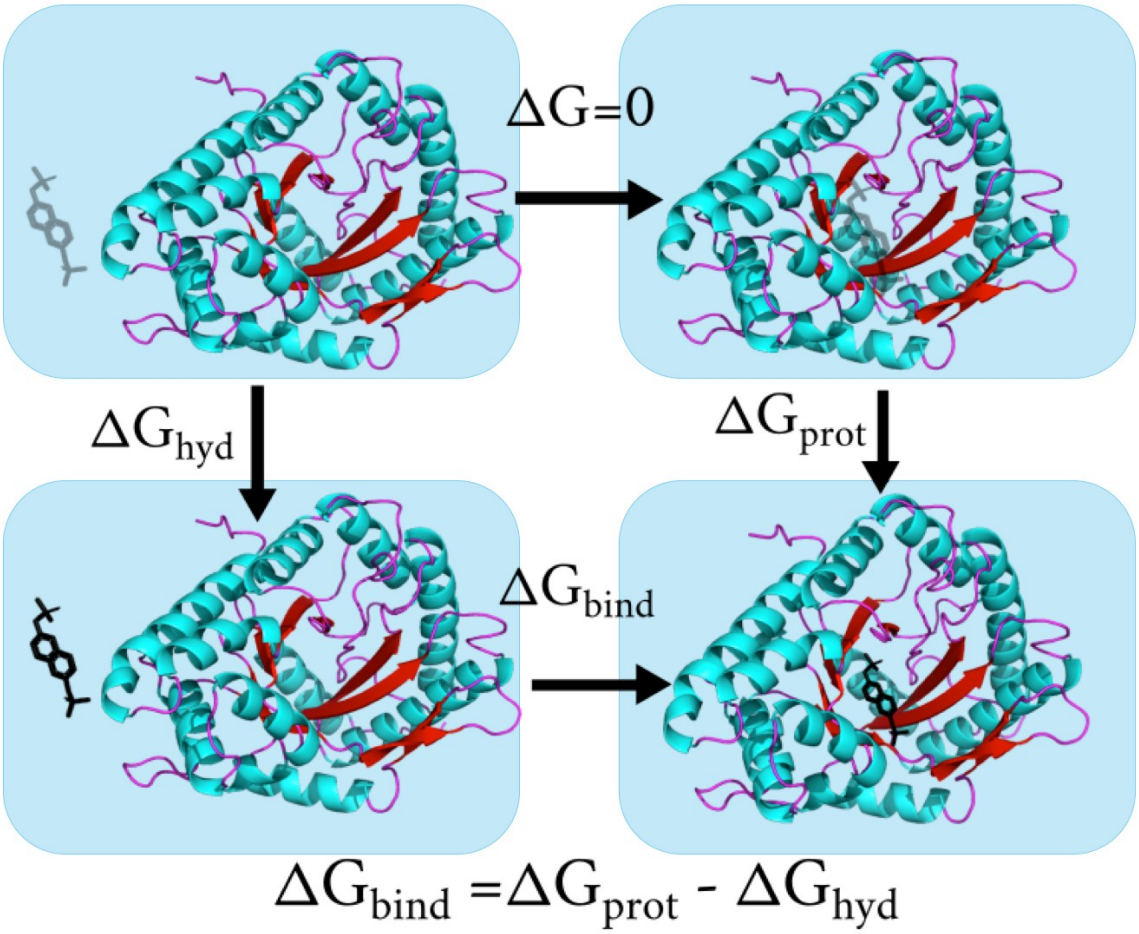


Figure S3. Thermodynamic cycle for calculating the binding free energy of ND1-ALDOLASE binding. The absolute binding free energy (ΔG_{bind}) of the ND1-ALDOLASE binding was calculated by the double-decoupling method. This involvers "disappearing" the ND1 in water and in the protein-ND1 complex.[8] Hence, the binding free energy can be defined as the difference between the decoupling free energies in water (ΔG_{hyd}) and protein environments (ΔG_{prot}).[9-10]

Table S1. Relative binding free energy ($\Delta\Delta G_{bind}$)of all ligands to ND1 including decoupling the ligand from water ($\Delta\Delta G_{hyd}$) and from protein environments ($\Delta\Delta G_{prot}$). All energies are in kcal/mol and the uncertainties are shown in ($\Delta\Delta G_{*}$_err) columns.

Ligands	$\Delta\Delta G_{bind}$	$\Delta\Delta G_{bind\_err}$	$\Delta\Delta G_{hyd}$	$\Delta\Delta G_{hyd\_err}$	$\Delta\Delta G_{prot}$	$\Delta\Delta G_{prot\_err}$	restrain energy
ND5	2.98	-0.04	117.50	-0.07	120.62	-0.02	-0.14
ND7	-1.69	0.17	51.56	-0.22	51.22	0.56	-0.23
NDA_1	-10.06	0.00	81.38	-0.09	71.52	0.09	-0.21

NDA_2	1.20	-0.01	81.38	-0.09	82.79	0.07	-0.21
NDB	-16.12	-0.02	152.52	-0.11	136.69	0.07	-0.29
NDC	0.49	-0.05	-19.58	-0.09	-19.04	-0.01	-0.05
NDD	1.28	0.01	-19.86	-0.06	-18.51	0.09	-0.07
NDE	8.51	0.04	-40.21	-0.07	-31.62	0.18	-0.07

Table S2. Relative energy decomposition results of the non-covalent inhibitors to ND1 including the relative enthalpy (ΔΔH), relative entropy (TΔΔS). All energies are in kcal/mol and the uncertainties are shown in ΔΔH_err column.

Ligands	TΔΔS	ΔΔH	ΔΔH_err
ND5	-14.50	-11.52	19.60
NDA_2	-11.67	-10.47	19.76
ND7	-11.52	-13.22	17.42
NDA_1	-28.26	-38.33	19.40
NDB	27.60	11.48	19.50

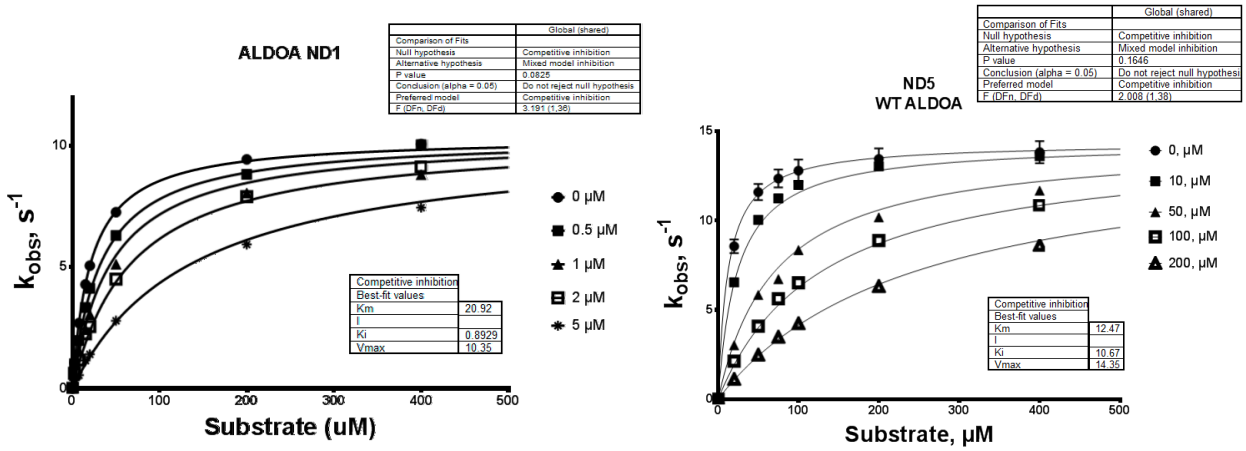


Figure S4. Experimental data of the Competitive inhibitors ND1 and ND5, with the Ki value of 0.89 and 10.67 µM, respectively.

Table S3. Mean distances of the conserved interactions across all derivatives in the Figure 4 left.

	ND1	NDB	ND7	ND1CRY	NDA_2	ND5	NDA_1
LYS229-NZ-O_P1	0.251	0.262	0.285	0.296	0.263	0.275	0.259
SER271-OG-O_P1	0.266	0.476	0.277	0.255	0.275	0.274	0.269
SER271-N-O_P1	0.307	0.302	0.333	0.277	0.298	0.367	0.333
GLY302-N-O_P1	0.289	0.511	0.294	0.303	0.286	0.286	0.284
SER271-CB-O_P1	0.337	0.386	0.331	0.334	0.325	0.357	0.336
SER35-OG-O_P2	0.345	0.313	0.313	0.285	0.310	0.476	0.366
SER35-CB-O_P2	0.341	0.331	0.354	0.315	0.378	0.402	0.329
LYS229-CE-O_P1	0.341	0.349	0.373	0.347	0.362	0.380	0.341
SER300-O-O_P1	0.326	0.350	0.377	0.333	0.382	0.376	0.364
ARG303-N-O_P1	0.298	0.583	0.308	0.458	0.374	0.310	0.329
LYS107-NZ-O_P2	0.456	0.560	0.456	0.311	0.341	0.274	0.294
SER271-CA-O_P1	0.388	0.399	0.411	0.352	0.361	0.416	0.383
LYS229-NZ-P1	0.384	0.345	0.393	0.437	0.395	0.412	0.387
GLY302-CA-O_P1	0.353	0.582	0.346	0.406	0.361	0.357	0.348
SER271-N-P1	0.407	0.422	0.409	0.379	0.374	0.419	0.395
TYR301-CA-O_P1	0.383	0.544	0.415	0.339	0.371	0.382	0.392
SER271-OG-P1	0.388	0.537	0.383	0.358	0.389	0.393	0.382
TYR301-C-O_P1	0.368	0.583	0.398	0.367	0.376	0.366	0.382
ALA31-CB-O_P1	0.321	0.354	0.436	0.422	0.513	0.408	0.453
GLY302-N-P1	0.409	0.554	0.410	0.361	0.390	0.410	0.404
SER35-N-O_P2	0.313	0.325	0.560	0.356	0.546	0.402	0.441
SER35-OG-P2	0.473	0.403	0.401	0.375	0.398	0.499	0.450
GLY302-C-O_P1	0.372	0.636	0.382	0.478	0.417	0.378	0.387
SER300-O-P1	0.387	0.431	0.472	0.413	0.474	0.462	0.444
SER35-CA-O_P2	0.389	0.404	0.498	0.397	0.509	0.448	0.440
LYS107-NZ-P2	0.431	0.547	0.553	0.354	0.424	0.386	0.405
SER38-OG-O_P2	0.279	0.276	0.505	0.300	0.557	0.647	0.577
LYS107-CE-O_P2	0.488	0.640	0.515	0.367	0.430	0.355	0.394
ARG303-CB-O_P1	0.381	0.616	0.385	0.582	0.474	0.417	0.438
ARG303-CA-O_P1	0.395	0.672	0.401	0.568	0.488	0.416	0.437
SER38-CB-O_P2	0.332	0.342	0.565	0.289	0.671	0.581	0.676
SER38-OG-P2	0.367	0.359	0.552	0.450	0.592	0.620	0.566
SER38-OG-G2	0.677	0.607	0.412	0.527	0.545	0.378	0.446
LYS107-NZ-G2	0.369	0.541	0.600	0.497	0.579	0.588	0.567

Table S4. Mean distances of the interactions that are not conserved among all derivatives in Figure4 right. 100 means interaction does not exist for a given derivative.

	ND1	NDB	ND7	ND1CRY	NDA_2	ND5	NDA_1
LYS107-NZ-OB_P2	0.254	0.392	0.509	0.280	0.415	100.000	0.492
ARG303-N-OB_P1	0.405	0.434	0.533	0.371	0.386	100.000	0.590
SER35-OG-OB_P2	0.523	0.554	0.352	0.493	0.435	100.000	0.485
LYS107-CE-OB_P2	0.334	0.493	0.507	0.416	0.533	100.000	0.573
GLY302-CA-OB_P1	0.507	0.519	0.567	0.334	0.356	100.000	0.590
GLY302-N-OB_P1	0.499	0.532	0.579	0.333	0.378	100.000	0.590
SER35-CB-OB_P2	0.535	0.605	0.390	0.523	0.492	100.000	0.414
GLY302-C-OB_P1	0.494	0.497	0.596	0.409	0.394	100.000	0.635
SER38-OG-OB_P2	0.527	0.603	0.398	0.505	0.545	100.000	0.448
ARG303-CB-OB_P1	0.392	0.451	0.560	0.494	0.497	100.000	0.632
ARG303-N-F_CF1	100.000	0.319	100.000	100.000	100.000	0.295	0.397
SER35-OG-F_CF2	100.000	0.366	100.000	100.000	0.356	0.305	100.000
GLY302-CA-F_CF1	100.000	0.366	100.000	100.000	100.000	0.328	0.400
ALA31-CB-F_CF1	100.000	0.417	100.000	100.000	100.000	0.298	0.410
ARG303-CB-F_CF1	100.000	0.348	100.000	100.000	100.000	0.348	0.436
GLY302-N-F_CF1	100.000	0.343	100.000	100.000	100.000	0.373	0.425
SER35-CB-F_CF2	100.000	0.430	100.000	100.000	0.387	0.328	100.000
GLY302-C-F_CF1	100.000	0.380	100.000	100.000	100.000	0.363	0.440
SER35-N-F_CF2	100.000	0.385	100.000	100.000	0.449	0.354	100.000
ARG303-CA-F_CF1	100.000	0.395	100.000	100.000	100.000	0.337	0.470
ARG303-N-CF1	100.000	0.415	100.000	100.000	100.000	0.370	0.457
SER35-OG-CF2	100.000	0.428	100.000	100.000	0.398	0.439	100.000
LYS107-NZ-F_CF2	100.000	0.307	100.000	100.000	0.434	0.566	100.000
LYS107-NZ-CF2	100.000	0.396	100.000	100.000	0.437	0.519	100.000
SER271-OG-F_CF1	100.000	0.364	100.000	100.000	100.000	0.486	0.520
SER38-OG-F_CF2	100.000	0.571	100.000	100.000	0.417	0.390	100.000
SER271-CB-F_CF1	100.000	0.315	100.000	100.000	100.000	0.538	0.552
SER300-O-F_CF1	100.000	0.368	100.000	100.000	100.000	0.509	0.550
SER38-CB-F_CF2	100.000	0.630	100.000	100.000	0.518	0.313	100.000
LYS107-CE-F_CF2	100.000	0.375	100.000	100.000	0.536	0.651	100.000
ALA31-CB-H_CH1	100.000	100.000	0.341	100.000	100.000	100.000	100.000
LYS229-NZ-H_CH1	100.000	100.000	0.359	100.000	100.000	100.000	100.000