ELECTROSTATIC ENHANCEMENT OF PROTEIN-LIGAND & PROTEIN-PROTEIN BINDING RATES

By

Wong Kwan Yin

A Thesis Presented to

The Hong Kong University of Science and Technology

In Partial Fulfillment

of the Requirements for the Degree of Master of Philosophy

in the Department of Biochemistry

Hong Kong, April 1998

Copyright © Wong Kwan Yin 1998
I hereby declare that I am the sole author of the thesis.

I authorize The Hong Kong University of Science and Technology to lend this thesis to other institutions or individuals for the purpose of scholarly research.

I further authorize The Hong Kong University of Science and Technology to reproduce this thesis by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.
ELECTROSTATIC ENHANCEMENT OF PROTEIN-LIGAND & PROTEIN-PROTEIN BINDING RATES

By

Wong Kwan Yin

Approved by

Dr. H. X. Zhou, Chairperson of
Supervisory Committee

Professor Jerry H.C. WANG, Head of Department

Department of Biochemistry

4 April 1998
ACKNOWLEDGMENTS

First of all, my gratitude goes to my supervisor, Dr. H. X. Zhou for his continuous encouragement, invaluable advice and guidance during the course of this research project. I would like to thank my colleagues, Dr. M. Vijayakumar and Dr. Qian Xie for their support and valuable suggestions on techniques for running computer simulations.

I also want to express my appreciation to Dr. David Keith Smith and Mr Geqing Chai for their stimulating discussions and assistance.

Finally, many thanks to my family members for continuous support over the period of M. Phil. study.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title Page</td>
<td>i</td>
</tr>
<tr>
<td>Authorization Page</td>
<td>ii</td>
</tr>
<tr>
<td>Signature Page</td>
<td>iii</td>
</tr>
<tr>
<td>Acknowledgments</td>
<td>iv</td>
</tr>
<tr>
<td>Table of Contents</td>
<td>v</td>
</tr>
<tr>
<td>List of Tables</td>
<td>viii</td>
</tr>
<tr>
<td>List of Figures</td>
<td>ix</td>
</tr>
<tr>
<td>Abstract</td>
<td>x</td>
</tr>
<tr>
<td><strong>Chapter 1 Introduction</strong></td>
<td>1</td>
</tr>
<tr>
<td>1.1 Chapter summary</td>
<td></td>
</tr>
<tr>
<td>1.2 Introduction</td>
<td>1</td>
</tr>
<tr>
<td>1.2.1 Why kinetics is so important</td>
<td>1</td>
</tr>
<tr>
<td>1.2.2 Physiological role of superoxide dismutase</td>
<td>2</td>
</tr>
<tr>
<td>1.2.3 Assay of metabolic SOD production in E. Coil</td>
<td>4</td>
</tr>
<tr>
<td>1.2.4 Barnase and barstar</td>
<td>5</td>
</tr>
<tr>
<td>1.2.5 Hirudin and thrombin</td>
<td>7</td>
</tr>
<tr>
<td>1.2.6 Local or global electrostatic enhancement</td>
<td>9</td>
</tr>
<tr>
<td>1.2.7 Format of thesis</td>
<td>10</td>
</tr>
<tr>
<td><strong>Chapter 2 Previous Theoretical Work</strong></td>
<td>11</td>
</tr>
<tr>
<td>2.1 Chapter summary</td>
<td></td>
</tr>
<tr>
<td>2.2 Brownian dynamics simulation</td>
<td>11</td>
</tr>
<tr>
<td>2.3 Analytical approach to diffusion influenced reactions</td>
<td>12</td>
</tr>
<tr>
<td>2.3.1 Rate constant calculations in force-free conditions</td>
<td>14</td>
</tr>
<tr>
<td>2.3.2 Introduction of electrostatic forces into Brownian dynamics calculations</td>
<td>14</td>
</tr>
<tr>
<td>2.3.3 Influence of chemical steps on stereospecific complex formation</td>
<td>14</td>
</tr>
<tr>
<td>2.3.4 Reactive patch model</td>
<td>14</td>
</tr>
<tr>
<td>2.4 Brownian dynamics approach to diffusion influenced reactions</td>
<td>15</td>
</tr>
<tr>
<td>2.5 Electrostatic calculations</td>
<td>18</td>
</tr>
<tr>
<td>2.5.1 Modeling of electrostatic interactions in biological systems</td>
<td>18</td>
</tr>
<tr>
<td>2.5.1.1 Coulomb’s law (Uniform dielectric constant)</td>
<td>18</td>
</tr>
<tr>
<td>2.5.1.2 Poisson Equation (Changing dielectric constant)</td>
<td>19</td>
</tr>
<tr>
<td>2.5.1.3 Poisson-Boltzmann Equation (Mobile ions in solution)</td>
<td>19</td>
</tr>
<tr>
<td>2.5.2 Numerical solution to Poisson-Boltzmann equation</td>
<td>20</td>
</tr>
<tr>
<td>2.5.2.1 Finite difference Poisson-Boltzmann method</td>
<td>20</td>
</tr>
<tr>
<td>2.5.2.2 Boundary element method</td>
<td>21</td>
</tr>
</tbody>
</table>
2.5.3 Electrostatic calculations on proteins ......................... 22
  2.5.3.1 Brownian dynamics simulations with force ............. 22
  2.5.3.2 Electrostatic interaction energy calculation .......... 23
2.5.4 Can the electrostatic rate enhancement be defined by a simple term? ....... 23
  2.5.4.1 Insight of electrostatic rate enhancement by average Boltzmann factor .......... 23
  2.5.4.2 Transition state theory .................................. 24

Chapter 3 Methodology ........................................... 25
  3.1 Chapter summary .............................................. 25
  3.2 Methods ...................................................... 25
    3.2.1 Protein-ligand association of human superoxide dismutase .......... 25
      3.2.1.1 Material preparation .................................. 25
      3.2.1.2 BD simulation procedures .............................. 26
    3.2.2 Barnase and barstar protein-protein association ............. 33
      3.2.2.1 Molecular modeling ................................... 33
      3.2.2.2 Calculation of protein-protein binding energy .......... 33
      3.2.2.3 Brownian dynamics simulation of force-free binding .......... 36
    3.2.3 Parallelization of BD simulation .......................... 38

Chapter 4 Results ................................................. 39
  4.1 Chapter summary .............................................. 39
  4.2 Human SOD .................................................. 39
    4.2.1 Surface charge properties of native superoxide dismutase and mutant ... 39
    4.2.2 Prediction of electrostatic enhancement by Boltzmann factor ........... 41
  4.3 Barnase-Barstar .............................................. 42
    4.3.1 Surface charge property ................................... 42
    4.3.2 BD simulation ........................................... 43
    4.3.3 Electrostatic rate enhancement ............................ 44

Chapter 5 Discussion .............................................. 45
  5.1 Chapter summary .............................................. 45
  5.2 Local versus global on rate enhancement ...................... 45
  5.3 Steering Mechanism .......................................... 47
  5.4 Minimal structural perturbation ................................ 48
  5.5 Insight of rational protein design from ABF approach .......... 49

Chapter 6 Conclusion .............................................. 51
  6.1 Chapter summary .............................................. 51
  6.2 Local electrostatics is dominant ................................ 51
  6.3 Electrostatic steering is misleading ........................... 51
  6.4 Provide a computation inexpensive alternative to traditional BD simulation .... 52
  6.5 Reducution of a kinetic problem to a thermodynamic problem ............ 52
  6.6 Future goals and prospects .................................. 52
    6.6.1 Effective charge method ................................... 52
    6.6.2 Validation of the average Boltzmann relationship .............. 53
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Enzymatic anti-oxidation by SOD</td>
<td>4</td>
</tr>
<tr>
<td>3.1</td>
<td>Mutation residues of SOD and their distance from the active site.</td>
<td>26</td>
</tr>
<tr>
<td>3.2</td>
<td>Parameter settings for calculations of electrostatic potential</td>
<td>34</td>
</tr>
<tr>
<td>4.1</td>
<td>ABFs of the mutants and native SOD</td>
<td>40</td>
</tr>
<tr>
<td>4.2</td>
<td>Comparison of the binding rate constants between experiment and simulation</td>
<td>42</td>
</tr>
<tr>
<td>4.3</td>
<td>Correlation of ABFs and the electrostatic rate enhancement of binding rate</td>
<td>42</td>
</tr>
<tr>
<td>6.1</td>
<td>A summary of the research project.</td>
<td>53</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>-----------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>1.1</td>
<td>The oxidative chain reaction of superoxide free radicals</td>
<td>3</td>
</tr>
<tr>
<td>1.2</td>
<td>Self-defence mechanism of <em>Bacillus amylo liquefaciens</em> towards Barnase</td>
<td>6</td>
</tr>
<tr>
<td>1.3</td>
<td>The cascade process in blood clotting</td>
<td>8</td>
</tr>
<tr>
<td>2.1</td>
<td>Smoluchowski’s model</td>
<td>13</td>
</tr>
<tr>
<td>2.2</td>
<td>Reactive patch model</td>
<td>15</td>
</tr>
<tr>
<td>2.3</td>
<td>Northrup, Allision and McCammon BD model</td>
<td>16</td>
</tr>
<tr>
<td>2.4</td>
<td>Analytical treatment of q-surface</td>
<td>17</td>
</tr>
<tr>
<td>2.5</td>
<td>A dielectric continuum model</td>
<td>19</td>
</tr>
<tr>
<td>2.6</td>
<td>Water probe for rendering the solvent accessible surface of proteins</td>
<td>21</td>
</tr>
<tr>
<td>2.7</td>
<td>Boundary element model</td>
<td>22</td>
</tr>
<tr>
<td>3.1</td>
<td>Definition of SOD active site</td>
<td>28</td>
</tr>
<tr>
<td>3.2</td>
<td>Illustration of collision-free distance</td>
<td>30</td>
</tr>
<tr>
<td>3.3</td>
<td>Partitioning of protein atoms &amp; the collision detection mechanism</td>
<td>31</td>
</tr>
<tr>
<td>3.4</td>
<td>Fitting of the time-dependent rate coefficient $k(t)$ of the native</td>
<td>32</td>
</tr>
<tr>
<td>3.5</td>
<td>Graphic representation of rotational angles</td>
<td>35</td>
</tr>
<tr>
<td>4.1</td>
<td>Surface potential of native SOD</td>
<td>40</td>
</tr>
<tr>
<td>4.2</td>
<td>Surface potential of native barnase</td>
<td>43</td>
</tr>
<tr>
<td>4.3</td>
<td>Surface potential of native barstar</td>
<td>43</td>
</tr>
<tr>
<td>4.4</td>
<td>Comparison of calculated and experimental results for the free energy of barnase-barstar “encounter” complex at transition state.</td>
<td>44</td>
</tr>
<tr>
<td>5.1</td>
<td>Graphic representation of electrostatic potential dominance on rate enhancement</td>
<td>46</td>
</tr>
<tr>
<td>5.2</td>
<td>The binding site of ${\text{Fe}^{3+}}$Trx-SOD</td>
<td>50</td>
</tr>
</tbody>
</table>
ELECTROSTATIC ENHANCEMENT OF PROTEIN-LIGAND & PROTEIN-PROTEIN BINDING RATES

By

Wong Kwan Yin

Abstract

The catalytic steps of a number of well-known enzymes have been perfected through evolution. The rates of the enzymatic reactions are thus determined by the rates of the diffusional encounter between the enzymes and their substrates. Provided that structural integrity is maintained, the diffusional encounter rate and thus the catalytic efficiency can be optimized by surface charge mutations near the active site. From past theoretical work, we have proposed that the electrostatic enhancement of protein-ligand or protein-protein association rates can be obtained from the following simple expression:

\[ k = k^0 \langle \exp(-U/k_B T) \rangle \]

where \( k \) and \( k^0 \) are the rate constants in the presence and absence of the interaction potential \( U \), and \( \langle \exp(-U/k_B T) \rangle \) is an average over the region in which the substrate can effectively bind to the enzyme to form a complex. In the present thesis project, this was used to study a number of enzyme-substrate and enzyme-inhibitor systems, including the binding of superoxide to superoxide dismutase and the association of barnase and barstar. Such studies have led to fundamental understanding on protein-ligand and protein-protein association and to designed enzymes that have optimized catalytic efficiency.
Chapter 1

INTRODUCTION

1.1 Chapter summary

In this chapter, the importance of kinetics in protein-ligand and protein-protein recognition processes and the role of electrostatic interactions in the rate enhancement are illustrated on several systems. The outline of the thesis is presented.

1.2 Introduction

1.2.1 Why kinetics is so important

Recognition between two proteins or an enzyme and its ligand is a central event in a diverse range of biological processes such as, antigen-antibody binding, muscle contraction, cellular signal transduction, protein-DNA binding and electron transport. Of fundamental importance is the specificity of the proper in forming a stereospecific complex via a diffusion process. The severe orientational constraints imposed by the specificity may decrease the binding rate by as much as four orders of magnitude (Zhou, 1997c). To compensate for this, biomolecules commonly use electrostatic interactions to enhance the association rate (Schreiber and Fersht, 1996).

High association rate constants are exhibited, for example, in the binding of the anti-oxidant enzyme superoxide dismutase (SOD), with its substrate. This reaction occurs at a rate of $2\times10^9\text{M}^{-1}\text{s}^{-1}$ (Getzoff et al., 1992). Similar association rate constants are observed in the association processes of the inhibitor, barstar, with the ribonuclease, barnase from Bacillus
*amylo liquefaciens* (Schreiber and Fersht 1995) and the intracellular inhibitor hirudrin with thrombin (Stone *et al.*, 1989). Other examples of fast association rates include the hydrolysis of acetylcholine by acetylcholinesterase (Quinn, 1987; Bazelyansky *et al.*, 1986; Hasinoff, 1982; Rosenberry, 1975), the recognition of antigens by antibodies (Shen *et al.*, 1996; Janin, J. & Chothia, C., 1990) and the dimerization of $\alpha$ and $\beta$ subunits of hemoglobin in the developing erythroblast (Bunn & McDonald, 1983).

The rationale behind these extremely high rate constants can be elucidated by a few well-known systems, for instance, the detoxification of the superoxide free radical by SOD, the inhibition of barnase by barstar, and hirudrin by thrombin.

**1.2.2 Physiological role of superoxide dismutase**

Superoxide free radicals are generated from the reduction of oxygen molecules into oxygen free radicals:

$$O_2 \xrightarrow{e^-} O_2^-$$

Superoxide free radicals, $O_2^-$, are formed primarily by autoxidation of components of the respiratory chain, like fumarate reductase (Imlay & Fridovich, 1991) and neutrophil respiratory burst (Tizard, 1995). Some superoxide free radicals will be further reduced to more reactive hydroxyl radical (OH') through Haber-Weiss (Haber & Willstatter, 1931; Haber & Weiss, 1934) and Fenton (Fenton & Jackson, 1899) reactions in the presence of ferrous iron.

$$O_2^- + H_2O_2 \xrightarrow{Fe^{III}} OH^- + OH \cdot \text{(Haber-Weiss reaction)}$$
\[ O_2^{-} \xrightarrow{\text{Fe(III)}} O_2 \]
\[ H_2O_2 \xrightarrow{\text{Fe(II)}} OH^- \]

(Fenton reaction)

Recently, evidence has been advanced to indicate that the free iron source can be obtained by superoxide induced iron leakage from storage proteins or [4Fe-4S] enzyme clusters. Bacteria under superoxide-stress will have a much higher level of free iron (Keyer, 1996).

Proteins, lipids and deoxyribonucleic acid are highly susceptible to oxidative stress (See Figure 1.1) even under SOD-controlled intracellular superoxide levels (10^{-10}M).

\[
\begin{align*}
\text{RX} + O_2^- & \rightarrow \text{RO} + X^- \\
\text{ROO}^- + O_2 & \rightarrow \text{ROO} + O_2 \\
\text{ROO}^- + \text{RX} & \rightarrow \text{ROOR} + X^- 
\end{align*}
\]

**Figure 1.1:** The oxidative chain reaction is initiated by a nucleophilic attack by superoxide free radical (O$_2^-$) on the biomolecules (RX).

SOD is widespread enzyme in oxygen-tolerant organisms. The SOD family including Cu/Zn- SODs, Mn-SODs and Fe-SODs and can protect cells from oxidative damage by the superoxide free radical (O$_2^-$) by catalyzing the dismutation of the superoxide free radical to hydrogen peroxide (H$_2$O$_2$).

\[ 2O_2^- + 2H^+ \rightarrow H_2O_2 + O_2 \]

Transient oxygen exposure experiments reveal that the SOD concentration of obligate anaerobes is correlated to oxidative stress tolerance (Kuby, 1991). Increases in levels of SOD were observed among facultative species like *Bacillus faecalis* (Gregory et al., 1973a, 1973b) and *Saccharomyces cerevisiae* (Gregory et al., 1974) upon exposure to oxygen. This shows that a higher level of SOD is responsible for stronger resistance to superoxide-toxicity. Improper functioning or a reduced level of SOD activity in
mutant strains of *E. coli*, yeast, and Drosophila caused severe growth impairment due to the accumulation of superoxide (Bilinski et al., 1985; van Loon et al., 1986; Carlioz et al., 1986; Philips et al., 1989).

### 1.2.3 Assay of metabolic SOD production in *E. coli*

Imlay and Fridovich (1991) demonstrated that the accumulation of superoxide in *E. coli* was essentially limited by SOD (See Table 1.1). At the steady state, the level of superoxide was 2.0×10⁻¹⁰M.

<table>
<thead>
<tr>
<th></th>
<th>Steady-state [O₂⁻⁻] M</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LB</td>
</tr>
<tr>
<td>Accumulation limited by:</td>
<td></td>
</tr>
<tr>
<td>Superoxide dismutase</td>
<td>2.0×10⁻¹⁰</td>
</tr>
<tr>
<td>Glutathione</td>
<td>4.9×10⁻⁷</td>
</tr>
<tr>
<td>Spontaneous dismutation</td>
<td>6.7×10⁻⁶</td>
</tr>
</tbody>
</table>

**Table 1.1:** Intracellular O₂⁻⁻ concentration is more effectively limited by enzymatic reaction (~30,000 fold) over spontaneous dismutation at physiological medium (LB: 10g of bacterotryptone, 5g of yeast extract, 10g of NaCl, 2g glucose, per liter, pH7.0). Source: Imlay & Fridovich, (1991) *J. Biol. Chem.* **266**, 6957-6965.

Since the concentration of superoxide is mainly under the control of SOD, the rate constant of SOD becomes a determining factor of [O₂⁻⁻] in vivo. If the rate constant, k, of SOD decreases ten fold to 2×10⁸M⁻¹s⁻¹, cells cultured in LB + glucose will have 360 units/ml of SOD activity. The pseudo-first order rate constant of 1 unit/ml SOD with O₂⁻⁻ is 7.8s⁻¹ and the rate of superoxide production is 5.7μM/s (Imlay & Fridovich, 1991).

As in the steady state, the rate of O₂⁻⁻ synthesis = the rate of dismutation, it follows that

\[ 5.7\mu M/s = 360 \text{ units/ml SOD} \times 7.8 \text{s}^{-1} \times [O_2^{\cdot-}] \]
Therefore the steady state \([O_2^-]\) (2.0x10^8M) in vivo would be 10 times higher than that with normal SOD.

If a stereospecific complex of SOD and \(O_2^-\) must be formed before dismutation of superoxide, the diffusional encounter rate becomes a limiting factor on the overall reaction rate of \(\sim 10^5 \text{ M}^{-1}\text{s}^{-1}\). However, in the presence of electrostatic enhancement, the rate constant \(k\) of SOD increases 4 order of magnitude at physiological ionic strength (\(\sim 0.15\text{M}\)) to 2.0x10^9 M\(^{-1}\)s\(^{-1}\) (Getzoff et al., 1983; Sharp et al., 1987; Sines et al., 1990; Buckle et al., 1994).

### 1.2.4 Barnase and barstar

Large association rate constants often play an important role in the protein-protein association processes, especially in competitive association processes, for example, the barnase-barstar association.

Barnase, a small (110 amino residues) extracellular ribonuclease produced by *Bacillus amyloliquefaciens* is involved in breaking down RNA molecules in the extracellular environment. It serves as a digestive enzyme or toxin for killing predators and competitors against the host (Hartley, 1989). Expression of the active form of barnase in the cytosol may result in cell death due to RNA cleavage. A direct competition between RNA and barstar for binding to newly synthesized barnase in the cytosol is involved in preventing this (See Figure 1.2). Co-evolution of barnase with a tightly binding inhibitor, barstar, is thus essential for the survival of barnase-producing cells (Hartley, 1989).
To avoid cell death caused by accumulation of active barnase in the cell, the preferential binding of barstar to barnase can be achieved by a large rate constant for the barnase-barstar association ($5 \times 10^9 M^{-1}s^{-1}$) (Schreiber et al., 1993). This essentially overcomes the damaging effect caused by barnase-RNA association ($\sim 10^5 M^{-1}s^{-1}$) (Bastyns et al., 1994).

Such preferential binding cannot be obtained under a basal diffusion limited association rate constant of $10^5 M^{-1}s^{-1}$ because of the comparable barnase-RNA association rate. Electrostatic interactions can enhance the rate of barnase-barstar association.

Meiering et al. (1992) suggested that the clustering of the positively charged residues in the binding site of barnase is important to its interaction with both RNA and barstar. This work also showed that these charged residues destabilize the folded structure of barnase. A similar was also observed in barstar by mutational studies (Asp→Ala39, Asp→Ala35, Glu→Ala76, Glu→Ala80) (Schreiber et al., 1994). These represent the compromise between structural stability and activity i.e. the charged residues are conserved in the course of evolution to allow electrostatic interactions to accelerate the binding rate.
1.2.5 Hirudin and thrombin

Association rate similar to those seen with barnase-barstar based on electrostatic attraction, have been found for thrombin-inhibitor association, such as thrombin-hirudin binding ($4.7 \times 10^8 \text{M}^{-1}\text{s}^{-1}$) (Braun et al., 1988).

Thrombin, a serine protease, plays a significant role in the blood-clotting cascade and prevents extensive blood loss via platelet activation and fibrin generation. In vivo, thrombin is produced by converting the inactive prothrombin in bloodstream to active thrombin through the activation of the prothrombinase complex. Upon generation, thrombin starts to cleave fibrinogens into fibrin monomers, which in turn polymerize to form a blood clot. In order to achieve homeostasis, blood-coagulation system should be kinetically efficient and this is achieved by means of an amplification cascade (See Figure 1.3).
**Figure 1.3:** Thrombin, a serine protease (**) is dominant in both intrinsic and extrinsic pathways of the blood-clotting cascade. The blood-clotting process is accelerated through a signal amplification mechanism (denoted as pine, curved arrows). Coagulation time is prolonged if thrombin* is inactivated by binding to hirudin. The active and inactive forms of blood-clotting factors are in pink and green respectively.

This amplification mechanism can increase thrombin concentration in blood to 140 nM upon vascular injury (Walz et al., 1985). However, uncontrolled production of thrombin results in thrombosis intravascular coagulation, a phenomenon often associated with major surgery. A clinical resolution of this problem is to find inhibitors, for example, heparin and hirudin with high potency towards thrombin inhibition. Traditionally,
the effectiveness of protein inhibition is evaluated by the inhibition equilibrium constant $K_i$. The smaller the $K_i$ value the more potent the inhibitor is. It turns out that $K_i$ alone cannot fully explain the higher potency of argatroban and NAPAP which have larger $K_i$ values than the boronic acid analog, Z-D-Phe-Pro-boroMpgC_{16}H_{16}. Tapparelli et al. (1993) concluded that thrombin must be inhibited rapidly in order to prevent platelet aggregation which is more effectively achieved by large rate constants of the inhibitors like hirudin.

Application of long-range electrostatic interactions, enhancing early non-stereospecific protein-protein binding, followed by non-diffusion influenced docking, a so called two-phase association, may be a common strategy used in nature (Schreiber & Fersht, 1995).

1.2.6 Local or global electrostatic enhancement

A question arises as to whether the rate of association is under the influence of the local or global electrostatic properties of the proteins. At physiological pH, the negatively charged (-4e) SOD should create a strong electrostatic barrier to incoming superoxides (-1e). This barrier should be weakened with increasing ionic strength so that the reaction rate will increase due to the screening effect. However, that the rate of reaction decreases with increasing ionic strength showing that local electrostatic attraction is dominant (Getzoff et al., 1992).

The dominance of local electrostatic effects was further elucidated by the effects of charge mutations at the active site and those away from the active site. Sharp et al. (1987) carried out simulations on bovine SOD with the electrostatic potential of the face of the molecule opposite to the active site switched off. This did not affect the reaction rate. Stronger support came from study of numerous charge mutations on bovine SOD far away from the active site of the enzyme, which had limited effect on the reaction
rate (Sines et al., 1990). In addition, charge mutation studies demonstrated that the mutations at the active site of human SOD have significant effect on the association rate constants (Getzoff et al., 1992).

1.2.7 Format of thesis

The remainder of this thesis is organized as follows. Chapter 2 reviews the previous theoretical work on the electrostatic enhancement of protein-ligand and protein-protein binding rates. Chapter 3 presents a newly developed relation for estimating the rate increment caused by electrostatic attraction (Zhou, 1996):

\[ \frac{k}{k^0} = \langle \exp(-\beta U) \rangle \]  \hspace{1cm} [1.1]

where \( k \) and \( k^0 \) are the binding rate constants in the presence and absence of an interaction potential \( U \). The term, \( \langle \exp(-\beta U) \rangle \), is the average Boltzmann factor over the region in space where substrate can bind to the protein or enzyme. Chapter 3 also describes the methods used for Brownian dynamics simulations (BDS) of enzyme-substrate binding and protein-protein association, evaluation of average Boltzmann factor (ABF) in the binding region and calculation of the binding rate constant. Chapter 4 summarizes the results of charge mutations on rate of diffusion-influenced reactions and judges the accuracy of evaluating electrostatic rate enhancement by equation 1.1. The remaining chapters discuss and provide conclusion on the consequences and implications of this study. Attention will be placed on analyzing the “electrostatic steering” mechanism for describing electrostational rate enhancement, and the recently developed “encounter complex” theory (Vijayakumar et al., in press) in accounting for electrostatic enhancement of the protein-protein associations.
Chapter 2

PREVIOUS THEORETICAL WORK

2.1 Chapter summary: This chapter reviews the major methods and results of Brownian dynamics simulations and introduces a new approach for evaluating association rate enhancement by electrostatic interactions.

2.2 Brownian dynamics simulation

For any specific biological recognition to occur, diffusion is the first step bringing the two interacting components together. Understanding the diffusion processes under the influence of electrostatic forces is critically important to gain an insight into the effect of electrostatics on the rate enhancement of protein-protein association.

Robert Brown, a British botanist, observed pollen grains undergoing erratic motion in water in 1827. It was not until 1905 that Albert Einstein and Marian von Smoluchowski, independently, explained this phenomenon. The erratic movement of pollen grains in water is rationalized by the random collision of water molecules on the grains. This irregular movement is frequently referred as Brownian motion, diffusion or a random walk. Robert Brown was able to show that the average linear displacement \( \langle x^2 \rangle \) of a diffusing particle is given by \( \langle x^2 \rangle = 2Dt \) where \( t \) is the time and \( D \) is the diffusion constant. By applying Newton’s second law of motion to diffusing particles, the Langevin equation is obtained:

\[
mx = F - \nu + f(t)
\]

[2.1]
where $\dot{m}$ is the resultant force acting on the diffusing particle, $F$ is the external force, $\gamma$ is the frictional coefficient, $v$ is the velocity and $f(t)$ is the random, statistical force.

Based on this equation, Ermak and McCammon (1978) presented an algorithm to simulate a Brownian particle:

$$\vec{F}(\vec{r}) = -\nabla U(\vec{r}) - \gamma \vec{v} + \vec{f}(t) \quad [2.2]$$

where $\vec{F}(\vec{r})$ is the vector force acting on the Brownian particle, $\nabla U(\vec{r})$ is the systemic force, $\gamma \vec{v}$ is the fictional force and $\vec{f}(t)$ is the random force. This allows one to propagate a three-dimension Brownian trajectory by the following expression:

$$\vec{r} = \vec{r}_0 + (k_b T)^{-1} D \vec{F}(\vec{r}_0) \Delta t + \vec{R} \quad [2.3]$$

where $\vec{r}$ is the new position on the diffusion pathway, $\vec{r}_0$ is the previous position, $\vec{R}$ is the random vector with a Gaussian distribution which has the properties $\langle R \rangle = 0$, $\langle R_i R_j \rangle = \delta_{ij}$ and $\Delta t$ is the time step in which the force remains constant.

2.3 Analytical approach to diffusion influenced reactions

2.3.1 Rate constant calculations in force-free conditions

In 1917, Smoluchowski presented a model consisting of two interacting spheres. Coagulation occurs once the two diffusing spherical particles come into contact. The diffusion rate constant estimated by this model is

$$k_d^{f0} = 4\pi(D_1 + D_2)(r_1 + r_2) = 4\pi DR \quad [2.4]$$

where $k_d^{f0}$ is the rate constant under the force-free diffusion-influenced reaction. $D = (D_1 + D_2)$ is the relative diffusion constant of these two
spheres and \( R = (r_1 + r_2) \) is the sum of the radii. The Stokes-Einstein equation gives:

\[
D = \frac{k_B T}{6\pi \eta r}
\]  

[2.5]

where \( k_B T \) is the product of Boltzmann constant and absolute temperature, \( \eta \) is the solvent viscosity and \( r \) is the solute hydrodynamic radius. This simple model was used to estimate the theoretical upper limit of protein-protein association rate constants \( k_a^{f0} = 8k_B T/3\eta \sim 10^8-10^{10} \text{M}^{-1}\text{s}^{-1} \) (Smoluchowski et al., 1918).

**Figure 2.1:** Smoluchowski’s model consists of two interacting spheres, reaction occurs instantaneously once the sphere-A comes in contact with reactive surface of sphere-B.
2.3.2 Introduction of electrostatic forces into Brownian dynamics calculations:

By introducing the centrosymmetrical interaction potential \( U(r) \) into the Smoluchowski’s rate equation, Debye (1942) was able to show that the association rate constant \( k_d \) in the presence of this force was

\[
k_d = 4\pi D \left\{ \int_0^\infty r^{-2} e^{U(r)/k_BT} \, dr \right\}^{-1}
\]  

[2.6]

2.3.3 Influence of chemical steps on stereospecific complex formation

By considering the involvement of chemical steps in the stereospecific complex formation, the Smoluchowski’s association rate equation becomes (Collins & Kimball, 1949)

\[
\begin{align*}
\frac{1}{k_d}^{(0)} &= \frac{1}{4\pi DR} + \frac{1}{k_{chem}} \\
\frac{1}{k_d} &= \int_0^\infty r^{-2} e^{U(r)/k_BT} \, dr + \frac{1}{k_{chem}}
\end{align*}
\]  

[2.7]

where \( k_{chem} \) is the bimolecular rate constant for the chemical steps when the diffusion is extremely fast and the reactants are all at equilibrium with respect to their spatial distribution.

2.3.4 Reactive patch model

In general, the surface of proteins is not totally reactive, but only a very restricted surface area, which function as the binding sites, for protein-protein associations to take place is reactive. The uniform reactive spherical models fail to consider orientational constraints as a factor of stereospecific complex formation. A more realistic model, is one the
reactive-patch model, in which only small areas of spherical surfaces scaled by angles $\theta_1$ and $\theta_2$ is reactive (See Figure 2.2).

![Figure 2.2: An association reaction occurs only when two heterogeneous reactive patches come together.](image)

Berg (1985) showed that $k_\alpha \approx 4\pi DR \sin(\theta_i/2)$ when $\theta_i \to 0$ (i.e. highly orientationally restricted binding).

### 2.4 Brownian dynamics approach to diffusion influenced reactions

To eliminate the shortcomings of the analytical approach for calculating rate constants of protein-protein associations, which are subjected to spatial and orientational constraints, Brownian dynamics simulations appear to be the most practical approach. In 1984, Northrup et al. proposed an algorithm to estimate the enzyme-ligand binding rate (See Figure 2.3).
Figure 2.3: A diagram illustrates the original model that used by Northrup, Allison and McCammon (1984) to do BD simulation.

All individual trajectories start at the $b$-surface and propagate, according to equation 2.3, until they cross the $q$-surface (truncated trajectories) or react with the target enzyme at its active site (reactive trajectories). By simulating a large numbers of trajectories, the capture probability, $\beta$, of the following rate equation can be obtained if the trajectories are truncated at some large separation $q$.

$$\beta_\infty = \frac{\beta}{1 - (1 - \beta) \left( k(b) / k(q) \right)}$$  \[2.8\]

where $k(b)$ is the rate at which the reactants first striking the $b$-surface and is defined as

$$k(b) = 4\pi D \left\{ \int_0^\infty r^{-2} e^{U(r)/k_b T} dr \right\}^{-1}$$  \[2.9\]

Therefore, the steady-state rate $k$ is given by

$$k = k(b) \beta_\infty$$  \[2.10\]
Zhou (1990a) later showed that the distance between the $b$-surface (See Figure 2.3) and the geometric center of the protein should be sufficiently large in order to have a centrosymmetrical mean potential. The $q$-surface, the trajectory cutoff surface, also has to be large enough to ensure the inward reactive flux at the given distance is centrosymmetric in the reaction.

Luty et al. (1992) replaced the uncertainty of the outer cutoff $q$-surface with an analytical treatment, in which a fraction $f_{\text{occ}}$ of trajectories is truncated once across the $m$-surface (See Figure 2.4) and the rest are put back to the $b$-surface with $w(\theta, \phi)$ distribution. The reacting fraction $(1 - f_{\text{occ}})$ of the trajectories results in $\beta_-$ and thus the rate constant $k$ from equation 2.8.

![Figure 2.4](image)

**Figure 2.4:** A schematic representation of diffusion influenced reaction system (Analytical treatment of $q$-surface in Northrup, Allision and McCammom algorithm). Here $(b, \theta, \phi)$ are the spherical coordinates of the new when the previous position on the $m$ surface is put on the polar axis. (A reproduced figure from Luty et al., 1992).

However, the prerequisite of a centrosymmetrical mean potential for the starting positions (i.e. a large starting distance from the protein binding site), as well as the small binding site of the target, guaranteed that the algorithm is very time-consuming for accurate simulations (Zhou, 1990a).
In order to overcome the practical difficulties of the Northrup *et al.* (1984) algorithm, Zhou (1990b) took an alternative approach, in which trajectories began at the reactive region of the target. A trajectory is terminated either when a given time limit is reached or when a reaction occurs. The time dependent rate constant \( k(t) \) is then retrieved from the survival fraction of the trajectories. The steady state rate constant \( k(\infty) \) is obtained from the asymptotic behavior of the rate coefficient. The details of this algorithm and its extension to cover rotational dynamics will be given in chapter 3.

2.5 Electrostatic calculations

2.5.1 Modeling of electrostatic interactions in biological systems

2.5.1.1 Coulomb’s law (Uniform dielectric constant)

The conventional description of electrostatic models are based the Poisson equation:

\[
\nabla^2 \Phi(\vec{r}) = \frac{-4\pi \rho(\vec{r})}{\epsilon} \quad [2.11]
\]

The first term \( \nabla^2 \Phi(\vec{r}) \) in the above equation describes the change of potential \( \Phi \) with respect to the radius \( \vec{r} \). \( \rho \) and \( \epsilon \) are the charge density and the dielectric constant of the medium respectively. If the charges are finite, the Poisson equation is reduced to Coulomb’s law:

\[
\Phi(\vec{r}) = \sum_i \frac{q_i}{\epsilon |\vec{r} - \vec{r}_i|} \quad [2.12]
\]
2.5.1.2 Poisson Equation (Changing dielectric constant)

Coulomb's law becomes invalid when the dielectric constant varies through space (the polarizability of regions in aqueous media) (See Figure 2.5). Now, the Poisson equation becomes

$$\nabla \cdot \varepsilon(\vec{r}) \nabla \Phi(\vec{r}) = 4\pi \rho(\vec{r})$$  \[2.13\]

reflecting the change of the dielectric property, $\varepsilon$, with respect to the radius $\vec{r}$.

![Figure 2.5: A practical approach to handle a variable dielectric constant in space is to model the protein system as a dielectric continuum, with two discrete dielectric regions, namely a low dielectric protein region ($\varepsilon=4$) and a high dielectric solvent region ($\varepsilon=78.5$).](image)

2.5.1.3 Poisson-Boltzmann Equation (Mobile ions in solution)

In 1923, Debye and Huckel included electrolyte (ions) in the description of electrostatic phenomena in solvents. This allows one to represent mobile ions in ionic solvents explicitly. Statistically, the ions are distributed in the liquid according to a Boltzmann distribution. Thus, the local ion concentration $C(\vec{r})$ can be expressed as
\[ C(\vec{r}) = C(\infty) \exp \left( \frac{q\Phi(\vec{r})}{k_B T} \right) \]  

where \( C(\infty) \) is the bulk ion concentration, \( q\Phi(\vec{r}) \) is the electrostatic contribution from the chemical energy of positive or negative ions at position \( \vec{r} \) and \( T \) is the absolute temperature. By putting the above mobile ion description into the Poisson equation, the Poisson-Boltzmann (PB) equation is obtained

\[ \nabla \cdot \epsilon(\vec{r}) \nabla \Phi(\vec{r}) - k^2 \epsilon \sinh[\Phi(\vec{r})] = -4\pi\rho(\vec{r}) \]  

where \( k \) is the reciprocal of Debye and Huckel's length.

2.5.2 Numerical solution to Poisson-Boltzmann equation

2.5.2.1 Finite difference Poisson-Boltzmann method

Analytical approaches have been attempted to solve the Poisson-Boltzmann equation (Tanford-Kirkwood 1957; Gurd 1985) based on simplified protein geometry. This limitation was overcome by the introduction of the finite difference Poisson-Boltzmann method (FDPB). Warwick and Watson (1982) were the first to apply this method to handle the calculation of the electrostatic surface of an irregular cavity. The method can be summarized as follows: the space enclosing the solvent and macromolecule(s) is divided into small and regular cubic lattices or grid points, so that each grid point represents a small portion of the system. Dielectric constants of 78.5 and 4 are usually assigned for the solvent and macromolecule, respectively, and the potential is evaluated at the center of each grid point. The boundary between macromolecule and solvent is usually defined by a water molecule probe (radius \( \sim 1.4\text{Å} \)) (See Figure 2.6). The spacing and number of grid points are selected based on a tradeoff of accuracy, computation time and memory size. Alternatively, the numerical
accuracy can be improved by applying a focusing techniques, in which new potential grid points are calculated from the boundary conditions of the last run (Gilson et al., 1988).

Figure 2.6: A water atom with radius (1.4 Å) is used as a probe for rendering the solvent accessible surface of proteins.

2.5.2.2 Boundary element method

Zauhar and Morgan (1985) proposed a new technique in the electrostatic calculation of macromolecules, which is called the boundary element method (BEM). The BEM provides an alternative approach with a much better description of the molecular surface than the finite difference method. This is because the integral equations can be solved by “discretizing” the macromolecule surface into small boundary elements. (See Figure 2.7)
Figure 2.7: The dielectric interface is divided into small elements of simple geometry (discretization) and which are joined together to form a smooth protein surface. The induced dipole is assumed to have little change over neighboring elements. Then, the electrostatic potential along the normal to the center of each element on surface can be calculated. \( E_s \) and \( E_b \) are the electrostatic fields acting on the protein-solvent interface due to the protein and solvent, respectively.

Surface matching, using a small triangle discretization scheme (Zauhar and Morgan (1985, 1988), was used by Yoon et al. (1990) to estimate the molecular electrostatic effect in ionic solvent. Recent advancements in discretization schemes were made by Zhou (1993b).

2.5.3 Electrostatic calculations on proteins

2.5.3.1 Brownian dynamics simulations in the presence of force

As pointed out above, the use of Coulomb’s law in force calculations assumes a uniform dielectric constant, is unrealistic in molecular modeling due to the neglect of both charge-dipole interactions and ion screening in solvent. Fortunately, either FDPB or BEM provide a better solution for electrostatic force calculation. By combining the techniques of BD
simulation (Northrup et al., 1984) and FDPB, it is feasible to handle rotational and translational diffusion simultaneously (Sharp et al., 1987b).

2.5.3.2 Electrostatic interaction energy calculation

The free energy change, \( \Delta G \), is coupled to all the binding processes and configurational changes of proteins. Since electrostatic interactions are the major factor contributing to the solvation energy, FDPB and BEM allows the free energy change to be calculated as

\[
\Delta G = \frac{1}{2} \sum q_i (\Phi_i^r + \Phi_i^s + \Phi_i^m) = \frac{1}{2} \sum q_i \Phi_i
\]

[2.16]

where \( \Phi_i \) is the resultant potential from three components, namely, the potential from interacting charges \( \Phi_i^r \), solvent \( \Phi_i^s \) and mobile ions \( \Phi_i^m \) (Sharp et al., 1990).

2.5.4 Can the electrostatic rate enhancement be defined by a simple term?

2.5.4.1 Insight of electrostatic rate enhancement by average Boltzmann factor

Advances in electrostatic calculations and BD techniques have enabled a wide range of protein binding phenomena, for example, electrostatic rate enhancement (ERE) by charge mutation, to be accounted for. However, the computation is costly for these simulations, and is still a challenge on current computing equipment. This is one of the motivations for developing a simple theory that can explain electrostatic phenomena like ERE. Based on previous extensive BD studies, Zhou (1993a, 1996a) proposed a simple relation (See Equation 1.1) to elucidate the role of electrostatic interactions in rate enhancement. This relation has been used
to predict the electrostatic rate enhancement on acetylcholinesterase-substrate binding (Zhou et al., 1996c).

2.5.4.2 Transition state theory

Generalizing the ABF approach, the original kinetic problem on ERE can be changed to a thermodynamic question. The transition state free energy of an "encounter complex" (containing a variety of intermediate configurations) is obtained from

\[ G^\ddagger = -k_B T \ln \langle \exp(-\beta U) \rangle \]

[2.17]

where \( \langle \exp(-\beta U) \rangle \) is the ABF from various transition state configurations that satisfy relative orientational constrains (to be explained in a later chapter) and have a layer of solvent between the binding interface (M. Vijayakumar et al., in press). Thus, if one can determine the magnitude of \( G^\ddagger \) at a transition state, the order of rate enhancement can also be found without difficulty by using the equation 1.1. This follows as ABF captures the electrostatic properties of the binding processes.
Chapter 3

METHODOLOGY

3.1 Chapter summary: This chapter focuses on the introduction of the methods used to model charge mutants and estimate the binding rate constant (based on BD techniques). We will also present a new way to estimate the binding rate based on the ABF approach.

3.2 Methods:

3.2.1 Protein-ligand association of human superoxide dismutase

3.2.1.1 Molecular modeling

X-ray structure

The X-ray structure of SOD was taken from the Brookhaven Protein Data Bank (Bernstein et al., 1977) entry 1sos. A dimer chains, A and F was cut from the penta-dimeric X-ray structure of SOD and the missing heavy atoms of the dimer were added using the Biopolymer program from Molecular Simulations, Inc.

Charge mutation

It is clear from the studies described in chapter 2 that any mutation away from the active site will have a negligible effect on the binding rate constant. Computer-aided site-direct mutagenesis was used to replace the negatively charged residues Glu-132, Glu-133, Glu-49, Asp-52, Glu-121 and Asp-11 to their neutral counterparts. (See Table 3.1) Though these residues are close to or within the catalytic groove (i.e. near the Cu-atom),

25
neutralizing these residues is expected to give a positive contribution to ABF without causing severe structural alterations in the binding site.

<table>
<thead>
<tr>
<th>Charge Residue(s) involved</th>
<th>Distance between Cu at active site and C_{\text{mutated residue}} (\text{Å})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glu133 \rightarrow Gln133</td>
<td>10</td>
</tr>
<tr>
<td>Glu49 \rightarrow Gln49</td>
<td>12</td>
</tr>
<tr>
<td>Glu121 \rightarrow Glu121</td>
<td>12</td>
</tr>
<tr>
<td>Arg-52 \rightarrow Asn-52</td>
<td>15</td>
</tr>
<tr>
<td>Arg-11 \rightarrow Asn-11</td>
<td>16</td>
</tr>
<tr>
<td>Glu132 \rightarrow Gln132</td>
<td>17</td>
</tr>
<tr>
<td>Glu-133 \rightarrow Gln-133 &amp; Glu-49 \rightarrow Gln-49</td>
<td></td>
</tr>
<tr>
<td>Glu-133 \rightarrow Gln-133 &amp; Arg-52 \rightarrow Asn-52</td>
<td></td>
</tr>
<tr>
<td>Glu-133 \rightarrow Gln-133 &amp; Glu-132 \rightarrow Gln-132</td>
<td></td>
</tr>
<tr>
<td>Glu-49 \rightarrow Gln-49 &amp; Arg-52 \rightarrow Asn-52</td>
<td></td>
</tr>
<tr>
<td>Glu-49 \rightarrow Gln-49 &amp; Glu-132 \rightarrow Gln-132</td>
<td></td>
</tr>
<tr>
<td>Arg-52 \rightarrow Asn-52 &amp; Glu-132 \rightarrow Gln-132</td>
<td></td>
</tr>
</tbody>
</table>

Table 3.1: A list of the residue(s) involved in each mutation and their distance between the active site of SOD (Cu atom).

3.2.1.2 BD simulation procedures

Algorithm

The Brownian simulation procedure used here is described in detail elsewhere (Zhou 1990, 1993; Potter et al. 1996). The rate dependent constant $k(t)$ is estimated by
\[
\begin{align*}
    k(t) &= k(0)S(t) \\
    k(0) &= k_i \langle \exp(-\beta U) \rangle V_{br} 
\end{align*}
\]  

where \( k_i \) is the first-order rate constant, \( \langle \exp(-\beta U) \rangle \) is the ABF over the binding region of the enzyme and \( V_{br} \) is the volume of the binding region. The steady state rate constant \( k(\infty) \) is determined by taking the long time limit of \( k(t) \) (Zhou, 1990b).

**Setup of the potential grid**

The potential under an ionic strength of 0.01M (Getzoff et al., 1992) was calculated using the UHBD package (Davis et al., 1991) which uses a finite difference method. The grid size was 140×140×140 with 1.0Å spacing in dimension. The dielectric constants for the protein \( \varepsilon_p \) and solvent \( \varepsilon_s \) were set to be 4 and 78.5 respectively (See APPENDIX A for detailed parameter settings).

**Generation of initial positions**

The catalytic site of SOD was defined as a concentric spherical shell with 7.5 Å outer and 7.0 Å inner radii. The center of the shell was located at the center of the copper ion (See Figure 3.1). The substrate was modeled as a sphere of radius 2.05Å. Forty thousand substrate starting positions were randomly generated and tested against the two SOD catalytic sites. However, all initial substrate positions that were in collision with the enzyme were excluded. This yielded 3,600 allowable positions. The electrostatic potential of these allowable positions was retrieved from the corresponding potential grid. The Boltzmann factor (BF) of individual positions was calculated according to
\[ BF = \exp(-\beta\Phi) \]  

where \( \Phi \) is the potential at that position. Then, the Boltzmann factor is averaged over the 3,600 positions (i.e. ABF).

All the 3,600 positions were further selected against their Boltzmann weight to give initial positions with a Boltzmann distribution. In short, each position with its Boltzmann factor was compared to a random number \( R \) uniformly distributed between 0 and 1. If \( R \) was smaller, the individual position was selected. This process will continued until a required number of initial substrate positions (12,000) was determined.

![Diagram of SOD active site](image)

**Figure 3.1:** A schematic representation of the SOD active site and the definition of the reactive spherical shell (in between the two red-dotted lines). The Cu atom is modeled as a (green in figure) 1.87Å sphere.

**Simulation of Trajectories**

Each trajectory started at the catalytic site of SOD and propagated according to the Ermak-McCammon algorithm (i.e. equation 2.3 in chapter 2). A tri-linear interpolation function was used to obtain potentials and
fields for given points enclosed by the grid. The electrostatic forces $f_x$, $f_y$ and $f_z$ among the XYZ directions were calculated as:

\[
\begin{align*}
F_x &= q \cdot \frac{dU}{dx} \\
F_y &= q \cdot \frac{dU}{dy} \\
F_z &= q \cdot \frac{dU}{dz}
\end{align*}
\]  

[3.3]

where $q$ is the charge and $dU/dx$ the potential gradient in X-direction. More specifically, the force along x-direction is obtained from following function:

\[
f_x = q \cdot (1 - a)(1 - b)(1 - c) f(i, j, k) \\
+ (a)(1 - b)(1 - c) f(i + 1, j, k) \\
+ (1 - a)(1 - b)(1 - c) f(i, j + 1, k) \\
+ (a)(b)(1 - c) f(i + 1, j + 1, k) \\
+ (1 - a)(1 - b)(c) f(i, j, k + 1) \\
+ (a)(1 - b)(c) f(i + 1, j, k + 1) \\
+ (1 - a)(b)(c) f(i, j + 1, k + 1) \\
+ (a)(b)(c) f(i + 1, j + 1, k + 1)
\]  

[3.4]

where $a$, $b$ and $c$ are the fractional distances along the grid axes and so the actual location of the charge $q$ (i.e. the location of the superoxide ion) and $(i, j, k)$ are the grid coordinates (Briggs et al., 1989).
Except where the catalytic sites of SOD were defined as an absorption region, the rest of the enzyme was a reflecting to the substrate, superoxide ion. Whenever the substrate crosses the reflective boundary, a new random position is reassigned. To increase the efficiency of the algorithm, collision detection and time-step procedures are optimized with the following strategies: 1) Collision detection is stopped once the center of the substrate moves beyond a spherical shell with radius $r_{surf} = 38.9 \text{ Å}$ (See Figure 3.2).

![Diagram](image)

**Figure 3.2:** Illustration of collision-free distance $r_{surf}$ ($r_{surf} =$ radius of superoxide + distance between protein geometric center and the SOD outermost atom + radius of the SOD outermost atom).

2) Before performing Brownian dynamic simulations, the immobile protein is partitioned according to the Cartesian coordinate system, so that its atoms are grouped into 8 octants according to their geometric location in the system. An atom which intersects with the planes or axes of the coordinate system will be put into its corresponding groups. Collision detection of the moving substrate only checks against the protein atoms which share the same octant(s) as substrate. (See Figure 3.3)
3) The time step is gradually increased as the center of substrate propagates \( (r_{\text{away}}) \) away from the geometric center \( (G_{\text{center}}) \) of the protein.

\[
\Delta t = \begin{cases} 
4 \times 10^{-5} ns & (r_{\text{away}} \leq r_{\text{surf}}) \\
4 \times 10^{-7} ns + (r_{\text{away}} - r_{\text{surf}})^2 / 200D & (r_{\text{away}} > r_{\text{surf}})
\end{cases}
\]

[3.5]

The trajectories are terminated when either a reaction occurs at the active site or a preset cutoff time (typically 20 – 40 ns) is reached. For each trajectory, the life span is recorded and the survival probability over time \( S(t) \) is estimated. The value of \( S(t) \) is in turn fitted into equation 3.1 to obtain the steady rate coefficient \( k(t) \).

**Definition of chemical reaction**

At each step within the catalytic site of the enzyme, a random number \( R \), uniformly distributed between 1 and 0, is generated. If
$R \geq \exp\left\{-\frac{\Delta t k(\tilde{r}_n)/\omega + \Delta t k(\tilde{r})/\omega}{2}\right\}$, the substrate is considered to have reacted and the trajectory is stopped immediately. The value $\omega$ is the thickness of the catalytic site, $\tilde{r}_n$ and $\tilde{r}$ are the original and new positions of the substrate, respectively.

**Fitting of time dependent $k(t)$ curves**

Once the required number of trajectories have been reached, the steady-state rate coefficient, $k$, is extracted by fitting to the long-time limit of $k(t)$ (Zhou et al., 1997b) (See Figure 3.4).

$$k(t) = k \left(1 + \frac{k}{4\pi D} \frac{1}{\sqrt{\pi Dt}}\right) \quad [3.6]$$

**Figure 3.4:** The precise fitting of the time-dependent rate coefficient $k(t)$ of the native SOD to its asymptotic behavior allows us to determine the steady-state rate constant $k$ accurately.
3.2.2 Barnase and barstar protein-protein association

3.2.2.1 Molecular modeling

X-ray structure

The X-ray complex of Barnase and barstar was obtained from Brookhaven Protein Data Bank (Bernstein et al., 1977) entry 1brs; the missing heavy atoms were added using the Biopolymer program from Molecular Simulations, Inc. Hydrogen atoms were then added and energy minimized by the Discover3 program (Molecular Simulations, Inc) while the heavy atoms of the protein were fixed during the minimization process. The same steps are also applied to the mutation studies of barnase and barstar.

3.2.2.2 Calculation of protein-protein binding energy

Calculation of the electrostatic potentials of barnase, barstar and the barnase-barstar bound complex was performed with the UHBD package (Davis et al., 1991) in which the solvent dielectric constant and that of protein were set to be 4 and 78.5 respectively.

The average electrostatic binding energy between barnase and barstar in various configurations was calculated as:

\[
U = \frac{1}{2} \left( \sum_{\text{barnase-barstar complex}} q_i \Phi_i - \sum_{\text{barnase}} q_i \Phi_i - \sum_{\text{barstar}} q_i \Phi_i \right) \quad [3.7]
\]

where \( q_i \) is the partial charge of the protein atom and \( \Phi_i \) is the corresponding potential.

The solvation energy was calculated with different grid sizes to act as an indicator for the point of convergence. Consequentially, an optimized grid
of 60×60×60 points with 1.5 Å spacing, followed by a focusing grid of 140×140×140 grid points with 0.35Å spacing, was used to estimate the electrostatic energy of both barnase and barstar. The same strategy was used to obtain the electrostatic energy of the barnase-barstar complex (See Table 3.2).

<table>
<thead>
<tr>
<th></th>
<th>Barnase or barstar</th>
<th>Barnase-barstar complex (with a layer of water in between)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary grid size and</td>
<td>60×60×60 1.5Å</td>
<td>100×100×100 1.5Å</td>
</tr>
<tr>
<td>spacing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>First focusing grid</td>
<td>140×140×140 0.35Å</td>
<td>140×140×140 0.5Å</td>
</tr>
<tr>
<td>size, spacing &amp; focal</td>
<td>N-atom of mutated</td>
<td>N-atom of Arg59</td>
</tr>
<tr>
<td>point</td>
<td>residue</td>
<td></td>
</tr>
<tr>
<td>Second focusing grid</td>
<td>Not applicable</td>
<td>60×60×60 0.5Å</td>
</tr>
<tr>
<td>size, spacing &amp; focal</td>
<td></td>
<td>N-atom of Arg59</td>
</tr>
<tr>
<td>point</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 3.2:** Parameter settings for the electrostatic potential calculations. (refer to Appendix B for details).

Systemic errors in estimating the electrostatic potential \( \Phi_i \) and thus the electrostatic energy, are minimized by referencing to the self-potential \( \Phi_i^{self} \) calculated with the same protein and solvent dielectric constants of 4 and zero ionic strength.

\[
\Phi_i - \Phi_i^{self} + \sum_{i \neq j} \frac{q_j}{\epsilon \rho r_{ij}} \tag{3.8}
\]

**Generation of transition state configurations**

To mimic the transition state of the associating barnase-barstar encounter complex, a series of configurations were constructed from the X-ray
structure of the bound complex by the insertion of a layer of water between the proteins.

More specifically, first, the binding interface between the two protein units is mapped out by extracting all protein atoms within 4Å of the interface. A plane passing through the geometric center of the interface atoms and having least squared distance to the atoms is defined as the least-squared-fit plane. Second, the major vector is found by an axis passing through the geometric center of the interface atoms and along the normal of the least-squared-fit plane. Third, the proteins are allowed to adjust randomly their respective rotational \((\theta_A, \theta_B, \psi_A, \psi_B, \omega_A, \omega_B)\) (See Figure 3.5) and translational positions within an allowed range to produce multiple configurations in the transition state. Since the angles \(\theta_A\) and \(\theta_B\) are dependent variables, the change of \(\theta_B\) has no effect on the overall complex configuration.

![Graphic representation of rotational angles.](image)

**Figure 3.5:** Graphic representation of rotational angles.
3.2.2.3 Brownian dynamics simulation of force-free binding.

Generation of initial configurations

Instead of generating starting points within the active site like SOD, a sample of 200,000 configurations are randomly from uniform distribution at the interface of the barnase-barstar complex within the allowable scope of $\theta_{A,B}$, $\psi_{A,B}$, $\omega_{A,B}$ and translational positions. As before, the configurations where the proteins overlap are excluded from the initial configuration list. Only 31,761 non-overlapping configurations were selected.

Definition of reactive volume in transition state

With $\theta_{A,B} < 3^\circ$ and $d_e = 23.4\,\text{Å}$ (the center-center distance of the two protein units), the accommodation of a water layer in the interface requires a separation of the proteins along the major vector by $d = 8\,\text{Å}$. For 31,761 non-overlapping configurations, the reactive volume in the transition state is defined as

$$V_{r_v} = \frac{4\pi}{3} (31.4^3 - 23.4^3) \times \frac{1 - \cos 3^\circ}{2} \times \frac{1 - \cos 3^\circ}{2} \times \frac{31.761}{200,000}$$

$$= 5.67 \times 10^{-3} \, \text{Å}^3$$

Extension of BD simulation to cover rotational dynamics

Using the Ermak and McCammon algorithm (1978) as the framework, Zhou and Szabo (1996b) derived a general algorithm for simulating rotational motion. The Euler angles $(\theta, \phi, \psi)$ specify the orientation of the molecule in space and by making $x_1 = \cos \theta$, $x_2 = (\theta + \psi)/2$ and $x_3 = (\theta - \psi)/2$, the rotational diffusion equation in terms of these
variables is shown below (See Zhou & Szabo, 1996b: Appendix B for details):

\[ x_1 = x'_1 - \left[ 2x'_1 + \beta (1 - x'_1^2) \frac{\partial U}{\partial x_1} \right] D\Delta t + \left[ 2(1 - x'_1^2) \right]^{1/2} C_1 \]  \hspace{1cm} [3.9]

\[ x_2 = x'_2 - \frac{\beta}{2(1 + x'_1)} \frac{\partial U}{\partial x_2} D\Delta t + \left[ \frac{D\Delta t}{(1 + x'_1)} \right]^{1/2} C_2 \]  \hspace{1cm} [3.10]

\[ x_3 = x'_3 - \frac{\beta}{2(1 - x'_1)} \frac{\partial U}{\partial x_3} D\Delta t + \left[ \frac{D\Delta t}{(1 - x'_1)} \right]^{1/2} C_3 \]  \hspace{1cm} [3.11]

where \( C_1, C_2 \) and \( C_3 \) are normally distributed random numbers.

**Propagation of trajectories**

Using a similar routine to that described in the 3.2.1.2 section, the initial configurations started at the reactive binding interface propagate according to the Zhou and Szabo algorithm (Zhou et al., 1996b). The translational diffusion constant is set to be 30Å²/ns and the rotational diffusion constants are 0.04 and 0.045 ns⁻¹ for barnase and barstar respectively. The first order rate constant in the binding region is fixed to be \( k_m = 160\text{ns}^{-1} \). Conceptually, \( k_m \) can be regarded as the chemical kinetic rate of bond formation at the barnase-barstar interface.

When a trajectory crosses the reflective boundary, the configuration before the collision is restored and a new attempt is executed. The trajectories are terminated when either the stereospecific barnase-barstar complex is formed or the cutoff time is reached. The survival probability \( S(t) \) over time \( t \) can then be estimated.
Estimation of association rate constant

The steady-state association rate constant $k^0$ for native barnase and barstar is again analyzed by fitting $S(t)$ to the analytical asymptotic behavior (Zhou and Szabo, 1996b)

$$S(t) = \frac{k^0}{k_{in} V_{rv}} \left(1 + \frac{k^0}{4\pi D} \frac{1}{\sqrt{\pi Dt}}\right)$$ \[3.12\]

where $k_{in}$ is the first-order rate coefficient with the binding region and $V_{rv}$ is the volume of the binding region.

3.2.3 Parallelization of BD simulation

Since the individual trajectories in the BD simulation are statistically independent of each other, a large number of trajectories may be performed simultaneously using a parallel algorithm. The implementation is similar to that used by Bagheri et al. (1993) and Kozack et al. (1995) and the steps can be summarized as follows (See Appendix C)

1. The electrostatic potential at different grid points and parameter settings is transformed into binary form.

2. The binary data is broadcast to all the processors, called nodes.

3. A random number seed is assigned to different processors according to their node numbers.

4. Trajectories are evenly distributed among processors.

The above procedure reduces the running time by as much as 40 to 50 fold compared to a SGI impact 10000 if 128 nodes of an Intel Paragon supercomputer are used.
Chapter 4

RESULTS

4.1 Chapter summary

In this chapter, we firstly summarize the results of simulation studies of charge mutants of SOD and secondly, we compare the calculated transition state free energy of the barnase-barstar “encounter” complex with that from experimental results.

4.2 Human SOD

4.2.1 Surface charge properties of native superoxide dismutase and mutant

The surface electrostatic potentials of native SOD and the charge mutants were rendered using GRASP (Nicholls et al., 1991). Native SOD is shown in Figure 4.1. The ABFs for the native and mutant proteins are listed in Table 4.1. As expected, the Boltzmann factor (local electrostatic enhancement factor) declines with the distance of the mutated charged residue from the active site of SOD.
Figure 4.1: The positive and negative electrostatic potential on the SOD surface is shown in blue and red respectively. Positive potential is increased by charge mutation(s) of the labeled residues around the active site of SOD.

<table>
<thead>
<tr>
<th>Mutant</th>
<th>Distance between Cu at active site and C/C_e of the mutated residue. (Å)</th>
<th>Average Boltzmann factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wild</td>
<td>10</td>
<td>95.2</td>
</tr>
<tr>
<td>Gln-133</td>
<td>12</td>
<td>464.0</td>
</tr>
<tr>
<td>Gln-49</td>
<td>12</td>
<td>156.2</td>
</tr>
<tr>
<td>Gln-121</td>
<td>12</td>
<td>151.1</td>
</tr>
<tr>
<td>Asn-52</td>
<td>15</td>
<td>151.5</td>
</tr>
<tr>
<td>Asn-11</td>
<td>16</td>
<td>151.1</td>
</tr>
<tr>
<td>Mutant Combination</td>
<td>$k_{native}$</td>
<td>$k_{mutant}$</td>
</tr>
<tr>
<td>-------------------</td>
<td>--------------</td>
<td>--------------</td>
</tr>
<tr>
<td>Gln-132</td>
<td>17</td>
<td>148.2</td>
</tr>
<tr>
<td>Gln-133 &amp; Gln-49</td>
<td></td>
<td>759.3</td>
</tr>
<tr>
<td>Gln-133 &amp; Asn-52</td>
<td></td>
<td>736.9</td>
</tr>
<tr>
<td>Gln-133 &amp; Gln-132</td>
<td></td>
<td>726.4</td>
</tr>
<tr>
<td>Gln-49 &amp; Asn-52</td>
<td></td>
<td>249.1</td>
</tr>
<tr>
<td>Gln-49 &amp; Gln-132</td>
<td></td>
<td>244.2</td>
</tr>
<tr>
<td>Asn-52 &amp; Gln-132</td>
<td></td>
<td>236.3</td>
</tr>
</tbody>
</table>

Table 4.1: A list of the mutants and native SOD and their corresponding ABFs.

4.2.2 Prediction of electrostatic enhancement by Boltzmann factor

In general, the binding rate constants determined by BD simulation were in good agreement with the experimental results obtained from Getzoff et al., 1992 (See Table 4.2). However, can we predict the rate enhancement by ABF alone without having to perform a computationally-costly force field based BD simulation? A traditional BD simulation approach was performed to estimate the rate constants of native SOD $k_{native}$, and of the 10 single and double mutants, $k_{mutant}$. Then, the ratios of $k_{mutant}/k_{native}$ from this study were compared with the corresponding ratios determined from the Boltzmann factors (See Table 4.3). Table 4.3 showed that the approach of using ABF for prediction was very successful. There is an exception for those mutants containing Gln-133; the ABFs overestimate the increase in binding rate by a factor of ~2. Nevertheless, the ABFs correctly rank all the mutants in order of their binding rate constants.
<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Binding rate constant from experiment (k_i \times 10^7 \text{M}^{-1} \text{s}^{-1})</th>
<th>Binding rate constant from BD simulation (k_i \times 10^5 \text{M}^{-1} \text{s}^{-1})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Native</td>
<td>2.3</td>
<td>2.3</td>
</tr>
<tr>
<td>Gln-132</td>
<td>4.3</td>
<td>3.4</td>
</tr>
<tr>
<td>Gln-133</td>
<td>6.7</td>
<td>6.3</td>
</tr>
<tr>
<td>Gln-132 &amp; Gln-133</td>
<td>7.2</td>
<td>8.6</td>
</tr>
</tbody>
</table>

*Table 4.2:* A comparison of the binding rate constants of experiment and simulation.

<table>
<thead>
<tr>
<th>Mutant</th>
<th>Average Boltzmann factor</th>
<th>Binding rate constant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gln-133</td>
<td>4.9</td>
<td>2.8</td>
</tr>
<tr>
<td>Gln-49</td>
<td>1.6</td>
<td>1.5</td>
</tr>
<tr>
<td>Asn-52</td>
<td>1.6</td>
<td>1.5</td>
</tr>
<tr>
<td>Gln-132</td>
<td>1.6</td>
<td>1.5</td>
</tr>
<tr>
<td>Gln-133 &amp; Gln-49</td>
<td>8.0</td>
<td>3.8</td>
</tr>
<tr>
<td>Gln-133 &amp; Asn-52</td>
<td>7.7</td>
<td>3.7</td>
</tr>
<tr>
<td>Gln-133 &amp; Gln-132</td>
<td>7.6</td>
<td>3.8</td>
</tr>
<tr>
<td>Gln-49 &amp; Asn-52</td>
<td>2.6</td>
<td>2.2</td>
</tr>
<tr>
<td>Gln-49 &amp; Gln-132</td>
<td>2.6</td>
<td>2.2</td>
</tr>
<tr>
<td>Asn-52 &amp; Gln-132</td>
<td>2.5</td>
<td>2.2</td>
</tr>
</tbody>
</table>

*Table 4.3:* Correlation of ABFs and the electrostatic rate enhancement of binding rate constants.

### 4.3 Barnase-Barstar:

#### 4.3.1 Surface charge property

The surface potential of the native and mutant proteins were rendered by GRASP (See Figure 4.2 & 4.3).
Figure 4.2: The surface potential of native barnase. The electrostatic color-code is: blue, positive; red, negative. The residues present in the interface are labeled.

Figure 4.3: The surface potential of native barstar. The interface atoms are labeled.

4.3.2 BD simulation

At infinite ionic strength, the rate constant for the barnase-barstar association obtained from BD simulations is of the same order of
magnitude as the experimental rate, \( k^0 = 1.6 \times 10^5 \text{M}^{-1}\text{s}^{-1} \) (Schreiber and Fersht, 1996).

### 4.3.3 Electrostatic rate enhancement

By definition, the transition state free energy is

\[
G^{\dagger}_{\text{cal}} = -k_B T \ln \langle \exp(-\beta U) \rangle
\]

[4.1]

where \( \langle \exp(-\beta U) \rangle \) is the ABF. To verify the usefulness of equation 1.1 directly, the free energies \( G^{\dagger}_{\text{cal}} \) for the native and mutants of barnase-barstar complexes were calculated using equation 4.1. The results were then compared to the experimental data \( G^{\dagger}_{\text{exp}} = -k_B T \ln \left[ k(I = 0)/k(I = \infty) \right] \) (See Figure 4.4). The \( k(I = 0) \) and \( k(I = \infty) \) were taken from Fersht and co-workers (Schreiber and Fersht, 1993, 1995, 1996; Sreiber et al., to be published).

![Comparison of calculated and experimental results for the free energy of barnase-barstar "encounter" complex at transition state.](image)

**Figure 4.4:** Comparison of calculated and experimental results for the free energy of barnase-barstar “encounter” complex at transition state.
Chapter 5

DISCUSSION

5.1 Chapter summary

This chapter discusses possible errors in the calculations and the implication of the equation 1.1 on understanding on protein-ligand and protein-protein association. This provides a basis to designed enzymes that have optimized catalytic efficiency.

5.2 Local versus global on rate enhancement

According to the two-phase association mechanism (Schreiber et al., 1995), the protein or substrate molecule should diffuse to near the binding surface of another protein to allow a reaction to occur. Schematically, the two-phase association mechanism can be symbolized as:

\[ X + Y \xrightleftharpoons[k_{-1}]{k_1} X \cdot Y \xrightarrow{k_m} XY \]

where \( k_m \) is the first-order reaction rate constant, \( X \cdot Y \) stands for the transient reactive complex \( XY \) for the final bound complex. The terms \( k_1 \) and \( k_{-1} \) are the association and dissociation constants, respectively. Thus, the overall rate of association at the steady state is

\[ k = \frac{k_{in}}{k_{-1} + k_{in}} k_1 \]  \hspace{1cm} [5.1]
If $k_{in}$ is small compared to $k_i$ and $k_{-1}$, the ratio of $k_i$ and $k_{-1}$ becomes an equilibrium constant with the following relation (Shoup and Szabo, 1982)

$$\frac{k_i}{k_{-1}} = V_{br} \langle \exp(-\beta U) \rangle$$ \[5.2\]

Therefore, equation 5.1 can be rewritten as

$$k = \frac{k_{-1}k_{in}}{k_i + k_{in}} \langle \exp(-\beta U) \rangle V_{br}$$ \[5.3\]

provided that a negligible difference between the interactive potential of the binding region and its immediate surroundings (i.e. the near region) (See Figure 5.1). Under such circumstances, $k_{-1}$ would show no significant change, as it is insensitive to the uniform interaction potential. For a finite $k_{in}$, the binding rate constant $k$ is scaled by the ABF over the binding region. The interpretation of equation 5.3 would be $k\alpha\langle \exp(-\beta U) \rangle$ which is physically equivalent to equation 1.1.

---

**Figure 5.1:** Graphic representation of electrostatic potential dominance on rate enhancement. The potential in the active site and near region is “dominant” whereas the potential in the far region is “recessive” on rate enhancement.

In summary, if the electrostatic potential at the binding region and its near region change gradually, the diffusion process is only weakly dependent on
distance. This allows us to ignore the potential in the far region in estimating the order of rate enhancement by electrostatic terms.

### 5.3 Steering Mechanism

The proposal of a steering mechanism was developed by Getzoff et al. (1983) from the study of bovine SOD. Getzoff et al. suggested that the repulsive potential of the negatively-charged Glu-131 (the analog to Glu-133 in human SOD) located at the opening of the active site of the SOD plays an active role by guiding the superoxide ions towards the active site. That is, the charged residues on the protein surface might act like guides to steer the diffusing substrate particles to the active site. If this were the case, one would expect that charge neutralization of Glu-131 of bovine SOD or Glu-133 of human SOD would decrease the binding rate based on the steering argument. This is because the steering effect of the electrostatic field towards the active site is weakened. Later experimental results showed, however, that the binding rate is increased (Getzoff et al., 1992).

In addition, if the steering mechanism is correct, the dominance of the local interactive potential on rate enhancement becomes hard to explain. According to this mechanism, charge mutations far away from active site should have a significant effect on the binding rate. Again, this is not the case.

The results obtained by the ABF approach correctly predict the effect of local electrostatic enhancement, for example, the charge neutralization of Glu-133. Mechanistically, the local potential at the neutralized Glu-133 site becomes less repulsive or even attractive to the negative substrate and enhances the rate of binding by a reduction-in-dimensionality of the substrate diffusion path (Adam et al., 1968).

Since all the rate enhancements were done by charge alternation at the active site of the enzyme, one might suspect that the increase in attractive
potential would deteriorate the release rate of the product from the active site. The deterioration of the release rate was found to be negligible compared to the beneficial effect on the rate enhancement of the charge mutation (Zhou, 1997). This can be rationalized as $k_{-1}$ is insensitive to the interactive potential (the release of product species is the analog of the release of substrate from the active site). In short, the rate enhancement of binding outweighs the effect of release rate retardation caused by the charge mutation.

5.4 Minimal structural perturbation

The minimal structural perturbation process (all structures of the mutants resemble the original X-ray structures) used in creating the mutated structures for this work assumes that there are no structural changes at the active site of SOD or in the binding interface of barnase and barstar. This assumption may introduce a possible source of errors into the potential calculations. Buckle et al. (1993) performed a structural superposition study of unbound barnase and bound barnase-barstar complexes. The results indicated that there were obvious structural deviations at positions Asp-35 and Asp-39 of barstar after binding. Presumably, disruption of the salt bridge between barstar Asp-35 and barnase Arg-59 after mutation disturbs the structure of the complex. The calculated binding energies of the mutations D35A and D39A may therefore be underestimated by as much as 2 kcal/mol (unpublished data) when compared to experimental data (Schreiber et al., 1995).

The minimal structural perturbation procedure also appears to be the major reason for the overestimation of the values of $G_{\text{cal}}^{\dagger}(I = 0)$ for barnase D54A, E60A and E73W.
5.5 Insight of rational protein design from ABF approach

We have tested the capability of ABF approach in designing faster enzymes in chapter 4. The rate enhancement is given by the ABF in the binding according to the equation 1.1. Recently, protein design methodologies have targeted at increasing the catalytic step, $k_{cat}$ of enzymatic reactions by introducing transition metal ions into proteins (Barrick, 1995; Higaki et al., 1992; Regal, 1995). Our work provides a complementary approach to increase the binding constant $k_{on}$ of the two-phased enzymatic reaction (See Equation 5.4).

\[
\begin{align*}
E + S & \overset{k_{on}}{\underset{k_{off}}{\rightleftharpoons}} E \cdot S \overset{k_{cat}}{\rightarrow} ES
\end{align*}
\]

[5.4]

Pinto et al. (1997) constituted *E. Coli* thorredoxin (Trx), a protein that naturally contains no transition metal-binding site, with incorporation of a high-affinity metal-ion-binding site. The resulting proteins, \(\{\text{Fe}^{3+}\}\)Trx-SOD resembled the dismutation capability of native \(\{\text{Fe}^{3+}\}\)SOD but catalyzes the dismutation reaction at a much lower reaction rate of \(10^3 \text{M}^{-1} \text{s}^{-1}\). On the other hand, pure apoprotein, Trx-SOD showed no dismutation capability in vitro. Pinto et al. (1997) concluded that one of major factor slowing down the overall reaction rate was caused by the unfavorable electrostatic interactions at the binding surface of the \(\{\text{Fe}^{3+}\}\)Trx-SOD (See Figure 5.2). The vicinity of \(\{\text{Fe}^{3+}\}\)Trx-SOD binding site is similar to that of native SOD. Therefore, ERE is expected to boost the binding rate, $k_{on}$ by charge mutations near the binding site of the \(\{\text{Fe}^{3+}\}\)Trx-SOD.
Figure 5.2: The positively charged binding site of $\{\text{Fe}^{2+}\}$Trx-SOD is surrounded by negatively charged residues (Asp9, Asp10, Asp15, Asp61 and Glu31). This resembles the electrostatic environment of native SOD. Charge neutralization of the negatively charged residues will enhance the local ABF and the overall reaction rate.
CONCLUSION

6.1 Chapter summary

This chapter summarizes this research project, suggests possible areas of improvement, and recommends further studies of the topic.

6.2 Local electrostatics is dominant

Local electrostatic interactions in the binding region of proteins have a dominant effect on electrostatic rate enhancement as shown in chapter 4. The ABF approach captures the essential feature of this phenomenon as demonstrated in the result section of this thesis and it can be used as a guide for protein engineering. Enzymes that are more efficient can be designed by charge optimization at the active site of the enzyme and promising species can be screened by the use of the ABF method.

6.3 Electrostatic steering is misleading

The previously proposed method of electrostatic rate enhancement by “electrostatic steering” or “electrostatic guidance” may be conceptually misleading (See section 5.3). This is because it fails to explain the rate increase from charge neutralization of Glu133 to Gln133 for human SOD (Getzoff et al., 1992).

In contrast, the agreement between the predicted rate enhancement by ABF and the experimental results clearly shows that the decrease in dimensionality plays the major role in enhancing the two-phase protein-
protein or protein-ligand binding processes. The substrate particles reduce their search time for an active site via electrostatic attraction to the protein surface first and then through a two-dimensional search for the active site on that surface.

6.4 Provide a computation inexpensive alternative to tradition BD simulation

All the studies of SOD and barnase-barstar described above using the ABF concept have definitely simplified the role of electrostatic interactions on rate enhancement. ABF provides a computationally inexpensive alternative to traditional BD simulations.

6.5 Reduction of a kinetic problem to a thermodynamic problem

The ABF data determined for the binding interface of barnase and barstar (See section 4.3.3) captures the electrostatic role in both the kinetic and thermodynamic processes of protein-protein binding. Provided that the electrostatic potential in the binding site and the near region changes gradually, if the free energy change in a transitional state is known, then the rate enhancement can be correctly predicted or vice versa.

6.6 Future goals and prospects

6.6.1 Effective charge method

The applicability of either the BD or ABF approach in accounting for the electrostatic enhancement depends highly on the accuracy of the potential obtained by solving the Poisson-Boltzmann equation. The newly introduced effective charge method (Gabdoulline et al., 1996) could provide an alternate way to calculate the electrostatic potential.
6.6.2 Validation of the average Boltzmann relationship

As an extension of fundamental theoretical works (Zhou 1993a, 1996a), ongoing efforts are focused on verifying the validity of equation 1.1 in real protein systems. The following table summarizes the work completed as well as possible foci for future research.

<table>
<thead>
<tr>
<th>Enzyme / Protein</th>
<th>Status</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acetylcholine esterase</td>
<td>Completed</td>
<td>Zhou et al., 1996c</td>
</tr>
<tr>
<td>Superoxide dismutase</td>
<td>Completed</td>
<td>Zhou, Wong and Vijayakumar, 1997</td>
</tr>
<tr>
<td>Barnase &amp; barstar</td>
<td>Completed</td>
<td>Vijayakumar, Wong and Zhou (in press)</td>
</tr>
<tr>
<td>Thrombin &amp; hirudrin</td>
<td>Future</td>
<td>Preliminary studies by Stone et al., 1989</td>
</tr>
<tr>
<td>Haemoglobin α &amp; β</td>
<td>Future</td>
<td>Preliminary studies by Bunn &amp; McDonald, 1983; Mrabet et al., 1986</td>
</tr>
</tbody>
</table>

Table 6.1: Summary of the research project.
APPENDIX A

! Compute the potential at each grid point of the native human SOD
! (Sample input file for UHBD program)

read
  mol 1 ! protein SOD-dimer
  file "wild.pdb" ! from file "pdb1sos.pdb"
  pdb ! in PDB free-format format
end

read
  mol 2 ! superoxide free radical geometry
  file "sofr.pdb" ! from file "sofr.pdb"
  pdb ! in PDB free-format format
end

set
  charge ! charges
  radii ! and radii
  param HUSOD ! using parameter set “HUSOD”
  file "sod.dat" ! from the file "sod.dat"
end

print
  apar mol1 ! print to standard output
  ! atomic parameters (coord, charge, radii)
end

print
  apar mol2 ! print to standard output
  ! atomic parameters (coord, charge, radii)
end

elec
  calc ! calculate
  mol 1 ! molecule 1 electrostatic potential
  pdie 4.0 ! set protein dielectric constant \( \varepsilon_p=4.0 \)
  sdie 78.5 ! set solvent dielectric constant \( \varepsilon_s=78.5 \)
  maxits 400 ! increase the number of iterations
  ! that can be used to solve the FDPBE
  temp 298.00 ! set temperature in K (default=300.00K)
  dime 140 140 140 ! define the grid points(Max=140x140x140)
  ionstr 10.0 ! ionic strength(mM)
spacing 1.0  ! using a grid spacing of #.# Å
smooth   ! use dielectric boundary smoothing module
center   ! put grid center at molecule 1
bcfl 0   ! using boundary condition 0 (phi=0)
nmap 0.0 ! radius of ion(solvent)
nsph 500 ! no. of surface points
end

write     ! write
  phi     ! electro-potential (phi.grd)
end
stop
APPENDIX B

! Compute the binding energy of Barstar to
! the active site of barnase focusing on the site of mutation
! (Sample input file for UHBD program)
!
! read in the coords., charges, and radii in CHARMM qcard
! format
!
read mol 1 file "wild_wild_cm_001.qcd" qcard end ! both barnase &
barstar
!
! Regular Runs:
!    coarse grid
!    molecule complex
!
elec calc mol 1
    pdie  4.00     ! internal dielectric=4
    sdie  78.00    ! solvent dielectric=78
    temp 300.00    ! temperature in K
    ions  0.00     ! ionic strength
    bcfl  2        ! boundary condition - each atom
                    ! is a Debye-Huckel sphere
    eflid .00      ! external electric field
    grid 1.50      !  grid spacing
    dime 100 100 100 ! grid dimension
    center        ! center the grid on the molecule
    maxit 200      ! maximum # of iteration for FDPBE
end

print elec phizero mol1 end ! zero out the phisite accumulator
print elec phisave mol1 end ! compute and store phi at atoms on grid
!
!    fine grid
!    molecule complex
!
elec calc mol 1
    pdie  4.00     ! internal dielectric=4
    sdie  78.0    ! solvent dielectric=78
temp 300.00  ! temperature in K
ions 0.00   ! ionic strength
bcfl 4     ! focusing - use the coarse grid to
            ! set the boundary potential of this
            ! focused grid.
efld 0.00   ! external electric field
grid 0.500  ! spacing for focused grid
dime 140 140 140! grid dimension
center
maxit 20    ! maximum # of iteration for FDPBE
end

print elec phisave mol1 end
!
!
!    fine grid
!    molecule complex
!
elec calc mol 1
   pdie 4.00   ! internal dielectric=4
   sdie 78.00  ! solvent dielectric=78
   temp 300.00 ! temperature in K
   ions 0.00   ! ionic strength
   bcfl 4      ! focusing - use the coarse grid to
                ! set the boundary potential of this
                ! focused grid.
efld .00     ! external electric field
   grid 0.250  ! spacing for focused grid
dime 60 60 60 ! grid dimension
gcenter -0.091 9.598 9.306
maxit 200    ! maximum # of iteration for FDPBE
end

print elec phisave mol1 end
print elec phinrg mol1 end
!
! Reference Runs:
!    coarse grid
!    molecule complex
!
elec calc mol 1
   pdie 4.00  ! internal dielectric=4
sdie 4.00   ! exterior dielectric=4. This is
          ! the computation for the self
          ! energy
temp 300.0  ! temperature in K
ions 0.00   ! ionic strength
bcfl 2      ! boundary condition - each atom
          ! is a Debye-Huckel sphere
efld .00    ! external electric field
grid 1.50   ! grid spacing
dime 100 100 100 ! grid dimension
center      ! center the grid on the molecule
maxit 200   ! maximum # of iteration for FDPBE
end

print elec phizero mol1 end    ! zero out the phisite accumulator
print elec phisave mol1 end     ! compute and store phi at atoms on
grid                           !

! fine grid
! molecule complex
!

elec calc mol 1

  pdie 4.00    ! internal dielectric=4
  sdie 4.00    ! exterior dielectric=4. This is
               ! the computation for the self
               ! energy
  temp 300.00  ! temperature in K
  ions 0.00    ! ionic strength
  bcfl 4       ! focusing - use the coarse grid to
               ! set the boundary potential of this
               ! focused grid.
efld .00      ! external electric field
  grid 0.500   ! spacing for focused grid
dime 140 140 140 ! grid dimension
center
maxit 200     ! maximum # of iteration for FDPBE
end

print elec phisave mol1 end

! fine grid
! molecule complex
!

58
elec calc mol 1

  pdie  4.00  ! internal dielectric=4
  sdie  4.00  ! exterior dielectric=4. This is
               ! the computation for the self
               ! energy
  temp 300.00  ! temperature in K
  ions  0.00  ! ionic strength
  bcfl  4  ! focusing - use the coarse grid to
            ! set the boundary potential of this
            ! focused grid.
  efld .00  ! external electric field
  grid  0.250  ! spacing for focused grid
  dime  60 60 60  ! grid dimension
  gcenter -0.091 9.598 9.306
  maxit 200  ! maximum # of iteration for FDPBE

end

print elec phisave mol1 end
print elec phinrg mol1 end

stop
**APPENDIX C**

*Paralleled version of Brownian Dynamic Simulation Program for* Intel Paragon supercomputer

*Copyright© 1996-98 Laboratory of Dr. H.-X. Zhou at HKUST.*
*All rights reserved.

*Program:

Main program pliff.f

*Function:

This code simulates an ion around a protein to obtain its lifetime (before reaction). A force is present.

*Input Files:

Unit 1---- 'tmp.int', input file
Unit 10---- 'tmp.fdat', electrostatic force data
Unit 12---- 'tmp.dat', initial substrate position data

*Output Files:

Unit 15---- 'tmp.liff', survival time

*Input Data:

NN ------ Maximum number of elements
Maxnas-- Maximum number of active sites
Nf ------ Maximum number of elements
in 1-D array of 3-D matrix
NN1 ----- Number of protein atoms
nas ------ Number of active sites
cent(Maxnas) --- Record of active site number
xav,yav,zav --- Average coordination of protein center
sepout --- Distance between outermost atoms and average center
rout ------ Range under the influence of force
rrad ------ Radius of active site
rion ------ Radius of ion
ntrj ------ Total number of trajectory
idum ------ Random seed, must be a -ve integer
D ------ Relative diffusion constant
lfcut ------ Cutoff time
dti ------ Time step in active site
kinhalf --- Half reactivity constant
qion ------ Net charge of ion
integer Maxnas
parameter (Maxnas=10)

integer Nf
parameter (Nf=140)

real beta
parameter (beta=1.6606E-27*1.0E+4/1.3807E-23/300.0)

integer NN
parameter (NN=5000)

integer MM
parameter (MM=50000)

integer NN1,nas,cent(Maxnas),ntrj,idum

real xav,yav,zav
real sepout,rout
real rrad, rion
real rcent(Maxnas)
real D,Lfcut,dti,kinhlf
real qion,qprtn
real phi(Nf*Nf*Nf),fx, fy, fz

class*12 chtime
common /set1/ NN1,nas,cent,xav,yav,zav,rcent
common /set2/ rrad
common /set3/ rion
common /fx/fx
common /fy/fy
common /fz/fz

integer inext,inextp,ma(55)
real ran1,ran2
real s0,s1,s2,s3
real xo,yo,zo,ro,x,y,z,r

integer bump,reacrgn,oldin
integer i,j,k,trj
real life
real dt,fac,fac1,reac

integer np,me,nptrj
real xxo(MM),yyo(MM),zzo(MM),llife(MM)
real xx(NN),yy(NN),zz(NN),rr(NN)
common /xxo/xxo,yyo,zzo
common /xx/xx,yy,zz,rr

C Program start.
open(1, file='tmp.int')
read(1,'(2I6)')NN1,nas
read(1,'(10I6)')(cent(i),i=1,nas)
read(1,'(3F15.4)')xav,yav,zav
read(1,'(2F15.2)')sepopout,rout
read(1,'(2F15.5)')rrad,rion
read(1,'(10F8.2)')(rcent(i),i=1,nas)
read(1,'(4E15.4))D,lfcut,dti,kinhlf
read(1,'(2F15.2)')qion,qrtn
read(1,'(2F6)')ntrj,idum
close(1)

np=numnodes()
me=mynode()
idum=idum*(me+1)

mytrj=ntrj[np]
if(ntrj.ne.mytrj*np) mytrj=mytrj+1

62
if (me.eq.0) then
  call runtime(chtime)
  print*, chtime, ntrj, D, sepout, lfcut, NN1, rrad,
  cent(1), cent(2), xav, yav, zav, rion
  print*, chtime, 'Total number of trajectories: ', ntrj
  print*, chtime, 'Total number of nodes: ', np
  print*, chtime, 'Number of trajectories on each node: '
  mytrj
endif

C Calculate useful quantities and assign variables

ran2=ran1(idum, inext, inextp, ma)
open(11, file='tmpbin.dat', form='UNFORMATTED', status='old'
read(11) (x(i), y(i), z(i), i=1, ntrj)
read(11) (x(i), y(i), z(i), r(i), i=1, NN1)
close(11)

call readcor()
call binpgrd(phi)
call gsync()

C Propagate Brownian trajectories

if (ntrj - me*mytrj.ge.mytrj) then
  nmytrj = mytrj
else
  nmytrj = ntrj - me*mytrj
endif

if (nmytrj.lt.0) nmytrj = 0

call gsync()
do 500 trj = 1, nmytrj
  xo = x(i) + trj
  yo = y(i) + trj
  zo = z(i) + trj
  ro = sqrt(x(i)**2 + y(i)**2 + z(i)**2)
  oldin = 1
dt = dti
  life = 0.0
life=life+dt
if (life .gt. lfcut) goto 400

call gauss(s0,s1,Idum,inext,inextp,ma)
call gauss(s2,s3,Idum,inext,inextp,ma)
call bdfrece(qion,phi,xo+xav,yo+yav,zo+zav,fx,fy,fz)
fac=sqrt(2.0*D*dt)
fac1=D*dt*beta
x=xo+fac1*fx+fac*s0
y=yo+fac1*fy+fac*s1
z=zo+fac1*fx+fac*s2
r=sqrt(x**2+y**2+z**2)

reac=exp(-dt*kinhlf)

if (r .lt. sepout) then
  if (bump(x,y,z) .eq. 1) then
    if (oldin .eq. 1) then
      ran2=ran1(Idum,inext,inextp,ma)
      if (ran2 .ge. reac**2) goto 400
    endif
    goto 333
  endif
endif

if (reacrgn(nas,x,y,z) .eq. 1) then
  if (oldin .eq. 1) then
    ran2=ran1(Idum,inext,inextp,ma)
    if (ran2 .ge. reac**2) goto 400
  else
    ran2=ran1(Idum,inext,inextp,ma)
    if (ran2 .ge. reac) goto 400
  endif
  oldin=1
else
  if (oldin .eq. 1) then
    ran2=ran1(Idum,inext,inextp,ma)
    if (ran2 .ge. reac) goto 400
  endif
  oldin=0
endif
dt=dti
else
  if (oldin .eq. 1) then
    ran2=ran1(Idum,inext,inextp,ma)
    if (ran2 .ge. reac) goto 400
  endif
endif
endif
oldin=0
dt=dti+1.0E-2*(r-sepout)**2/(2.0*D)
endif

xo=x
yo=y
zo=z
ro=r
goto 333

400  llife(trj)=life
if(me.eq.0) then
call runtime(chtime)
endif

500  continue

call gsync()

if(me.ne.0) then
call csend(me,nmytrj,4,0,0)
call csend(me,llife(1),4*nmytrj,0,0)
elsethe
  do i=1,np-1
    call crecv(i,nmytrj,4)
    call crecv(i,llife(mytrj*i+1),4*nmytrj)
  enddo
endif

call gsync()

if(me.eq.0) then
  open(15,file='tmp.liff')
do i=1,ntrj
  write(15, '(E15.5)') llife(i)
enddo
close(15)
call runtime(chtime)
print*,chtime,'Program ended.'
endif

call gsync()

stop
end

subroutine gauss(z1,z2,idum,inext,inextp,ma)
Generate Gaussian distribution.

real z1,z2,z
integer idum,inext,inextp,ma(55)
real ran1,ran2

ran2=ran1(idum,inext,inextp,ma)
z1=2.0*ran2-1.0
ran2=ran1(idum,inext,inextp,ma)
z2=2.0*ran2-1.0
z=z1*z1+z2*z2
if (z.ge. 1.0) go to 1
z=sqrt(-2.0*alog(z)/z)
z1=z1*z
z2=z2*z

return
end

subroutine readcor()
This code reads in coordinates of surface atoms of the two proteins.

integer NN
parameter (NN=5000)

integer Maxnas
parameter (Maxnas=10)

integer NN1,nas,cen(t(Maxnas)
real xav,yav,zav,rcen(t(Maxnas)
common /set1/ NN1,nas,cen,xav,yav,zav,rcen

real x1(NN),y1(NN),z1(NN),r1(NN)
common /array1/ x1,y1,z1,r1

real x2(NN),y2(NN),z2(NN),r2(NN)
common /array2/ x2,y2,z2,r2
real x3(NN),y3(NN),z3(NN),r3(NN)
common /array3/ x3,y3,z3,r3

real x4(NN),y4(NN),z4(NN),r4(NN)
common /array4/ x4,y4,z4,r4

real x5(NN),y5(NN),z5(NN),r5(NN)
common /array5/ x5,y5,z5,r5

real x6(NN),y6(NN),z6(NN),r6(NN)
common /array6/ x6,y6,z6,r6

real x7(NN),y7(NN),z7(NN),r7(NN)
common /array7/ x7,y7,z7,r7

real x8(NN),y8(NN),z8(NN),r8(NN)
common /array8/ x8,y8,z8,r8

real asx(Maxnas),asy(Maxnas),asz(Maxnas)
common /asite/ asx,asy,asz

integer n1,n2,n3,n4,n5,n6,n7,n8
common /params/ n1,n2,n3,n4,n5,n6,n7,n8

real xx(NN),yy(NN),zz(NN),rr(NN)
common /xx/xx,yy,zz,rr

real x,y,z,r

real xt,yt,zt

integer xpi,ypi,zpi,xpf,ypf,zpf

integer i,j,k,l

n1=0
n2=0
n3=0
n4=0
n5=0
n7=0
n8=0

do j=1,NN1
  x=xx(j)
  y=yy(j)

67
\( z = z z(j) \)
\( r = r r(j) \)
\( x = x - x a v \)
\( y = y - y a v \)
\( z = z - z a v \)

\[ \text{do i=1,nas} \]
\[ \text{if (j .eq. cent(i)) then} \]
\[ \text{asx(i)=x} \]
\[ \text{asy(i)=y} \]
\[ \text{asz(i)=z} \]
\[ r=rcent(i) \]
\[ \text{endif} \]
\[ \text{enddo} \]

\( K=-1 \)
\[ \text{do l=0,6} \]
\[ x t=x \]
\[ y t=y \]
\[ z t=z \]
\[ \text{if (l .eq. 0) then} \]
\[ x p i=0 \]
\[ y p i=0 \]
\[ z p i=0 \]
\[ \text{if (xt .ge. 0) xpi=1} \]
\[ \text{if (yt .ge. 0) ypi=1} \]
\[ \text{if (zt .ge. 0) zpi=1} \]
\[ \text{endif} \]

\[ \text{if ((l .eq. 1) .or. (l .eq. 2)) then} \]
\[ x p f=0 \]
\[ \text{if (l .eq. 1) xt=x+r} \]
\[ \text{if (l .eq. 2) xt=x-r} \]
\[ \text{if (xt .ge. 0) xpf=1} \]
\[ \text{endif} \]

\[ \text{if ((l .eq. 3) .or. (l .eq. 4)) then} \]
\[ y p f=0 \]
\[ \text{if (l .eq. 3) yt=y+r} \]
\[ \text{if (l .eq. 4) yt=y-r} \]
\[ \text{if (yt .ge. 0) ypf=1} \]
\[ \text{endif} \]

\[ \text{if ((l .eq. 5) .or. (l .eq. 6)) then} \]
\[ z p f=0 \]

68
if (l.eq. 5) zt=z+r
if (l.eq. 6) zt=z-r
if (zt.ge. 0) zpf=1
endif

C  Check any overlapping of elements in octacts

if ((l.gt. 0) .and. (l.le. 2) .and. (xpi.eq. xpf)) goto 222
if ((l.gt. 2) .and. (l.le. 4) .and. (ypi.eq. ypf)) goto 222
if ((l.gt. 4) .and. (l.le. 6) .and. (zpi.eq. zpf)) goto 222
k=k+1

C  (x,y,z +,+,+)

if (((xt.ge. 0) .and. (yt.ge. 0) .and. (zt.ge. 0)) then
   n1=n1+1
   x1(n1)=x
   y1(n1)=y
   z1(n1)=z
   r1(n1)=r
endif

C  (x,y,z +,-,+)

if (((xt.ge. 0) .and. (yt.lt. 0) .and. (zt.ge. 0)) then
   n2=n2+1
   x2(n2)=x
   y2(n2)=y
   z2(n2)=z
   r2(n2)=r
endif

C  (x,y,z -,-,+)

if (((xt.lt. 0) .and. (yt.lt. 0) .and. (zt.ge. 0)) then
   n3=n3+1
   x3(n3)=x
   y3(n3)=y
   z3(n3)=z
   r3(n3)=r
endif

C  (x,y,z -,+,+)

if (((xt.lt. 0) .and. (yt.ge. 0) .and. (zt.ge. 0)) then
   n4=n4+1
x4(n4)=x
y4(n4)=y
z4(n4)=z
r4(n4)=r
endif

C (x,y,z +,+,-)

if ((xt .ge. 0) .and. (yt .ge. 0) .and. (zt .lt. 0)) then
  n5=n5+1
  x5(n5)=x
  y5(n5)=y
  z5(n5)=z
  r5(n5)=r
endif

C (x,y,z +,-,-)

if ((xt .ge. 0) .and. (yt .lt. 0) .and. (zt .lt. 0)) then
  n6=n6+1
  x6(n6)=x
  y6(n6)=y
  z6(n6)=z
  r6(n6)=r
endif

C (x,y,z -,-,-)

if ((xt .lt. 0) .and. (yt .lt. 0) .and. (zt .lt. 0)) then
  n7=n7+1
  x7(n7)=x
  y7(n7)=y
  z7(n7)=z
  r7(n7)=r
endif

C (x,y,z -,+,+)

if ((xt .lt. 0) .and. (yt .ge. 0) .and. (zt .lt. 0)) then
  n8=n8+1
  x8(n8)=x
  y8(n8)=y
  z8(n8)=z
  r8(n8)=r
endif
integer function bump(x,y,z)
This code tests whether the ion bumps into the protein.

integer NN
parameter (NN=5000)

real rion
common /set3/ rion

real x,y,z
real xt,yt,zt
real x1(NN),y1(NN),z1(NN),r1(NN)
common /array1/ x1,y1,z1,r1
real x2(NN),y2(NN),z2(NN),r2(NN)
common /array2/ x2,y2,z2,r2
real x3(NN),y3(NN),z3(NN),r3(NN)
common /array3/ x3,y3,z3,r3
real x4(NN),y4(NN),z4(NN),r4(NN)
common /array4/ x4,y4,z4,r4
real x5(NN),y5(NN),z5(NN),r5(NN)
common /array5/ x5,y5,z5,r5
real x6(NN),y6(NN),z6(NN),r6(NN)
common /array6/ x6,y6,z6,r6
real x7(NN),y7(NN),z7(NN),r7(NN)
common /array7/ x7,y7,z7,r7
real x8(NN),y8(NN),z8(NN),r8(NN)
common /array8/ x8,y8,z8,r8

integer n1,n2,n3,n4,n5,n6,n7,n8
common /params/ n1,n2,n3,n4,n5,n6,n7,n8
integer xpi,ypi,zpi,xpf,ypf,zpf

integer 1

do l=0,6
   xt=x
   yt=y
   zt=z
   if (l.eq.0) then
      xpi=0
      ypi=0
      zpi=0
      if (xt.ge.0) xpi=1
      if (yt.ge.0) ypi=1
      if (zt.ge.0) zpi=1
   endif

   if ((l.eq.1).or.(l.eq.2)) then
      xpf=0
      if (l.eq.1) xt=x+rion
      if (l.eq.2) xt=x-rion
      if (xt.ge.0) xpf=1
   endif

   if ((l.eq.3).or.(l.eq.4)) then
      ypf=0
      if (l.eq.3) yt=y+rion
      if (l.eq.4) yt=y-rion
      if (yt.ge.0) ypf=1
   endif

   if ((l.eq.5).or.(l.eq.6)) then
      zpf=0
      if (l.eq.5) zt=z+rion
      if (l.eq.6) zt=z-rion
      if (zt.ge.0) zpf=1
   endif

   if ((l.gt.0).and.(l.le.2).and.(xpi.eq.xpf)) goto 444
   if ((l.gt.2).and.(l.le.4).and.(ypi.eq.ypf)) goto 444
   if ((l.gt.4).and.(l.le.6).and.(zpi.eq.zpf)) goto 444

   C (x,y,z +,+,+)

   if ((xt.ge.0).and.(yt.ge.0).and.(zt.ge.0)) then
      if (crash(x,y,z,n1,x1,y1,z1,r1).eq.1) then
bump=1
goto 333
endif

C (x,y,z +,-,+)
if ((xt .ge. 0) .and. (yt .lt. 0) .and. (zt .ge. 0)) then
  if (crash(x,y,z,n2,x2,y2,z2,r2) .eq. 1) then
    bump=1
    goto 333
  endif
endif

C (x,y,z -,-,+)
if ((xt .lt. 0) .and. (yt .lt. 0) .and. (zt .ge. 0)) then
  if (crash(x,y,z,n3,x3,y3,z3,r3) .eq. 1) then
    bump=1
    goto 333
  endif
endif

C (x,y,z -,-,+)
if ((xt .lt. 0) .and. (yt .ge. 0) .and. (zt .ge. 0)) then
  if (crash(x,y,z,n4,x4,y4,z4,r4) .eq. 1) then
    bump=1
    goto 333
  endif
endif

C (x,y,z +,+,-)
if ((xt .ge. 0) .and. (yt .ge. 0) .and. (zt .lt. 0)) then
  if (crash(x,y,z,n5,x5,y5,z5,r5) .eq. 1) then
    bump=1
    goto 333
  endif
endif

C (x,y,z +,-,-)
if ((xt .ge. 0) .and. (yt .lt. 0) .and. (zt .lt. 0)) then
  if (crash(x,y,z,n6,x6,y6,z6,r6) .eq. 1) then
    bump=1
goto 333
endif

dendif

C
(x,y,z -,-,-)

if (((x .lt. 0) .and. (y .lt. 0) .and. (z .lt. 0))) then
    if (crash(x,y,z,n7,x7,y7,z7,r7) .eq. 1) then
        bump=1
        goto 333
    endif
endif

C
(x,y,z -,+,-)

if (((x .lt. 0) .and. (y .ge. 0) .and. (z .lt. 0))) then
    if (crash(x,y,z,n8,x8,y8,z8,r8) .eq. 1) then
        bump=1
        goto 333
    endif
endif

444
enddo
bump=0

333
return
end

C
C
integer function crash(x,y,z,n1,xArray,yArray,zArray,rArray)
C
This code tests whether the ion collides with protein atom(s).
C
integer NN
parameter (NN=5000)

real rion
common /set3/ rion

real x,y,z
real xArray(NN),yArray(NN),zArray(NN),rArray(NN)
integer n1,n
real dis

n=1
crash=0
do while ((n .le. n1) .and. (crash .eq. 0))
dis=(xArray(n)-x)**2+(yArray(n)-y)**2+(zArray(n)-z)**2
if (dis .lt. (rArray(n)+rion)**2) then
    crash=1
endif
n=n+1
enddo
return
end

C

C

integer function reacrgrn(nas,x,y,z)
This code test whether the ion is in the reactive region.

C

C

integer     NN
parameter (NN=5000)

integer     Maxnas
parameter (Maxnas=10)

real       rrad
common /set2/ rrad

integer     nas
real       x,y,z
real       dis

real       asx(Maxnas),asy(Maxnas),asz(Maxnas)
common /asite/ asx,asy,asz

integer     i

i=1
reacrgrn=0

do while ((i .le. nas) .and. (reacrgrn .eq. 0))
    dis=(x-asx(i))**2+(y-asy(i))**2+(z-asz(i))**2
    if (dis .le. rrad**2) then
        reacrgrn=1
    endif
    i=i+1
enddo
return
end
GLOSSARY

**Average Boltzmann factor (ABF)**
A local electrostatic term for description of the association rate enhancement due to electrostatic interaction between biomolecules.

**Boundary element method (BE)**
A technique used to calculate the electrostatic potential of molecules

**Brownian dynamic simulation (BDS)**
A computational technique that used to simulate the random movement of particles.

**Deoxyribonucleic Acid (DNA)**
A molecule which bears the genetic information of the individual organisms.

**Electrostatic rate enhancement (ERE)**
A increase in the association rate of two biomolecules caused by intermolecular electrostatic interaction

**Escherichia Coli (E. Coli)**
A bacterium commonly found in the intestine of animals.

**Facultative species**
Bacterial species possess the anti-oxidation enzymes even they are not aerobes.

**Finite difference Poisson-Boltzmann method (FDFB)**
An alternative method used to calculate the electrostatic potential of molecules

**Respiratory burst**
An event of immune response in which oxidative free radicals are released by neutrophils and macrophages to kill the invasing germs.

**Superoxide dismutase (SOD)**
An enzyme for converting superoxide free radical into less reactive hydrogen peroxide.
REFERENCES


84


