UC Berkeley
UC Berkeley Previously Published Works

Title

The Influence of Protein Dynamics on the Success of Computational Enzyme Design

Authors

Ruscio, Jory Z
Kohn, Jonathan E
Ball, K Aurelia
et al.

The Influence of Protein Dynamics on the Success of Computational Enzyme Design

Jory Z. Ruscio[1], Jonathan E. Kohn[1], K. Aurelia Ball[2], Teresa Head-Gordon[1,2*]
[1]*Department of Bioengineering and* [2]*Graduate Group in Biophysics*
University of California, Berkeley
Berkeley, California 94720 USA

We characterize the molecular dynamics of a previously described computational *de novo* designed enzyme optimized to perform a multi-step retrol-aldol reaction when engineered into a TIM barrel protein scaffold. The molecular dynamics simulations show that the protein dynamics under physiological conditions of temperature and aqueous environment distorts the designed geometric factors of the substrate-enzyme reaction intermediates, such that catalysis is limited by the primary retrol-aldol step of proton abstraction from the covalently bound substrate and its interactions with a histidine-aspartate dyad. These results emphasize that computational enzyme designs will benefit from considerations of dynamical fluctuations when optimizing active site geometries.

*Corresponding author
Stanley Hall 274
*510-666-2744 (V)*
TLHead-Gordon@lbl.gov

INTRODUCTION

Recently, the *de novo* computational design of enzymes took a major step forward when Baker and co-workers engineered several motifs of novel catalytic activity towards non-natural substrates into existing protein scaffolds[1,2], and subsequently verified them to be active in the laboratory. Although the experimentally measured activity of these designed enzymes is much lower than that of natural enzymes, the *de novo* structural-based design combined with directed evolution for the single step Kemp elimination reaction[1] realized a five order of magnitude improvement over the uncatalyzed reaction, and comparable catalytic efficiency to the catalytic antibody 34E4[3].

The much more difficult *de novo* design case was for the multi-step retrol-aldol reaction culminating in the carbon-carbon bond breaking of the non-naturally occurring substrate 4-hydroxy-4-(6-methoxy-2-naphthyl)-2-butanone[2] (Figure 1a). In this case the Baker group's structural-based approach was to define an active site geometry composed of protein functional groups in orientations optimized to describe a composite structure of reaction intermediates and transition states for the multi-step reaction[2]. With this approach they determined four different retro-aldol reaction motifs designed to meet the static geometric criteria, and which were engineered into various TIM barrel and jelly-roll protein scaffolds[2]. For example, the so-called Motif III manifested in a TIM barrel construct involved a kinetic mechanism of a nucleophilic lysine attack on the substrate to generate a covalently bound substrate-protein complex (Figure 1b), followed by a histidine-aspartate dyad acting as a general base to abstract a proton from the substrate during the carbon-carbon bond breaking step (Figure 1c). Combined with other mutations relevant for optimizing active site sidechain packing, the resulting so-called RA22 designed enzyme displayed a rate enhancement of ~3 orders of magnitude over the uncatalyzed reaction, but underperformed relative to a comparable catalytic antibody by ~2-3 orders of magnitude[4,5]. The kinetics of RA22 were reported to be complex, showing a burst phase, which was taken as evidence that product inhibition was a limiting feature of the intended design, given that an X-ray crystal structure on a virtually identical RA22 sequence was within ~1Å RMSD of the intended active site geometry[2].
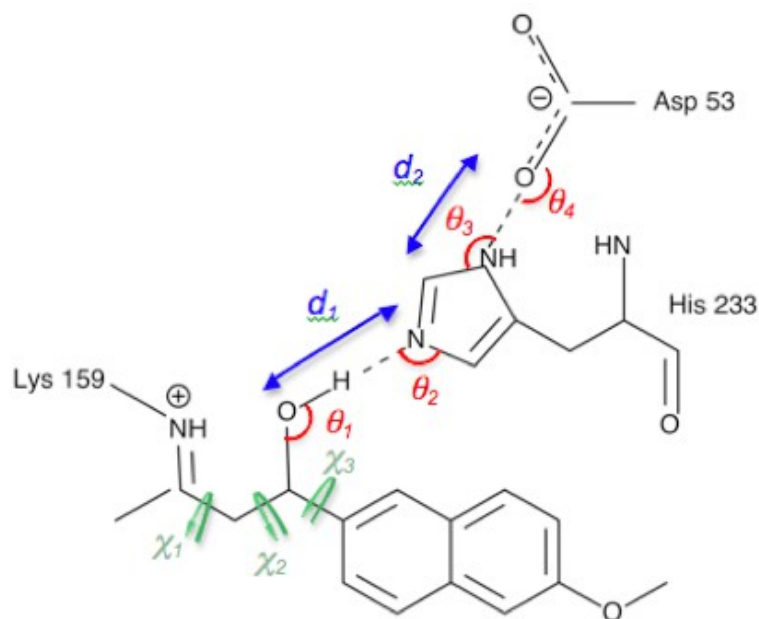
A

RA22 Enzyme

B

Lys 159 —NH₂
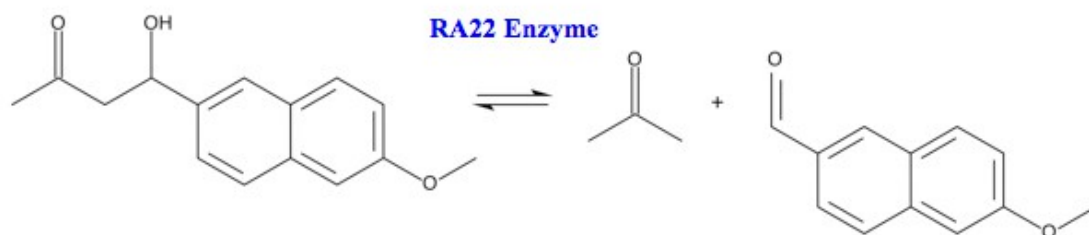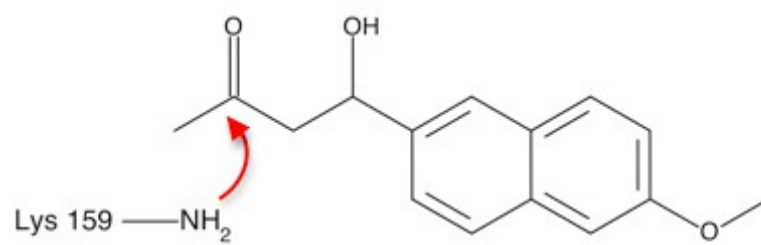
c

Figure 1. *The retro-aldol reaction catalyzed by de novo design enzyme RA22.* (a) The overall carbon-carbon bond-breaking of the substrate 4-hydroxy-4-(6-methoxy-2-naphthyl)-2-butanone, (b) the nucleophilic attack step to create the imine, (c) the geometric criteria for the His-Asp dyad base to abstract the proton to cleave the carbon-carbon bond.

In this work, we focus on a completely different design criterion than static considerations of proximity alone, by asking whether the geometric criteria of the RA22 active site and its variants are satisfied *temporally* during a molecular dynamics simulation. We want to assay the influence of protein structural fluctuations executed at physiological temperatures and under aqueous solvent conditions on realizing the active site architecture and its intended substrate and intermediate interactions. Since our own experiments on RA22 fail to reproduce the burst phase reported by Jiang and co-workers[2], it appears that product inhibition is not limiting the catalytic efficiency. Instead we find that catalysis is inhibited by the primary retrol-aldol step of proton abstraction from the covalently bound substrate, due to dynamical fluctuations that fail to meet the geometric criteria of substrate interactions with the histidine-aspartate dyad on a temporal basis.

EXPERIMENTAL METHODS

RA22 sequences. In our computational and experimental studies we used two sequences reported by the Baker group on RA22. The first was the intended design based on creation of 13 mutations

into the TIM barrel protein scaffold as reported[2], which we refer to as RA22. In showing that the RA22 geometric design criteria in the engineered active site were met with these 13 mutations, the crystal structure of RA22 was reported in 3B5V, with the understanding that an additional 14th mutation involving S210A was to be ignored as that sequence did not have catalytic activity. However, 3B5V contained an additional 15 alanine mutations of scaffold residues lysine, asparagine, and glutamic acid on the surface of the TIM protein; we call this sequence Ala-RA22.

Expression and purification. Codon optimized genes for RA22, and its variant Ala-RA22, were synthesized by GenScript (Piscataway, NJ). The genes were directionally subcloned between *NdeI* and *XhoI* sites in a pET29b+ vector (Novagen, Madison, WI) encoding a carboxy-terminal 6xHis tag. The plasmids were transformed into BL21(DE3) (Stratagene, Inc.) and cultured in LB-broth at 37ºC. Protein expression was induced with isopropyl-β-D-thiogalactopyranoside at $A_{600} = 0.6$ and growth was permitted to continue for an additional 8 hours. The cells were then pelleted by centrifugation for 10 minutes at 500 rpm and resuspended in 50mM potassium phosphate buffer (pH 7) and 200mM NaCl (loading buffer) with an additional 10mM imidazole. The cells were then lysed by French press and the lysate cleared by centrifugation at 15,000 rpm for 1 hour. The cleared lysates were loaded onto a Ni-NTA agarose (Qiagen, Inc.) affinity columns, and washed with 10 column volumes of loading buffer with an additional 40mM imidazole and eluted with a 40-250mM imidazole gradient. The purity of the protein, estimated at greater than 90%, was evaluated by SDS-page. Protein concentration was determined by absorbance at 280nm using a Lambda 650 UV/VIS spectrometer (PerkinElmer, Fremont, CA). The extinction coefficient of 24410 M^{-1} cm^{-1} was calculated from the amino acid sequence using the ProtParam Tool (Swiss Institute of Bioimformatics).

Enzyme kinetics and data analysis. Evaluation of RA22 and Ala-RA22 kinetics was performed as previously described[2]. The retro-aldol reaction of 4-hydroxy-4-(6-methoxy-2-anphthyl)-2-butanone (AsisChem, Cambridge, MA), was carried out in 2.7% acetonitrile, 25mM HEPES, 100mM NaCl (pH 7.5) at 28ºC; 180µl buffer containing 16.44 µM protein and 5µl substrate (1.25, 2.5, 5, 10, 20, and 40mM) in acetonitre. The progress of the reaction was followed using a Spectramax M2 plate reader (Molecular Devices, Sunnyvale, CA) in Costar 96 Well Black Round Bottom Plystyrene Non-Binding Surface Microplates (Cornnig, Lowell, MA). Reactions were monitored over 5 hours,

with λ_{ex} of 330nm and λ_{em} of 452nm and a cutoff filter at 435nm. Measurements of known product concentrations were used to quantify the reaction fluorescence curves.

Reaction rates for RA22 and Ala-RA22 were determined using a linear fit to the fluorescence data. The fluorescence intensity at 452nm was converted to product concentration values using a linear calibration equation, which was based on data collection at known product concentrations. The reaction rate fit was applied over 300 minutes of data for each of seven initial substrate concentrations, ignoring the initial 10 minutes to omit a lag phase we observed for both proteins at most concentrations. The reaction rate data at different substrate concentrations were fit to the Michaelis-Menten equation[21] to obtain the K_m and K_{cat} values for the reaction.

THEORETICAL METHODS

Computational Models. The initial structures for the apo-enzymes were obtained from the RCSB for Ala-RA22 (3B5V.pdb) and the supplementary material from Jiang and co-workers.[2] The protein was modeled with the AMBER ff99SB protein force field[22], and the TIP4P-Ew model[23] was used to describe the molecular solvent. Charges and parameters for the substrate were generated with the Antechamber module of the AMBER package (version 10)[24] using the AM1-BCC charge model and the general AMBER Force Field (GAFF)[25]. Charges and parameters for the COV2 complex were generated with the Antechamber module using the AM1-BCC charge module and the lysine portion of the molecule is described with standard lysine parameters while the substrate portion of the molecule is described with the GAFF. Additionally, we reparameterized the force constant for the dihedral angle of the iminium bond, as the force constant assigned by Antechamber was not strong enough to maintain the double bond character.

We used the sander module in the AMBER package for all of the molecular dynamics simulations[24]. The enzymes were solvated in a box of water and neutralized before undergoing four stages of minimization: three stages with loosening restraints (5.0 kcal/mol/Å^2, 1.0 kcal/mol/Å^2, and 0.1 kcal/mol/Å^2) on the solute and one stage with no restraints. The system then underwent five steps of equilibration. In the first step, under constant volume, the temperature of the system was raised from 0 to 300K using the Anderson thermostat with a 10.0 kcal/mol/Å^2 harmonic restraint on the solute. The next three steps were under NVT conditions while the harmonic restraints were loosened

(5.0 kcal/mol/Å^2, 1.0 kcal/mol/Å^2, and 0.5 kcal/mol/Å^2). Finally, we simulated the system under NPT conditions using a weak barostat coupling at 1 bar for 1 ns. For the MD simulations, we used a 2 fs timestep, with the long-range electrostatic interactions calculated using the Particle Mesh Ewald method and a cutoff of 9.0 Å for real space electrostatics and LJ interactions.

RESULTS

Experimental results

We analyzed the kinetics of the retro-aldol reaction for the RA22 and Ala-Ra22 enzymes as described in Methods. We obtained K_m and K_{cat} values which were within the (large) error bars of the steady state rate reported by Jiang and co-workers[2] (Figure 2a), but not within error of the K_{cat} that they obtained for the burst phase. In fact, for RA22 the K_{cat}/K_m value we obtained, 0.017 ± 0.010 M^{-1} s^{-1}, is very close to the steady-state value reported by Jiang and co-workers, 0.018 ± 0.006 M^{-1} s^{-1}. However, we did not observe a burst phase in the first 20 minutes of our progress curve (Figure 2b). The characteristic double-exponential behavior of the previously reported burst phase can be clearly seen as an abrupt change in slope after 20 minutes in the inset of Figure 3a of the original study[2]. In contrast, we observe a brief lag-phase followed by single-exponential behavior throughout the remainder of the experiment. Therefore we conclude that the RA22 enzyme does not exhibit burst phase kinetics, and thus product inhibition is not a limiting factor for catalysis by this enzyme. Consequently we focused on the designed catalysis steps, and the effect of thermal protein motion, as the likely source of the under performance for RA22 and its variants.
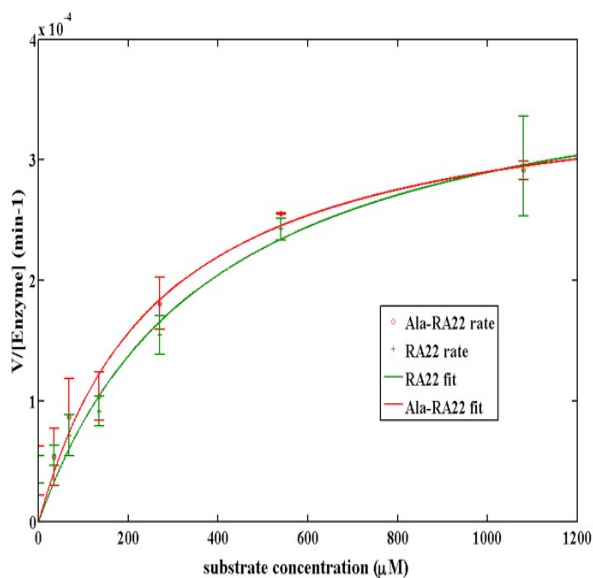
a.
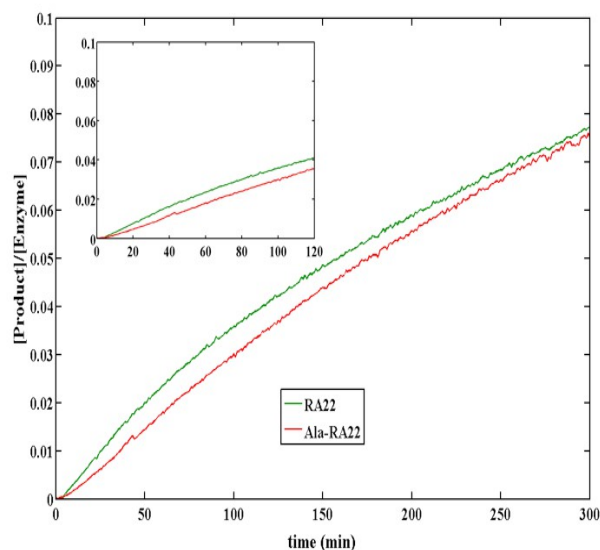
b.

Figure 2. *Retro-aldol reaction catalysis by RA22 (green) and Ala-RA22 (red).* (a) Reaction rate dependence on initial substrate concentration for RA22 and Ala-RA22. The RA22 reaction rate dependence on substrate concentration yields $K_{cat} = 0.4 \pm 0.2$ x 10^{-3} min^{-1} and $K_m = 380 \pm 50$ μM, while Ala-RA22 has $K_{cat} = 0.4 \pm 0.03$ x 10^{-3} min^{-1} and $K_m = 270 \pm 60$ μM. The Michaelis-Menten fit for RA22 and Ala-RA22 reaction rates are indistinguishable within error. (b) Progress curve for RA22 and Ala-RA22 with 540 μM initial substrate concentration. Shown is an average of nine independent progress curves taken at this substrate concentration for RA22 and three progress curves taken for Ala-RA22. Neither protein exhibits a burst phase.

Computational results

We next conducted numerous molecular dynamics simulations of RA22 and Ala-RA22 to gain insight into how the protein and substrate atomic motions affect the performance of this designed enzyme. Primarily we are interested in seeing if the geometric parameters optimized for the retro-aldol reaction under Motif III are satisfied during the time course of molecular dynamics of the RA22 enzymes (Figure 1). Although there are numerous geometric criteria we can follow based on the supplementary material given for RA22[2], it appears that certain geometric criteria exert a more critical influence than others on the reaction rate. These include: (1) the distance between the nucleophilic lysine and the ketone carbon of the unbound substrate; (2) the hydrogen-bond distance and angles between the substrate and His233 in the subsequent covalently bound complex and 3) the hydrogen-bond distance and angles between His233 and Asp53 to form the His-Asp dyad to manifest the general catalytic base. The first geometric criterion is necessary for the nucleophilic attack step that allows the covalent complex to form (Figure 1b), while the latter two geometric criteria are important for the critical second step of proton abstraction preceeding the carbon-carbon bond-breaking of the substrate (Figure 1c).

Nucleophilic attack step. The design of RA22 empolyed quantum chemistry calculations to optimize the orientation of the substrate's napthyl ring to be perpendicular to the imidazole ring of His233 (which we designate O1 in Figure 3a) for the imine formation, since it is consistent with the ligand orientation of the subsequent covalent intermediate complex in the carbon-carbon bond-breaking step. This is an aspect of the static composite structure approach that we question when we consider starting the substrate orientation with the plane of its aromatic groups parallel to the

His233 imidazole ring to optimize π-π stacking (labeled O2 in Figure 3b). We use the criteria of the near attack conformation (NAC) hypothesis[6-8] to assess the frequency of likely nucleophilic attack conformations, and to assess the relative favorability between the O1 and O2 orientations of step 1 in the RA22 design. Specifically, a successful nucleophilic attack for bond formation requires that the lysine ϵ-amino group and the substrate ketone carbon are at a van der Waals contact distance of $d_{NAC} < 3.25$Å, resembling the bond to be formed in the transition state; we also consider a more relaxed criteria of $d_{NAC} < 3.5$Å to account for potential discrepancies of the empirical force field.

The molecular dynamics simulations indicate that the parallel O2 orientation with respect to His233 is optimal for the nucleophilic attack by Lys159. In fact when the simulations are initiated with the O1 conformation, the napthyl ring of the substrate eventually rotates 90° and goes from the initial T-shaped π-π interaction to a parallel O2-like π-π interaction. As this rotation occurs, the substrate hydroxyl group engages in a hydrogen bond with the lysine ϵ-amino, which in turn pulls the lysine amine group further away from the substrate ketone (Figure 3c). The result is that the ϵ-amino group is never within nucleophilic attack distance of the substrate ketone over the course of the simulation when started in the O1 orientation (Table 1). On the other hand, when started in the O2 orientation, the substrate remains as initially positioned throughout the simulation. The O2 orientation is stabilized not only by π-π stacking with His233, but also through additional π-π stacking with Phe211. These combined interactions position the substrate near the lysine amine, such that it forms hydrogen-bonds with the substrate hydroxyl group and ketone oxygen, and satisfying the NAC condition (Figure 3d). By our temporal assay, we find that ~50% of the time the dynamics allows formation of the NAC geometry in the O2 orientation for the less stringent $d_{NAC} < 3.5$Å and ~15-20% of the time with the more stringent $d_{NAC} < 3.25$Å (Table 1).
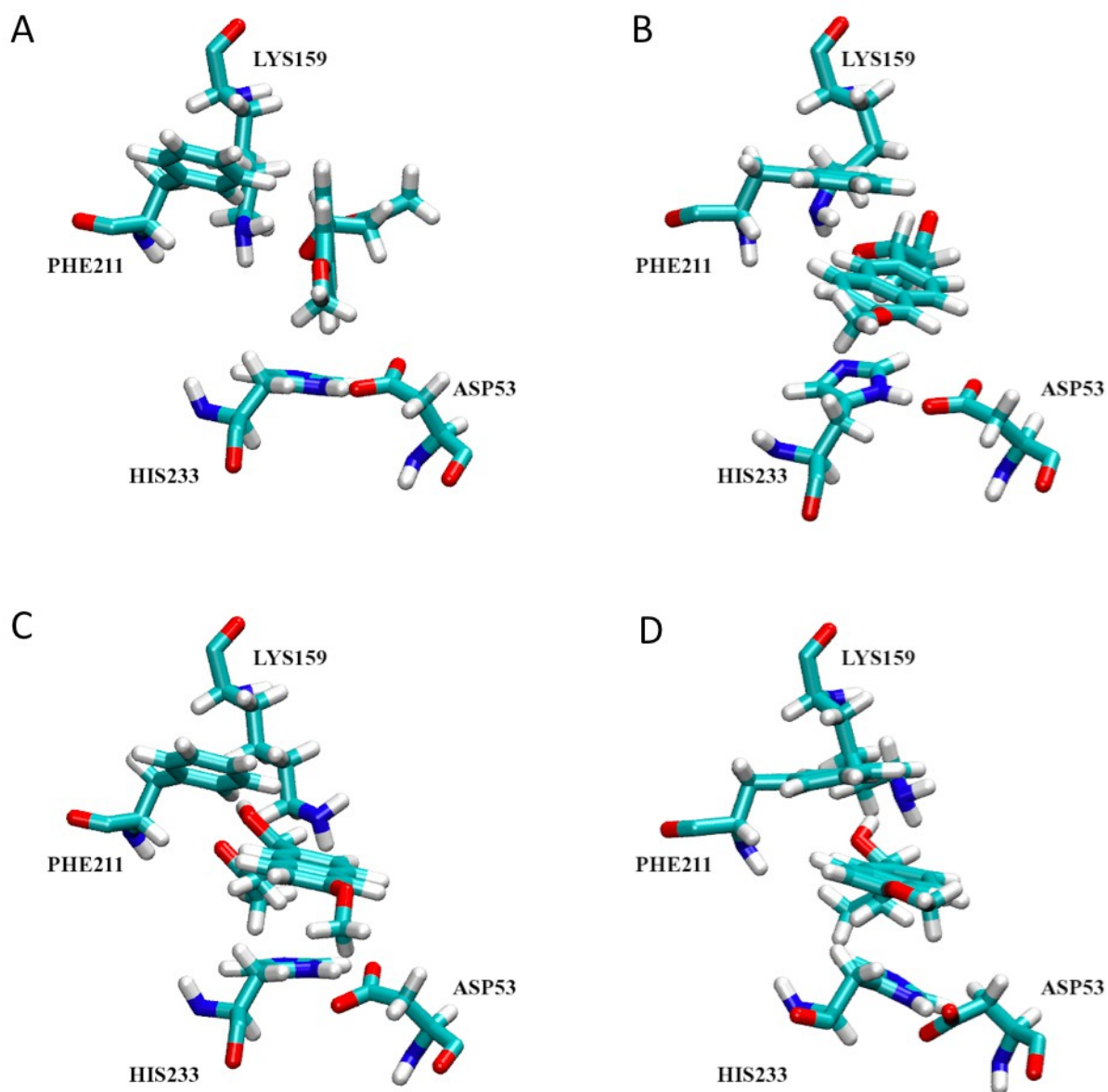
Figure 3. *Orientation of the plane of the napthyl ring of the substrate in the active site pocket of RA22.* Initial orientation at the beginning of the molecular dynamics trajectory for the (a) O1 orientation and (b) O2 orientation. Final orientation at the end of the molecular dynamics trajectory when started from the (c) O1 orientation and (d) O2 orientation.

TABLE 1. The percentage of frames in which the nucleophilic attack conformation is met out of multiple 10ns trajectories (10,000 frames each)

Simulation	Distance < 3.5A	Distance < 3.25A

RA22		
Substrate-O1	3.0%	0.2%
Substrate-O2	46.4%	15.3%
Ala-RA22		
Substrate-O1	1.01%	0.1%
Substrate-O2	54.2%	20.7%

Proton abstraction step. The critical catalysis step in the RA22 enzyme design is the formation of a His233-Asp53 dyad which acts as a general base, that in turn aids the proton abstraction from the covalently bound substrate through its hydrogen-bonded interaction with His233 (Figure 1c). This reaction intermediate sets up the transition state that subsequently breaks the carbon-carbon bond to yield the ketone and napthyl moiety products. Because the composite transition structure was used to optimize both the nucleophilic-attack and the proton abstraction steps of the reaction, a single ligand orientation (O1) guided the design process. Although the O1 substrate orientation is highly disfavored in the nucleophilic attack step, nonetheless it is the orientation that was optimized in the composite transition structure that also accounts for the proton abstraction step[2]. Given that we found that the O2 orientation is the productive nucleophilic attack geometry, we therefore consider both the O1 and O2 substrate orientations of the covalently bound substrate-enzyme complex. In either case, the MD simulations indicate that the RA22 active site fluctuations distorts the optimized geometric factors of this substrate-enzyme reaction intermediate, albeit in different ways.

The formation of the His-Asp dyad is satisfied when $d_2 < 3.0$Å, $110° < \theta_3 < 130°$, and $115° < \theta_4 < 135°$ (Figure 1c). Table 2 shows that the His-Asp dyad geometric criteria are rarely met with the O2 ligand orientation, while they are met ~25% of the time with the O1 orientation. This is because for the majority of the simulation Asp53 is not hydrogen bonded with His233, but is instead interacting with water molecules. The inability of the active site to conform to the geometric constraints of the His-Asp dyad with any frequency leaves the His233 side-chain much floppier, so that it is never able to form the optimal $\theta_1 \sim 110°$ and $\theta_2 \sim 125°$ angles or $d_1 \sim 2.8$Å distance necessary for hydrogen-bonding with the covalently bound substrate. The lack of a hydrogen-bonding constraint means that although individually the χ_1, χ_2, and χ_3 geometric criteria for the lysine-imine complex are met with reasonable frequency, they are rarely met simultaneously. This failure to satisfy the geometric criteria in a concerted fashion - in particular the formation of the his-asp dyad- would account for the poor kinetic performance observed experimentally. It is interesting to note that the

same qualitative conclusions are reached whether considering RA22 or Ala-RA22, which suggests that the scaffold dynamics are very similar or possibly irrelevant to the observed active site dynamics, although, experimentally, Ala-RA22 is extremely prone to aggregation.

TABLE 2. The percentage of frames in which the proton abstraction step is met out of multiple 10ns trajectories (10,000 frames each)

Simulation	His233-Asp59 Dyad	χ_1	χ_2	χ_3	χ_{123}	Substrate-His233 H-bond
RA22						
COV2-O1	26.1%	22.5%	6%	32.6%	0.08%	0.5%
COV2-O2	4.0%	36.4%	18%	38.1%	3.0%	0.3%
Ala-RA22						
COV2-O1	24.2%	19.7%	44.6%	61.8%	4.3%	0.1%
COV2-O2	0.9%	36.0%	0%	6.8%	0%	0.0%

DISCUSSION AND CONCLUSIONS

We have used molecular dynamics to analyze the performance of a *de novo* designed enzyme, RA22, to catalyze a multi-step retrol-aldol reaction acting on a non-naturally occurring substrate. Although the reported evidence of burst phase kinetics would imply that product inhibition is a contributing source of the poor catalytic efficiency for RA22, our own experiments find no evidence of a burst phase and thus we conclude that the intended active site design is not being met for other reasons. In fact, we find that most aspects of the static design are distorted from the optimized composite active site structure, comprising the substrate-enzyme reaction intermediates and transition states, by the aqueous protein fluctuations at room temperature. By contrast, recent quantum and molecular mechanics studies on mechanistic details of the Kemp elimination reaction determined that a majority of the designed enzymes met the design criteria[9], which may be a result of greater design control on a reaction with single step mechanism.

While the intended design positioned the substrate napthyl ring edge to interact with the face of the His233 ring (O1 orientation) to set up the nucleophilic attack by lysine on the ketone portion of the substrate, we found that the substrate finds better energetics through π-π stacking with not just Phe211 (as the original design intended) but with π-π stacking with His233 (O2 orientation). While *ab initio* studies have reached consensus that, at room temperature, the face-to-face and face-edge

geometries are essentially isoenergetic for benzene-benzene interactions[10], the relative energetics of these conformers can change when considering ring heteroatoms such as nitrogen[11], ring substituents[12], or a heterogeneous protein environment[13]. These are highly subtle effects for fixed charge force fields to model with confidence; however, empirical force fields nonetheless do a reasonable job of representing aromatic side chain interactions in the folded state of proteins[14] as well as model systems[15]. This would suggest that reoptimization of the composite active site structure with respect to the O2 orientation may be beneficial for improving the catalytic design.

While the O2 orientation permits a near attack distance $3.25Å < d_{NAC} < 3.5Å$ to be met with reasonable frequency in the nucleophilic attack step, we find that the O1 orientation is preferred over O2 for the proton abstraction step from the covalently bound substrate, consistent with the intended design. However even in the O1 orientation, competing interactions with water and the active site dynamics prevented Asp53 from forming the His233-Asp53 dyad, which in turn propagates through the covalently bound complex so that His233 rarely forms a hydrogen-bond with the imine-lysine complex. We attribute the under performance of the RA22 enzyme to the dynamical distortions of the needed base for proton abstraction at this critical enzymatic step.

Naturally occurring enzymes accelerate rates over uncatalyzed reactions by 8-9 orders of magnitude, with catalytic efficiencies as measured by K_{cat}/K_m that are typically 10^7-10^8 with some exceptional enzymes yielding K_{cat}/K_m upwards of 10^{14}. Introduction of novel activity and enhancement of native activity in naturally occurring enzymes is difficult because native enzyme catalysis and specificity are already well optimized through natural selection. Computational *de novo* protein design[16-19], when applied to the development of enzyme catalysis of chemical reactions that are not part of nature's repertoire, resets the performance bar to something more manageable- by serving as a plausible beginning for subsequent laboratory directed evolution that will enhance performance through rapid exploration of sequence space using mutation and recombination. This was demonstrated explicitly for a single step Kemp elimination reaction, which took the underperforming computational *de novo* designed enzymes with efficiencies that ranged over $K_{cat}/K_m \sim 10^1$-10^2, and under directed evolution increased their performance by an order of magnitude[1].

Because the laboratory directed evolution mutations primarily occurred in the near satellite regions of the designed active site for Kemp elimination[1,5], it implicitly suggests that the non-designed mutations were able to fine tune the ability of the active site geometries to meet the design criteria on a temporal basis. This may be an optimal strategy for future protein engineering strategies[20]. However, we believe that directed evolution may be inhibited inherently by the performance of an enzyme built from scratch based on a static design that only optimizes spatial proximity but not dynamical congruence. For example recent NMR and MD studies of native TIM showed that the active site loop residue dynamics were highly correlated and necessary for catalytic activity[26]. We suggest that computational models and methods are up to the task to incorporate dynamics into the early design stage, an alternative approach to protein engineering that deserves further consideration.

SUPPORTING INFORMATION AVAILABLE. Complete reference 24 is given. This information is available free of charge via the Internet at http://pubs.acs.org/.

REFERENCES

1. Röthlisberger, D.; Khersonsky, O.; Wollacott, A. M.; Jiang, L.; DeChancie, J.; Betker, J.; Gallaher, J. L.; Althoff, E. A.; Zanghellini, A.; Dym, O.; Albeck, S.; Houk, K. N.; Tawfik, D. S.; Baker, D. *Nature* **2008**, 453, 190-195.

2. Jiang, L.; Althoff, E. A.; Clemente, F. R.; Doyle, L.; Röthlisberger, D.; Zanghellini, A.; Gallaher, J. L.; Betker, J. L.; Tanaka, F.; Barbas III, C. F.; Hilvert, D.; Houk, K. N.; Stoddard, B. L.; Baker, D. *Science* **2008**, 319, 1387–1391.

3. Hu, Y.; Houk, K. N.; Kikuchi, K.; Hotta, K.; Hilvert, D. *J. Am. Chem. Soc.* **2004**, 126, 8197–8205.

4. List, B.; Barbas III, C. F.; Lerner, R. A. *Proc. Natl. Acad. Sci. USA* **1998**, 95, 15351–15355.

5. Sterner, R.; Merkl, R.; Raushel, F. M. *Chemistry & Biology* **2008**, 15, 421-423.

6. Hur, S.; Bruice, T. C. *Proc. Nat. Acad. Sci. USA* **2003**, 100, 12015–12020.

7. Bruice, T. C. *Acc. Chem. Res.* **2002**, 35, 139–148.

8. Scheiner, S.; Lipscomb, W. N.; Kleier, D. A. *J. Am. Chem. Soc.* **1976**, 98, 4770-4777.

9. Alexandrova, A. N.; Röthlisberger, D.; Baker, D.; Jorgensen, W. L. *J. Am. Chem. Soc.* **2008**, 130 15907-15915.

10. Sinnokrot, M. O.; Valeev, E. F.; Sherrill, C. D. *J. Am. Chem. Soc.* **2002**, 124, 10887-10893.

11. Hohenstein, E. G.; Sherrill, C. D. *J. Phys. Chem. A* **2009**, 113, 878-886.

12. Sinnokrot, M. O.; Sherrill, C. D. *J. Am. Chem. Soc.* **2004**, 126, 7690-7697.

13. Chelli, R.; Gervasio, F. L.; Procacci, P.; Schettino, V. *J. Am. Chem. Soc.* **2002**, 124, 6133-6143.

14. Morozov, A. V.; Misura, K. M. S.; Tsemekhman, K.; Baker, D. *J. Phys. Chem. B* **2004**, 108, 8489-8496.

15. Berka, K.; Laskowski, R.; Riley, K. E.; Hobza, P.; Vondrek, J. *J. Chem. Theor. Comp.* **2009**, 5, 982-992.

16. Bolon, D. N.; Mayo, S. L. *Proc. Nat. Acad. Sci. USA* **2001**, 98, 14274-14279.

17. Bolon, D. N.; Voigt, C. A.; Mayo, S. L. *Curr. Opin. Chem. Biol.* **2002**, 6, 125-129.

18. Kaplan, J.; DeGrado, W. F. *Proc. Nat. Acad. Sci. USA* **2004**, 101, 11566–11570.

19. Kuhlman, B.; Dantas, G.; Ireton, G. C.; Varani, G.; Stoddard, B. L.; Baker, D. *Science* **2003**, 302, 1364 – 1368.

20. Nanda, V. *Nature Chem. Bio.* **2008**, 4, 273 – 275.

21. Michaelis, L.; Menten, M. L. *Biochemische Zeitschrift* **1913**, 49, 333-369.

22. Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. *Proteins-Structure Function and Bioinformatics* **2006**, 65, 712-725.

23. Horn, H. W.; Swope, W. C.; Pitera, J. W.; Madura, J. D.; Dick, T. J.; Hura, G. L.; Head-Gordon, T. *J. Chem. Phys.* **2004**, 120, 9665-9678.

24. Case, D.A.; Darden, T.A.; Cheatham III, T.E.; Simmerling, C.L.; Wang, J.; et al. *AMBER 10*, University of California, San Francisco (2008).

25. Wang, J.; Wolf, R.M.; Caldwell, J.W.; Kollman, P. A.; Case, D. A. *J. Comput. Chem.* **2004**, 25, 1157-1174.

26. Massi, F.; Wang, C.; Palmer, III, A. G. *Biochemistry* **2006**, 45, 10787–10794.