Chapter I

Introduction

Enzymes are awesome

Enzymes are extremely efficient catalysts, accelerating chemical reaction rates up to 10^{19} times that of the uncatalyzed reaction.[1] In addition to the large rate enhancements that can be observed, enzymes can also carry out reactions with extreme regio- and stereospecificity, eliminating the need for protecting groups and reliably producing a single product.[2,3] In the face of concern over the environmental impact of chemical synthesis, enzymes have emerged as attractive alternatives to chemical catalysts because they work under mild, aqueous conditions, reducing the generation of hazardous wastes that are often associated with organic synthesis.[2] The enzymes themselves are, of course, biodegradable and can usually be produced in large quantities via recombinant expression in bacteria or fungi.

Despite their promise, significant challenges prevent the widespread use of enzymes as industrial catalysts. The applicability of enzymes can be limited by their instability in conditions appropriate for industrial processes, including high temperatures, extreme pHs, and organic solvents.[2] In addition, the scope of reactions that can be catalyzed by enzymes is limited to those found in natural metabolic pathways, although some enzymes, including lipases, have been shown to be somewhat promiscuous in their substrate specificity.[4] Directed evolution has been used to improve stability, optimize efficiency, and modify the substrate specificity of many enzymes.[5-7] However, this

method requires existing activity toward the reaction of interest, so it cannot be used to introduce truly novel chemistries.[5]

Catalytic antibodies have shown promise in the catalysis of a wide variety of chemical transformations including stereoselective and novel reactions.[8] However, the catalytic efficiencies of antibodies have traditionally been modest due in part to selections that are based on binding to a synthetic transition state analog instead of on enzymatic turnover.[8,9] Reactive immunization has been used to overcome these limitations, producing catalytic antibodies with reaction rates approaching those of the wild-type enzyme, but this method can only be applied to reactions whose transition states are known and can be readily mimicked with a reactive transition state analog that is accessible by current synthetic methods.[10,11] In addition, Xu *et al*. has suggested that the immunoglobulin scaffold itself may limit the scope of reactions amenable to catalysis with antibodies.[8]

While directed evolution and catalytic antibodies have both been successfully used for enzyme engineering, both have features that keep them from being applied generally for the engineering of novel enzymatic activities. In contrast, computational protein design does not suffer from these limitations, and can be envisioned as a solution to many complex synthetic organic chemistry problems.

Computational protein design

Computational protein design has shown great promise for developing novel functions in proteins. The general approach to solving a computational design problem is cyclic (Figure 1-1). Beginning with the backbone coordinates of a high-resolution

protein crystal structure, an optimization algorithm is used to search through combinations of side-chain identities and conformations for the sequence and geometries of amino acids that will best stabilize the protein fold.[12,13] The extent to which a particular sequence might stabilize the desired fold is evaluated through a force field scoring function, which calculates an energy for the sequence that should correlate to the protein's free energy of folding. The sequence and conformation of amino acids that will best stabilize the backbone structure (i.e., fold into the desired conformation) is assumed to be the one with the lowest energy. This global minimum energy conformation (GMEC) must then be experimentally validated by structural and/or thermodynamic comparison of the designed sequence to that of the native protein.[12-14] Information gained from evaluating the deviations of theory from experiment can then be used to readjust the force field parameters, thus completing the design cycle.[12]

For a small 59-residue protein, there are about 10^{77} possible sequences (assuming that all 20 amino acids are allowed at all positions). If a single molecule of each of these sequences were to actually be synthesized, their combined mass would be approximately 7.6×10^{56} g, which according to some estimates, approaches the mass of the observable universe. When the conformational flexibility of the side chains is also taken into account, the number of possible solutions explodes even further. To reduce the combinatorial complexity of the problem to a reasonable size, we limit our designs to use a library of discrete sidechain conformations called rotamers, which represent the statistically significant amino acid sidechain conformations found in protein crystal structures.[15] In results described later in this text, we used rotamer libraries based on

those developed by Dunbrack and Karplus,[16] as well as sidechain conformer libraries developed in the Mayo lab.[17]

The ORBIT (Optimization of Rotamers By Iterative Techniques) software suite is a computational protein design package developed in the Mayo lab.[13] Standard implementations of ORBIT use a scoring function based on physical principles and apply the DREIDING force field,[18] which incorporates four empirically based potential functions to calculate the total energy (E_{total}) of a structure:

$$E_{total} = E_{vdw} + E_{h\text{-}bond} + E_{elect} + E_{as} \quad .$$

(1) A van der Waals (VDW) interaction energy (E_{vdw}) is calculated for each pair of rotamers using a Lennard-Jones 12-6 potential.[18]

(2) A hydrogen bond potential ($E_{h\text{-}bond}$) is used that is angle-, distance-, and hybridization-dependent.[19]

(3) Electrostatic interactions (E_{elect}) are calculated based on Coulomb's Law incorporating a distance-dependent dielectric of 40r, where r is the interatomic distance.[19]

(4) A solvation term (E_{as}) is used that employs a solvation potential based either on the protein's surface area or the occlusion of one atom by another. Both of these solvation models give an energy benefit to buried nonpolar regions of the protein and penalize exposed nonpolar and buried polar regions.[20,21]

The optimization algorithms provided by ORBIT apply a variety of methods to establish the optimal sequence or set of sequences to stabilize a given fold. Algorithms based on the Dead-End Elimination theorem (DEE)[22-25] are used to quickly identify and remove amino acid rotamers and pairs of rotamers that cannot exist in the GMEC.

ORBIT also supplies stochastic methods of sequence optimization such as those based on the Fast and Accurate Side-Chain Topology and Energy Refinement (FASTER) algorithm[15] and Monte Carlo[26,27] either alone or in combination with DEE to both decrease calculation time and to sample sequence space around the GMEC.

ORBIT has been used to design proteins in a wide variety of systems. Successful implementations include the full sequence design of a protein that adopts a zinc finger fold independent of zinc binding,[28] the redesign of calmodulin to increase its binding specificity for a single target peptide,[29] and the *de novo* design of a protein-protein interface.[30]

Computational enzyme design

Promisingly, ORBIT has also been used to design an enzyme with modest catalytic activity.[31] This "protozyme" with p-nitrophenol acetate hydrolysis activity with a k_{cat}/k_{uncat} of 10^2 was one of the first examples of a *de novo* computationally designed enzyme. Other labs have also employed computational tools to design enzymes, including transplanting reactive metalloenzyme active sites into inert proteins.[32,33] In addition, computational enzyme design methodologies have been used to switch the specificity of existing enzymes.[34,35] More recent dramatic successes from one of these labs include the *de novo* design of retroaldolases as well as enzymes that catalyze the Kemp elimination, a reaction for which no natural enzyme exists.[36,37]

The promise of computational enzyme design has been clearly established. However, the generalizability of these methods for other chemical transformations has not yet been demonstrated. In addition, while some of the enzymes designed so far have

had impressive catalytic activity, the computational design of enzymes with true native-like efficiency still presents a challenge.[36,37] The goal of engineering "designer enzymes" for any reaction remains extremely attractive as it will allow the scope of enzymatic reactions to extend beyond the limits of natural cellular metabolism, broadening the range of possible substrates and products, especially those with stereogenic centers. Once this challenge has been met and the generality of computational protein design techniques has been established, rapid, on-demand engineering of enzymes will be possible for many important chemical reactions.

Towards this goal, my work has focused on the introduction and evaluation of new enzyme design capabilities both in ORBIT and in a related program, Phoenix. Force field parameterization in protein design has historically been carried out with respect to protein stability and overall fold without regard to specific function. In the case of ORBIT, parameters for the DREIDING force field were optimized through the sequence design of a small protein and the subsequent experimental evaluation of changes in overall thermodynamic stability.[12,38] The resulting parameters were weighted to emphasize VDW contacts and buried hydrophobic surface area. While these parameters were successfully used to design many hyperstable proteins, they are not necessarily well suited for designing enzymes because most natural enzymes are not evolved for optimum stability.[39,40]

According to the transition state theory, enzymes achieve such large rate enhancements through specific tight binding and stabilization of the reaction transition state (TS).[41] Computational simulations of enzyme active sites have suggested that polar and nonpolar residues that contact the TS but are not directly involved in the reaction

chemistry can provide additional stabilization to the TS through electrostatics and VDW interactions, thus promoting catalysis.[42] To design an effective enzyme *de novo*, we must first determine the nature of the interactions between the protein and the rate limiting TS that promote catalysis; information about these interactions (e.g., sidechain functional groups and contact geometries) can be gained from *ab initio* calculations, analogy to existing enzymes, or chemical intuition.[17,43] These protein-TS contacts can then be incorporated into the scoring function.[17,44]

The major changes made to ORBIT to accommodate enzyme design include the introduction of a geometry biasing term that allows an energetic benefit to be added to those sequences that can make specified stabilizing contacts to a TS model present in the active site. In addition, we have implemented various methods for the creation of libraries of transition state poses within the active site that can be sampled during the sequence search.[17]

Chapter II of this thesis is a journal article that I co-authored describing our computational enzyme design methodology in detail. Here, our methods were evaluated through recapitulation of the active site configurations of three natural enzymatic/binding protein systems: *Escherichia coli* chorismate mutase, *Saccharomyces cerevisiae* triosephosphate isomerase, and *Streptomyces avidinii* streptavadin.

As a result of our previous *de novo* design experience, we chose to focus on a well-studied chemical system: the general base-catalyzed Kemp elimination (KE) of 5-nitrobenzisoxazole (Figure 1-2). The KE has been used since the 1970s as a physical organic model for proton transfer from carbon[45,46] and more recently as a model system for enzymatic proton transfer reactions.[47-49] Other attractive features of this reaction are

that it is irreversible with a single transition state and that it has a product that can be observed spectrophotometrically (λ_{max} = 405 nm).[45] In addition, multiple catalytic antibodies have been created that can catalyze this reaction with rate accelerations up to 10^6 times faster than the background reaction.[50,51] The crystal structure of one of these antibodies has given us clues as to how transition state stabilization can be achieved for this reaction (Figure 1-3).[52] In this catalytic antibody, stabilization occurs through a combination of a carboxylate general base, extensive π-stacking above and below the plane of the ring system, and hydrogen bond contacts to the base. In addition, there is precedent for the amenability of this reaction to computational enzyme design methodologies, as Röthlisberger *et al.* were able to computationally introduce catalytic activity for the KE into three separate inert scaffolds, creating multiple active enzymes.[37]

In Chapter III, the computational and experimental details of this system are described and one of the resulting inactive designs is discussed. In the protein design cycle (Figure 1-1), the design procedure cannot be adjusted without some information from the initial inactive design indicating the possible cause of inactivity. To complete the design cycle, we first had to determine why this initial design was inactive. This chapter also includes details of crystallographic analysis carried out in collaboration with the Molecular Observatory at Caltech and molecular dynamic (MD) simulation studies carried out in collaboration with Ken Houk's lab at the University of California, Los Angeles that were used to analyze the inactive design. The X-ray crystal structure of this inactive design confirmed that the actual active site of the design was very similar to the predicted structure. Thus, the inactivity was not due to gross misplacement of the active site residues or disruption of the overall protein fold. MD analysis of the design helped

us to determine possible causes of the inactivity including an active site that was too flexible and solvent exposed.

The lessons learned from our first KE designs pointed us in the direction of more buried, less polar active sites. Using the same scaffold as the initial design, we moved the active site away from the natural, solvent-exposed active site and located it farther into the barrel of the protein. In Chapter IV, I discuss this new design, HG-2, which was predicted to have activity by the blind MD simulations due to the drier, less flexible active site. This activity was confirmed by experimental characterization of this enzyme, also discussed in Chapter IV.

Because of the expense and time associated with experimental evaluation of designed enzymes, strategies for the *a priori* differentiation of active designs from inactive ones are needed to make the process of enzyme design more efficient. In Chapter IV, additional designs are described which were carried out using two scaffolds that have been used to produce successful KE designs in David Baker's lab at the University of Washington.[37] Of the six enzymes synthesized, four showed activity and three of these resulted from the evaluation our enzyme design methods through redesigning the active site of scaffolds used in the active KE designs from the Baker lab. The fourth active design is unique. In most cases, blind MD analysis of these designs was successful in distinguishing active designs from inactive ones. MD analysis could thus serve as an important tool in the computational design procedure, providing an initial screen of sequences predicted by the design procedure to help us determine the designs on which to focus our experimental efforts.

Appendix A describes an early attempt at *de novo* design of enzymatic activity. Here, our goal was to design an enantioselective enzyme for the kinetic resolution of of N-benzoyl-L-phenylalanine through the selective hydrolysis of L-2-phenyl-4-benzylphenyloxazolin-5-one (FOX). Lessons learned from this first unsuccessful attempt at enzyme design led us to focus on less flexible scaffolds and chemical systems that have a smaller background reaction rate. In addition to the *de novo* enzyme design project that has spanned my entire graduate career, I have had the opportunity to work on other computational design projects related to enzymes and binding proteins. In Appendix B, I discuss our ongoing efforts to alter the specificity of an existing thermophilic xylanase. Appendix C presents computational efforts toward changing the specificity of an androgen receptor as part of a collaboration with the Fletterick lab at the University of California, San Francisco. Appendix D is the first description of the recombinant over-expression and purification of a thermophilic xylanase from *Thermoascus aurantiacus* (TAX). A recombinant version of this enzyme was necessary to allow genetic manipulation in the creation of new designs and TAX was used as the scaffold for the initial inactive KE design and one of the subsequent active designs. Because of its ease of expression, thermostability, and ability to tolerate multiple mutations, this enzyme proved to be a useful scaffold for computational design.

In sum, the work presented in this thesis shows that by iterative structural and theoretical evaluation of active and inactive designs, adjustment of our enzyme design procedure, and subsequent redesign, we can identify and address deficiencies in our design methodology, resulting in *de novo* designed enzymes with significant activity for the reaction of interest. The field of *de novo* computational enzyme design is extremely

promising and the work presented here is a significant step towards the goal of a general method for the computational design as enzymes in the Mayo lab. This work will serve as the foundation for future studies, which will be undertaken to obtain reaction rate accelerations and efficiencies comparable to those of natural enzymes and to generalize these methods for a wide variety of chemistries.

References

1. Wolfenden, R.; Snider, M. J., The depth of chemical time and the power of enzymes as catalysts. *Acc. Chem. Res.* **2001**, *34*, 938-945.

2. Wong, C.-H.; Whitesides, G. M., *Enzymes in Synthetic Organic Chemistry*. Elsevier Science & Technology Books: Tarrytown, NY, **1994**; Vol. 12.

3. Schmid, A.; Dordick, J. S.; Hauer, B.; Kiener, A.; Wubbolts, M.; Witholt, B., Industrial biocatalysis today and tomorrow. *Nature* **2001**, *409*, 258-268.

4. Bornscheuer, U.; Kazlauskas, R., Catalytic promiscuity in biocatalysis: using old enzymes to form new bonds and follow new pathways. *Angew. Chem. Int. Ed.* **2004**, *43*, 6032-6040.

5. Shao, Z.; Arnold, F. H., Engineering new functions and altering existing functions. *Curr. Opin. Struct. Biol.* **1996**, *6*, 513-518.

6. Williams, G. J.; Nelson, A. S.; Berry, A., Directed evolution of enzymes for biocatalysis and the life sciences. *Cell. Mol. Life Sci.* **2004**, *61*, 3034-3046.

7. Horsman, G. P.; Liu, A. M. F.; Henke, E.; Bornscheuer, U. T.; Kazlauskas, R. J., Mutations in distant residues moderately increase the enantioselectivity of Pseudomonas fluorescens esterase towards methyl- 3-bromo-2-methylpropanoate and ethyl 3-phenylbutyrate. *Chem. Eur. J.* **2003**, *9*.

8. Xu, Y.; Yamamoto, N.; Janda, K. D., Catalytic antibodies: hapten design strategies and screening methods. *Biorg. Med. Chem.* **2004**, *12*, 5247-5268.

9. Hilvert, D., Critical analysis of antibody catalysis. *Annu. Rev. Biochem.* **2000**, *69*, 751-793.

10. Wirshing, P.; Ashley, J. A.; Lo, C. H. L.; Janda, K. D.; Lerner, R. A., Reactive immunization. *Science* **1995**, *270*, 1775-1782.

11. Barbas III, C. F.; Heine, A.; Zhong, G.; Hoffmann, T.; Gramatikova, S.; Bjornestedt, R.; List, B.; Anderson, J.; Stura, E. A.; Wilson, I. A.; Lerner, R. A., Immune versus natural selection: Antibody aldolases with enzymatic rates but broader scope. *Science* **1997**, *278*, 2085-2092.

12. Dahiyat, B. I.; Mayo, S. L., Protein design automation. *Protein Sci.* **1996**, *5*, 895-903.

13. Street, A. G.; Mayo, S. L., Computational protein design. *Structure* **1999**, *7*, 105-109.

14. Gordon, D. B.; Marshall, S. A.; Mayo, S. L., Energy functions for protein design. *Curr. Opin. Struct. Biol.* **1999**, *9*, 509-513.

15. Dunbrack, R. L. J., Rotamer libraries in the 21st century. *Curr. Opin. Struct. Biol.* **2002**, *12*, 431-440.

16. Dunbrak, R. L.; Karplus, M., Backbone-dependent rotamer library for proteins. *J. Mol. Biol.* **1993**, *230*, 543-574.

17. Lassila, J. K.; Privett, H. K.; Allen, B. D.; Mayo, S. L., Combinatorial methods for small-molecule placement in computational enzyme design. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 16710-16715.

18. Mayo, S. L.; Olafson, B. D.; Goddard III, W. A., DREIDING: A generic force field for molecular simulations. *J. Phys. Chem.* **1990**, *94*, 8897-8909.

19. Dahiyat, B. I.; Gordon, B.; Mayo, S. L., Automated design of the surface positions of protein helices. *Protein Sci.* **1997**, *6*, 1333-1337.

20. Street, A. G.; Mayo, S. L., Pairwise calculation of protein solvent-accessible surface areas. *Fold Des.* **1998**, *3*, 253-258.

21. Lazaridis, T.; Karplus, M., Effective energy functions for protein structure prediction. *Curr. Opin. Struct. Biol.* **2000**, *10*, 139-145.

22. Desmet, J.; De Maeyer, M.; Hazes, B.; Lasters, I., The dead-end elimination theorem and its use in protein side-chain positioning. *Science* **1992**, *356*, 539-542.

23. Gordon, D. B.; Mayo, S. L., Radical performance enhancements for combinatorial optimization algorithms based on the dead-end elimination theorem. *J. Comput. Chem.* **1998**, *19*, 1505-1514.

24. Goldstein, R. F., Efficient rotamer elimination applied to protein side-chains and related spin-glasses. *Biophys. J.* **1994**, *66*, 1335-1340.

25. Pierce, N. A.; Spriet, J. A.; Desmet, J.; Mayo, S. L., Conformational splitting: a more powerful criterion for dead-end elimination. *J. Comput. Chem.* **2000**, *21*, 9999-1009.

26. Metropolis, N.; Rosenbluth, A. W.; Rosenbluth, M. N.; Teller, A. H., Equation of state calculations by fast computing machines. *J. Chem. Phys.* **1953**, *21*, 1087-1092.

27. Kirkpatrick, S.; Gelatt, C. D.; Vecchi, M. P., Optimization by simulated annealing. *Science* **1983**, *220*, 671-675.

28. Dahiyat, B. I.; Mayo, S. L., De novo protein design: fully automated sequence selection. *Science* **1997**, *278*, 82-87.

29. Shifman, J. M., Mayo, S. L., Modulating calmodulin binding specificity through computational protein design. *J. Mol. Biol.* **2002**, *323*, 417-423.

30. Huang, P. S.; Love, J. J.; Mayo, S. L., A de novo designed protein protein interface. *Protein Sci.* **2007**, *16*, 2770-2774.

31. Bolon, D. N.; Mayo, S. L., Enzyme-like proteins by computational protein design. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 14274-14279.

32. Benson, D. E.; Wisz, M. S.; Hellinga, H. W., Rational design of nascent metalloenzymes. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 6292-6297.

33. Benson, D. E.; Haddy, A. E.; Hellinga, H. W., Converting a maltose receptor into a nascent binuclear copper oxygenase by computational design. *Biochemistry* **2002**, *41*, 3262-3269.

34. Ashworth, J.; Havranek, J. J.; Duarte, C. M.; Sussman, D.; Monnat, R. J.; Stoddard, B. L.; Baker, D., Computational redesign of endonuclease DNA binding and cleavage specificity. *Nature* **2006**, *441*, 656-659.

35. Chen, C. Y.; Georgiev, I.; Anderson, A. C.; Donald, B. R., Computational structure-based redesign of enzyme activity. *Proc. Natl. Acad. Sci. USA* **2009**.

36. Jiang, L.; Althoff, E. A.; Clemente, F. R.; Doyle, L.; Röthlisberger, D.; Zanghellini, A.; Gallaher, J. L.; Betker, J. L.; Tanaka, F.; Barbas, C. F.; Hilvert, D.; Houk, K. N.; Stoddard, B. L.; Baker, D., De novo computational design of retro-aldol enzymes. *Science* **2008**, *319*, 1387-1391.

37. Rothlisberger, D.; Khersonsky, O.; Wollacott, A.; Jiang, L.; Dechancie, J.; Betker, J.; Gallaher, J. L.; Althoff, E. A.; Zanghellini, A.; Dym, O.; Albeck, S.; Houk, K. N.; Tawfik, D.; Baker, D., Kemp elimination catalysts by computational enzyme design. *Nature* **2008**, *453*, 190-U194.

38. Dahiyat, B. I.; Mayo, S. L., Probing the role of packing specificity in protein design. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 10172-10177.

39. Arnold, F. H.; Wintrode, P. L.; Miyazaki, K.; Gershenson, A., How enzymes adapt: lessons from directed evolution. *Trends Biochem. Sci.* **2001**, *26*, 100-106.

40. Giver, L.; Gershenson, A.; Freskgard, P. O.; Arnold, F. H., Directed evolution of a thermostable esterase. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 12809-12813.

41. Lienhard, G. E., Enzymatic catalysis and transition-state theory. *Science* **1973**, *180*, 149-154.

42. Villa, J.; Warshel, A., Energetics and dynamics of enzymatic reactions. *Journal of Chemical Physics B* **2001**, *105*, 7887-7907.

43. Tantillo, D. J.; Chen, J. G.; Houk, K. N., Theozymes and compuzymes: theoretical models for biological catalysis. *Curr. Opin. Chem. Biol.* **1998**, *2*, 743-750.

44. Zanghellini, A.; Jiang, L.; Wollacott, A.; Cheng, G.; Meiler, J.; Althoff, E.; Rothlisberger, D.; Baker, D., New algorithms and an in silico benchmark for computational enzyme design. *Protein Sci.* **2006**, *15*, 2785-2794.

45. Casey, M. L.; Kemp, D. S.; Paul, K. G.; Cox, D. D., The physical organic chemistry of benzisoxazoles. I. The mechanism of the base-catalyzed decomposition of benzisoxazoles. *J. Org. Chem.* **1973**, *38*, 2294-2301.

46. Kemp, D. S.; Casey, M. L., Physical organic chemistry of benzisoxazoles. II. Linearity of the Bronsted free energy relationship for the base-catalyzed decomposition of benzisoxazoles. *J. Am. Chem. Soc.* **1973**, *95*, 6670-6680.

47. Hollfelder, F.; Kirby, A. J.; Tawfik, D. S.; Kikuchi, K.; Hilvert, D., Characterization of proton-transfer catalysis by serum albumins. *J. Am. Chem. Soc.* **2000**, *122*, 1022-1029.

48. Hu, Y.; Kouk, K. N.; Kikuchi, K.; Hotta, K.; Hilvert, D., Nonspecific medium effects versus specific group positioning in the antibody and albumin catalysis of the base-promoted ring-opening reactions of benzisoxazoles. *J. Am. Chem. Soc.* **2004**, *126*, 8197-8205.

49. Hilvert, D.; Seebeck, F. P., Positional ordering of reacting groups contributes significantly to the efficiency of proton transfer at an antibody active site. *J. Am. Chem. Soc.* **2005**, *127*, 1307-1312.

50. Thorn, S. N.; Daniels, R. G.; Auditor, M. T. M.; Hilvert, D., Large rate accelerations in antibody catalysis by strategic use of haptenic charge. *Nature* **1995**, *373*, 228-230.

51. Mueller, R.; Debler, E. W.; Steinmann, M.; Seebeck, F. P.; Wilson, I. A.; Hilvert, D., Bifunctional catalysis of proton transfer at an antibody active site. *J. Am. Chem. Soc.* **2007**, *129*, 460-461.

52. Debler, E. W.; Ito, S.; Seebeck, F. P.; Heine, A.; Hilvert, D.; Wilson, I. A., Structural origins of efficient proton abstraction from carbon by a catalytic antibody. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 4984-4989.

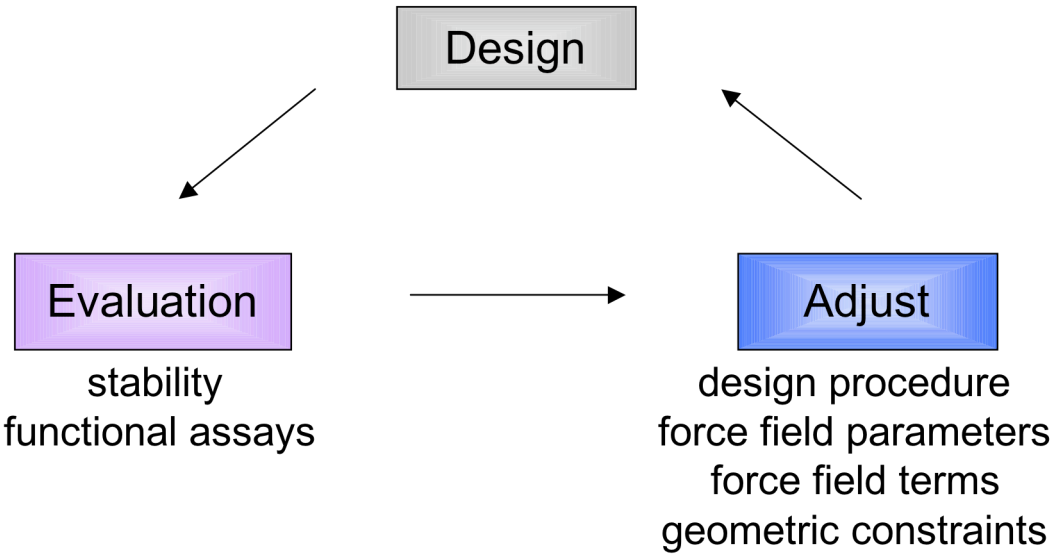

Figure 1-1. The protein design cycle. After design, the sequences predicted by the algorithm are synthesized and experimentally evaluated for their desired characteristics. The correlation of experiment and theory is used to adjust the design procedure for future designs. Adapted from Dahiyat *et al.* 1996.[12]

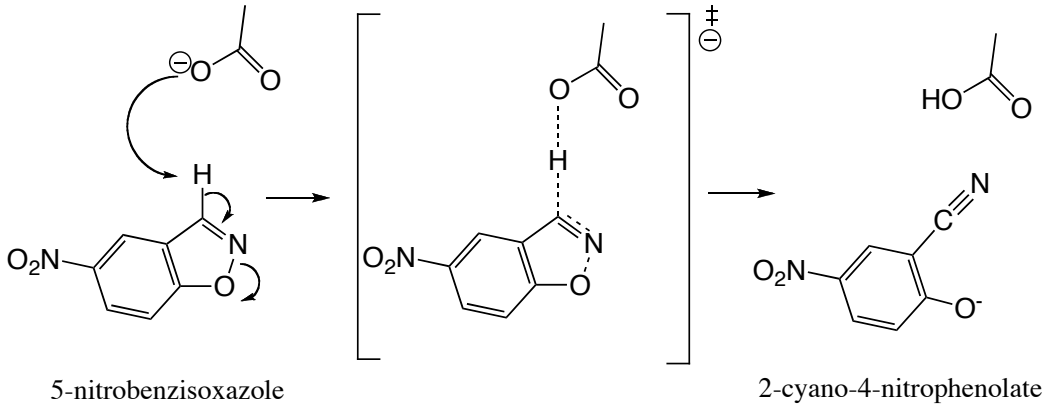

Figure 1-2. The Kemp elimination of 5-nitrobenzisoxazole.

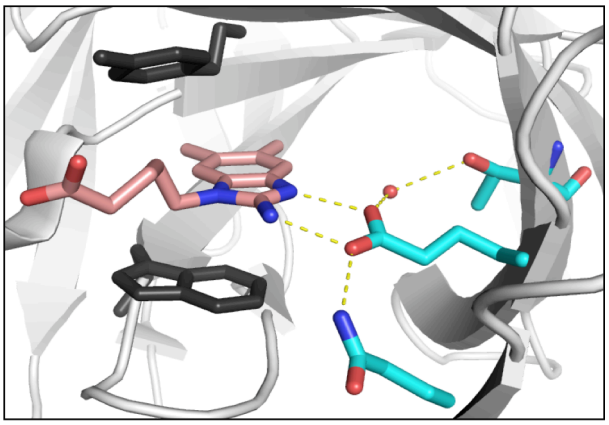

Figure 1-3. Kemp elimination catalytic antibody 34E4.[52] The co-crystallized hapten is shown in pink, the general base is shown in cyan along with two supporting contacts. Hydrogen bonds are indicated with dotted lines.