INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA

UMI
800-521-0600
NOTE TO USERS

This reproduction is the best copy available

UMI
BROWNIAN DYNAMICS STUDY OF THE INTERACTION BETWEEN CYTOCHROME F AND MOBILE ELECTRON TRANSFER PROTEINS

DISSERTATION

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By

Douglas C. Pearson Jr., B.S.

The Ohio State University 1999

Dissertation Committee:
Professor Elizabeth Gross, Adviser
Professor James Alben
Professor Charles Brooks
Professor Richard Swenson

Approved by

Adviser
Biophysics Graduate Program
Copyright by
Douglas C. Pearson Jr.
1999
ABSTRACT

Cytochrome f (cyt f), a member of the membrane-bound cytochrome b_{fj} complex, is a participant in electron transport in photosynthetic organisms. It receives an electron from the Rieske Fe-S protein, also within the cyt b_{fj} complex, and passes that electron on to plastocyanin (PC) in higher plants and algae, a beta-barrel copper protein, or to cytochrome c_{6} (c_{6}) in algae and cyanobacteria, an alpha-helical heme protein.

Despite the differences in the proteins, the orientation of negatively charged residues on the surface of PC and c_{6} generates a similar potential field. There is an interaction between those negative charges and positive charges on cyt f in complex formation between the two pairs of molecules.

It was postulated, based on electrostatic modeling and available biochemical data, that electrostatic interactions draw the acidic patch of PC and the basic patch of cyt f into a pre-docking configuration in which there is no interaction between any potential electron-transfer surfaces, after which rearrangement results in a electron-transfer active complex in which H87 on PC, a ligand to the copper atom, would contact Y1 on cyt f, the sixth ligand to the heme. Electrostatic interactions would not be sufficient to form such a complex; hydrophobic interactions would likely be required.
Brownian dynamics (BD), simulating Brownian motion of PC in the presence of electrostatic attraction or repulsion to cyt f, was then used to model the interaction between PC and cyt f quantitatively and qualitatively and evaluate this hypothesis. Electrostatic interactions tended to draw PC into a single dominant orientation with respect to cyt f; such a orientation was observed to be similar to the postulated electron-transfer active complex mentioned above. Experiments between poplar PC and turnip cyt f agreed with later spinach PC/turnip cyt f and *Chlamydomonas* PC/cyt f experiments; the tendency to form the same dominant complex remains.

BD simulations were then used to model the action of mutants of PC and cyt f residues for comparison with previous results and to provide predictions for future experiments, and to simulate the weaker electrostatic interaction between *Chlamydomonas* c₆ and cyt f.
For Amelia.
("Daddy, is the thesis done yet?")
ACKNOWLEDGMENTS

The first acknowledgment I ever received in a scientific publication came out of the laboratory of Dr. Alex Green at the University of Florida; Dr. Green served as a mentor on a summer science training program project the summer before my senior year in high school. It was out of that experience, sitting in front of a Zenith personal computer (monochrome monitor, 5 ¼ inch floppy drive, and this wonderful piece of hardware I’d never worked with before called a "hard disk"), running Fortran code that simulated the function of a waste incinerator, that a life-long love for computer simulation (in all of its forms, from macromolecular function to soccer management!) was born. So perhaps it’s fitting that Dr. Green’s name is the first mentioned here.

But I digress. I was acknowledged along with a couple of other high school seniors who had worked in the lab on other projects; the acknowledgement was made out to those "whose youthful enthusiasm helped renew us on our quest." I always thought that was poetic, and it’s because I received benefit of such a salute that I feel like any list of acknowledgment that I give for this dissertation and the five years of work that went into it is horribly incomplete if I do not first thank the multitude of undergraduates at Ohio State who have been a part of this lab over the past five years and the "youthful enthusiasm" they have shown for this project and the ideas I’ve had.
Two names stand out among this lot. I shared my initial publication credit along with my adviser, Dr. Elizabeth Gross, and an undergraduate, Erico David. In this document, he is directly responsible for the representations of the electrostatic field of cytochrome \( f \) as a function of ionic strength and responsible for some of the surrounding work as it relates to ionic strength effects. Nathan Nelson will get his publication credit shortly for the fine work he supplied regarding "mutagenesis" study of spinach plastocyanin. Both of these individuals not only displayed enthusiasm, but gave me and Dr. Gross a lot of hours of hard work, and I appreciate both of these guys immensely.

Other undergraduates helped a great deal with surrounding projects. Michael Farrow was responsible for some early homology modeling work that allowed us to have preliminary ideas about how the interaction between cyanobacterial PCs and cyt \( f \)'s did (or did not) work. William Weily and Marcus Goodman performed electrostatic modeling and contributed ideas about trends in the electrostatic potential field of PCs from different species of plant and algae. Two graduate students also contributed ideas while passing through our lab: George Bernier helped work out some of the kinks in the analysis of Brownian dynamics results, and Mike Farkas performed some statistical analysis on PC structures and provided ideas that proved useful in retrospect.

As for others who worked in our lab - while they did not help with this project, provided useful feedback, encouragement, and friendship - I would like to thank Baocheng Pan, Baifan Li, Daisy Sinha, Cecilia Longoria, Samantha Stanko, Matthew Butler, and especially Amiee Wagner, who joined the laboratory a couple of years after I did and who I was always happy to work with as a fellow graduate student. (I know I
forgot some names; forgive me, it’s late and this needs to be completed now...!

I cannot conclude my mention of this lab without offering every measure of praise and thanks to our adviser, Dr. Elizabeth Gross, and the overwhelming amount of grace and inspiration she has shown me in her guidance, both through this research and through my growth as a scientist overall. I am forever indebted to her for her generosity.

I offer thanks to the rest of my committee, Dr. James Alben, Dr. Richard Swenson, and Dr. Charles Brooks, for the time and insight they were able to give me as I pursued my work and what it means to be a scientist. I would also like to thank Dr. Robert Ross and Dr. Lawrence Berliner for early guidance.

In this project, two software packages were used beyond all others, and I would be remiss if I did not thank Dr. Anthony Nicholls of Columbia University, the chief author of GRASP, and Dr. Scott Northrup of Tennessee Tech, the author of MacroDox, for the time and help they offered in support of my work. The early stages of this work was performed on the hardware of the OSU College of Biological Sciences Computational Biology Facility; thanks to Dave Stutes and Will Ray for their assistance. Later, we moved our work to hardware owned by the OSU Department of Biochemistry; thanks to the Department for the funds to purchase that hardware.

Finally, and most importantly, thanks to my wife, Kristin, and our daughters, Amelia and Anna, for their neverending love and support, and somehow tolerating the nights when Daddy was working on campus in a feverish effort to finish. Yes, Amelia, the thesis is done...!
VITA

December 6, 1971............................... Born - Tampa, Florida.

1993....................................................... B.S. Physics, Rose-Hulman Institute of Technology, Terre Haute, IN.

1993 - present ...................................... Graduate Teaching and Research Associate, The Ohio State University.

PUBLICATIONS

Research Publication


FIELDS OF STUDY

Major Field: Biophysics
TABLE OF CONTENTS

Dedication ...................................................................................................................................iv

Acknowledgments .......................................................................................................................v

Vita .............................................................................................................................................viii

List of Tables ..............................................................................................................................xii

List of Figures ............................................................................................................................xiv

Chapters:

1. Introduction.............................................................................................................................1

2. Electrostatic properties of plastocyanin and cytochrome *f* ...........................................17
   Methods..........................................................................................................................18
   Structures and sequences .........................................................................................18
   Electrostatics theory and computation ..............................................................19
   Manual docking ....................................................................................................21
   Results and discussion.............................................................................................22
   Sequences..................................................................................................................22
   Plastocyanin structures and electrostatics ...................................................29
   Cytochrome *f* structures and electrostatics ...........................................38
   Original hypotheses of complex formation................................................47

ix
3. Initial Brownian dynamics experiments ................................................................. 57
   Methods .................................................................................................................. 60
   Structures .............................................................................................................. 60
   Brownian dynamics theory and computation ...................................................... 60
      Charge assignment ............................................................................................. 60
      Electrostatic potential calculation .................................................................... 61
      Brownian dynamics simulation ........................................................................ 62
      Analysis of simulation results ......................................................................... 66
   Results and discussion .......................................................................................... 67
      Standard runs - demonstrating minimized metal center distances............... 67
      Standard runs - characterizing closest complexes ....................................... 74
      Other control experiments ............................................................................. 81
      Standard runs - defining the most reasonable reaction distance ............... 83
      Runs at different ionic strengths ..................................................................... 84

4. Brownian dynamics modeling of spinach PC mutant effects ................................ 95
   Methods .................................................................................................................. 98
   Structures .............................................................................................................. 98
   "Mutation" of residues ........................................................................................ 99
   Brownian dynamics computation ...................................................................... 99
   Results and discussion ........................................................................................ 100
      Standard run and ionic strength comparisons ............................................ 100
      Mutations of spinach PC .............................................................................. 105
      The effect of structure on complex formation between PC and cyt f ......... 112

5. The PC/cyt f interaction vs. the c₆/cyt f interaction: the Chlamydomonas system .. 117
   Methods .................................................................................................................. 120
   Structures .............................................................................................................. 120
   Electrostatic computation ................................................................................... 121
   Brownian dynamics computation .................................................................... 121
   Results and discussion ........................................................................................ 122
      Cytochrome c₆ sequence, structure, and electrostatics .................................. 122
      Brownian dynamics of Chlamydomonas PC and cyt f ............................. 124
      Brownian dynamics of Chlamydomonas c₆ and cyt f ............................... 131
      Low ionic strength .......................................................................................... 137
      Brownian dynamics with cyt f mutants ....................................................... 140
      Interaction between PC and cyt f dimer ...................................................... 142
6. Discussion, context and conclusions...............................................................................144

Electrostatics modeling of PC, cyt f and c6.............................................................144
Manual docking of PC and cyt f..............................................................................146
Quantitative results of BD.........................................................................................148
Qualitative results of BD............................................................................................149
Modeling of mutants..................................................................................................155
Comparison of c6/cyt f interaction with PC/cyt f interaction................................157
Overall conclusions....................................................................................................157

Bibliography..................................................................................................................161
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>The 38 currently available plastocyanin sequences, aligned</td>
</tr>
<tr>
<td>2.2</td>
<td>The 23 currently available cytochrome f sequences, aligned</td>
</tr>
<tr>
<td>3.1</td>
<td>Best seven triplet contacts between cyt f and PC at 100 mM ionic strength</td>
</tr>
<tr>
<td>3.2</td>
<td>Successes and rates as a function of distance between metal centers for 9 sets of 1000 MacroDox trajectories between PC and cyt f at an ionic strength of 100 mM and pH 7</td>
</tr>
<tr>
<td>3.3</td>
<td>Successes and rates for MacroDox runs as a function of ionic strength and reaction distance (35 Å, 17 Å, or 16 Å)</td>
</tr>
<tr>
<td>3.4</td>
<td>Best seven triplet contacts between cyt f and PC at 50 mM ionic strength</td>
</tr>
<tr>
<td>3.5</td>
<td>Best seven triplet contacts between cyt f and PC at 300 mM ionic strength</td>
</tr>
<tr>
<td>3.6</td>
<td>Best seven triplet contacts between cyt f and PC at 10 M ionic strength</td>
</tr>
<tr>
<td>4.1</td>
<td>Negative-patch residues on PC and mutations formed</td>
</tr>
<tr>
<td>4.2</td>
<td>Comparison of electrostatic interaction as measured through BD simulations between mutant spinach plastocyanins and turnip cyt f. The &quot;triple&quot; mutant is E43N/E59K/E60Q, as in Kannt et al. (1996)</td>
</tr>
<tr>
<td>5.1</td>
<td>The 17 currently available cytochrome c₆ sequences, aligned</td>
</tr>
</tbody>
</table>
5.2 Best seven triplet contacts between *Chlamydomonas* cyt f and PC at 100 mM ionic strength.................................................................130

5.3 Best seven triplet contacts between *Chlamydomonas* cyt f and c₈ at 100 mM ionic strength........................................................................134

5.4 Best seven triplet contacts between *Chlamydomonas* cyt f and PC at 10 mM ionic strength..................................................................138

5.5 Best seven triplet contacts between *Chlamydomonas* cyt f and c₈ at 10 mM ionic strength.................................................................138

5.6 Comparison of electrostatic interaction as measured through BD simulations between *Chlamydomonas* cyt f and *Chlamydomonas* PC. All runs at 100 mM. Double, triple, and quintuple mutants as in Figure 5.10.........................................................140

6.1 Table of contact residues in characteristic complexes between PC and cyt f from different species shown graphically in Figure 6.1. Contact residues must be within 6 Å of one another center-to-center. "trnp" = turnip cyt f; "Chls" = *Chlamydomonas* cyt f........................................152
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 Molecular graphic representation of turnip cytochrome $f$ (Martinez et al., 1996) as prepared by GRASP (Nicholls et al., 1991). Residues of key importance are labeled.</td>
<td>2</td>
</tr>
<tr>
<td>1.2 Molecular graphic representation of poplar plastocyanin (Guss et al., 1992) as prepared by GRASP. Residues of key importance are labeled.</td>
<td>4</td>
</tr>
<tr>
<td>2.1 Electrostatic potential representation of oxidized poplar plastocyanin. Electrostatic potential was calculated through solution of the linearized Poisson-Boltzmann equation, with the program GRASP. The ionic strength used in the calculation is 100 mM. 3-dimensional contour represents $-1 \text{kT/e}$; 2-dimensional contour represents zero potential. Residues colored as in Figure 1.2.</td>
<td>31</td>
</tr>
<tr>
<td>2.2 Electrostatic potential representation of french bean plastocyanin (Moore et al., 1991). Electrostatic contours as in Figure 2.1.</td>
<td>32</td>
</tr>
<tr>
<td>2.3 Electrostatic potential representation of parsley plastocyanin (Bagby et al., 1994). Electrostatic contours as in Figure 2.1.</td>
<td>33</td>
</tr>
<tr>
<td>2.4 Electrostatic potential representation of <em>Chlamydomonas</em> plastocyanin (Redinbo et al., 1993). Electrostatic contours as in Figure 2.1.</td>
<td>34</td>
</tr>
<tr>
<td>2.5 Electrostatic potential representation of <em>Synechocystis</em> plastocyanin (Romero et al., 1998) with the three mutated residues in the crystal structure restored to the wild type. Electrostatic contours as in Figure 2.1.</td>
<td>35</td>
</tr>
</tbody>
</table>
2.6 Electrostatic potential representation of *Anabaena* plastocyanin (Badsberg et al., 1996). Electrostatic contours as in Figure 2.1.................................................................................................36

2.7 Electrostatic potential representation of *Prochlorothrix* plastocyanin (Babu et al., 1999). Electrostatic contours as in Figure 2.1.......................................................................................................................37

2.8 Electrostatic potential representation of reduced turnip cytochrome *f*. Electrostatic potential was calculated through solution of the linearized Poisson-Boltzmann equation, with the program GRASP. The ionic strength used in the calculation is 100 mM. 3-dimensional contour represents +1 kT/e; 2-dimensional contour represents zero potential. Residues colored as in Figure 1.1..............................................................................39

2.9 Electrostatic potential representation of *Chlamydomonas* cytochrome *f* (E.A. Berry, unpublished results). Electrostatic contours as in Figure 2.8............................................................................................41

2.10 Electrostatic potential representation of *Synechocystis* cytochrome *f* as homology modeled for this work. Electrostatic contours as in Figure 2.1.....................................................................................................................43

2.11 Electrostatic potential field of reduced cyt *f* for ionic strength of 2 mM. Key charged residues (K187, R209, K65, K66, K58) colored black; heme colored black; Y1 shaded. All other conditions besides ionic strength same as Figure 2.8..................................................................................................................45

2.12 Electrostatic potential field of reduced cyt *f* for ionic strength of 25 mM. All other conditions besides ionic strength same as Figure 2.8.....................................................................................................................46

2.13 Electrostatic potential field of reduced cyt *f* for ionic strength of 200 mM. All other conditions besides ionic strength same as Figure 2.8................................................................................................................47

2.14 A plot of electrostatic potential against distance for a series of points extending up to 10 Å away from a porphyrin oxygen (atom O1D), with a spacing of up to 0.5 Å, representing the cross-over of potential over the heme from negative to positive at lower ionic strengths..............................................................................................................48
2.15 Hypothetical "Complex 1" between PC and cyt f, representing electrostatic attraction between cyt f positive residues and PC negative residues. Representation prepared with GRASP. Proteins were docked manually as described in Methods, and then separated 15 Å to show electrostatic interactions. Potential calculation was performed at an ionic strength of 100 mM; two-dimensional interpolation plane contours are of values of -1, -0.5, 0, 0.5, and 1 kT/e; positive and negative sides of zero contour marked.

2.16 Hypothetical "Complex 2" between PC and cyt f, representing a dock with close contacts between H87 (PC) and Y1 (cyt f) and D44 (PC) and K187 (cyt f). Representation prepared as Figure 2.15.

2.17 Hypothetical "Complex 2" between PC and cyt f, representing a dock with close contact between H87 (PC) and Y1 (cyt f) while maximizing hydrophobic interactions. Representation prepared as Figure 2.15.

3.1 A representation of the Brownian dynamics trajectory at its onset. PC, the mobile ligand, is set on an imaginary sphere with radius \( b = 89 \) Å with cyt f, the fixed receptor, set at the center of that imaginary sphere. PC is subject to two forces: electrostatic attraction/repulsion between the two proteins, and random (Brownian) force set to model diffusion. The trajectory concludes when PC escapes a concentric imaginary sphere \( c = 200 \) Å, with the position and orientation of PC having been recorded when it reaches its closest approach to cyt f by a preset measure (e.g., distance between metal centers).

3.2 The distribution of distances between metal centers for all complexes with metal centers of less than 35 Å for 9 sets of 1000 trajectories each at 100 mM and 10 M, where complexes were recorded on the basis of minimized Cu-Fe distances.

3.3 The distribution of distances between metal centers for all complexes with metal centers of less than 35 Å for 9 sets of 1000 trajectories each at 100 mM and 10 M, where complexes were recorded on the basis of minimized distances of centers of mass.

xvi
3.4 The distribution of distances between metal centers for all complexes with metal centers of less than 35 Å for 9 sets of 1000 trajectories each at 100 mM and 10 M, where complexes were recorded on the basis of minimized distance between the delta-1 oxygen atom of D44 and the zeta nitrogen on K187 of cyt f.

3.5 Alpha-carbon traces of 12 complexes between poplar PC and turnip cyt f. These 12 complexes were selected randomly from a single run of 1000 trajectories of pH 7 and 100 mM ionic strength from the four best contact triplets listed in Table 3.1. The complexes are superimposed so that the coordinates of cyt f in each complex are coincident with one another, showing the positioning of PC relative to cyt f in each of the 12 different trajectories. B is a 90° rotation of A. The figure was prepared using the software package Quanta (MSI, 1996).

3.6 A wide view of a characteristic complex from Figure 3.5 between PC and cyt f, composed in GRASP with more atomic detail in order to highlight some of the proposed electrostatic contacts between the two proteins. Note the close proximity of H87(PC) to Y1(cyt f) (and, by connection, the heme of cyt f as well).

3.7 A 180° rotation of Figure 3.6.

3.8 A closer view of the characteristic complex of Figure 3.6, highlighting some of the proposed hydrophobic/van der Waals contacts between the two proteins.

3.9 A 180° rotation of Figure 3.8.

3.10 The distribution of distances between metal centers from the original set of BD experiments (given in Figure 3.2) compared to two control cases, interactions between Anabaena PC and Alcaligenes azurin and cyt f, both run at 100 mM and pH 7, demonstrating the unique nature of the electrostatic interaction between PC and cyt f to generate close contacts between the two proteins.
3.11 Calculated rates, based on number of BD successes at reaction criteria of 16 Å and 17 Å, for the interaction between PC and cyt f as a function of ionic strength.................................................................87

3.12 Average Fe-Cu distance for all BD successes at a reaction distance of 35 Å for several ionic strengths..........................................................88

3.13 Alpha-carbon traces of 12 complexes between poplar PC and turnip cyt f at an ionic strength of 50 mM. Traces were prepared as in Figure 3.5.................................................................92

3.14 Alpha-carbon traces of 12 complexes between poplar PC and turnip cyt f at an ionic strength of 300 mM. Traces were prepared as in Figure 3.5.................................................................93

3.15 Alpha-carbon traces of 12 complexes between poplar PC and turnip cyt f at an ionic strength of 10 M. Traces were prepared as in Figure 3.5.................................................................94

4.1 The distribution of distances between metal centers from the BD experiments on the interaction between spinach PC and turnip cyt f, presented in identical fashion to data on the interaction between poplar PC and turnip cyt f in Figure 3.10 and compared to the same *Anabaena* PC and *Alcaligenes* azurin control cases.........................................................101

4.2 The distribution of successful trajectories for a number of different ionic strengths, 100 mM and below, as a function of Cu-Fe distance.........................................................102

4.3 The number of trajectories falling within a particular distribution level for a number of different distances (e.g., number at 16 Å represents number of successful trajectories with closest approach between 15 and 16 Å Cu-Fe distance) as a function of ionic strength.........................................................103

4.4 Alpha-carbon trace of 10 complexes between spinach PC and turnip cyt f generated by MacroDox for a single run of 1000 trajectories at pH 7 and 100 mM ionic strength, selected randomly from all complexes with Cu-Fe distance less than 16 Å. The figure was prepared as in Figure 3.5.........................................................106
4.5 Alpha-carbon trace of 10 complexes between spinach PC and turnip cyt f generated by MacroDox for a single run of 1000 trajectories at pH 7 and 10 mM ionic strength, selected randomly from all complexes with Cu-Fe distance less than 16 Å. The figure was prepared as in Figure 3.5 ........................................................107

4.6 BD success probability for the interaction of spinach PC mutants with cyt f, as plotted against net charge change for that mutant ...................................................................................................................111

4.7 Original 11 spinach NMR structures from A. Bergkvist (personal communication), superimposed on one another. Key regions of deviation within the ensemble of structures are labeled ..........................................................................................................................114

5.1 Electrostatic potential representation of oxidized Chlamydomonas cₕ. Electrostatic potential was calculated through solution of the linearized Poisson-Boltzmann equation, with the program GRASP. The ionic strength used in the calculation is 100 mM. 3-dimensional contour represents -1 kT/e; 2-dimensional contour represents zero potential. Important residues shaded and labeled in upper left .............................................125

5.2 Electrostatic potential representation of oxidized Monoraphidium cytochrome cₕ. Residue and electrostatic potential contour coloring as in Figure 5.1 ..................................................................................................................126

5.3 Electrostatic potential representation of oxidized Synechococcus cytochrome cₕ. Residue and electrostatic potential contour coloring as in Figure 5.1 ..................................................................................................................127

5.4 The distribution of distances between metal centers for all Chlamydomonas PC/cyt f complexes with metal center distances of less that 35 Å for 9 sets of 1000 trajectories each at 10 mM, 100 mM, and 10 M, where complexes were recorded on the basis of minimized Cu-Fe distances ..................................................................................129

5.5 Alpha-carbon trace of 12 complexes between Chlamydomonas PC and cyt f. These 12 complexes were selected randomly from a single run of 1000 trajectories of pH 7 and 100 mM ionic strength from the four best contact triplets listed in Table 5.2. The figure was prepared as in Figure 3.5 ..........................................................130
5.6 The distribution of distances between metal centers for all *Chlamydomonas c*~cyt f complexes with metal center distances of less than 35 Å for 9 sets of 1000 trajectories each at 10 mM, 100 mM, and 10 M, where complexes were recorded on the basis of minimized Fe-Fe distances .......................................................... 132

5.7 The distribution of distances between metal centers for all *Chlamydomonas c*~cyt f complexes with metal center distances of less than 35 Å for 9 sets of 1000 trajectories each at 100 mM and 10 M, where complexes were recorded on the basis of minimized distance between centers-of-mass .............................................. 133

5.8 Alpha-carbon trace of 12 complexes between *Chlamydomonas c*~cyt f These 12 complexes were selected randomly from a single run of 1000 trajectories of pH 7 and 100 mM ionic strength from the four best contact triplets listed in Table 5.3. The figure was prepared as in Figure 3.5 .......................................................... 136

5.9 Alpha-carbon trace of 12 complexes between *Chlamydomonas c*~cyt f These 12 complexes were selected randomly from a single run of 1000 trajectories of pH 7 and 10 mM ionic strength from the four best contact triplets listed in Table 5.5. The figure was prepared as in Figure 3.5 .......................................................... 139

5.10 The distribution of distances between metal centers for complexes between *Chlamydomonas cyt f* and various mutant *Chlamydomonas* PCs, with 9 sets of 1000 trajectories run for each mutant at 100 mM. Double mutant is K188N/K189Q; triple mutant is K58Q/K65S/K66E; quintuple mutant is a combination of above double and triple mutants ........................................ 141

5.11 A comparison of the distribution of distances between metal centers for complexes between *Chlamydomonas PC* and monomeric and dimeric cyt f, where complexes were recorded on the basis of minimized Cu-Fe distances. Monomeric data is taken directly from Figure 5.4. Dimeric data is based on single runs of 10000 trajectories at 10 mM .......................................................... 144

6.1 Alpha-carbon traces of characteristic complexes between poplar PC and turnip cyt f, spinach PC and turnip cyt f, and *Chlamydomonas* PC and cyt f. The key point in this image is the coincidence of PC's orientation regardless of the species of PC and cyt f interacting .......................................................... 151

xx
6.2 Alpha-carbon trace comparison of the positioning of the characteristic BD complex between spinach PC and turnip cyt f (light lines) and a characteristic complex between the same proteins from the NMR study and rigid-body MD simulation of Ubbink et al. (1998; PDB ID 2PCF) (dark lines). The NMR/MD complex appears to favor electrostatic interactions, particularly between residues 42-45 on PC and residues K187 and R209 on cyt f, far more than the BD complex does.
Cytochrome f (cyt f; see Hope, 1993 for review) is an electron transport protein in the photosynthetic electron transport chain, participating in the process of extracting electrons from water and fixing them to high-energy electron carriers such as NADPH for the reduction of glyceraldehydes to glucose. Cyt f is a member of the cytochrome b_{6f} membrane-bound complex, facing the lumen of the thylakoid and anchored to the membrane by a single membrane-spanning helix. It receives an electron from the Rieske Fe-S protein, a member of the same complex, and depending on the organism, cyt f passes that electron on to either cytochrome c_{6} (cyt c_{6}, or simply c_{6}) or plastocyanin (PC).

Cyt f (Figure 1.1) was originally believed to have structural similarity to the other members of the cytochrome c family of proteins, based on its similar spectroscopic properties and strong positive charge (Prince and George, 1995). However, the solution of the structure of cyt f by Martinez et al. (1995) revealed a unique two-domain, beta-sheet protein, not the alpha-helical structure found in most c-type cytochromes. The structure of cyt f, on reflection, is more reminiscent of a
Figure 1.1. Molecular graphic representation of turnip cytochrome f (Martinez et al., 1996) as prepared by GRASP (Nicholls et al., 1991). Residues of key importance are labeled.
hematopoietic receptor, such as the hGH receptor, than that of a cytochrome c. In addition, the heme ligation was unique - while the sixth ligand to the heme of most such proteins is a methionine, the sixth ligand of cyt f is the protein's N-terminus. This places a tyrosine parallel to the plane of the heme's porphyrin ring in those cyt f's for which structures are known.

In higher plants and algae, cyt f passes an electron to plastocyanin (PC, Figure 1.2, see Gross, 1996 for a review), with the reaction occurring in the lumen of the thylakoid. PC is a beta-barrel copper protein of approximately 100 amino acids, with the copper bound to four invariant ligands in a distorted tetrahedral geometry. The strong overall negative charge of PC is widely believed to play a role in its interactions with cyt f and with Photosystem I, PC's other reaction partner. Cytochrome c₆ (Frazao et al., 1995; Kerfield et al., 1995) has a similar overall negative charge, but a quite dissimilar structure. Instead of being a beta-sheet protein, c₆ is an alpha-helical protein of about 90 amino acids. It stores electrons not on a copper center, but (as the name implies) in a heme pocket. c₆ participates in electron transfer with cyt f in algae and cyanobacteria. Most organisms utilize either PC or c₆ in photosynthetic electron transport; however, a few algae (such as *Chlamydomonas reinhardii*) possess the ability to express and utilize both (Navarro et al., 1997).

When association between two proteins (such as PC and cyt f, or c₆ and cyt f) is required for them to fulfill their function; the problem of how that association takes place is fundamental. Forces guiding interaction between proteins are in many respects the same forces as those that guide the proteins' folding process; electrostatic
Figure 1.2. Molecular graphic representation of poplar plastocyanin (Guss et al., 1992) as prepared by GRASP. Residues of key importance are labeled.
interactions between charged residues and hydrophobic interactions between nonpolar residues. Hydrogen bonds can be important in highly specific interactions between molecules but are, on balance, not as significant in bimolecular interactions as in the secondary structure of proteins, in which the portions of the protein backbone involved in hydrogen bonding are much more organized than side chains on the exterior of the protein.

The specificity and time scale of such interactions are dependent on the number of residues involved and the topography of those residues; for example, the interaction between human growth hormone (hGH) and its receptor (Clackson and Wells, 1995) involves several hydrophobic residues on both proteins, a few of which contribute significant energy to the resulting complex. Such an interaction is very specific and tends to have a long lifetime, aiding the biological role of the complex (which is to allow the activation of the receptor).

On the other hand, the lifetime of complexes between electron-transfer proteins, whose biology is dependent on the speedy association and dissociation for purposes of cycling electrons about the system, should not be as long as the lifetime of complexes between hormones and receptors. So one would expect the binding between two electron-transfer proteins (e.g., PC and cyt f) to be less specific; instead of a larger binding epitope on the surface of both proteins, perhaps two or three key interactions between residue pairs on these proteins are important.

The interaction between c₆ and cyt f has not been well characterized at all, due to difficulty in designing an experiment that will describe an electron transfer reaction.
between two c-type cytochromes - spectroscopic observation of such a reaction is essentially impossible (Navarro et al., 1997). On the other hand, several experiments concerning the interaction between PC and cyt f have been carried out.

Electron transfer in PC is likely to take place through one of two sites, as characterized by NMR studies (Cookson et al., 1980; Handford et al., 1980). One possibility is a site centered at H87, the only PC copper ligand that is surface-exposed. The residues surrounding H87 - L12, F35, P36, P86, etc. - are hydrophobic in nature, and a potential interaction between PC and cyt f at this site was not believed to be electrostatic in nature, despite the slight positive potential surrounding the site due to the effect of the buried copper atom observed according to the electrostatic analysis of Durell et al. (1990). The other is a site surrounded by Y83, which would transfer electrons in a through-bond pathway (as opposed to the outer-sphere pathway for H87) through Y83’s adjacent copper ligand, C84. In all higher plant and algal PCs, Y83 is surrounded by negatively charged residues, and therefore was naturally pinpointed as a target for early experiments because of the known preponderance of positively charged residues on cyt f. Chemical modification with tetranitromethane to place a nitrotyrosine at position 83 (Gross and Curtis, 1991; Christensen et al., 1992) was observed to have no effect on the interaction between PC and cyt f; however, He et al. (1991) observed in stopped-flow experiments a 40-fold in the second-order rate constant with cyt f for a Y83L mutant as compared to the wild type. He constructed a kinetics scheme to deconvolute rate constants for each step of the PC-cyt f electron transfer reaction, and in calculations based on this scheme asserted a 15-fold decrease in intrinsic electron
transfer rate for the Y83L mutant. A similar experiment with a Y83F mutant resulted in no significant change in the calculated intrinsic electron transfer rate value when compared to the wild type; those two observations implied that a aromatic residue was necessary at position 83 for electron transfer to take place.

Experiments involving the surrounding negative patch and its impact on the interaction between PC and cyt f are also significant. Beoku-Betts et al. (1985), found that chromium(III) neutralization of negative residues 42-45 significantly decreased interaction between the two proteins. More significant was ethylenediamine chemical modification of negative patch residues, performed by Takabe et al. (1984 and 1986) and Anderson et al. (1987). An ethylenediamine modification of a negatively charged residue has the effect of reversing the charge on that residue. The Anderson et al. (1987) studies demonstrated significant inhibition of the interaction between spinach PC and turnip cyt f when the 42-45 patch or the 59-61 patch of PC were so modified, and some inhibition when residue 68 was modified.

Mutagenesis studies on negatively charged residues produced mixed results. Nordling et al. (1991) and Modi et al. (1992) mutagenized D42 specifically, but observed no effect on the interaction with cyt f. These observations were later challenged independently by Lee et al. (1995) in studies of Silene PC, and Kannt et al. (1996) in studies of spinach PC - both demonstrated that mutations on the 42-45 patch had significant effects on the interaction, and Kannt demonstrated effects on the 59-61 patch as well. However, mutagenesis studies on PC negative patch residues have never systematically dealt with the effects of the two patches in comparison, because of
difficulties in the genetics of assembling mutant proteins. For example, in Kannt's study, comparison of the importance of the two patches was made in part on the basis of single mutations of residues in one patch as compared to double mutants of the opposite patch. (These studies in particular, in addition to the Anderson et al. (1987) chemical modification studies, are discussed in much greater detail later.)

The most significant experiment involving the interaction between PC and cyt f prior to Martinez' (1994) turnip cyt f structure determination was the chemical cross-linking study of Morand et al. (1989). Amino groups on cyt f were first cross-linked to carboxyl groups on PC utilizing a carbodiimide (usually 1-ethyl-3-[3-(dimethylamino)propyl]carbodiimide, or EDC) by Davis and Hough (1983), and cross-linking experiments were carried out by Takabe et al. (1989) and Qin and Kostic (1993) in addition to Morand's experiment. The significance of Morand's accomplishment was the determination that K187 on turnip cyt f and D44 on spinach PC were directly and uniquely cross-linked, in addition to the implication of E59 and E60 on PC with positive (but indeterminate) residues on cyt f.

The question about the cross-linked complex concerns the ability for electrons to transfer within the complex. The consensus in literature describing complex formation (as referred to in the later works of Ullmann et al. (1997) and Ubbink et al. (1998), among others) is derived from the results of Qin and Kostic (1993), who observed no electron transfer when studying a carbodiimide cross-link through laser flash spectroscopy, with riboflavin semiquinone as reductant. This would imply that the electrostatic complex between PC and cyt f, as observed by Morand et al. (1989), is not
electron-transfer competent. However, earlier experiments from Takabe et al. (1989) did yield observations of electron-transfer activity in the cross-linked complex. In addition, D.J. Davis (personal communication; reported in Gross, 1996) has observed electron transfer rates from a cross-link between ruthenium-labeled PC and cyt f that were only inhibited 20% when compared to a non-covalently bound complex.

This question of the electron-transfer competence of the covalently cross-linked adduct encroaches on a debate that was starting as the cyt f structure was released and the study described here began and has continued to evolve since - the question of whether PC could diffuse into a single, electron-transfer competent complex, or whether a predocking complex involving one type of interaction (presumably electrostatic) would be required before a final complex, with a different type of interaction (presumably hydrophobic) would take over. The idea that the electrostatic complex, represented by the covalent cross-link, is not (as Qin and Kostic would argue) electron-transfer competent would imply that substantial rearrangement of the complex from the electrostatic complex to an electron-transfer competent complex would be necessary.

Two other groups of residues on PC had been investigated with respect to PC's interaction with cyt f. The first was from the aforementioned Modi et al. (1992) study. With respect to the H87 interaction site on PC (and in addition to several other mutations, including the aforementioned D42N mutant and Y83F and F35Y mutants, with little effect), L12 was mutated to asparagine and showed a significant increase in binding, and was mutated to glutamate and showed an equally significant decrease in
binding. Modi postulated the simplest explanation possible at the time - that cyt f contained an acidic residue at the binding region that could act as a hydrogen bond acceptor for a N12, or would repel a E12.

The other residues studied were lysines, in chemical modification studies of Gross et al. (1990). 4-chloro-3,5-dinitrobenzoic acid (CDNB) replaces positive charges (either the N-terminal or lysines) with a negative charge. Some inhibition of the PC-cyt f interaction was observed when the N-terminal or K54 of PC were modified; almost total inhibition was observed when K77 was modified. The results of K77 modification was quite surprising to these investigators; their only substantial argument to describe that result was to argue that K77's positive electrostatic field serves as a steering factor to place negatively charged residues in the right position.

The questions I personally became interested in upon joining the laboratory involved the means of interaction between PC and cyt f. The work of Morand et al. (1989) seemed to imply some role for electrostatic interactions in complex formation between the two proteins. With the availability of Martinez' (1995) solution for the turnip cyt f crystal structure and appropriate software, it would be relatively straightforward to calculate the extent of electrostatic fields of the two proteins and demonstrate the means by which electrostatic interactions guide PC into a docked complex with cyt f, and further to enquire whether the electrostatically-favored complex could be a viable electron-transfer complex. The mathematics of other types of interactions between proteins (hydrogen bonds, hydrophobic interactions, etc.) are nowhere near as straightforward. Therefore, a computational analysis of the interaction
between two proteins would be limited to the electrostatic interaction.

Northrup et al. (1988) had demonstrated an electron-transfer system for which such an analysis was valid - the interaction between cytochrome c and cytochrome c peroxidase. Electrostatic interactions between these two proteins are well-studied, and the biology of the two proteins is dependent on their electrostatics (Margoliash and Bosshard, 1983). The similar evidence of the importance of electrostatic interactions between PC and cyt f, in addition to the computational difficulty of considering other types of interactions, justified studying the interaction between PC and cyt f in a similar manner.

In addition to manual docking studies to sort out possibilities for docking between the two proteins, we had two methods in mind to study electrostatic interactions between proteins. The first was a "static" study - looking at single protein structures with software for the representation of electrostatic fields, GRASP in particular. The second was a "dynamic" study - modeling the diffusion of the mobile electron carriers with respect to cyt f through Brownian dynamics (BD) methods, with MacroDox being the software of choice.

The study of the electrostatic fields of proteins had precedent in this laboratory, and the work of Durell et al. (1990) characterizing the electrostatic potential field of PC served as a model once we obtained the structure of cyt f. Durell used an early version of the DelPhi software from Honig's laboratory at Columbia University (Klapper et al., 1986; Gilson et al., 1987; Nicholls and Honig, 1991) in a substantially quantitative study. We used the graphical descendant of DelPhi, GRASP (Graphical Representation
and Analysis of Surface Properties) to engage in a more qualitative study of cyt f’s electrostatic potential fields and in ongoing comparisons of the potential fields of different PCs, cyt f’s, and cₐ’s. Both programs are written to solve the linear Poisson-Boltzmann equation numerically, with the charges and potential results mapped to a three-dimensional grid. DelPhi utilizes a single, user-defined grid; GRASP uses a pair of 33x33x33 grids sized automatically to envelop the macromolecule being studied.

A word should be said here about the macroscopic view of electrostatics, as opposed to the microscopic view. Macromolecular electrostatics have been reviewed by Harvey (1989), and the underlying philosophy of that review stands up very well. In that review, the two dominant views of electrostatics are addressed: the macroscopic approach, in which the molecules to be analyzed are treated as large regions of low dielectric in a medium of high dielectric, with the molecules themselves treated only implicitly; and the microscopic approach, in which each atom in the molecules to be analyzed (even waters) are treated explicitly, in terms of the polarizability of each atom.

In the best of all possible circumstances, the microscopic approach would be ideal for all problems. Not only is the treatment of all atoms in all molecules likely to lead to a more accurate solution, it also lends itself to the treatment of the molecules as dynamic bodies better than the macroscopic approach, since the latter requires the predetermined determination of the molecular surface, so all inside that surface can be declared as low-dielectric. In addition, the very idea of a molecular method based on dielectric constants is dubious - how do you describe a microscopic thing (a molecule) with a macroscopic property (dielectric constant)?
However, an all-atom treatment of electrostatic properties in computationally quite intense, and the computational hardware we had at our disposal at the time of this work was not sufficient for quick execution of a microscopic analysis. GRASP was designed for electrostatics calculation using the macroscopic/continuum approach, and such calculations are very fast - on low-end SGI workstations, calculation of a potential map for cyt f took 20-30 seconds. In addition, we had no reason to believe at the time (and we still do not believe) that conformational changes or significant motion of atoms (or water molecules) within the proteins play a role in complex formation between the two proteins, thereby eliminating one of the primary benefits of microscopic analysis.

Our primary goal in use of GRASP was to develop hypotheses, based on the electrostatics, of the interaction between PC and cyt f. We also used GRASP to describe electrostatic effects of changes in ionic strength, pH, and oxidation state. Results from these studies are summarized in Chapter 2.

To model diffusion and docking of PC or c₅ to cyt f, we used the MacroDox software developed in the laboratory of Northrup at Tennessee Technological University (Northrup et al., 1987; Northrup et al., 1988). Northrup’s software has its genesis in the work of Ermak and McCammon (1977) who derived equations to describe motion of particles (molecules) subject to Brownian forces in the presence of a potential, with a view to describing Brownian dynamics of molecules "searching" for their reactive site on a receptor protein. McCammon’s laboratory developed BD software of its own (UHBD, for University of Houston Brownian Dynamics - Davis et al., 1990 and 1991), in addition to Northrup’s software. The reason for selecting MacroDox for use in these
studies was less philosophical and more practical. UHBD is written to describe the motion of a group of particles, independently defined and characterized, about a receptor protein. For the diffusion of small molecules, UHBD is ideal; it considers the flexibility of the particle ensemble and provides for a great deal of variety in the characterization of particles. However, UHBD is not well-designed for the treatment of even small proteins; either the protein must be treated as a single sphere with charge and dipole, with all of the obvious weaknesses of that approach, or the protein must be treated as several atomic-sized particles (or even residue-sized particles) with flexibility considered, and the computational time for such a process (considering not only the diffusion of the protein, but internal motions within the protein as well) would be prohibitive even for a small number of diffusion trajectories.

Macrodox, on the other hand, is designed from the simulation of the diffusion of small-to-moderate sized proteins (and electron transfer proteins in particular!) towards a receptor. The motion of the mobile protein is considered as a rigid-body, which is a significant weakness if conformational changes in the protein are significant in the protein’s biological function. No significant evidence exists to suggest that such is the case for PC in its interaction with cyt f. Because internal motions within the protein are not considered, the computational time for the trajectories is quite reasonable - ranging from ten minutes to an hour for a standard run on various Silicon Graphics hardware.

One should keep in mind that both packages are designed to deal solely with electrostatic effects on the diffusion of the mobile particle towards its receptor, and not with other intermolecular forces (hydrophobic interactions) that might be involved in
the interaction between the two particles. Again, this is as much a matter of convenience as anything else; the theory of electrostatic fields is reasonably well established, and forces due to electrostatic effects can be easily computed with a small workstation. There is no such easily-definable thing as a "hydrophobic field"; the thermodynamic nature of such an effect does not lend itself to easy computation. So these programs deal with interactions (electrostatic, Brownian) that are easily computable, and models based on simulation with such programs are limited by this narrower scope.

Our primary goal in utilizing Brownian dynamics methods was to generate possible electrostatic complexes between PC and cyt f, and to determine if any of those electrostatic complexes might be electron-transfer competent. These BD experiments would serve to test hypotheses previously generated through electrostatic field studies and through manual docking of the proteins. As this work began, we studied the interaction between poplar PC and turnip cyt f, the two structures of interest that were available at the time, and the results of those studies are outlined in Chapter 3.

As more structures came available throughout this work, we set up BD experiments for those structures as well. In addition, we were able to compare the experimental results of chemical modification and mutagenesis work on various PCs (Anderson et al., 1987; Lee et al., 1995; Kannt et al., 1996) with PC "mutants" we modeled, and extend those mutagenesis studies to get a more complete view of single mutant affects on the negative patch of PC. Results of these studies on spinach PC are outlined in Chapter 4.
Finally, with the availability of crystal structures of PC, c₆ and cyt f from a single organism (Chlamydomonas reinhardtii), a comparison and contrast of docking between the PC/cyt f pair and the c₆/cyt f pair (the latter of which has been only sparingly studied experimentally to date) could be made. The electrostatics of the two proteins of homologous function (PC and c₆) could be compared, and BD for the two pairs could be performed and analyzed with respect to one another. These experiments are described in Chapter 5.

To summarize, the aim of this work is to describe with computational models, as completely as is practical, the electrostatic interaction between cyt f and its physiological electron transfer partners, PC and c₆, with the goal of predicting the nature of the electrostatic complex that these proteins form and making a first guess as to the ability of these complexes to undergo electron transfer.
CHAPTER 2

ELECTROSTATIC PROPERTIES OF PC AND CYT F

Durell et al. (1990) published the first analysis of the electrostatic field of plastocyanin, with observations on the electrostatic properties of the protein, electrostatic effects on redox potential, and analysis of the pKa of Y83 and a nitrotyrosine chemical modification of that residue. Durell's conclusions as they concerned molecular association with PC, however, were limited to low-molecular-weight reagents such as cobalt phenanthroline and ferricyanide, since very little was known about the structure of cyt f at the time.

Electrostatic analysis of the interaction between PC and cytochrome c was published by Roberts et al. (1991) when it was assumed that the structure of cyt f was going to be reasonably similar to that of cyt c. Software was developed to search for "precollision orientations" between PC and cyt c, and docked complexes were generated based on the interaction of charged side chains, most of which implicated Y83 as the nearest residue for purposes of electron transfer from cyt c. It was noted, however, that a complex that minimized the distance between the metal centers would implicate H87.

When Martinez et al (1994) did solve the structure of turnip cyt f, the two-
domain, beta-sheet structure of the protein (as opposed to the alpha-helical structure of most c-type cytochromes) was surprising to many, as was the ligation of the N-terminal tyrosine residue to the iron atom (Prince and George, 1994). A unique analysis of the electrostatic properties of the newly characterized protein was essential.

The goal of the original work that would serve as the genesis for later ideas within this thesis was to model the electrostatic fields surrounding PC and cyt f (and, later, cytochrome c₆ as well), to get an idea (and, hopefully, to predict successfully) how the two proteins might interact. These studies would form the basis of a paper presented at the 1995 International Photosynthesis Congress (Pearson and Gross, 1995) and a paper published in the Biophysical Journal (Pearson et al., 1996).

METHODS

Structures and sequences. Originally, we obtained the coordinates of reduced cyt f, solved from crystallography to a resolution of 1.96 angstroms, directly from Martinez et al. (1996). It has since been released to the Protein Data Bank, (http://www.rcsb.org/pdb/ - Bernstein et al., 1977) as ID 1HCZ, in addition to the crystal structure of cyt f from Chlamydomonas reinhardtii (ID 1CFM; Berry et al., manuscript in preparation). For purposes of comparison here, a homology model of cyt f from Synechocystis SCC 6803 was prepared based on the structures of turnip and Chlamydomonas cyt f using the fully-automated comparative modeling module within
the software package MODELLER-4 (Sali et al., 1997).

Poplar plastocyanin (ID I PLC; Guss et al., 1992) was the structure of choice for the published works (Pearson and Gross, 1995; Pearson et al., 1996) because of the strength of its structural characterization to date. However, plastocyanin structures from parsley (1PLA and 1PLB; Bagby et al., 1994), french bean (9PCY; Moore et al., 1991), Chlamydomonas reinhardtii (2PLT; Redinbo et al., 1993) and Enteromorpha prolifera (7PCY; Collyer et al., 1990) were also available at the onset of this work, and many more structures (including structures from Anabaena variabilis (ID 1NIN; Badsberg et al., 1996), a Synechocystis triple mutant (ID 1PCS; Romero et al., 1998), and Prochlorothrix hollandica (ID 1B3I; Babu et al., 1999), used in this work for comparative purposes) have come available since. All of these structures were obtained from the Protein Data Bank (http://www.rcsb.org/pdb/ - Bernstein et al., 1977).

Sequences of different plastocyanins and cytochromes f were obtained through the National Center for Biotechnology Information's ENTREZ (http://www.ncbi.nlm.nih.gov/Entrez/) database. The sequences were aligned using the "Protein Design" tool within MSI's QUANTA97 package (MSI, 1997).

Electrostatics theory and computation. We qualitatively analyzed the electrostatic fields of these proteins with the software package GRASP (Nicholls et al., 1991). GRASP is designed primarily as a visualization tool, and our most relevant use of the program is to make comparisons and predictions regarding the extent and role of electrostatic fields in the interaction between PC and cyt f, not in the calculation of
specific values for field intensity at any given point. GRASP utilizes a macroscopic approach to molecular electrostatics, mapping the charges applied to individual atoms within the macromolecule to a pair of 33 x 33 x 33 grids - one just enveloping the molecule for fine calculation of effects close to the surface, the other with twice the spacing between grid points to allow for a sampling of electrostatic effects in the surrounding solution - and then solving for the electrostatic potential at those grid points using a finite-difference numerical approach to solve the Poisson-Boltzmann equation (popularly known as the finite-difference Poisson-Boltzmann, or FDPB, method - Nicholls and Honig, 1991).

In our study of the electrostatics of these molecules, we established a standard case of pH 7 and 100 mM ionic strength, with dielectric constant set to 2 within the molecule, and 80 for the solvent. We placed charges on the molecules simply: charges on oxygen atoms within acidic residues (including the porphyrin carboxyls on cyt f and cyt c6) were set to -0.5e, and charges on nitrogen atoms on basic residues were set to +1.0e. We tested this against results for the molecules in question from the MacroDox Tanford-Kirkwood pKa solver (Matthew, 1985; Matthew and Gurd, 1986) at pH 7 and generally found agreement within 0.01e for all charge assignments. The lone exception was for H142 on cyt f, which is predicted to take a charge around 0.2e at pH 7 by the MacroDox TK module. For qualitative purposes, neglecting this charge had no real effect; when quantitative studies began, the MacroDox-solved charge set was used and therefore this charge was included.

Whenever the potential grid map calculated by GRASP needed to be mapped to
a specific point in Cartesian space for purposes of predicting electrostatic effects away from the protein surface, a program in the (related) DelPhi package (Nicholls and Sharp, 1989) entitled PHITOPDB was used to estimate the potential at that given point (given in PDB format) based on the saved GRASP/DelPhi potential map (named in format *.phi).

**Manual docking.** Based on the conclusions drawn from the electrostatics of the proteins and known biochemical data from the literature (particularly results from Morand et al, 1989 and Modi et al, 1992b) docked complexes between PC and cyt f were manually generated in an attempt to give structure to our hypotheses at the time about their interaction. We maintained two criteria for the formation of these "docks": one, that key residues of interest had atoms within 5 Å of one another; two, that "severe steric overlap" (more than five pairs of atoms closer than 4 Å of one another) was eliminated. Electrostatic fields were qualitatively calculated for these docked complexes using GRASP. The exposure of atoms to the protein surface was calculated, using the MacroDox implementation of a Richards surface algorithm (Lee and Richards, 1971; Richards, 1977), both for the proteins apart and for the docked complexes. From this, an estimation of solvent-excluded hydrophobic surface area was made on the following basis: all aliphatic or aromatic carbon atoms or sulfur atoms that lost exposure to the surface in the docked complexes were considered solvent-excluded surface area. The total solvent-excluded hydrophobic surface area in a given docked complex was estimated to be twice the solvent-excluded hydrophobic surface area on the **more-**
exposed (less-covered) protein. A hydrophobic interaction energy was estimated based on these values of solvent-excluded surface area and predictions from the literature (Janin and Chothia, 1990; Rose and Wolfenden, 1993; Sharp et al, 1991) of 25-47 cal/mol contribution to the hydrophobic interaction energy per square angstrom of covered surface area.

RESULTS AND DISCUSSION

Sequences. Tables 2.1 and 2.2 give the currently available sequences of plastocyanin and cytochrome f. 38 different plastocyanin sequences from 33 different organisms are available presently; 23 unique cyt f sequences are available.

Plastocyanins have been broken down into four classes by Gross (1996). Type I plastocyanins come from the cyanobacteria; Type II from the green algae; Type III and IV from the higher plants. Type IV plastocyanins are very highly conserved, 99 amino acid proteins. Type III and Type II plastocyanins are less conserved and are 97 or 98 residues in length, with an omission of residues at positions 57 and 58 (relative to Type IV PCs) characterizing them. Type I plastocyanins are predictably diverse, and anywhere from 85 (Synechococcus) to 102 (Anabaena) residues in length.

Thirteen residues are conserved in all plastocyanins - G6, P16, G24, H37, N38, G78, Y80, C84, P86, H87, A90, M92, and V98. Four of those residues - H37, C84, H87, and M92 - are ligands necessary for the formation of PC's characteristic copper
Table 2.1. The 38 currently available plastocyanin sequences, aligned.
Table 2.1 (continued).

| pcpoplarA    | pcpoplarB    | pcspinach    | pctobaccoA' | pctobaccoA'' | pctobaccoB' | pctobaccoB'' | pclettuce    | cm mercury   | cpotato      | ccpotato     | cpumpkin     | pceler       | pcbitter     | pckidney     | pctomato     | cfava        | cpcsilene    | cpcurse      | cpcarabi     | pcphyscom    | pculva       | pscened            | pcenter     | pchlor        | pchlamy      | pcdryopter    | pcsynech    | pccoccus    | pcana       | pcana7120    | pcpormi     | pcpprochlo   |
|--------------|--------------|--------------|-------------|--------------|-------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| DAV-----------KISMSEEDLLNAAGETYSVTLT---KGTYFCAHPAGMVQKVTVN
| DAV-----------KISMSEEDLLNAAGETYSVTLT---KGTYFCAHPAGMVQKVTVN
| DAV-----------KISMSEEDLLNAAGETYSVTLT---KGTYFCAHPAGMVQKVTVN
| DAV-----------KISMSEEDLLNAAGETYSVTLT---KGTYFCAHPAGMVQKVTVN
| DAV-----------KISMSEEDLLNAAGETYSVTLT---KGTYFCAHPAGMVQKVTVN
| DAV-----------KISMSEEDLLNAAGETYSVTLT---KGTYFCAHPAGMVQKVTVN
| DAV-----------KISMSEEDLLNAAGETYSVTLT---KGTYFCAHPAGMVQKVTVN
| DAV-----------KISMSEEDLLNAAGETYSVTLT---KGTYFCAHPAGMVQKVTVN
| DAV-----------KISMSEEDLLNAAGETYSVTLT---KGTYFCAHPAGMVQKVTVN
| DAV-----------KISMSEEDLLNAAGETYSVTLT---KGTYFCAHPAGMVQKVTVN
| DAV-----------KISMSEEDLLNAAGETYSVTLT---KGTYFCAHPAGMVQKVTVN
| DAV-----------KISMSEEDLLNAAGETYSVTLT---KGTYFCAHPAGMVQKVTVN
| DAV-----------KISMSEEDLLNAAGETYSVTLT---KGTYFCAHPAGMVQKVTVN
| DAV-----------KISMSEEDLLNAAGETYSVTLT---KGTYFCAHPAGMVQKVTVN
| DAV-----------KISMSEEDLLNAAGETYSVTLT---KGTYFCAHPAGMVQKVTVN
| DAV-----------KISMSEEDLLNAAGETYSVTLT---KGTYFCAHPAGMVQKVTVN
| DAV-----------KISMSEEDLLNAAGETYSVTLT---KGTYFCAHPAGMVQKVTVN
| DAV-----------KISMSEEDLLNAAGETYSVTLT---KGTYFCAHPAGMVQKVTVN
| DAV-----------KISMSEEDLLNAAGETYSVTLT---KGTYFCAHPAGMVQKVTVN
| DAV-----------KISMSEEDLLNAAGETYSVTLT---KGTYFCAHPAGMVQKVTVN
| DAV-----------KISMSEEDLLNAAGETYSVTLT---KGTYFCAHPAGMVQKVTVN
| DAV-----------KISMSEEDLLNAAGETYSVTLT---KGTYFCAHPAGMVQKVTVN
| 5 6 7 8 8 8 9 9
| 1 0 0 0 4 7 2 9

* complete conservation

>90% conservation

< all residues of same type

< significant trend in this region
Table 2.2. The 23 currently available cytochrome f sequences, aligned.
Table 2.2. (continued).

<p>| | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>cytffava</td>
<td>cytffmaize</td>
<td>cytffscybean</td>
<td>cytffpea</td>
<td>cytffspinach</td>
<td>cytftobacco</td>
<td>cytftturnip</td>
<td>cytffwheat</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>cytffpicea</th>
<th>cytffpine</th>
<th>cytffprimrose</th>
<th>cytfftort</th>
<th>cytffchlamy</th>
<th>cytffchorella</th>
<th>cytffodorentella</th>
<th>cytffphormidium</th>
<th>cytffguillardia</th>
</tr>
</thead>
<tbody>
<tr>
<td>QKYSEITFPILSPPATKRDVYFLKYPYVGGNRGQVYQDGSKSNVNYNATAGVQV</td>
<td>QKYSEITFPILSPPATKRDVYFLKYPYVGGNRGQVYQDGSKSNVNYNATAGVQV</td>
<td>QKYSEITFPILSPPATKRDVYFLKYPYVGGNRGQVYQDGSKSNVNYNATAGVQV</td>
<td>QKYSEITFPILSPPATKRDVYFLKYPYVGGNRGQVYQDGSKSNVNYNATAGVQV</td>
<td>QKYSEITFPILSPPATKRDVYFLKYPYVGGNRGQVYQDGSKSNVNYNATAGVQV</td>
<td>QKYSEITFPILSPPATKRDVYFLKYPYVGGNRGQVYQDGSKSNVNYNATAGVQV</td>
<td>QKYSEITFPILSPPATKRDVYFLKYPYVGGNRGQVYQDGSKSNVNYNATAGVQV</td>
<td>QKYSEITFPILSPPATKRDVYFLKYPYVGGNRGQVYQDGSKSNVNYNATAGVQV</td>
<td>QKYSEITFPILSPPATKRDVYFLKYPYVGGNRGQVYQDGSKSNVNYNATAGVQV</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>cytffcyanophora</th>
<th>cytfnostoc</th>
<th>cytffphormidium</th>
<th>cytffpicea</th>
<th>cytffpine</th>
<th>cytffprimrose</th>
<th>cytfftort</th>
<th>cytffchlamy</th>
<th>cytffchorella</th>
</tr>
</thead>
<tbody>
<tr>
<td>QKYSEITFPILSPPATKRDVYFLKYPYVGGNRGQVYQDGSKSNVNYNATAGVQV</td>
<td>QKYSEITFPILSPPATKRDVYFLKYPYVGGNRGQVYQDGSKSNVNYNATAGVQV</td>
<td>QKYSEITFPILSPPATKRDVYFLKYPYVGGNRGQVYQDGSKSNVNYNATAGVQV</td>
<td>QKYSEITFPILSPPATKRDVYFLKYPYVGGNRGQVYQDGSKSNVNYNATAGVQV</td>
<td>QKYSEITFPILSPPATKRDVYFLKYPYVGGNRGQVYQDGSKSNVNYNATAGVQV</td>
<td>QKYSEITFPILSPPATKRDVYFLKYPYVGGNRGQVYQDGSKSNVNYNATAGVQV</td>
<td>QKYSEITFPILSPPATKRDVYFLKYPYVGGNRGQVYQDGSKSNVNYNATAGVQV</td>
<td>QKYSEITFPILSPPATKRDVYFLKYPYVGGNRGQVYQDGSKSNVNYNATAGVQV</td>
<td>QKYSEITFPILSPPATKRDVYFLKYPYVGGNRGQVYQDGSKSNVNYNATAGVQV</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>cytffpine</th>
<th>cytffprimrose</th>
<th>cytffphormidium</th>
<th>cytffpicea</th>
<th>cytffpine</th>
<th>cytffprimrose</th>
<th>cytfftort</th>
<th>cytffchlamy</th>
<th>cytffchorella</th>
</tr>
</thead>
<tbody>
<tr>
<td>QKYSEITFPILSPPATKRDVYFLKYPYVGGNRGQVYQDGSKSNVNYNATAGVQV</td>
<td>QKYSEITFPILSPPATKRDVYFLKYPYVGGNRGQVYQDGSKSNVNYNATAGVQV</td>
<td>QKYSEITFPILSPPATKRDVYFLKYPYVGGNRGQVYQDGSKSNVNYNATAGVQV</td>
<td>QKYSEITFPILSPPATKRDVYFLKYPYVGGNRGQVYQDGSKSNVNYNATAGVQV</td>
<td>QKYSEITFPILSPPATKRDVYFLKYPYVGGNRGQVYQDGSKSNVNYNATAGVQV</td>
<td>QKYSEITFPILSPPATKRDVYFLKYPYVGGNRGQVYQDGSKSNVNYNATAGVQV</td>
<td>QKYSEITFPILSPPATKRDVYFLKYPYVGGNRGQVYQDGSKSNVNYNATAGVQV</td>
<td>QKYSEITFPILSPPATKRDVYFLKYPYVGGNRGQVYQDGSKSNVNYNATAGVQV</td>
<td>QKYSEITFPILSPPATKRDVYFLKYPYVGGNRGQVYQDGSKSNVNYNATAGVQV</td>
</tr>
</tbody>
</table>

<p>| | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>cytffpine</td>
<td>cytffprimrose</td>
<td>cytffphormidium</td>
<td>cytffpine</td>
<td>cytffpine</td>
<td>cytffpine</td>
<td>cytffpine</td>
<td>cytffpine</td>
<td>cytffpine</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>cytffpine</th>
<th>cytffprimrose</th>
<th>cytffphormidium</th>
<th>cytffpine</th>
<th>cytffpine</th>
<th>cytffpine</th>
<th>cytffpine</th>
<th>cytffpine</th>
<th>cytffpine</th>
</tr>
</thead>
<tbody>
<tr>
<td>QKYSEITFPILSPPATKRDVYFLKYPYVGGNRGQVYQDGSKSNVNYNATAGVQV</td>
<td>QKYSEITFPILSPPATKRDVYFLKYPYVGGNRGQVYQDGSKSNVNYNATAGVQV</td>
<td>QKYSEITFPILSPPATKRDVYFLKYPYVGGNRGQVYQDGSKSNVNYNATAGVQV</td>
<td>QKYSEITFPILSPPATKRDVYFLKYPYVGGNRGQVYQDGSKSNVNYNATAGVQV</td>
<td>QKYSEITFPILSPPATKRDVYFLKYPYVGGNRGQVYQDGSKSNVNYNATAGVQV</td>
<td>QKYSEITFPILSPPATKRDVYFLKYPYVGGNRGQVYQDGSKSNVNYNATAGVQV</td>
<td>QKYSEITFPILSPPATKRDVYFLKYPYVGGNRGQVYQDGSKSNVNYNATAGVQV</td>
<td>QKYSEITFPILSPPATKRDVYFLKYPYVGGNRGQVYQDGSKSNVNYNATAGVQV</td>
<td>QKYSEITFPILSPPATKRDVYFLKYPYVGGNRGQVYQDGSKSNVNYNATAGVQV</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>cytffpine</th>
<th>cytffpine</th>
<th>cytffpine</th>
<th>cytffpine</th>
</tr>
</thead>
<tbody>
<tr>
<td>QKYSEITFPILSPPATKRDVYFLKYPYVGGNRGQVYQDGSKSNVNYNATAGVQV</td>
<td>QKYSEITFPILSPPATKRDVYFLKYPYVGGNRGQVYQDGSKSNVNYNATAGVQV</td>
<td>QKYSEITFPILSPPATKRDVYFLKYPYVGGNRGQVYQDGSKSNVNYNATAGVQV</td>
<td>QKYSEITFPILSPPATKRDVYFLKYPYVGGNRGQVYQDGSKSNVNYNATAGVQV</td>
</tr>
</tbody>
</table>
Table 2.2. (continued).

<table>
<thead>
<tr>
<th></th>
<th>2</th>
<th>2</th>
<th>2</th>
<th>2</th>
<th>2</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>6</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Protein</th>
<th>Sequence 1</th>
<th>Sequence 2</th>
<th>Sequence 3</th>
<th>Sequence 4</th>
<th>Sequence 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cytffava</td>
<td><code>DAEIVLQDPLRVLQGLLFLASIIIAQILFLVLKKQFEKQVQALSEMNF</code></td>
<td><code>DAEIVLQDPLRVLQGLLFLASIIIAQILFLVLKKQFEKQVQALSEMNF</code></td>
<td><code>DAEIVLQDPLRVLQGLLFLASIIIAQILFLVLKKQFEKQVQALSEMNF</code></td>
<td><code>DAEIVLQDPLRVLQGLLFLASIIIAQILFLVLKKQFEKQVQALSEMNF</code></td>
<td><code>DAEIVLQDPLRVLQGLLFLASIIIAQILFLVLKKQFEKQVQALSEMNF</code></td>
</tr>
<tr>
<td>Cytfsyrac</td>
<td><code>DAEIVLQDPLRVLQGLLFLASIIIAQILFLVLKKQFEKQVQALSEMNF</code></td>
<td><code>DAEIVLQDPLRVLQGLLFLASIIIAQILFLVLKKQFEKQVQALSEMNF</code></td>
<td><code>DAEIVLQDPLRVLQGLLFLASIIIAQILFLVLKKQFEKQVQALSEMNF</code></td>
<td><code>DAEIVLQDPLRVLQGLLFLASIIIAQILFLVLKKQFEKQVQALSEMNF</code></td>
<td><code>DAEIVLQDPLRVLQGLLFLASIIIAQILFLVLKKQFEKQVQALSEMNF</code></td>
</tr>
<tr>
<td>Cytfsynech</td>
<td><code>DAEIVLQDPLRVLQGLLFLASIIIAQILFLVLKKQFEKQVQALSEMNF</code></td>
<td><code>DAEIVLQDPLRVLQGLLFLASIIIAQILFLVLKKQFEKQVQALSEMNF</code></td>
<td><code>DAEIVLQDPLRVLQGLLFLASIIIAQILFLVLKKQFEKQVQALSEMNF</code></td>
<td><code>DAEIVLQDPLRVLQGLLFLASIIIAQILFLVLKKQFEKQVQALSEMNF</code></td>
<td><code>DAEIVLQDPLRVLQGLLFLASIIIAQILFLVLKKQFEKQVQALSEMNF</code></td>
</tr>
<tr>
<td><em>complete conservation</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>90% conservation</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>all residues of same type</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>significant trend in this region</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
center. Thirteen additional residues are conserved in 90% of all plastocyanins, including several glycines (positions 10, 67, 89, 91, and 94). The negative charge of higher plant and algal plastocyanins is key to the function of the protein.

Given that, it seems surprising that none of the conserved residues (or even the nearly-conserved residues) are charged; indeed, the only residue that is consistently an acidic residue is in position 25 and (as results later will imply) does not appear to have a functional role. However, when only type II, III and IV plastocyanins are considered, the trends become far more apparent.

In higher plant and algal plastocyanins, residues 42, 43, 44 are completely conserved. There is always a negatively charged residue in either position 51 or 53, and in all cases except parsley (which only has E61) there are at least two negative charges in positions 58-61. The positioning of the positive charges is far less conserved - all Type IV PCs have lysines in positions 30 and 54, and most have lysines at positions 27, 77, and 95 as well. But few of the Type III's and fewer of the Type II's follow those trends. And, of course, the diversity of the Type I PCs equate to few trends with respect to their electrostatics being obvious.

Similar trends are apparent when one looks at the sequences of cytochrome f. 75 of the 286 residues in the full protein - and 64 of the roughly 250 residues in the soluble portion of the protein - are conserved across all 23 of the known cyt f sequences. Of those 64 residues in the soluble domains, only H25, E34, K39, K145, R154, R156, and E243 are titrable. (H25 is part of the characteristic CANCH heme-binding motif, and nothing further should be read into its conservation.) However, when the seven
cyanobacterial cyt f's are eliminated from consideration, interesting trends emerge again. All higher plant and algal cyt f's have a positive charge somewhere in position 55, 57, or 58. An absolutely conserved stretch from positions 62-66 ends with a pair of lysines. Positively charged residues are found at positions 94 and 96. At least one, often two lysines are found in positions 121 and 122. Higher plant cyt f's conserve an arginine at position 184 and lysines at 185 and 187; the algal f's either have two lysines at positions 188 and 189 (Chlamydomonas, Chlorella) or a single lysine at position 187 (Odontella).

The trends in the sequences of the two proteins imply that there is some role for electrostatic attraction to play in the interaction between plastocyanin and cytochrome f in higher plants and algae. Moreover, it is clear that whatever that role is, it does not carry over to the interaction of cyanobacterial plastocyanins and cytochrome f's - the lack of conservation in those same charged residues implies that, if there is any electrostatic role for PC and cyt f in cyanobacteria, it is vastly different than that in higher plants and algae.

Plastocyanin structures and electrostatics. Qualitative electrostatic analysis, through the software package GRASP, should reinforce these conclusions. Durell (1990), through experiments with a similar software package, DelPhi, had previously made note of the dominance of the negative charge on the so-called "east face" of the protein, encompassing residues 42-45 and 59-61, and argued about the importance of these residues in the interaction with a positively charged patch on cyt f, based on evidence
from previous experiments (Beoku-Betts, 1983; Takabe, 1984; Takenaka, 1984; Anderson et al., 1987).

Figures on the next several pages depict the electrostatic fields of several plastocyanins. The dominant negative charge Durell described clearly carries over to other higher plant and algal plastocyanins, and equally clearly is not conserved in the cyanobacterial plastocyanins.

In poplar plastocyanin, the negative patch on the east face is composed most notably of residues 42-44 on the lower left (as presented in the figure), D51 on the lower right, and residues 59-61 on the upper right. Y83 holds a central position around this negative charge, which leads to its postulated interaction with the heme in the complex formed with cytochrome c (Roberts et al., 1991) and, when the structure of cyt f was unknown, made for natural speculation about Y83's physiological role in electron transfer with cyt f. While these negative charges are not specifically conserved, especially in Type II and III plastocyanins, there is typically replacement on the surface of the protein for those residues not conserved. For instance, in parsley plastocyanin there is a negative charge at position 53 (glutamate) instead of 51 (aspartate), and while there is only one negative charge at position 61 instead of poplar's three between 59 and 61, parsley PC adds a glutamate at position 85; this negative charge at position 85 is conserved throughout the Type II and III plastocyanins. This begs an idea of functional homology, instead of strict sequential homology. The electrostatic properties of the higher plant and algal plastocyanins are remarkably similar for the measure of diversity there is in those sequences.
Figure 2.1: Electrostatic potential representation of oxidized poplar plastocyanin. Electrostatic potential was calculated through solution of the linearized Poisson-Boltzmann equation, with the program GRASP. The ionic strength used in the calculation is 100 mM. 3-dimensional contour represents -1 kT/e; 2-dimensional contour represents zero potential. Key negative charges shaded.
Figure 2.2. Electrostatic potential representation of french bean plastocyanin (Moore et al., 1991). Electrostatic contours as in Figure 2.1.
Figure 2.3. Electrostatic potential representation of parsley plastocyanin (Bagby et al., 1994). Electrostatic contours as in Figure 2.1.
Figure 2.4. Electrostatic potential representation of *Chlamydomonas* plastocyanin (Redinbo et al., 1993). Electrostatic contours as in Figure 2.1.
Figure 2.5. Electrostatic potential representation of *Synechocystis* plastocyanin (Romero et al., 1998) with the three mutated residues in the crystal structure restored to the wild type. Electrostatic contours as in Figure 2.1. Note the distinct change in negative potential distribution as compared to higher plant and algal plastocyanins.
Figure 2.6. Electrostatic potential representation of *Anabaena* plastocyanin (Badsberg et al., 1996). Electrostatic contours as in Figure 2.1.
Figure 2.7. Electrostatic potential representation of *Prochlorothrix* plastocyanin (Babu et al., 1999). Electrostatic contours as in Figure 2.1.
Very little such conservation happens in the cyanobacterial plastocyanins. Only a single negative charge in the 42-44 patch is conserved in all cyanobacteria, and the dominant negative patch is all but nonexistent in cyanobacterial plastocyanins. PC from *Synechocystis* does have a significant negative patch, but it is shifted far down on the east face and results primarily from aspartates in positions 44 and 47 and a glutamate at position 76. The only other negative charge on the east face is E85, at the upper right - and it is surrounded by a lysine in position 58 and an arginine in position 88, covering the east face tyrosine in positive potential instead of its usual negative.

*Synechocystis*, as it turns out, has the most notable negative charge of all the cyanobacteria. The negative potential over the tyrosine in *Anabaena* PC (due to D44 and E90) is surrounded by positive potential above due to the copper and R93, to the left due to K20 and K51, and to the right due to K57 and K62. *Prochlorothrix* PC is almost the total reverse of the higher plant and algal plastocyanins due to lysines in positions 17, 33, 43, and 57 and arginine 84 in addition to the copper. D42 and E48 provide for a small negative patch homologous to that on *Synechocystis*.

**Cytochrome f structures and electrostatics.** While the electrostatics of plastocyanin has already been well-investigated by Durell (1990), at the onset of this work the solution of the cytochrome f crystal structure remained very new, and begged a unique investigation. The 1996 publication deals with the electrostatics of poplar cytochrome f in detail.
Figure 2.8: Electrostatic potential representation of reduced turnip cytochrome f. Electrostatic potential was calculated through solution of the linearized Poisson-Boltzmann equation, with the program GRASP. The ionic strength used in the calculation is 100 mM. 3-dimensional contour represents $+1 \text{kT/e}$; 2-dimensional contour represents zero potential. Key positive charges shaded.
Turnip cytochrome $f$ (Figure 2.8), the first cyt $f$ to be solved, has a large positive patch to complement the negative patch on PC. That patch encompasses K185, K187, R209, K65, K66, K58, and K122 along the spine of the molecule, bridging the small domain and the large, as well as R156 immediately above the heme. The positive charge of the iron center also contributes to the positive potential, surrounding the heme edge and only interrupted by the negative charge on the carboxylic acid side chains on the porphyrin. There is a substantial extent of positive potential away from the surface of cyt $f$, which we hypothesized in the 1996 article would make for a "beacon" to attract negatively-charged PC. Negative charges on cyt $f$ are of less consequence. The most substantial region of negative potential on turnip cyt $f$ results most predominantly from E200, a residue occurring only in turnip cyt $f$. A smaller negative patch on the large domain results from residues D32, E34, E79, E95, D241, and E243 - those residues are conserved in a majority of cyt $f$'s, both eukaryotic and prokaryotic.

The other available crystal structure for cyt $f$ comes from *Chlamydomonas reinhardtii* (Figure 2.9), where the dominant positive patch is made up of many of the same residues (albeit in different positions; K188 and K189 in *Chlamydomonas* as opposed to K185 and K187 in turnip; K121 and K122 in *Chlamydomonas* as opposed to solely K122 in turnip; R209 in turnip is not present in *Chlamydomonas*), and the less prevalent negative patch in *Chlamydomonas*’ small domain extends the positive "beacon" of potential as observed in turnip cyt $f$ even wider.

While the aforementioned negative surface charges in cyt $f$ are largely conserved in all cyt $f$'s, however, the charges involved in the positive patches are not. Using the
Figure 2.9. Electrostatic potential representation of *Chlamydomonas* cytochrome *f* (Berry et al., unpublished results). Electrostatic contours as in Figure 2.8.
structures of turnip and *Chlamydomonas* cyt f, a homology model of cyt f from *Synechocystis* SCC 6803 was constructed and electrostatically modeled. The only positive charges retained from the eukaryotic cyt f's is a single lysine (position 67) along the spine and the arginine (position 157) above the heme. There are glutamates in positions 57, 64, 122, and 187 - near where there would be positive charges in higher plant or algal cyt f's. The result, according to the model (Figure 2.10), is a cyt f with substantial negative charge in the same region where the positive charge on higher plant and algal cyt f's would be. (An aside: the purpose of dealing with the cyanobacterial proteins so thoroughly has been to demonstrate the substantial differences between their electrostatic properties and the electrostatic properties of plastocyanin and cytochrome f from higher plants and algae. As previously stated, experimental evidence points to the importance of the electrostatic interaction between PC and cyt f for the function of the two proteins. The stark contrast of the electrostatics of the cyanobacterial proteins implies that the electrostatic interactions between those proteins, if present at all, are clearly divergent from the electrostatic interactions of the corresponding eukaryotic proteins. Therefore, the study of cyanobacterial plastocyanins and cytochromes f will not be a substantial part of this work. We will concentrate from this point forward on the study of plastocyanins and cytochromes f from higher plants and algae.)

One of the key points of Durell's (1990) work with plastocyanin was the "collapsing" of the potential field closer to the protein surface with increasing ionic strength, implying that long-range interactions between proteins would decrease in importance as ionic strength increased. This was demonstrated for plastocyanin with
Figure 2.10. Electrostatic potential representation of *Synechocystis* cytochrome *f* as homology modeled for this work. Electrostatic contours as in Figure 2.8.
simple 2-D contours surrounding a slice through the protein. The 3-D graphical
representation of electrostatic contours with GRASP allow one to observe this
collapsing effect more starkly, as seen in Figures 2.11-2.13.

An additional effect of ionic strength changes that was not expected at the time
was the disappearance of the negative potential over the heme at low ionic strength. At
more moderate ionic strengths, there is a negative potential directly over the exposed
heme edge due to the propionates on the porphyrin. However, at very low ionic
strengths, the positive potential due to the basic residues (K187, K65, K58, etc.)
overwhelm this negative potential and the region over the heme is all but totally covered
by positive potential. A crude computation of potential values away from the surface-
exposed carboxyl from the 1996 paper (Figure 2.14) demonstrates this quantitatively.

We speculated at the time that the lower the ionic strength, the better the
attraction cyt f would have for plastocyanin, particularly because of the negative
potential over the heme (which would repel the negatively charged residues on PC)
would be obscured. We also used this effect to help justify a view of how PC and cyt f
would form a complex that would later prove to be somewhat misguided.

*Original hypotheses of complex formation.* The initial goal of this research (which
remains one of the primary goals) was to generate ideas about how PC and cyt f form
complexes for purposes of electron transfer. As such, when I joined the Gross lab, my
first task was to assemble putative complexes between PC and cyt f, which would later
serve as the basis for hypotheses about the docking mechanism and electron transfer
Figure 2.11. Electrostatic potential field of reduced cyt f for ionic strength of 2 mM. Key charged residues (K187, R209, K65, K66, K58) colored black; heme colored black, Y1 shaded. All other conditions besides ionic strength same as Figure 2.8.
Figure 2.12. Electrostatic potential field of reduced cyt f for ionic strength of 25 mM. Residue colors and all other conditions besides ionic strength same as Figure 2.11.
Figure 2.13. Electrostatic potential field of reduced cyt f for ionic strength of 200 mM. Residue colors and all other conditions besides ionic strength same as Figure 2.11.
Figure 2.14. A plot of electrostatic potential against distance for a series of points extending up to 10 Å away from a porphyrin oxygen (atom O1D), with a spacing of up to 0.5 Å, representing the cross-over of potential over the heme from negative to positive at lower ionic strengths.
pathway. The basic ideas surrounding this manual docking were:

- Based on the results of Handford et al. (1980) and Cookson et al. (1980), H87 or Y83 needed to be involved in an electron transfer pathway to PC;
- Based on the fact that there really did not appear to be a better option, Y1 or the exposed portion of the heme should be involved in an electron transfer pathway from cyt f (we still needed to consider other options, but no other realistic option that minimized the distance between metal centers and therefore increased the probability of electron transfer (Marcus and Sutin, 1985; Tollin et al., 1986; Moser et al., 1995) presented itself);
- Available biochemical data, if possible, had to be satisfied.
- Electrostatic interactions had to be favorable.
- Steric hindrance needed to be eliminated.

In the end, it was impossible to satisfy all of these conditions with any one manually docked complex. A set of three manually docked complexes was generated that we believed best represented the range of possibilities. Figures 2.15-2.17, describing those complexes and residues on each protein in close contact in those complexes, are found on the following pages.

Detailed descriptions of these complexes are found in the 1995 and 1996 publications. However, an overall context for these complexes and their role in a hypothesized reaction mechanism is useful.
Complex 1 (Figure 2.15) was designed to maximize electrostatic interactions between the two proteins. In it, PC is oriented with respect to cyt f so that the negatively charged residues on PC line up with the positively charged residues on cyt f. Specifically in mind with the design of this complex was the study of Morand et al. (1993) where chemical cross-linking of charge interactions between PC and cyt f implicated a pairing between D44 on PC and K187 on cyt f. Morand also observed that E59 or E60 on PC interacted with a positively charged residue on cyt f, but the nature of the peptide hydrolysis used to identify residues did not permit the identification of the positively charged residue. Based on the topography of the two proteins, we concluded that the basic residue was either K58 or K65, or possibly both in different circumstances.

While electrostatic interactions were favorable, the distance between metal centers was not. A distance of 31 Å is hardly suitable for electron transfer, and the fact that the only "possible" electron transfer pathway was a contact between Y83 on PC and the buried water chain within cyt f required a search for other possibilities.

Complex 2 (Figure 2.16) minimized the distance between metal centers (to 15.6 Å) by placing H87 on PC immediately adjacent to Y1 on cyt f. This could be done while maintaining a reasonably low distance between D44 and K187, as previously paired. An added point of interest was the position of L12 on PC. When H87 and D44 were paired with Y1 and K187, respectively, L12 was oriented facing Y1 and Y160 on cyt f. Experiments from Modi et al. (1992b) demonstrated an increase in binding affinity between PC and cyt f when L12 was mutated to an asparagine. An L12N mutant
Figure 2.15. Hypothetical "Complex 1" between PC and cyt f, representing electrostatic attraction between cyt f positive residues and PC negative residues. Representation prepared with GRASP. Proteins were docked manually as described in Methods, and then separated 15 Å to show electrostatic interactions. Potential calculation was performed at an ionic strength of 100 mM; two-dimensional interpolation plane contours are of values of -1, -0.5, 0, 0.5, and 1 kT/e; positive and negative sides of zero contour marked.
Figure 2.16. Hypothetical "Complex 2" between PC and cyt f, representing a dock with close contacts between H87(PC) and Y1(cyt f) and D44(PC) and K187(cyt f). Representation prepared as Figure 2.15.
could conceivably hydrogen-bond with the hydroxyls on those tyrosine residues or with the porphyrin carboxyl which is also adjacent to L12 in the dock. Finally, there appeared to be some possibility for hydrophobic interactions to form between residues on PC (L12, Q88 aliphatic carbons) and cyt f (F4, A62, P161, Y1 and Y160 aromatic carbons). We estimated the solvent-excluded hydrophobic surface area in this complex as 256 Å² and estimated the hydrophobic interaction energy to be between -6.4 and -12.0 kcal/mol.

The major problem with this complex seemed to be the poor electrostatic interactions between the heme region on cyt f and the H87 region of PC, as demonstrated by the repulsive (2-D) electrostatic contours in the published figure. There was also an issue of steric repulsion due to the interference of the loop at positions 88-91; however, we postulated that the presence of glycines in that loop would allow for sufficient flexibility to get offending residues out of the way in complex formation.

The conclusion was reached that any complex between PC and cyt f that involved a H87/Y1 interaction was going to have poor electrostatics, so an effort was made to improve the hydrophobic interactions, leading to Complex 3 (Figure 2.17). The D44/K187 contact was broken to involve hydrophobic residues such as F35, P36 and L62 on PC and I3, F4 and P117 more fully in the complex. We were able to maximize the solvent-excluded hydrophobic surface area at an estimated 385 Å², or a hydrophobic interaction energy of -9.6 to -18.1 kcal/mol; however, this maximized repulsion was at a cost at greater electrostatic repulsion.
Figure 2.17. Hypothetical "Complex 3" between PC and cyt f, representing a dock with close contact between H87(PC) and Y1(cyt f) while maximizing hydrophobic interactions. Representation prepared as Figure 2.15.
Because all of these complexes involved the interaction of H87 with the heme, we looked to form a complex between Y83, the other postulated electron transfer site on PC, and the heme. At this time, NMR evidence (He et al., 1991) pointed to the likelihood that Y83 was the electron transfer residue of greater importance as it concerned electron transfer from cyt f. However, no complex between poplar PC and turnip cyt f allowed Y83 to get any closer than 10 Å to the surface of cyt f without significant steric hindrance, due to the partial burial of Y83 behind the negatively charged residues surrounding it on poplar PC. We were able to create such a complex using parsley PC and turnip cyt f; however, such a complex did not satisfy any of the known biochemical data the way Complex 2 in particular did, and at this point a "gut feeling" among the group was forming that H87 was the site for electron transfer from cyt f to PC. In the 1996 publication, we stated that while conformational changes could allow for a satisfactory complex involving a Y83/Y1 link to form, such large-scale conformational changes had not been observed to date in PC and the general trend of similarity across all known PC structures implied that such large-scale conformational changes would not be observed.

Therefore, the view of the reaction mechanism we took was the following. On PC’s diffusion to interact with cyt f, an initial, pre-docking complex similar to Complex 1 would form through electrostatic interactions. This complex might or might not lead to the formation of Complex 2, pivoting about the tight link between the charged residues (particularly the D44/K187 link) and pulled by hydrophobic attraction of the H87 face of PC and the Y1/heme region of cyt f. If Complex 2 was formed, the
"forward momentum" of the proteins would push the proteins together close enough for electron transfer to happen, likely as a transition to Complex 3, where the Cu-Fe distance we observed was (what we thought was) a minimum at 15.3 Å. In the Complex 2/Complex 3 situation, the instability due to the repulsive electrostatic interactions would be what eventually drove the proteins apart, as they must for the electron transfer chain to move forward. In other words:

\[ \text{PC}_1 + \text{cyt f}_1 \leftrightarrow [\text{PC-cyt f}_1]_{1-2} \rightarrow [\text{PC-cyt f}_1]_{1-3} \rightarrow \text{PC}_3 + \text{cyt f}_3 \]

The hypothesis in place, the next step was to test it through a less biased method.
CHAPTER 3

INITIAL BROWNIAN DYNAMICS EXPERIMENTS

Manual docking experiments are useful for the design of hypotheses and visualization of possibilities for complex formation between two proteins, but not for much more. In order to test these hypotheses we developed computationally, we required "as unbiased a means as possible to determine interaction sites and model the means of interaction between the two proteins" (Pearson and Gross, 1998).

We were especially intrigued by the study of Northrup et al. (1987; also Northrup et al., 1988) concerning the interaction between cytochrome c and cytochrome c peroxidase. Northrup applied a computational technique called Brownian dynamics (BD) to the study of these proteins, with methodology developed by he and McCammon (Ermak and McCammon, 1978; Northrup et al., 1984). Using this method, Northrup made estimates of bimolecular rate constants, determined a range of positions of closest approach where cyt c was most likely to make its closest approach to cyt c per, modeled the history of a single diffusive encounter, and calculated ionic strength dependence of the interaction between the proteins, among other results.

We determined that we could use Brownian dynamics to study the interaction
between plastocyanin and cytochrome f. Our primary goal was, again, to get an idea of possible complexes that the two proteins might form during electron transfer. However, we were also interested in estimates of bimolecular rate constants, and extending those estimates into a model of ionic strength dependence. Previous experiments from our group (Gross et al., 1990) and other research groups (Qin and Kostic, 1992; Kannt et al., 1996) determined that the bimolecular rate constant of PC interaction with cyt f is around $10^8 \text{M}^{-1}\text{s}^{-1}$. We wondered whether Brownian dynamics could replicate those results, and if so, how reliably.

Despite our goal of a "more unbiased method", we could not make the argument that Brownian dynamics simulation was truly unbiased. Because the computation of rate constants through this method, as a statistical process, requires some definition of "success" and "failure", those definitions as supplied by the end user biased the calculation of rate constants and the determination of closest complexes towards what the user perceives as success.

Further, Brownian dynamics is a method that depends on the calculation of forces that might influence the diffusion of the mobile protein (PC) towards the stationary protein (cyt f). Electrostatic forces can easily be computed through electromagnetic theory, if one knows the potential field the two molecules (through solution of the Poisson-Boltzmann equation). Hydrophobic interactions, however, are not so easily quantifiable, and certainly a "hydrophobic field" of the same type as the electrostatic potential field can’t be easily calculated with present knowledge. Therefore Brownian dynamics methods deal only with electrostatic interactions between
proteins. This works very well for proteins like cyt c and cyt c per, where electrostatic interactions are far and away most important in the interaction (Kang et al., 1978; Margoliash and Bosshard, 1983); this works very poorly for proteins like human growth hormone and its receptor, where hydrophobic interactions are anticipated to have a dominant role (Clackson and Wells, 1995; Wells, 1996). For PC and cyt f, where the importance of hydrophobic interactions is postulated but has not been experimentally demonstrated, the magnitude of the effect of not considering hydrophobic interactions is unknown.

Software for BD simulation is available from several groups; we looked seriously at two packages, UHBD (Davis et al., 1991) and MacroDox (Northrup et al., 1996). The design of UHBD turned out to be better applied to small-molecule/large-molecule interaction problems (e.g., Luty et al., 1993; Elcock and McCammon, 1996) as well as for more theoretical studies (e.g., Antosiewicz and McCammon, 1995). MacroDox was designed as a result of Northrup's cyt c/cyt c per studies, and later applied to the interaction of cyt c and cytochrome b₅₆₇ (Northrup et al., 1993) - both electron-transfer systems. We saw MacroDox as the best possible program for studying the interaction between PC and cyt f.

The hope here, therefore, was for the development of a plausible model that would describe complex formation between PC and cyt f as completely as possible. This model would then be used to describe how complex formation might change as a function of ionic strength, and would serve as the basis for future study.
METHODS

Structures. The molecular structures used for this study were those of turnip cytochrome f (PDB ID: 1HCZ; Martinez et al., 1996) and poplar plastocyanin (ID: 1PLC; Guss et al., 1992), obtained from the Protein Data Bank (http://www.rcsb.org/pdb/ - Bernstein et al., 1977). *Anabaena variabilis* plastocyanin (ID 1N1N; Badsberg et al., 1996) and azurin from *Alcaligenes denitrificans* (ID 2AZA, Baker 1988) were used as controls against the behavior of poplar PC.

Brownian dynamics theory and computation. MacroDox (Northrup, 1996) was used for Brownian dynamics simulation. The overall process used in a single BD simulation was as follows:

- Charges were assigned to the two proteins.
- Electrostatic potential maps were calculated for the two proteins.
- The BD simulation was prepared and executed.
- The results of the simulation were analyzed.

Charge assignment. A program based very closely on the Tanford-Kirkwood algorithm (Tanford and Kirkwood, 1957) for determining pK\textsubscript{a} values, based on the algorithms of Matthew and Gurd (1986), is part of the MacroDox package and was used for charge assignment for both proteins. For the age of the algorithm, the Tanford-
Kirkwood formalism works very well; however, for the sake of simplicity of the method, approximations are made (for example, considering each charge as being placed on the surface of a sphere) that affect the accuracy of the method.

For proteins being analyzed at pH 7 (and almost all of the experiments from this point forward will assume pH 7), the Tanford-Kirkwood method gives essentially the same answers as the simple charge assignments (+1 for all basic residues, -1 for all acidic) made for the studies with GRASP described previously. Histidine residues are subject to partial charges; however, both histidine residues on PC and H25 on cyt f are not charged, because of their status as metal ligands. H142/cyt f, the only other histidine on either protein, is on the side of cyt f reverse to the relevant interactions between PC and cyt f and winds up having negligible effect on the binding of PC.

Metal centers are charged as they would be previous to biological electron transfer: PC is oxidized, with a Cu$^{2+}$ center; cyt f is reduced, with an Fe$^{2+}$ center.

The outstanding advantage of the Tanford-Kirkwood pKₐ solver is speed. An entire solution for pKₐ values on the larger of the two proteins, cyt f, takes less than 15 seconds on an SGI O2 workstation.

**Electrostatic potential calculation.** On charge assignment, potential maps for both proteins were computed using a FDPB method similar, but not identical, to that used in the electrostatics studies. The program used (again, designed specifically for MacroDox) utilizes the Klapper's (1986) original adaptation of Warwicker and Watson's (1982) methodology for solving the Poisson-Boltzmann equation. While the MacroDox FDPB algorithm might be cruder than the Nicholls and Honig (1991) FDPB
algorithm utilized in GRASP, the implementation of the MacroDox algorithm is actually more detailed - MacroDox employs a 51 x 51 x 51 outer grid and a 61 x 61 x 61 inner grid, as opposed to GRASP’s 33 x 33 x 33 inner and outer grids.

*Brownian dynamics simulation.* The BD simulation implemented in MacroDox is based on the Ermak-McCammon model (1978):

\[ \mathbf{r} = \mathbf{r}_0 + \beta D \mathbf{F}(\mathbf{r}_0) \Delta t + \mathbf{R} \]

This equation describes the motion of a mobile particle in space, guided by some external (here, electrostatic) force \( \mathbf{F}(\mathbf{r}_0) \). \( \mathbf{r} \) and \( \mathbf{r}_0 \) refer to the position of the mobile particle after and before, respectively, the time step \( \Delta t \). \( \beta \) is a constant, \( (k_B T)^{-1} \). \( D \) is the diffusion coefficient for the mobile particle. \( \mathbf{R} \) is a random (“Brownian”) vector with the following properties:

\[ \langle \mathbf{R} \rangle = 0 \]
\[ \langle \mathbf{R}^2 \rangle = 2 D \Delta t \]

The process that is being described in a Brownian dynamics simulation is the diffusion of this mobile particle (here, PC) towards a stationary receptor (cyt f). The above equation represents a single step in a "trajectory" that simulates the motion of PC near to and far from the protein.

The overall scheme in which BD simulations are carried out is as follows. MacroDox calculates a maximum distance at which electrostatic interactions between the two proteins might have a significant impact relative to random diffusion - for the
PC/cyt f simulation, that distance is 89 Å. The mobile protein is then, at the start of the trajectory, placed at a random location on an imaginary sphere of that distance (89 Å), with the stationary receptor at that sphere’s center. The trajectory begins; at each step the electrostatic force and the random force are calculated anew, and the mobile particle is moved in that direction. The size of the step is variable; it is determined dynamically at each iteration, and it is never so large that the forces or torques on the mobile protein vary by more than 5% from beginning to end of the step. This continues until the mobile particle becomes unlikely to return to interact with the stationary receptor (here, when PC moves outside of an "escape sphere" of radius 200 Å away from cyt f), at which point the trajectory is concluded. (See Figure 3.1.)

In practice, step sizes are calculated two different ways depending on the position of the mobile protein. Inside the inner (89 Å) sphere, where electrostatic interactions are considered relevant, smaller steps (of time duration on the order of 10-100 ps) are taken to ensure careful calculation of the effect electrostatic interactions have on the interaction between the two proteins. Should the mobile protein get "stuck" in a single position during the trajectory, a small boost of ~1 Å is provided. Outside that sphere, where electrostatic interactions are not considered relevant, larger steps (of time duration ~1-10 ns) are allowed on the assumption that Brownian forces are the only forces guiding the protein.

In this process, the mobile protein may or may not reach a point where it is near enough to PC for electron transfer. Ideally, one would define a "successful" trajectory as one where electron transfer would definitely occur, and an "unsuccessful" trajectory
Figure 3.1. A representation of the Brownian dynamics trajectory at its onset. PC, the mobile ligand, is set on an imaginary sphere with radius $b=89 \, \text{Å}$ with cyt f, the fixed receptor, set at the center of that imaginary sphere. PC is subject to two forces: electrostatic attraction/repulsion between the two proteins, and random (Brownian) force set to model diffusion. The trajectory concludes when PC escapes a concentric imaginary sphere $c=200 \, \text{Å}$, with the position and orientation of PC having been recorded when it reaches its closest approach to cyt f by a preset measure (e.g., distance between metal centers).
as one where electron transfer would definitely not occur. However, the only data available here is the position of PC relative to cyt f, which is not in and of itself sufficient to determine probability of electron transfer. (In addition, the only dynamics of the protein being modeled are those of the two proteins with respect to one another; the internal dynamics of the two proteins are ignored. The drawback of this will be discussed in detail shortly; here we will simply say that this further demonstrates the inability to estimate electron transfer probability.) Therefore, success rates are simply measured based on the distance between key atoms in the proteins.

(MacroDox was designed with heme proteins in mind, and therefore has a preset reaction-criterion scheme involving distances and angles between atoms in the heme plane for mobile protein and receptor. However, this scheme had to be rethought with the mobile protein being plastocyanin; the distorted tetrahedral geometry of the copper center is most certainly not homologous to the heme plane. The decision was made to think of the criterion for establishing a successful trajectory as simply a distance between a single atom on PC and a single atom on cyt f, as described below.)

The obvious best way to determine success would be as a function of distance between the two metal centers on the proteins, as those are the places where the electrons are carried. However, this borders on circular reasoning - we intend to demonstrate that PC and cyt f dock at a place where electron transfer is reasonable by minimizing the distance between the metal centers where electrons are carried. To alleviate those concerns, success was determined in separate experiments by minimizing the distance between the center of mass of the two proteins, and also between the
residues on each protein implicated in a chemical cross-link, D44/PC and K187/cyt f.

Sets of 1000 trajectories were prepared and executed. Out of those 1000 trajectories, a fraction (p) of those trajectories returned successes. For a single experimental setup, groups of 1000 trajectories were executed (four or nine, depending on the situation). Each run was executed using a different randomly-generated random number seed.

Analysis of simulation results. Northrup (1987) derived equations from the theories of Smoluchowski (1916) to convert the success fractions (p) to estimates of rate constants (k):

\[ k_D(s) = 4\pi s D N_A \]

\[ k = p k_D(b) / [1 - (1 - p) k_D(b) / k_D(c)] \]

\( b \) and \( c \) are the radii of the two imaginary spheres surrounding the stationary receptor (89 Å and 200 Å, respectively). \( D \), as before, is a diffusion coefficient for the mobile protein (PC). \( N_A \) represents Avogadro's number (6.022 x 10^23).

\( p \) and \( k \) results were recorded for individual runs of 1000 trajectories. For each set of experimental criteria, the multiple runs with different random number seeds were summed and averaged, with standard errors calculated, to provide "more reliable" predictions for \( p \) and \( k \). The goal was to pinpoint MacroDox' prediction for probabilities and rate constants for each experimental criterion as accurately as possible and with as few total runs as possible. Errors calculated were obviously not the real
range of error in the prediction of probabilities of rate constants (that would require considering the effect of eliminating hydrophobic interactions from consideration, and other consideration of fallacies of the model), but merely theoretical deviations of MacroDox' measurements.

For each successful trajectory, the closest approach of PC to cyt f was also recorded. We modified FORTRAN code from Northrup to sort these closest approaches on the basis of the closest pairs of charged residues. The sorting was based on the three closest pairs, hence our terminology "triplet contacts" that we have used frequently in our work. This allowed us to get a first clear view of where PC was most likely to dock with cyt f. Later, we selected complexes between the two proteins for analysis based merely on closeness of metal centers. We must emphasize that these were complexes outputted by MacroDox without refinement, either in terms of relaxation within the individual proteins or optimization of the interaction between proteins. The resulting complexes were simply the closest approach of the two rigid bodies to one another. No energy minimization was performed, nor was any other optimization procedure performed, to idealize the structure of the resulting complexes.

RESULTS AND DISCUSSION

*Standard runs - demonstrating trend towards minimized metal center distance.* Initially, we set up nine standard runs of 1000 trajectories each, determining success for each
trajectory as a function of distance between metal centers, with the success distance being placed at 35 Å between metal centers. These monitor atoms (the copper atom on PC, the iron atom on cyt f) were selected for their biological relevance - it only makes sense that we determine where these proteins make their "closest approach" on the basis of the atoms that matter most in determining whether electron transfer takes place (Marcus and Sutin, 1985; Tollin et al., 1986; Moser et al., 1995). The problem with this view is that part of what we are attempting to determine is the site on PC and cyt f where the two proteins will make their closest approach, and using distance between metal centers to make that determination biases the experiment. We will address this issue shortly.

Figure 3.2 describes the results of two series of runs: one with the ionic strength for the electrostatic calculations set at 100 mM to approximate physiological conditions, the other set at 10 M to eliminate electrostatic interactions (i.e., to measure the effect of diffusion alone). The results are presented as a distribution of successful complexes (i.e., the position of closest approach of PC towards cyt f for each trajectory) as a function of metal center distance. (In this plot, the number at each distance \( n \) reflects the number of closest approaches per 1000 trajectories that were observed at distances between \( n-1 \) and \( n \) angstroms; for example, the number of complexes recorded at 20 Å represents all complexes with a metal center distance between 19 Å and 20 Å.)

These experiments were done to see if electrostatic interactions favored an interaction that would minimize the distance between metal centers. In fact, there is a clear tendency for a greater number of complexes to form through electrostatic
Figure 3.2. The distribution of distances between metal centers for all complexes with metal center distances of less than 35 Å for 9 sets of 1000 trajectories each at 100 mM and 10 M, where complexes were recorded on the basis of minimized Cu-Fe distances.
interactions and diffusion than from diffusion alone through a distance of 21 Å. For the 1998 paper, we calculated an average probability of complex formation at a distance less than 21 Å of 0.059 ± 0.005 at an ionic strength of 100 mM (electrostatic interactions case) as opposed to 0.005 ± 0.002 at an ionic strength of 10 M (diffusion case), and similar probabilities at distances less than 17 Å of 0.011 ± 0.001 at an ionic strength of 100 mM as opposed to 0.0001 ± 0.0001 at an ionic strength of 10 M.

So clearly, by this standard, there is some preference for electrostatic interactions to steer PC into a complex with cyt f where distance between metal centers is minimized, where that tendency does not exist when it is left solely to diffusion.

However, it is very easy to criticize this experiment by noticing that the experiment searches for closest complexes on the basis of minimization of distance between metal centers. In order to provide confidence that the effect of electrostatic interactions leading to close contact between metal centers is not merely an artifact of the experiment, we need to do other experiments where that search criterion is eliminated.

Two experiments were performed. In the first, the reaction criterion was the distance between the atoms closest to the center of mass in the two proteins (the γ2 carbon of 139 on PC; the β carbon of N153 on cyt f). The goal here is to eliminate bias by considering only the geometry of the protein, instead of biologically relevant portions of the protein, in determining whether or not interactions have occurred. Even here, though, bias is not eliminated completely; the irregular shape of cyt f means complexes towards the center of cyt f’s oblong shape will be favored. However, since complexes
forming at either end of cyt f is highly unlikely (and unimportant for electron transfer since the distance between metal centers would be prohibitively large in such a complex), the desired effect will still be achieved.

The second experiment utilized a reaction criterion between the δ1 oxygen atom of D44 on PC and the ζ nitrogen on K187 of cyt f. The contact between D44 and K187 is well-established from Morand’s (1989) cross-linking studies; the goal is to see if such contacts when electrostatic interactions are "turned on" have a stronger tendency to form close contacts between metal centers as opposed to when the interaction is solely diffusion-controlled. Here, we have an obvious (experimentally-demonstrated) bias in the experiment; we want to observe the nature of complexes forming with that bias assumed.

In both cases, as demonstrated in Figures 3.3 and 3.4, the tendency for electrostatic interactions to favor complexes with smaller metal-center distances remains clear and evident, although not as pronounced as previously. In the case where center of mass distance is minimized, the probability of complexes with a Cu-Fe distance of less than 21 Å is 0.025 ± 0.004 at 100 mM and 0.005 ± 0.002 at 10 M; at distances of less than 17 Å, the probabilities are 0.005 ± 0.001 at 100 mM and 0.0003 ± 0.0003 at 10 M. The case where the distance between D44 and K187 is minimized yields far less satisfying results; for Cu-Fe distances less than 21 Å, the probabilities are 0.005 ± 0.002 at 100 mM and zero at 10 M, and the results for distances less than 17 Å are meaningless. However, at distances less than 25 Å the probability at 10 M remains only 0.001 ± 0.001; the probability at 100 mM is 0.022 ± 0.005. Recall the thinking in
Figure 3.3. The distribution of distances between metal centers for all complexes with metal center distances of less than 35 Å for 9 sets of 1000 trajectories each at 100 mM and 10 M, where complexes were recorded on the basis of minimized distance between centers of mass.
Figure 3.4. The distribution of distances between metal centers for all complexes with metal center distances of less than 35 Å for 9 sets of 1000 trajectories each at 100 mM and 10 M, where complexes were recorded on the basis of minimized distance between the δ1 oxygen atom of D44 on PC and the ζ nitrogen on K187 of cyt f.
the manual docking work that complexes between PC and cyt f that featured strong electrostatic attraction (i.e., Complex 1) would not necessarily result in close copper-iron distance, and indeed the model of Complex 1 had a Cu-Fe distance of 31 Å; this would represent a strong fraction of Complex 1-type closest approaches achieving Cu-Fe distances of less than 25 Å, a positive result indeed.

The bottom line here is that regardless of the criterion used to determine successful trajectories, electrostatic interactions favor an approach between the two proteins that yields significantly smaller distances between metal centers than would be yielded through diffusion alone.

*Standard runs - characterizing closest complexes.* Having established that PC shows a preference for minimizing the distance between its copper center and the iron center in cyt f, we now used those standard runs where we searched for complexes based on metal center distance to determine a preferred site for PC/cyt f binding, if one existed. To do this, we used our "triplet contacts" utility to determine the three closest *charged* contacts for each successful complex, and classified complexes thus. Results of that classification is given in Table 3.1 along with the average Cu-Fe distances for each group of complexes given by a triplet. For all 9000 trajectories (9 runs of 1000 trajectories each) executed in these runs, the average Cu-Fe distance of all (2156) successes (reaction distance of 35 Å) was 25.6 Å. It again is immediately evident that the most populated of the triplet contact groups have a significantly lower average Cu-Fe distance than the overall average.
Table 3.1. Best seven triplet contacts between cyt f and PC and 100 mM ionic strength.

<table>
<thead>
<tr>
<th>Triplet (cyt f res - PC res)</th>
<th># successes</th>
<th>$P$ (#/9000)</th>
<th>Avg. Cu-Fe distance (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>K65 -E60 K187-E43 R209-E59</td>
<td>39</td>
<td>0.0043</td>
<td>18.0</td>
</tr>
<tr>
<td>K65 -E60 K187-E43 heme-E59</td>
<td>24</td>
<td>0.0027</td>
<td>17.2</td>
</tr>
<tr>
<td>K65 -E43 K187-D44 heme-E59</td>
<td>16</td>
<td>0.0018</td>
<td>18.2</td>
</tr>
<tr>
<td>K187-D44 R209-E43 heme-E59</td>
<td>12</td>
<td>0.0013</td>
<td>18.4</td>
</tr>
<tr>
<td>K65 -E60 K185-D44 K187-E43</td>
<td>11</td>
<td>0.0012</td>
<td>19.1</td>
</tr>
<tr>
<td>K156-E59 K187-E43 heme-E60</td>
<td>11</td>
<td>0.0012</td>
<td>18.7</td>
</tr>
<tr>
<td>K58 -D61 K65 -E60 K122-K68</td>
<td>10</td>
<td>0.0011</td>
<td>24.5</td>
</tr>
</tbody>
</table>

Figure 3.5. Alpha-carbon traces of 12 complexes between poplar PC and turnip cyt f. These 12 complexes were selected randomly from a single run of 1000 trajectories at pH 7 and 100 mM ionic strength from the four best contract triplets listed in Table 3.1. The complexes are superimposed so that the coordinates of cyt f in each complex are coincident with one another, showing the positioning of PC relative to cyt f in each of the 12 different trajectories. B is a 90° rotation of A. The figure was prepared using the software package Quanta (MSI, 1996).
Again, please note that hydrophobic interactions are not being considered here. The effect being observed is simply due to electrostatic interactions, and the conclusion that can be drawn from this data alone is that the effect of electrostatic interactions to draw PC into a potentially electron-transfer competent complex is far greater than first thought.

This idea was supported to a startling extent when we chose 12 complexes from the top four triplet contacts in the table and superimposed the complexes on one another graphically using the molecular modeling package Quanta (MSI, 1996). The resulting image is Figure 3.5. The striking observation that one makes from this picture is the similarity of all of the complexes, and the subsequent implication that electrostatic interactions steer PC toward a single dominant orientation with respect to cyt f; moreover, this orientation is very likely an electron-transfer competent one based on the minimized Cu-Fe distance.

Further selection of complexes from these same runs supported this idea. Moreover, it became increasingly apparent that the electrostatically-favored complex was not similar to Complex 1 from the manual docking studies, as originally predicted, but rather to Complex 2, a complex we had postulated could not form without hydrophobic interactions. A characteristic complex was prepared graphically (as seen in Figures 3.6-3.9) and characterized for purposes of comparison with the manually docked complexes; the close approach of H87 to Y1 (within 6 Å, just like Complexes 2 and 3), the close electrostatic interactions between both patches, especially D44 and K187, and potential interaction between hydrophobic residues (as presented in Figures 3.8 and 3.9)
Figure 3.6. A wide of a characteristic complex from Figure 3.5 between PC and cyt f, composed in GRASP with more atomic detail in order to highlight some of the proposed electrostatic contacts between the two proteins. Note the close proximity of H87(PC) to Y1(cyt f) (and, by connection, the heme of cyt f as well).
Figure 3.7. A 180° rotation of Figure 3.6.
Figure 3.8. A closer view of the characteristic complex of Figure 3.6, highlighting some of the proposed hydrophobic/van der Waals contacts between the two proteins.
Figure 3.9. A 180° rotation of Figure 3.8.
further confirmed these similarities.

Among these favored complex types, there were no complexes similar to Complex 1 detected for this standard run (or for many subsequent runs), nor were there any complexes where Y83 on PC made any reasonable approach to Y1 on cyt f. We were rapidly coming to the conclusion at this point that the favored complex between PC and cyt f was a Complex 2-type dock.

Other control experiments. In order to confirm the importance of specific electrostatic interactions to the interaction between PC and cyt f, two proteins were substituted for poplar PC and docked with cyt f under similar conditions as in the standard 100 mM run. In the first case, PC from *Anabaena variabilis* was substituted for poplar PC. As discussed in the previous chapter, *Anabaena* PC is characteristic of the cyanobacterial plastocyanins in that it has several positive residues where there are normally negatively charged residues in the higher plant and algal plastocyanins. Its net charge, as set up on MacroDox, is +1.13 as opposed to -7.97 for poplar PC. In the second, azurin (from *Alcaligenes denitrificans*) was substituted for poplar PC. Azurin retains the same type of copper center found in the blue copper proteins, but with substantially different folding motifs surrounding the copper center and the east-face negative charges absent.

The results of these runs are shown graphically in Figure 3.10. At distances of less than 35 Å, *Anabaena* PC showed 84 successes for every 1000 trajectories attempted; azurin showed 64 successes for every 1000 trajectories. At distances of less than 20 Å, successes were nonexistent for both. The average Cu-Fe distance for all
Figure 3.10. The distribution of distances between metal centers from the original set of BD experiments (given in Figure 3.2) compared to two control cases, interactions between *Anabaena* PC or *Alcaligenes* azurin and cyt f, both run at 100 mM and pH 7, demonstrating the unique nature of the electrostatic interaction between PC and cyt f to generate close contacts between the two proteins.
Anabaena PC successes was 29.7 Å; the corresponding average for azurin was 30.8 Å.

In all cases, the results for these control runs were even worse than those for poplar PC at 10 M, implying no specificity for interaction between these proteins and cyt f. This implies that electrostatic interactions between those proteins actually hinder their interaction with cyt f, instead of being the strong driving force as in the case of poplar PC.

Standard runs - defining the most reasonable reaction distance. Up to this point, we have been considering the cutoff metal-center distance for BD trajectory "successes" as 35 Å. For the 9000 total trajectories run for interactions between PC and cyt f, 2156 were successful at this standard, giving a fraction $p$ of 0.240±0.004. When this value is converted into a Smoluchowski rate constant $k$, the value returned is $(6.58 ± 0.08) \times 10^9$ M$^{-1}$ s$^{-1}$. This is clearly a severe overestimate of the experimentally observed rate constant on the order of $10^8$ M$^{-1}$ s$^{-1}$ (Gross et al., 1990; Qin and Kostic, 1992; Kannt et al., 1996).

Fortunately, because we have observed that electrostatic interactions guide PC into a complex where distance between metal centers is minimized, we are able to ask questions about limiting this cutoff distance between metal centers and modeling the rate constant more successfully by making a statement of how close the metal centers would have to come to provide better agreement with experimentally observed values.

To do this, we make a rather substantial assumption about the interaction between PC and cyt f: we assume that the only determining factor in whether or not PC
passes an electron to cyt f is the electrostatically-guided diffusion of PC towards cyt f. We assume that when the two proteins dock, that there will be no dissociation of that dock before electron transfer occurs. For the reaction scheme

$$PC_{\text{ox}} + \text{cyt f}_{\text{red}} \stackrel{k_i}{\rightarrow} [PC_{\text{ox}}\text{cyt f}_{\text{red}}] \stackrel{k_{et}}{\rightarrow} PC_{\text{red}} + \text{cyt f}_{\text{ox}}$$

we are assuming that $k_{et} >> k_i$ and $k_{et} >> k_{-1}$.

Of course, this is not strictly true. But we can very reasonably expect electrostatic interactions to have a more substantial effect on the diffusion and binding of PC to cyt f than on the reorganizations (such as nuclear motions) that will occur once the complex has formed and will thus influence $k_{et}$.

With that in mind, Table 3.2 was prepared listing several reaction distances and the values of $p$ and $k$ that those rate constants return. It is interesting to note that the smallest integer reaction distances that return values for $p$ and $k$ are 16 Å and 17 Å; those are also the two distances that most closely estimate the experimental bimolecular rate constant for PC/cyt f interaction. This leads to a very predictable proposal: only those complexes where PC and cyt f have their metal-center distance minimized as far as possible will undergo the electron transfer reaction. The statistics given indicate that this will happen 3 to 10 times every 1000 "attempts".

*Runs at different ionic strengths.* For ionic strengths other than 100 mM, four runs of 1000 trajectories each were executed, with results analyzed at reaction distances of 35
Table 3.2. Successes and rates as a function of distance between metal centers for 9 sets of 1000 MacroDox trajectories between PC and cyt f at a ionic strength of 100 mM and pH 7.

<table>
<thead>
<tr>
<th>Fe-Cu Distance (Å)</th>
<th>~Successes</th>
<th>P (#/Trajectories)</th>
<th>k (M⁻¹s⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td>35</td>
<td>2156</td>
<td>0.2536</td>
<td>6.58 x 10⁻³</td>
</tr>
<tr>
<td>20</td>
<td>406</td>
<td>0.0451</td>
<td>1.43 x 10⁻³</td>
</tr>
<tr>
<td>18</td>
<td>191</td>
<td>0.0212</td>
<td>6.84 x 10⁻³</td>
</tr>
<tr>
<td>17</td>
<td>100</td>
<td>0.0111</td>
<td>3.61 x 10⁻³</td>
</tr>
<tr>
<td>16</td>
<td>30</td>
<td>0.0033</td>
<td>1.09 x 10⁻³</td>
</tr>
<tr>
<td>15</td>
<td>0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>

Å, 17 Å, and 16 Å for each. The ionic strength was varied between 50 mM and 500 mM in these runs. (At the time of these studies, the computational power we had available to us was insufficient to allow for runs of ionic strength below 50 mM in reasonable amounts of CPU time. We have since analyzed runs at lower ionic strength; nothing significant was added to these results by them. The trends of behavior for decreasing ionic strength remain fundamentally the same below 50 mM.) These results were compared to those from the 10 M control set, where electrostatic interactions were eliminated. Average values for the Cu-Fe distances from all successful trajectories (at 35 Å) from each of these runs were calculated. These results are presented in Table 3.3 and Figures 3.11 and 3.12.

The results here are not entirely unexpected. There is a clear decrease in the rate constants that are returned at the 17 Å and 16 Å reaction distances as a function of increasing ionic strength, with a much stronger dependence for these reaction distances.
<table>
<thead>
<tr>
<th>Ionic Strength (mM)</th>
<th>Number of Successes</th>
<th>Number of Successes</th>
<th># Successes/Trajectories</th>
<th>k (M⁻¹s⁻¹)</th>
<th>Distance (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>69</td>
<td>5.57 x 10⁻⁷</td>
<td>23.34</td>
<td></td>
<td></td>
</tr>
<tr>
<td>60</td>
<td>51</td>
<td>4.14 x 10⁻⁷</td>
<td>23.74</td>
<td></td>
<td></td>
</tr>
<tr>
<td>70</td>
<td>31</td>
<td>2.52 x 10⁻⁷</td>
<td>24.34</td>
<td></td>
<td></td>
</tr>
<tr>
<td>80</td>
<td>21</td>
<td>1.96 x 10⁻⁷</td>
<td>24.64</td>
<td></td>
<td></td>
</tr>
<tr>
<td>90</td>
<td>16</td>
<td>1.31 x 10⁻⁷</td>
<td>25.27</td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>36*</td>
<td>1.09 x 10⁻⁷</td>
<td>25.45</td>
<td></td>
<td></td>
</tr>
<tr>
<td>110</td>
<td>12</td>
<td>9.81 x 10⁻⁸</td>
<td>25.45</td>
<td></td>
<td></td>
</tr>
<tr>
<td>125</td>
<td>9</td>
<td>7.36 x 10⁻⁸</td>
<td>26.14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>150</td>
<td>6</td>
<td>4.91 x 10⁻⁸</td>
<td>26.96</td>
<td></td>
<td></td>
</tr>
<tr>
<td>175</td>
<td>3</td>
<td>2.46 x 10⁻⁸</td>
<td>27.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>1</td>
<td>8.19 x 10⁻⁸</td>
<td>27.66</td>
<td></td>
<td></td>
</tr>
<tr>
<td>300</td>
<td>5</td>
<td>4.09 x 10⁻⁸</td>
<td>27.85</td>
<td></td>
<td></td>
</tr>
<tr>
<td>400</td>
<td>1</td>
<td>8.19 x 10⁻⁸</td>
<td>28.33</td>
<td></td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>1</td>
<td>8.19 x 10⁻⁸</td>
<td>28.33</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10000</td>
<td>0</td>
<td>0</td>
<td>28.93</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Number of successes based on 9000 attempted trajectories, as opposed to 4000 for all other ionic strengths.

Table 3.3. Successes and rates for MacroDox runs as a function of ionic strength and reaction distance (35 Å, 17 Å, or 16 Å).
Figure 3.11. Calculated rates, based on number of BD successes at reaction criteria of 16 Å and 17 Å, for the interaction between PC and cyt f as a function of ionic strength.
Figure 3.12. Average Fe-Cu distance for all BD successes at a reaction distance of 35 Å for several ionic strengths.
than for the 35 Å case. (We argued in the 1998 paper that this suggests a predominance for Brownian forces over electrostatic interactions for larger reaction distances.) In addition, with stronger electrostatic interactions the overall average of Cu-Fe distances decreases markedly with decreasing ionic strength, with the drop-off getting steeper as ionic strength gets lower.

A triplet contact analysis similar to that performed for the 100 mM standard case was repeated for the 50 mM, 300 mM, and 10 M cases in order to get a better view of the role ionic strength (or the strength of electrostatic interactions) plays in the interaction between the protein. Tables 3.4-3.6 describe the nature of the resulting complexes from each of these runs. When all four of these "triplet" tables (Tables 3.1 and 3.4-3.6) are studied in comparison with one another, the striking observation is that the more specific complexes show up with greater frequency at the lower ionic strengths. The two most frequently occurring types of complexes (by triplets) at 50 mM are the same as the two most frequently occurring types at 100 mM; at 300 mM, these are two of the top four, and at 10 M these complexes do not show up in significant number at all. In addition, when the average Cu-Fe distances for the single most common type of complex (by triplets) are compared for 50 mM, 100 mM, and 300 mM, the loss of specificity is observed even more clearly - the average Cu-Fe distance for the \((K65/K60), (K187-E43), (K209-E59)\) triplet is 16.8 Å at 50 mM, 18.0 Å at 100 mM, and 18.5 Å at 300 mM. Again, this type of complex does not show up at all at 10 M.

Again, complexes from each of the "more specific" set of triplets were selected quasi-randomly and superimposed on one another, and α-carbon traces of those
<table>
<thead>
<tr>
<th>Triplets (cyt f res - PC res)</th>
<th># successes</th>
<th>P (#/4000)</th>
<th>Avg. Cu-Fe distance (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>K65 -E60 K187-E43 R209-E59</td>
<td>50</td>
<td>0.0125</td>
<td>16.8</td>
</tr>
<tr>
<td>K65 -E60 K187-E43 heme-E59</td>
<td>25</td>
<td>0.0063</td>
<td>17.2</td>
</tr>
<tr>
<td>K65 -E60 R106-D9 K187-E43</td>
<td>19</td>
<td>0.0048</td>
<td>16.2</td>
</tr>
<tr>
<td>K58 -E60 K65 -E59 K187-D44</td>
<td>13</td>
<td>0.0033</td>
<td>17.2</td>
</tr>
<tr>
<td>K156-E59 K187-E43 heme-E59</td>
<td>13</td>
<td>0.0033</td>
<td>17.3</td>
</tr>
<tr>
<td>K187-D44 R209-E43 heme-E59</td>
<td>13</td>
<td>0.0033</td>
<td>16.9</td>
</tr>
<tr>
<td>K65 -E43 K187-D44 heme-E59</td>
<td>12</td>
<td>0.0030</td>
<td>17.2</td>
</tr>
</tbody>
</table>

Table 3.4. Best seven triplet contacts between cyt f and PC and 50 mM ionic strength.

<table>
<thead>
<tr>
<th>Triplets (cyt f res - PC res)</th>
<th># successes</th>
<th>P (#/4000)</th>
<th>Avg. Cu-Fe distance (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>K65 -E60 K187-E43 R209-E59</td>
<td>7</td>
<td>0.0018</td>
<td>18.5</td>
</tr>
<tr>
<td>R156-E59 K187-E43 heme-E60</td>
<td>5</td>
<td>0.0013</td>
<td>19.4</td>
</tr>
<tr>
<td>E10 -D61 K187-E43 D162-K66</td>
<td>3</td>
<td>0.0008</td>
<td>30.3</td>
</tr>
<tr>
<td>K65 -E60 K187-E43 heme-E50</td>
<td>3</td>
<td>0.0008</td>
<td>19.9</td>
</tr>
</tbody>
</table>

Table 3.5. Best seven triplet contacts between cyt f and PC and 300 mM ionic strength.

<table>
<thead>
<tr>
<th>Triplets (cyt f res - PC res)</th>
<th># successes</th>
<th>P (#/4000)</th>
<th>Avg. Cu-Fe distance (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>K10 -D9 R106-D8 D162-K66</td>
<td>4</td>
<td>0.0010</td>
<td>30.2</td>
</tr>
<tr>
<td>K58 -E68 K65 -D61 K122-K66</td>
<td>2</td>
<td>0.0005</td>
<td>34.1</td>
</tr>
<tr>
<td>R106-D9 D162-Cu K187-E69</td>
<td>2</td>
<td>0.0005</td>
<td>29.8</td>
</tr>
<tr>
<td>R106-K66 D162-D8 heme-D9</td>
<td>2</td>
<td>0.0005</td>
<td>24.5</td>
</tr>
</tbody>
</table>

Table 3.6. Best seven triplet contacts between cyt f and PC and 10 mM ionic strength.
superpositions are seen in Figures 3.13-3.15. Again, this provides the most striking analysis of the results, and the conclusions from these images are clear as they concern the decreasing specificity of the interaction between PC and cyt f as ionic strength increases.

These ionic strength studies strongly reinforce previous work on the interaction between the two proteins, both theoretical work from our lab (Durell et al., 1990; previous chapter/Pearson et al., 1996) and experimental observations from others (Takabe et al., 1986; Anderson et al., 1987). We saw no evidence in our work to answer the observations of Meyer et al. (1993) of a decrease in the PC/cyt f bimolecular rate constant with decreasing ionic strength, thereby reinforcing the idea that their observed effect is not a simple electrostatic effect but a more complex effect, involving hydrophobic interactions or issues with the off-rate of PC from cyt f after electron transfer takes place.
Figure 3.13. Alpha-carbon traces for 12 complexes between poplar PC and turnip cyt f at an ionic strength of 50 mM. Traces were prepared as Figure 3.5.
Figure 3.14. Alpha-carbon traces for 12 complexes between poplar PC and turnip cyt f at an ionic strength of 300 mM. Traces were prepared as Figure 3.5.
Figure 3.15. Alpha-carbon traces for 12 complexes between poplar PC and turnip cyt f at an ionic strength of 10 M. Traces were prepared as Figure 3.5.
While initial BD studies were carried out with poplar PC, most of the available experimental studies of the interaction between PC and cyt f were accomplished with spinach PC. One of the original studies of the role of the negative patch of residues on spinach PC was that of Anderson et al. (1987), who chemically modified the negative charges to positive (with ethylenediamine) to determine degrees of effect. The covalently cross-linked complex between PC and cyt f (Davis and Hough, 1983; Morand et al., 1989) was created using spinach PC. Modi et al. (1992) mutated several residues on both reactive patches of PC, including D42.

For the BD model of the interaction between PC and cyt f to be considered a successful model, the BD methods described in the previous chapter have to be tested with spinach plastocyanin and turnip cytochrome f, for comparison with the aforementioned chemical modification work (Anderson et al., 1987) and mutagenesis work (Kannt et al., 1996) that had been previously performed. Computational "mutagenesis" of PC residues can be compared with the results from the aforementioned studies, in addition to the results of the study on Silene PC acidic patch residues by Lee.
et al. (1995). *Silene* PC is identical in net charge but with some differences in positions of those charges outside of the negative patches - positions 26 (glutamate in spinach, lysine in *Silene*), 69 (threonine in spinach, glutamate in *Silene*), and 71 (lysine in spinach, serine in *Silene*).

The mutants and modifications studied in these papers, however, not only provide opportunity to compare BD results to previous research but also to perform unique experiments. Table 4.1 outlines the key negative patch residues and the mutations made to those residues in experimental work. Only three of the negative patch residues - D42, E43, and E59 - have been singly mutated, and only two of those - D42 and E43 - were mutated to neutral residues. In only one study (Lee et al., 1995) is a pair of equivalent mutants compared between the 42-45 acidic patch and the 59-61 acidic patch, and in that case the result Lee observes concerning the E59K/E60K *Silene* mutant - that a net charge change of +4 results in practically no change in the bimolecular rate constant when compared to the wild type - is dubious and totally opposed to results from Kannt concerning their E59K/E60Q mutant on spinach PC.

When the previous experimental work is considered as a whole, it seems that all groups agree that the 42-45 patch is important in the electrostatic interaction between the two proteins - one would expect, therefore, that results from BD modeling of 42-45 mutants would agree with this conclusion. The controversy is in the role of the 59-61 patch. Kannt (on the basis of one double mutant and one triple mutant) asserts that this patch is equal in importance to the 42-45 patch. Lee (on the basis of a single and double mutant) observes no importance for the 59-61 patch in the interaction (although
Wild type rate constants

<table>
<thead>
<tr>
<th>Study</th>
<th>Rate Constant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anderson et al. (1987)</td>
<td>$7.8 \times 10^7$ M$^{-1}$ s$^{-1}$</td>
</tr>
<tr>
<td>Modi et al. (1992)</td>
<td>$4.2 \times 10^7$ M$^{-1}$ s$^{-1}$</td>
</tr>
<tr>
<td>Lee et al. (1995)</td>
<td>$9.8 \times 10^7$ M$^{-1}$ s$^{-1}$</td>
</tr>
<tr>
<td>Kannt et al. (1996)</td>
<td>$1.8 \times 10^8$ M$^{-1}$ s$^{-1}$</td>
</tr>
</tbody>
</table>

D42
- Modi - mutated to N, no significant effect on rate constant
- Lee - mutated to N, rate constant 37% of wild type
- Kannt - mutated to N, rate constant 33%-41% of wild type
- Lee - double mutant D42N/E43K, rate constant 10% of wild type

E43
- Lee - mutated to K, rate constant 13% of wild type
- Kannt - mutated to K, rate constant 12%-16% of wild type
- Kannt - mutated to N, rate constant 22%-30% of wild type
- Lee - double mutant D42N/E43K, rate constant 10% of wild type
- Lee - double mutant E43K/D44K, rate constant 5% of wild type
- Kannt - double mutant E43Q/D44N, rate constant 13%-15% of wild type
- Kannt - triple mutant E43N/E59K/E60Q, rate constant 3% of wild type

D44
- not singly mutated
- Lee - double mutant E43K/D44K, rate constant 5% of wild type
- Kannt - double mutant E43Q/D44N, rate constant 13%-15% of wild type

E45
- not mutated

D51
- not mutated or singly modified
- Anderson - EDA modification of D51 and E68, rate constant 18% of wild type

E59
- Lee - mutated to K, no significant effect on rate constant
- Lee - double mutant E59K/E60K, rate constant >90% of wild type
- Kannt - double mutant E59K/E60Q, rate constant 7%-8% of wild type
- Kannt - triple mutant E43N/E59K/E60Q, rate constant 3% of wild type

E60
- not singly mutated
- Lee - double mutant E59K/E60K, rate constant >90% of wild type
- Kannt - double mutant E59K/E60Q, rate constant 7%-8% of wild type
- Kannt - triple mutant E43N/E59K/E60Q, rate constant 3% of wild type

D61
- not mutated

E59-61
- Anderson - EDA modification in this patch, rate constant 31% of wild type

Table 4.1. Negative-patch residues on PC and mutations formed.
crystallography observations from this group (Sugawara et al., 1999) concerning structural differences between Silene PC and spinach PC may be relevant). These conclusions fall on either side of the observations of Anderson, who found that modifying a single residue in the 59-61 region had slightly (but significantly) less effect than an equivalent single modification among positions 42-45, although because the position of the chemical modification was determined by peptide digestion it was not possible to pinpoint exact effects for individual residues. A criticism of the undertaking of further mutagenesis studies on the acidic patch of PC is that enough studies have already been completed to settle the question of the acidic patches' role in the interaction; on this evidence, the question hardly appears settled.

Therefore, a study of spinach mutants - including the first systematic study of individual acidic patch residues and their impact on the electrostatic interaction between spinach PC and turnip cyt f - would not only verify the method's power on comparison with those experimental results on which there is general agreement, but also predict results for those cases in which there is not agreement.

METHODS

Structures. The turnip cytochrome f structure (PDB ID: 1HCZ; Martinez et al., 1996) is the same as before.

The spinach PC structure used was the first one available from the Protein Data...
Bank (http://www.rcsb.org/pdb/ - Bernstein et al, 1977) - a structure from the PC/cyt f complex ensemble of Ubbink et al. (1998; ID: 2PCF). Comparison of this structure's performance against that of the spinach PC crystal structure yielded questions about the behavior of plastocyanin as a function of conformational changes, as will be addressed shortly.

"Mutation" of residues. MacroDox was used to "mutate" spinach PC residues. An acidic residue was replaced with its corresponding nonpolar residue (aspartate replaced with asparagine; glutamate replaced with asparagine) in as identical a conformation as possible, without minimization after the replacement. This was accomplished so that the only alteration in the protein was the change in charge: considering that the most we can hope to accomplish in these studies is to demonstrate the electrostatic affects of these mutants on PC behavior, any additional adjustments would have only served to confuse matters.

Brownian dynamics computation. BD simulations were executed exactly as described in the previous chapter.
RESULTS AND DISCUSSION

*Standard run and ionic strength comparisons.* As before, standard runs of 9 sets of 1000 trajectories were executed to set benchmarks to compare with ionic strength variants and mutants. Using the same standards as used in the poplar PC/turnip cyt f studies, an average probability of complex formation for distances less than 21 Å was calculated as 0.046 \(\pm 0.006\) for 100 mM ionic strength, against 0.007 \(\pm 0.002\) for 10 M ionic strength. For distances less than 17 Å, the corresponding probabilities were 0.017 \(\pm 0.002\) for 100 mM and 0.0003 \(\pm 0.0003\) for 10 M. A comparison of complex distributions for spinach PC (as given in Figure 4.1) with that of poplar PC (as given in Figure 3.10) demonstrate more clearly spinach PC’s tendency (or, more correctly, this spinach PC structure’s tendency) to form closer, tighter complexes with turnip cyt f as opposed to poplar PC.

As we began these studies, we gained access to new computational hardware that allowed us to execute BD runs much more quickly (100 mM runs that had been taking on the order of an hour to complete now took us on the order of minutes). Because we had been limited in exploring low ionic strength effects on interactions between PC and cyt f as a function of computational power, we decided to take this opportunity to explore ionic strength effects for the interaction between spinach PC and turnip cyt f further. Figures 4.2 and 4.3 demonstrate the clear monotonic dependence of the electrostatic interaction between PC and cyt f on ionic strength, even for the lowest of ionic strengths. This is contrary to the observations of Meyer et al. (1993) who
Figure 4.1. The distribution of distances between metal centers from BD experiments on the interaction between spinach PC and turnip cyt f, presented in identical fashion to data on the interaction between poplar PC and turnip cyt f in Figure 3.10 and compared to the same *Anabaena* PC or *Alcaligenes* azurin control cases.
Figure 4.2. The distribution of successful trajectories for a number of different ionic strengths, 100 mM and below, as a function of Cu-Fe distance.
Figure 4.3. The number of trajectories falling within a particular distribution level for a number of different distances (e.g., number at 16 Å represents number of successful trajectories with closest approach between 15 and 16 Å Cu-Fe distance) as a function of ionic strength.
demonstrated a unique decrease in the rate constant of PC/cyt f interaction and electron transfer with decreasing ionic strength for ionic strengths of about 30 mM and below.

Meyer postulated on observation of this "anomalous behavior" that the poor electron transfer behavior at low ionic strengths was because the strong electrostatic interactions that result would prohibit rearrangement (possibly involving hydrophobic interactions) that would allow PC to form a close, electron-transfer competent complex with cyt f. However, these BD experiments (where only electrostatic interactions and diffusion are important) would support the idea that the lower the ionic strength is (and by extension, the stronger the electrostatic interactions are), the more likely PC is to enter a complex where metal-center strength is minimized and electron-transfer probability is maximized. If electrostatic interactions are involved at all in Meyer's effect, then the involvement is much more complex and subtle. (Again, remember that MacroDox treats proteins as rigid bodies and does not allow their structures to optimize given the environment of the resulting complex; any legitimate consideration of rearrangement needs to involve the conformational changes that accompany real complex formation.) More likely is a more complex series of effects, including changes in the electrostatic interactions between proteins post-electron transfer that might make electron transfer less energetically favorable.

Also interesting when the low ionic strength dependence on association probability is observed is the strong effect of ionic strength on forming close complexes (distribution points at 15 Å and 16 Å) and the nonexistent effect of ionic strength on formation of complexes with more moderate Cu-Fe distances (19 Å and above). The
plot in Figure 4.3 deals with ionic strengths up to 100 mM - similar trends can be observed for higher ionic strengths as well. This implies that electrostatic forces, in the PC/cyt f interaction, do not merely serve a role in the long-range attraction between the proteins - the high number of complexes at close distances in the low ionic strength cases imply that electrostatics serve to fine-tune the interaction between the molecules.

As in the poplar PC work in the previous chapter, a collection of the most frequently occurring complexes was made and superimposed. This was done for the standard 100 mM case and also for the 10 mM case, where electrostatic interactions should be at their strongest. Results of this superposition are given in Figures 4.4 and 4.5. As before in the experiments with poplar PC, there is a single dominant orientation that spinach PC prefers in its interaction with turnip cyt f. In both the 100 mM and 10 mM cases, all complexes present prefer this same orientation (with a somewhat startling coincidence, given the difference in strength of electrostatic interactions), and as in the poplar PC work the orientation tends towards the Complex 2 hypothesis proposed in Chapter 2.

Mutations of spinach PC. Mutants of spinach PC were made with two goals in mind - one, to compare BD results to the major experimental results concerning the importance of the negative patch (Anderson et al., 1987; Lee et al., 1995; Kannt et al., 1996); two, to systematically compare the importance of the negative charges to arrive at as a complete a conclusion as possible about each residue's role in the electrostatic interaction between the two proteins.
Figure 4.4. Alpha-carbon trace of 10 complexes between spinach PC and turnip cyt f generated by MacroDox for a single run of 1000 trajectories at pH 7 and 100 mM ionic strength, selected randomly from all complexes with Cu-Fe distance less than 16 Å. The figure was prepared as in Figure 3.5.
Figure 4.5. Alpha-carbon trace of 10 complexes between spinach PC and turnip cyt f generated by MacroDox for a single run of 1000 trajectories at pH 7 and 10 mM ionic strength, selected randomly from all complexes with Cu-Fe distance less than 16 Å. The figure was prepared as in Figure 3.5.
Each negative charge on the east face - D42, E43, D44, E45, D51, E59, E60, and D61 - was replaced by its corresponding polar residue (aspartate was replaced with asparagine, glutamate with glutamine). E25 was also mutated in the same manner, to serve as a control.

The results of these mutant studies are given in Table 4.2 for 10 mM and 100 mM ionic strength. At 100 mM ionic strength, the control mutant (E25Q) had no significant change in reaction probability from the wild type protein, as one might expect for a negatively-charged residue far isolated from the negative patch. The eight negative charges on the negative patch had varying magnitudes of affect, ranging from 21% to 63% inhibition. D61N (reaction probability at 17 Å at 79% of wild type) and E45Q (70% of wild type) were the least affected of the east face mutants. The most affected (37% of wild type), surprisingly to us and apparently contrary to results from Kannt et al. (1996), was the D42N mutant. Given the relative lack of involvement (as compared to E43 and D44) of D42 in complexes between PC and cyt f, we have no good explanation for this result.

The other mutants cut the probability of complex formation relative to the wild type roughly in half. Overall, while the difference between the 42-45 negative patch and the 59-61 negative patch is on the very edge of significance, the 42-45 mutants do have a slightly greater effect. When electrostatic effects are amplified by running BD simulations at 10 mM, the difference between the two patches is amplified as well.

D51, an acidic residue on the east face whose role in the interaction between PC and cyt f has been studied only sparingly to date, also appears to play a significant role.
<table>
<thead>
<tr>
<th>Mut</th>
<th>successes /1000 traj</th>
<th>% of 10 m M</th>
<th>successes /1000 traj</th>
<th>% of 100 m M</th>
<th>successes /1000 traj</th>
<th>% of 1000 m M</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>97.0±2.6</td>
<td>100</td>
<td>161.8±4.2</td>
<td>100</td>
<td>17.4±1.0</td>
<td>100</td>
</tr>
<tr>
<td>E25Q</td>
<td>76.0±1.1</td>
<td>87</td>
<td>152.0±2.3</td>
<td>94</td>
<td>16.8±1.3</td>
<td>96</td>
</tr>
<tr>
<td>D42N</td>
<td>25.7±1.3</td>
<td>32</td>
<td>63.8±2.2</td>
<td>39</td>
<td>6.6±0.7</td>
<td>37</td>
</tr>
<tr>
<td>E43N</td>
<td>35.0±3.1</td>
<td>40</td>
<td>94.0±5.8</td>
<td>58</td>
<td>8.1±0.8</td>
<td>47</td>
</tr>
<tr>
<td>D44N</td>
<td>29.8±5.8</td>
<td>34</td>
<td>92.5±3.5</td>
<td>57</td>
<td>7.8±0.7</td>
<td>45</td>
</tr>
<tr>
<td>E45Q</td>
<td>55.0±1.6</td>
<td>63</td>
<td>114.5±0.8</td>
<td>71</td>
<td>12.2±0.7</td>
<td>70</td>
</tr>
<tr>
<td>D51N</td>
<td>39.8±4.1</td>
<td>46</td>
<td>110.5±5.0</td>
<td>68</td>
<td>10.1±1.0</td>
<td>58</td>
</tr>
<tr>
<td>E59Q</td>
<td>55.5±3.4</td>
<td>64</td>
<td>116.5±2.7</td>
<td>72</td>
<td>8.1±0.7</td>
<td>47</td>
</tr>
<tr>
<td>E60Q</td>
<td>51.0±1.5</td>
<td>59</td>
<td>104.8±3.8</td>
<td>65</td>
<td>8.9±0.7</td>
<td>51</td>
</tr>
<tr>
<td>D61N</td>
<td>56.5±2.9</td>
<td>65</td>
<td>122.0±2.4</td>
<td>75</td>
<td>13.7±0.7</td>
<td>79</td>
</tr>
<tr>
<td>E43K</td>
<td>10.8±0.8</td>
<td>12</td>
<td>40.8±3.1</td>
<td>26</td>
<td>6.6±0.5</td>
<td>38</td>
</tr>
<tr>
<td>D44N/E59Q</td>
<td>11.8±2.1</td>
<td>14</td>
<td>54.2±1.2</td>
<td>34</td>
<td>4.2±0.4</td>
<td>24</td>
</tr>
<tr>
<td>E43N/E59Q</td>
<td>1.0±0.3</td>
<td>1</td>
<td>51.5±3.6</td>
<td>32</td>
<td>3.8±0.5</td>
<td>22</td>
</tr>
<tr>
<td>E59Q/E60Q</td>
<td>9.5±2.1</td>
<td>11</td>
<td>46.7±3.8</td>
<td>30</td>
<td>5.7±0.3</td>
<td>33</td>
</tr>
<tr>
<td>D42N/E43K</td>
<td>3.3±0.9</td>
<td>3</td>
<td>25.0±1.5</td>
<td>15</td>
<td>1.2±0.4</td>
<td>7</td>
</tr>
<tr>
<td>Triple</td>
<td>0</td>
<td>0</td>
<td>3.2±0.6</td>
<td>2</td>
<td>0.7±0.2</td>
<td>4</td>
</tr>
</tbody>
</table>

Table 4.2. Comparison of electrostatic interaction as measured through BD simulations between mutant spinach plastocyanins and turnip cyt f. The "triple" mutant is E43N/E59K/E60Q, as in Kannt et al. (1996).
in the interaction. At an inhibition of 42% as compared to the wild type, D51 is less important than most of the main acidic patch residues but more important than both E45 and D61.

When a 43-44 double mutant is compared with a 59-60 double mutant (at 100 mM), this conclusion is further confirmed. Here, there is a significant difference at 100 mM between the two mutants, and the E43Q/D44N mutant clearly has the greater effect - 22% of wild type vs. 33% of wild type for E59Q/E60Q. Observing the behavior of these double mutants at 10 mM yield mixed results - at the 17 Å reaction distance, the double mutants behave effectively the same, but at a 15 Å reaction distance, the difference between the double mutants is tenfold, with the E43Q/D44N mutant having the far greater impact; this might imply, in light of the role electrostatic interactions have in promoting formation of the closest, most specific contacts, that the 42-45 patch has a more important role to play in tight complex formation than the 59-61 patch.

There are then two key take-home-lessons from these mutants:

- The effect of the spinach PC mutants as observed in BD simulations is comparable to the effect of spinach PC mutants observed in experimental situations. A plot of success probability as a function of difference in mutant charge (Figure 4.6) takes a similar shape to plots from Kannt et al. (1996) and Lee et al. (1995). When ΔΔG values for mutants are calculated as in Kannt et al. (1996; $ΔΔG = RT \ln \left( \frac{k_{WT}}{k_{mutant}} \right)$), values are returned that are significantly but consistently lower than those Kannt measured. For example, the experimentally-

110
Figure 4.6. BD success probability for the interaction of spinach PC mutants with cyt f, as plotted against net charge change of that mutant.
obtained $\Delta G$ for the E43N mutant was 2.2 kJ/mol at pH 6.0 and 2.8 kJ/mol at pH 7.5; BD simulations return a value of 1.9 kJ/mol for that mutant.

There is a discernable and significant trend for mutants of the 42-45 negative patch to have a greater effect than mutants of the 59-61 negative patch in BD simulations, in agreement with the chemical modification studies of Anderson et al. and contrary to the conclusions of Kannt et al. (1996). The difference, however, is not to the extent of that observed by Lee et al. (1995) in their Silene mutants, where a (lysine!) mutant to the 59-61 patch was observed to have no discernable affect on the interaction between the proteins.

*The effect of structure on complex formation between PC and cyt f.* Shortly after we began experiments using the Ubbink et al. (1997) structure for spinach PC, a crystal structure for a mutant spinach PC was released to the Protein Data Bank by Xue et al. (1998). The mutant, G8D, was put in place to make spinach PC more poplar-like for crystallization purposes. We back-mutated this structure (as with our previous mutations on the Ubbink structure) and used this for BD experiments as a check. Surprisingly, the activity for this structure was significantly less than that for the NMR structure; at the 17 Å reaction distance, the success probability was 0.005 ± 0.001 for the crystal as opposed to 0.017 ± 0.002 for the Ubbink structure. At a 16 Å reaction distance, the difference was even more stark: 0.0010 ± 0.0005 for the crystal structure as opposed to 0.011 ± 0.001 for the Ubbink structure, or an eleven-fold difference.
This was obviously disturbing and intriguing to us. One would hope that differing positions of the same titrable residues on proteins with the same amino acid sequence would not affect affects from BD experiments so drastically. We sought out reasons for the differences.

In Ubbink's (1997) description of the NMR complex ensembles their research generated, one point immediately stood out: the rigid-body molecular dynamics docking of the structures, based on constraints drawn from their NMR experiments and the expected charge interactions, followed by energy minimization of the complex. The final step begs a question: are there conformational changes induced by complex formation that make PC more favorable electrostatically for interaction with cyt f? And, moreover, can one pinpoint a specific conformational change in PC as being key for complex formation?

To study this further, we contacted the provider of the NMR solution of spinach PC used in Ubbink's studies, A. Bergkvist, to obtain the structure. Bergkvist (personal communication) sent us first an ensemble of 11, and then later an ensemble of 100 NMR solutions to the structure of spinach PC. The initial ensemble of 11 structures is displayed as superimposed wire-frame structures in Figure 4.7.

In searching these structures for differences that might result in changes in the electrostatic interaction or steric hindrance between the two proteins, three variations stand out: deviations in the loop between positions 7 and 11 that result in changes in the orientation in D9, differences in the orientation of the backbone around K77, and (most strikingly) significant differences in orientation of the loop surrounding glutamates 59
Figure 4.7. Original 11 spinach NMR structures from A. Bergkvist (personal communication), superimposed on one another. Key regions of deviation within the ensemble of structures are labeled.
and 60, causing up to a 90° deviation in the position of those residues on the east face on top of changes in the orientation of the individual side chains. There are also significant differences in the position of E59 and E60 between the Xue crystal structure and the Ubbink energy-minimized structure.

We performed BD studies with the goal of connecting a specific structural fluctuation to the reactivity of a given PC structure. In so doing, we aimed to exploit one of the primary weaknesses of MacroDox - the fact that the Brownian dynamics simulations employed are rigid-body simulations. In real experiments, it is not possible to directly probe the impact of a single macromolecular conformation on the interaction between a pair of molecules because of the dynamic nature of macromolecules. MacroDox does not simulate the internal motions of the proteins involved in the interaction, and therefore can be used (in principle) to simulate the interaction of several different conformations of PC with cyt f.

BD simulations using the different structures yielded variable results. At the 17 Å reaction distance, reaction probabilities between 0.004 and 0.021 were observed. For 16 Å, the range was between 0.002 and 0.014. However, relationships between different structural qualities of PC and the reactivity of the given PC showed no hints of correlation. Specific structural qualities probed included absolute position of fluctuating residues, distances between charged residues on the east face, and backbone properties (phi and psi angles) in relevant loops. None showed any direct relationship to the reactivity of a given PC structure.

Studies of the relationship between PC structure and conformational changes
and PC reactivity are still at this time incomplete. Investigation separate from this work (E.L. Gross, personal communication) is focusing on the differing dipole moment of the different structures, resulting from differing spatial distribution of charges with changes in PC structure, and the dipole-dipole interactions between different PC structures and cyt f. Observations of Sugawara et al. (1999) concerning the newly-solved Silene PC structure concerning significant structural differences between the Silene and spinach PC structures in the main chains of residues 59 and 60, in addition to the previously discussed differences in behavior of mutants of Silene and spinach PC, give weight to the idea that the orientation of residues 59 and 60 is important in the interaction between PC and cyt f. Further work and better experimental design is necessary for this problem.
CHAPTER 5

THE PC/CYT F INTERACTION VS. THE C6/CYT F INTERACTION: THE

CHLAMYDOMONAS SYSTEM

One of the long-standing problems in studying the interaction of mobile electron carriers with cytochrome f has been the lack of molecular structures available from a single system. Until recently, the only structure determination for cytochrome f has been from turnip, and no turnip PC sequence is available for comparison. The recent release of the Chlamydomonas reinhardtii cyt f structure (E. A. Berry, manuscript in preparation) allows us for the first time to study the interaction of cyt f with PC and cytochrome c₆ structures available from the same organism.

Comparisons between PC and c₆ have been previously reported. Gross et al. (1995) first analyzed the similarities between the electrostatic fields of different PCs and cyt c₆ from Monoraphidium braunii, as obtained from the work of Frazao et al. (1995). Frazao's work also contained some discussion of the likely electron transfer pathways in the two proteins. Ullmann et al. (1997) have recently more rigorously analyzed the electrostatic properties of PC and cyt c₆ and identified homologous residues between the two proteins. In all cases, the fundamental conclusion has been
the same. Both proteins have significant negatively-charged patches in corresponding positions on the surface of the two proteins. *Chlamydomonas* PC's negative charges are D42, E43, D44, D53, D59, D61, and E85. The loop between residues 42 and 44 on PC corresponds to glutamates 69-71 on c6, D53 on PC corresponds to E47 on c6, D59 and D61 on PC to D41 on c6, and E85 to E54 on c6. (Ullmann et al., 1997.) A related study on the *Synechocystis* cyt c6 and comparison to PC (de la Cerda et al., 1999) implies a similar relationship between those two proteins as well, indicating that the similar electrostatics isn't simply a feature of the *Chlamydomonas* system.

In addition, concurrently to our studies on PC and cyt f, Soriano et al. pursued the role of electrostatic interactions in the electron transfer reaction between *Chlamydomonas* PC and cyt f through more traditional means. Flash photolysis experiments *in vivo* (Soriano et al., 1996) and stopped-flow kinetic experiments *in vitro* (Soriano et al., 1998) were used to describe the behavior of PC and cyt f in the wild type and when cyt f was mutated to remove most key charges.

Soriano's 1996 experiments took strains of *Chlamydomonas* algae expressing mutant forms of cyt f and analyzed electron transfer between the mutants and PC and c6 within the *Chlamydomonas* algae through measuring absorbance at the alpha-band of cyt f, at 554 nm following flash photolysis. For three mutant strains - a double mutant neutralizing charges on K188 (to N) and K189 (to Q), a triple mutant K58Q-K65S-K66E changing the net charge by -4, and a quintuple mutant combining the first two, the detectable impact on electron transfer activity turned out to be almost completely negligible. In the subsequent work (Soriano et al., 1998), a very slight decrease in
oxidation half-times was observed for a mutant strain of *Chlamydomonas* with a K58E-K65E-K66E mutant cyt f. Similar results were observed both in the presence and absence of copper, equating to reactions with PC and c₅₇ respectively. These are stunning observations considering other results (Anderson et al., 1987; Lee et al., 1995; Kannt et al., 1996) predicting a significant impact of the negative charges on PC on the interaction between the proteins, and the established links between the positive patch on cyt f and the negative patch on PC (esp. Morand et al. 1989).

The study that ensued (Soriano et al., 1998) dealt with the behavior of PC and cyt f, and corresponding mutants of cyt f, in solution. Stopped-flow spectrophotometry was used to obtain kinetic data, with absorbance changes measured at the Soret band of cyt f, 420 nm. Pseudo-first order kinetics obtained by reacting cyt f with an excess of PC yielded the expect electrostatic behavior, not the lack of effect the *in vivo* studies yielded. The double mutant’s rate constant at 100 mM was $3 \times 10^7 \text{ M}^{-1}\text{s}^{-1}$, and the triple mutant’s rate constant at 100 mM was $10^7 \text{ M}^{-1}\text{s}^{-1}$, as opposed to the classical $10^8 \text{ M}^{-1}\text{s}^{-1}$ for the wild type at 100 mM. In addition, both the wild type and the double mutant experienced the monotonic drop in rate constant with increasing ionic strength that is to be expected when the interaction between two molecules is electrostatic in nature. (The triple mutant showed no such behavior - its rate constant with PC stayed around $10^7 \text{ M}^{-1}\text{s}^{-1}$ independent of its ionic strength.)

Because the nature of the Brownian dynamics simulation, in that only one mobile ligand and one receptor are modeled and many successive trajectories are used to estimate a rate constant, and the volume in which the mobile molecule can diffuse is
spherical and large with respect to the extent of electrostatic interactions (Northrup et al., 1987), it is expected that BD simulations would be similar to the *in vitro* studies. However, the rather stunning discrepancy between Soriano’s *in vivo* and *in vitro* studies demanded the experiments be performed as part of a complete analysis of the *Chlamydomonas* system.

Thus, our goals here are threefold: to determine if the interaction between PC and cyt f from the same organism (*Chlamydomonas*) resembles that between PC and cyt f from different organisms (poplar and spinach PC vs. turnip cyt f), to determine whether BD-modeled PC/cyt f interactions are similar to the *in vivo* case or the *in vitro* case as in Soriano et al. (1996, 1998), and to determine similarities and differences between the PC/cyt f interaction and the c^6/cyt f interaction.

**METHODS**

*Structures.* The structures for all of the *Chlamydomonas* proteins were taken from the Protein Data Bank (http://www.rcsb.org/pdb/ - Bernstein et al., 1977), with the exception of a dimer structure of cyt f (E. A. Berry, personal communication). *Chlamydomonas* PC is structure 2PLT; the c^6 used is lCY1 (Kerfield et al., 1995), and the cyt f used is the second cyt f record within 1CFM (Berry et al., manuscript in preparation). Structures for mutant plastocyanins and cyt f's were generated using MacroDox, with all mutant structures kept identical to their wild types with the
exception of the mutant (i.e., no energy minimization was performed on the mutants, so that only affects of removing a given charged residue were tested).

Electrostatic computation. Electrostatic analysis was performed using GRASP as before. For comparison with *Chlamydomonas* c₆, structures were obtained for *Monoraphidium braunii* c₆ (PDB ID 1CTJ; Frazao et al., 1995) and *Synechococcus elongatus* c₆ (PDB ID 1C6S; Beissinger et al., 1998); those structures were analyzed with GRASP as well.

Brownian dynamics computation. Brownian dynamics runs between PC or c₆ and cyt f were performed as in the original poplar PC/tumip cyt f experiments. Obviously, adjustments to the method were necessary for c₆ because of the heme structure. For simplicity's sake, we initially chose to retain the method of evaluating trajectory success as a function of metal center distances, as opposed to the more complex evaluation for heme proteins programmed into MacroDox. We did test these routines, however, and compared all results against results obtained from using center-of-mass distances to determine success as part of determining the importance of electrostatic interactions in the interaction between c₆ and cyt f.
RESULTS

*Cytochrome c*\(_6\) *sequence, structure, and electrostatics.* Because this is our first serious discussion of *c*\(_6\), it is necessary to similarly discuss the available data on sequences and structure, and our own electrostatic analysis. *c*\(_6\) sequences are found in Table 5.1.

$cyt\ c_6$ is found only in the algae and cyanobacteria. Perhaps because of this, the protein is considerably more diverse than plastocyanin - only 15 residues are conserved in the (on average) 89-amino acid protein, with only 17 sequences of the protein known. None of those residues are electrostatically important: three are glycines, one is a proline, and three more are part of the CXXCH pocket characteristic of heme proteins. The remaining eight residues are F10, N13, L31, Q52, M60, L67, 172, and W88 - two polar, six hydrophobic.

The most significant electrostatic trend in the available sequences is the presence of at least two (usually more) acidic residues in positions 69, 70, 71, 73, and 74. Other acidic residues are present in widely varying positions in most *c*\(_6\)'s. Again, these acidic amino acids are in no way conserved in any of the *c*\(_6\)'s the way that they are in the higher plant (Type III and IV) plastocyanins, because of the diversity of algae and cyanobacteria from which they come. On the other hand, 17 additional residues are conserved in the green algal *c*\(_6\)'s between positions 42 and 60.

Structures of the three different cytochrome *c*\(_6\)'s for which structures are available (*Chlamydomonas, Monoraphidium, Synechococcus*) and electrostatic analysis of those structures emphasize the diversity in *c*\(_6\)'s from different organisms. In the two
<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>c6chlamy</td>
<td>-ADLALGAEVNGCAACHMGGRNSVPEKTLKDAALEQY--LDGFVE|</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c6bryopsis</td>
<td>GDLIEGADVFTGCAACHAGGANSVPEKTLKDAALEQY--LDGGLSIE|</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c6monora</td>
<td>EADLALGAEVNGCAACHAGGANSVPEKTLKDAALEQY--LDGGLSIE|</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c6euglena-grac</td>
<td>-----GGADVFADGCTCHVNGGTVSACVGLSKTAIEY--LDGGYTK</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c6euglena-viri</td>
<td>-----SGAEVNGCSSCHVNGGTVSACVGLSKTAIEY--LDGGYTK</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c6bumilleri</td>
<td>ADIENGERTPTANACACHAGGNVIMPEKTIKKALEANEEADG|</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c6porphyra</td>
<td>ADLDNGKGFANACACHAGGNVMPDKTLKVDLEYEADG|</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c6alaria</td>
<td>IDINGERTPTANACACHAGGNVIMPEKTIKKALEANEEADG|</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c6petalonia</td>
<td>VDINGERTPTANACACHAGGNVIMPEKTIKKALEANEEADG|</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| c6aphani| AQTVNGKGAMPAGAELRSEEEEIQAVAYFEEQATRAWKKY| -complete conservation-
| c6microcystis| TVQVNGKGAMPAGAELRSEEEEIQAVAYFEEQATRAWKKY| -complete conservation-
| c6anabaena| AQTVNGKGAMPAGAELRSEEEEIQAVAYFEEQATRAWKKY| -complete conservation-

Table 5.1. The 17 currently available cytochrome c₆ sequences, aligned.
Chlamydomonas and Monoraphidium; Figures 5.1 and 5.2) there is similarity in the positioning of the key negative charges on the "east face" of the proteins, and often identical residues or residue types in key positions: acidic residues in positions 69-71, conserved glutamates in positions 47 and 54, and a similarly-positioned acidic residue around positions 2 and 65. However, Monoraphidium c₆ has an additional negative charge at the N-terminal end of the protein, an additional negative charge at position 68, and is missing Chlamydomonas' lysine at position 45. These changes give Monoraphidium c₆ a net charge of around -7, as opposed to Chlamydomonas' net charge of approximately -3. These differences are observed electrostatically by comparing the zero-potential lines in the c₆ structures - the extent of Chlamydomonas c₆'s positive potential is clearly greater than that of Monoraphidium c₆. Variability of this sort is a step beyond the slight differences in charge among the algal plastocyanins.

Synechococcus, as would be expected for a cyanobacterial c₆, has a significantly different charge distribution. The region of negative potential on the east face is confined to the lower portion of the protein (as represented in Figure 5.3) - negative charges in positions 2, 47, and 69-71 are conserved, but acidic residues in positions 54 and 65 are no longer present, and with the addition of a lysine in position 8, the positive potential field around the heme extends well into the east face, into space otherwise reserved for negative charges in algal c₆'s.

Brownian dynamics of Chlamydomonas PC and cyt f. The methods of Chapter 3 were
Figure 5.1. Electrostatic potential representation of oxidized *Chlamydomonas* cytochrome c₆. Electrostatic potential was calculated through solution of the linearized Poisson-Boltzmann equation, with the program GRASP. The ionic strength used in the calculation is 100 mM. 3-dimensional contour represents -1 kT/e; 2-dimensional contour represents zero potential. Important residues shaded and labeled in upper left.
Figure 5.2. Electrostatic potential representation of oxidized Monoraphidium cytochrome c. Residue and electrostatic potential contour coloring as in Figure 5.1.
Figure 5.3. Electrostatic potential representation of oxidized *Synechococcus* cytochrome c₆. Residue and electrostatic potential contour coloring as in Figure 5.1.
repeated to the end of confirming the similar nature of the interaction between the interaction between PC and cyt f in *Chlamydomonas*. Upon completion of nine runs of 1000 trajectories each for cases of 100 mM (physiological) ionic strength and 10 M (electrostatic-interaction eliminating) ionic strength, a distribution of closest contacts for all successful trajectories was plotted, as a function of Cu-Fe distance, exactly as before. If not as clearly, Figure 5.4 again shows the strong preference for complexes with small Cu-Fe distance to form in the presence of electrostatic interactions. The probability of complex formation at distances less than 21 Å is 0.041 ± 0.004 in the 100 mM case, as opposed to 0.008 ± 0.001 in the 10 M case. At distances less than 17 Å, the probability is 0.010 ± 0.002 at 100 mM; at 10 M, no complexes with a Cu-Fe distance of less than 17 Å were observed.

Triplet-contact analysis performed in an identical manner to that in Chapter 3 (Table 5.2) yields very similar results to that for poplar PC and turnip cyt f. Again, the overall trends are not as pronounced as in the higher-plant studies, but the trend is still for the most popular types of complexes to have 3 to 8 Å smaller average Cu-Fe distance than the overall average (of 26.6 Å for 2027 successes out of 9000 attempted trajectories).

When, as previously, twelve of the most popular complexes were selected and superimposed upon one another (Figure 5.5), the uniformity of complex formation between PC and cyt f is again evident. Like complexes formed between poplar or spinach PC and turnip cyt f, there is a strong electrostatic component forged through interaction between the east patch acidic residues on PC (particularly D54 and
Figure 5.4. The distribution of distances between metal centers for all *Chlamydomonas* PC/cyt f complexes with metal center distances of less than 35 Å for 9 sets of 1000 trajectories each at 10 mM, 100 mM, and 10 M, where complexes were recorded on the basis of minimized Cu-Fe distances.
### Table 5.2. Best seven triplet contacts between *Chlamydomonas* cyt f and PC and 100 mM ionic strength.

<table>
<thead>
<tr>
<th>Triplets (cyt f res - PC res)</th>
<th># successes</th>
<th>( P ) (#/9000)</th>
<th>Avg. Cu-Fe distance (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>K65 -D61 K188-D53 K189-E59</td>
<td>36</td>
<td>0.0040</td>
<td>19.1</td>
</tr>
<tr>
<td>K65 -E85 K188-D53 K189-E59</td>
<td>29</td>
<td>0.0032</td>
<td>20.5</td>
</tr>
<tr>
<td>K65 -E85 K188-D53 K189-E43</td>
<td>26</td>
<td>0.0029</td>
<td>21.0</td>
</tr>
<tr>
<td>K65 -E43 K188-D42 K189-D44</td>
<td>19</td>
<td>0.0021</td>
<td>23.7</td>
</tr>
<tr>
<td>K65 -E59 K188-D53 K189-E85</td>
<td>19</td>
<td>0.0021</td>
<td>17.8</td>
</tr>
<tr>
<td>K65 -E85 K188-D44 K189-E43</td>
<td>14</td>
<td>0.0016</td>
<td>18.3</td>
</tr>
<tr>
<td>K65 -E59 K189-D53 heme-E85</td>
<td>11</td>
<td>0.0012</td>
<td>17.6</td>
</tr>
</tbody>
</table>

**Figure 5.5.** Alpha-carbon traces of 12 complexes between *Chlamydomonas* PC and cyt f. These 12 complexes were selected randomly from a single run of 1000 trajectories at pH 7 and 100 mM ionic strength from the four best contact triplets listed in Table 5.2. The figure was prepared in the same way as Figure 3.5.
E59/D61) and the positive inter-domain patch of cyt f (particularly K188 and K189), and that electrostatic interaction apparently drives the northern, more hydrophobic patch of PC (with H87 at its center) into close contact with the heme region of cyt f. Electrostatic interactions do seem to be slightly more intimate here than for the interactions of higher plant PCs with turnip cyt f, and the involvement of centrally-located D54 to a much greater extent in Chlamydomonas PC/cyt f complexes than that of D51 in other PC/cyt f complexes might play a role.

Brownian dynamics of Chlamydomonas cyt c₅ and cyt f. When moving to study through Brownian dynamic simulations the interaction between cytochrome c₅ and cyt f, we decided to initially retain the same basic method of determining success as before - that is, as distance between metal centers. The iron center of cyt c₅ is set farther back in the protein than is the copper center of PC; it therefore would make sense that the probability of complex formation at certain Fe-Fe distances for c₅/cyt f would be lower than the probabilities at corresponding Cu-Fe distances for PC/cyt f. This does, in fact, turn out to be the case: as can be demonstrated by Figure 5.6, for Fe-Fe distances less than 21 Å, the probability of complex formation is 0.007 ± 0.001 in the 100 mM case and 0.0010 ± 0.0005 in the 10 M case; for Fe-Fe distances less than 17 Å, no complexes were formed in either case. (Perhaps more telling is the 25 Å criterion, in which the probability of complex formation is 0.036 ± 0.006 in the 100 mM case and 0.018 ± 0.003 in the 10 M case.)

There does appear to be a difference in complex formation rates between the 100
Figure 5.6. The distribution of distances between metal centers for all Chlamydomonas c/o/cyt f complexes with metal center distances of less than 35 Å for 9 sets of 1000 trajectories each at 10 mM, 100 mM, and 10 M, where complexes were recorded on the basis of minimized Fe-Fe distances.
Figure 5.7. The distribution of distances between metal centers for all *Chlamydomonas* $c_0$/cyt f complexes with metal center distances of less than 35 Å for 9 sets of 1000 trajectories each at 100 mM and 10 M, where complexes were recorded on the basis of minimized distance between centers-of-mass.
mM case and the 10 M case, although not as pronounced as for the PC/cyt f interactions; this implies that complex formation might be less dependent on electrostatic interactions in the c₉/cyt f interaction than in the PC/cyt f interaction. A test of metal-center distance distribution using a center-of-mass reaction criterion (again, as in Chapter 3; see Figure 5.7) does in fact offer confirmation of this trend, as does the MacroDox method of determining success in heme protein trajectories by evaluating distance between any of the carbons immediately surrounding the iron in the porphyrin ring (figure not shown).

It is when the triplet contacts table (Table 5.3, from the initial metal-center reaction distance studies: again, other studies give similar results) is analyzed that we observe the nature of interaction between c₉ and cyt f significantly diverge from that of PC and cyt f. For the same number of trajectories, and a similar number of successes at 35 Å (1959 for c₉/cyt f, as opposed to 2027 for PC/cyt f) the most popular complex type

<table>
<thead>
<tr>
<th>Triplet</th>
<th># successes</th>
<th>P (#/9000)</th>
<th>Avg. Fe-Fe distance (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>K65-E54</td>
<td>5</td>
<td>0.0006</td>
<td>30.0</td>
</tr>
<tr>
<td>K65-E54</td>
<td>4</td>
<td>0.0004</td>
<td>22.4</td>
</tr>
<tr>
<td>E108-K89</td>
<td>4</td>
<td>0.0004</td>
<td>28.8</td>
</tr>
<tr>
<td>E108-K89</td>
<td>4</td>
<td>0.0004</td>
<td>27.0</td>
</tr>
<tr>
<td>E108-K29</td>
<td>3</td>
<td>0.0003</td>
<td>26.9</td>
</tr>
<tr>
<td>D162-E28</td>
<td>3</td>
<td>0.0003</td>
<td>27.0</td>
</tr>
</tbody>
</table>

Table 5.3. Best seven triplet contacts between *Chlamydomonas* cyt f and c₉ at 100 mM ionic strength.
only occurs 5 times, as opposed to 36 in the PC/cyt f case. Only seven triplet types occur more often than twice; the top seven triplet types for PC/cyt f each occur at least eleven times over the 9000 trajectories, and no fewer than 33 triplet types occur more often than twice.

In the poplar PC/turnip cyt f study, when triplet contact tables of the cases at 100 mM and 10 M ionic strength were compared, the fact that the most popular complexes at 10 M had average Cu-Fe distances close to or over the overall average for successful (35 Å) complexes at that ionic strength (as opposed to the 100 mM case, where the most popular complexes routinely had average Cu-Fe distances 6-8 Å below the overall average) was pointed to as a sign that specificity in complex formation is lost between PC and cyt f at higher ionic strengths. Likewise here, only one of the more popular complexes between c₆ and cyt f shows an average Fe-Fe distances significantly less than the overall average of 29.0 Å for successful complexes; the four complexes characterized by the triplet {K65-E54, K122-D41, K189-E69} had an average Fe-Fe distance of 22.4 Å. On the other hand, the five complexes with the triplet contact {K65-E54, K68-D65, K189-E69} had an average Fe-Fe distance of 30.0 Å, a full angstrom above the overall average.

The clearest evidence of the poor specificity between c₆/cyt f complexes is in the α-carbon trace of the twelve "most popular" complexes between c₆ and cyt f. With figures overall so low, however, the selection of these complexes is nearly the selection of 12 random complexes. The execution of this was: from a single run, complexes were selected that occurred in any of the other eight runs. For the run that was selected, this
Figure 5.8. Alpha-carbon traces of 12 complexes (with 90° rotations arranged vertically, not horizontally as previously) between *Chlamydomonas* c₃ and cyt f. These 12 complexes were selected randomly from a single run of 1000 trajectories at pH 7 and 100 mM ionic strength from the four best contract triplets listed in Table 5.3. The figure was prepared in the same way as Figure 3.5.
added up to exactly 12 complexes.

The superposition of the twelve complexes upon one another (Figure 5.8) reflects a non-specific pattern, although not as non-specific as one might think. The resulting complexes fan out through nearly a 90° angle around cyt f near the heme, around roughly a single axis of rotation. This implies some role for electrostatic interactions, possibly in a two-step mechanism similar to that proposed for the PC/cyt f interaction previously - with electrostatic interactions bringing c₅ into proximity of the cyt f heme, and hydrophobic interactions providing the specific interaction between the two proteins.

Low ionic strength. In order to hypothetically determine how the proteins would interact under conditions where electrostatic interactions are maximized, a similar regimen of BD runs between both PC and cyt f and c₅ and cyt f were executed with ionic strength set to 10 mM.

Tables 5.4 and 5.5 describe the results of those runs in terms of triplet-contact analysis; in previous Figures 5.4 and 5.6, 10 mM success distribution is charted alongside 100 mM and 10 M data for comparison. Again, it is clear that there is much more electrostatic preference for an interaction with PC than for an electrostatic interaction with c₅, and the numbers for triplet contacts between PC and cyt f follow earlier trends. However, some electrostatic interaction between c₅ and cyt f is apparent; the number of complexes for each type and the differential between the overall average metal center distance and each triplet contact type’s average metal center distance is the
<table>
<thead>
<tr>
<th>Triplets <em>(cyt f res - c, res)</em></th>
<th># successes</th>
<th>( P (#/9000) )</th>
<th>Avg. Fe-Fe distance (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>K65 -D61 K188-D53 K199-E59</td>
<td>135</td>
<td>0.0150</td>
<td>17.4</td>
</tr>
<tr>
<td>K65 -E85 K188-D53 K199-E43</td>
<td>99</td>
<td>0.0110</td>
<td>19.7</td>
</tr>
<tr>
<td>K65 -E85 K188-D53 K199-E59</td>
<td>89</td>
<td>0.0099</td>
<td>17.9</td>
</tr>
<tr>
<td>K65 -E85 K188-D44 K199-E43</td>
<td>63</td>
<td>0.0070</td>
<td>17.5</td>
</tr>
<tr>
<td>K65 -E59 K188-D53 K199-E85</td>
<td>55</td>
<td>0.0061</td>
<td>19.0</td>
</tr>
<tr>
<td>K188-D53 K199-E85 heme-Cu</td>
<td>46</td>
<td>0.0051</td>
<td>16.0</td>
</tr>
<tr>
<td>K65 -E43 K188-D42 K199-D44</td>
<td>34</td>
<td>0.0038</td>
<td>21.1</td>
</tr>
</tbody>
</table>

Table 5.4. Best seven triplet contacts between *Chlamydomonas* cyt f and PC at 10 mM ionic strength.

<table>
<thead>
<tr>
<th>Triplets <em>(cyt f res - c, res)</em></th>
<th># successes</th>
<th>( P (#/9000) )</th>
<th>Avg. Fe-Fe distance (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>D68 -D65 K189-E71 heme-R66</td>
<td>34</td>
<td>0.0038</td>
<td>20.4</td>
</tr>
<tr>
<td>D162-R66 K164-E71 heme-D65</td>
<td>28</td>
<td>0.0031</td>
<td>22.3</td>
</tr>
<tr>
<td>K65 -E71 K189-E70 heme-R66</td>
<td>27</td>
<td>0.0030</td>
<td>19.0</td>
</tr>
<tr>
<td>D68 -D65 K189-E70 heme-R66</td>
<td>19</td>
<td>0.0021</td>
<td>22.6</td>
</tr>
<tr>
<td>K65 -E54 K189-E69 heme-D65</td>
<td>16</td>
<td>0.0018</td>
<td>24.3</td>
</tr>
<tr>
<td>K188-E70 K189-E71 heme-R66</td>
<td>15</td>
<td>0.0017</td>
<td>21.0</td>
</tr>
<tr>
<td>K164-D65 K189-E69 heme-K57</td>
<td>14</td>
<td>0.0016</td>
<td>20.0</td>
</tr>
</tbody>
</table>

Table 5.5. Best seven triplet contacts between *Chlamydomonas* cyt f and c\(_a\) at 10 mM ionic strength.
Figure 5.9. Alpha-carbon traces of 12 complexes between *Chlamydomonas* c₆ and cyt f. These 12 complexes were selected randomly from a single run of 1000 trajectories at pH 7 and 100 mM ionic strength from the four best contract triplets listed in Table 5.5. The figure was prepared in the same way as Figure 3.5.

same for the c₆/cyt f interaction at 10 mM as for the PC/cyt f interaction at 100 mM.

Specificity is obviously sustained in the PC/cyt f interaction at lower ionic strengths; Figure 5.9 (as compared to Figure 5.8) demonstrates the increased specificity in the c₆/cyt f interaction with the lower ionic strength.

The specific nature of these complexes will be dealt with in the following chapter.
Table 5.6. Comparison of electrostatic interaction as measured through BD simulations between mutant *Chlamydomonas cyt* f and *Chlamydomonas* PC. All runs at 100 mM. Double, triple, and quintuple mutants as in Figure 5.10.

<table>
<thead>
<tr>
<th>Mutant</th>
<th>Successes /1000 traj</th>
<th>% of WT</th>
<th>Successes /1000 traj</th>
<th>% of WT</th>
<th>Successes /1000 traj</th>
<th>% of WT</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>4.1±1.3</td>
<td>100</td>
<td>9.3±2.2</td>
<td>100</td>
<td>33.2±1.2</td>
<td>100</td>
</tr>
<tr>
<td>K58Q</td>
<td>1.8±1.5</td>
<td>43</td>
<td>4.6±1.5</td>
<td>49</td>
<td>21.2±3.0</td>
<td>64</td>
</tr>
<tr>
<td>K65Q</td>
<td>1.1±0.9</td>
<td>27</td>
<td>3.2±1.4</td>
<td>34</td>
<td>15.0±3.8</td>
<td>45</td>
</tr>
<tr>
<td>K188Q</td>
<td>1.4±1.1</td>
<td>34</td>
<td>3.5±1.9</td>
<td>38</td>
<td>17.3±3.0</td>
<td>52</td>
</tr>
<tr>
<td>K189Q</td>
<td>0.8±0.6</td>
<td>20</td>
<td>2.6±0.8</td>
<td>28</td>
<td>14.1±2.8</td>
<td>42</td>
</tr>
<tr>
<td>Double</td>
<td>0.3±0.4</td>
<td>7</td>
<td>1.2±0.7</td>
<td>13</td>
<td>8.3±2.2</td>
<td>25</td>
</tr>
<tr>
<td>Triple</td>
<td>0.4±0.5</td>
<td>10</td>
<td>1.2±0.8</td>
<td>13</td>
<td>7.9±2.2</td>
<td>24</td>
</tr>
<tr>
<td>5-uple</td>
<td>0.1±1.7</td>
<td>2</td>
<td>0.3±0.4</td>
<td>3</td>
<td>2.3±0.9</td>
<td>7</td>
</tr>
</tbody>
</table>

*BD with cyt f mutants.* Similar to the mutants of spinach PC in the previous chapter, a set of seven mutants - four single lysine-to-glutamine mutants, and three mutants mimicking the double, triple, and quintuple mutants of Soriano et al. (1996 and 1998). Preliminary Brownian dynamics simulations were executed to evaluate the behavior of these mutants. Figure 5.10 represents the distribution of complexes between PC and the cyt f variants as a function of Cu-Fe distance, and Table 5.6 lists the effect of each mutant on the interaction between PC and cyt f.

As expected, the behavior of these mutants paralleled that of the PC/cyt f interaction in solution (Soriano et al., 1998). All mutants had a significant electrostatic impact on the
Figure 5.10. The distribution of distances between metal centers for complexes between Chlamydomonas cyt f and various mutant Chlamydomonas PCs, with 9 sets of 1000 trajectories run for each mutant at 100 mM. Double mutant is K188N/K189Q; triple mutant is K58Q/K65S/K66E, quintuple mutant is a combination of above double and triple mutants.
interaction between the two molecules - the mutant with the least impact, K58Q, formed complexes at 50% of the frequency of the wild type, and the mutant with the most impact, K189Q, formed complexes at only 25% the frequency of the wild type. Double and triple mutants were strongly inhibited (~10% of the wild type) and the quintuple mutant was almost completely inhibited. This reinforces the idea that a unique effect in the *in vivo* experiments is responsible for the lack of observed electrostatic impact on the interaction between the molecules in the cell.

**Interaction between PC and cyt f dimer.** In solving the structure of *Chlamydomonas* cyt f, Berry et al. (manuscript in preparation) observed an especially intimate dimeric structure of cyt f, with more than 10% of the surface area on each monomer buried in the contact. The lysine patches and the heme region were exposed in this contact, however, and it would hypothetically be possible for PC to dock at the heme face on either monomer within this structure. We obtained a crystal structure for this cyt f dimer (E. A. Berry, personal communication).

In electrostatic representation of the dimer (not shown), the key point of interest is positive potential generated by both proteins surrounding the heme region, making attraction of PC to that region favorable. The two primary regions of positive potential in the dimer, in fact, are the regions surrounding the two hemes, providing support for Berry’s hypothesis that PC could dock with either or both sides of the dimer assuming the dimer’s position with respect to the rest of the cyt *b* complex and the membrane was favorable.
Brownian dynamics simulations were carried out between the dimer and Chlamydomonas PC. Single runs of 10000 trajectories were performed at 100 mM and 10 M ionic strength, to compare with the runs with similar runs simulating interaction between the cyt f monomer and PC. Reaction criteria were set so that the distance between either iron center in the two hemes within the dimer and the copper atom in PC would be minimized when evaluating the success of trajectories.

A plot mirroring Figure 5.4 was prepared to compare interaction probabilities between the cyt f monomer and the cyt f dimer; this plot is Figure 5.11. Overall, the monomer shows a far greater tendency towards electrostatic interactions than the dimer: at distances greater than 17 Å, the dimer shows less probability for interaction with PC at 10 mM than the monomer shows at 100 mM.

However, in neither ionic strength case does the monomer show any complexes with Cu-Fe distances less than 14 Å. The dimer shows such complexes in both ionic strength cases; 3 such complexes form at 100 mM out of 10000 trajectories, and 20 such complexes form at 10 mM. This is possibly due to structural changes around the heme in the dimer structure; however, since both the monomer and dimeric structures were solved from the same crystal (E. A. Berry, personal communication), those structural changes are not likely to be that severe. It is more likely that cooperative effects in the electrostatic interactions between both cyt f's allow for the formation of such tight complexes in the event of a close approach (even though such close approaches are less likely in the dimer case than the monomer.)
Figure 5.11. A comparison of the distribution of distances between metal centers for complexes between *Chlamydomonas* PC and monomeric and dimeric cyt f, where complexes were recorded on the basis of minimized Cu-Fe distances. Monomeric data is taken directly from Figure 5.4. Dimeric data is based on single runs of 10000 trajectories at 10 mM and 100 mM.
CHAPTER 6

DISCUSSION, CONTEXT, AND CONCLUSIONS

Electrostatics modeling of PC, cyt f, and c₆. To the extent that electrostatic interactions do dominate complex formation between PC or c₆ and cyt f, a single general region of charged residues can be pinpointed on each protein as the most important region in electrostatic interactions. On (turnip) cyt f, that region most prominently includes K58, K65, and K187; other positively charged residues, such as K66, K122, R156, K185, and R209, have peripheral roles. The iron center of the heme also contributes to the large positive potential field that attracts PC or c₆.

Of all the above positive charges, K65, K66, R156, and K187 are absolutely conserved in all green plant and algal cyt fs, and K58, K122 and K187 are conserved in a vast majority of cyt fs. Only R209 is unique to turnip cyt f. However, when one looks at the cyanobacterial cyt f sequences, only K66 and R156 are conserved, and negative residues are inserted around positions 122 and 187.

Plastocyanin has a so-called "east face" with a surface full of negative residues; there are two dominant patches, one consisting of D42, E43, and D44, all of which are conserved throughout all green plant and algal plastocyanins, and one in positions 59-61.
in which usually (but not always, especially in cyanobacterial PCs) two of those three residues are acidic. In addition, most plastocyanins have a negative charge in either position 51 or position 53, and the third negative charge in the 59-61 group found in type IV plastocyanins is substituted with a negative charge in position 85 in type II and III plastocyanins. Again, the greatest deviation from this overall pattern regarding the distribution of negative charges is in the cyanobacterial (type I) PCs.

Cyt f has a homologous region of negative potential, despite that protein’s radically different structure (alpha-helical instead of beta sheet, heme protein instead of copper). Homologous to the conserved D42/E43/D44 patch in PC is a sequence of negative charges in positions 68-71; in all green algal c₅’s, at least three of those four positions are negative charges. This correlation was also observed by Ullmann et al. (19976).

The homology (in the type II, or green algal, PCs) to the negative charge in positions 53, 59 and 61 is less clear, but those PC residues appear to correlate to E47, E54 and D65 in Chlamydomonas c₅, however, only E47 is conserved across all green algal c₅’s. (E54 and D65 also turn up in Monoraphidium c₅, the other algal c₅ studied in this work.)

Again, none of these residues are conserved in algae outside of the green algae nor in the cyanobacteria, repeating the same theme that has dominated consideration of the key residues generating the charged patches in these three proteins. This is mere circumstantial evidence, but the weight of this evidence is nearly overwhelming: there are key conserved charged residues in PCs, cyt f’s, and c₅’s in higher plants and green algae that are much less reliably conserved in other organisms. The statement made in Pearson et al. (1996) is: "Either the electrostatic model of interactions between PC and
cyt f does not apply to cyanobacteria, or the electrostatic interactions are quite different from those for the higher plants and algae." Not only is that conclusion supported by the additional information presented here, one can also consider the same to be true for the interaction between c₆ and cyt f.

Manual docking of PC and cyt f. The goal of manual docking of PC and cyt f was to build hypotheses of complex formation between the two proteins. This discussion of the manual docking will focus on two complexes described graphically in figures 2.15 and 2.16, intended to represent two distinct scenarios: one in which the interaction between the two proteins results exclusively from electrostatics, and one in which the distance between metal centers was minimized.

Complex 1 (Figure 2.15) is a complex in which contacts between charged residues are maximized. However, distance between metal centers is not minimized in this complex; the measured Cu-Fe distance is 31.4 Å. It is very unlikely that this complex is electron-transfer active; proposed electron transfer pathways for such a complex are no better than highly speculative scenarios in which electrons would be transferred from the heme through the conserved water chain in cyt f (described in Martinez et al., 1996) and to Y83 on PC (Pearson et al. 1996) or through a cation-pi interaction between K65 on cyt f and Y83 on PC (Ullmann et al., 1997a). No significant experimental evidence exists for either scenario outside of the experiments and calculations of He et al. (1991) implying that an aromatic residue at position 83 is necessary for electron transfer to take place; further attempts to repeat those
experiments to obtain similar results (J.C. Gray, personal communication) have failed, calling He's conclusions into question. The conclusion of Pearson et al. (1996) that the major difficulty of such a complex was the lack of a suitable electron transfer pathway from heme to copper is further supported by the lack of repeatability in He's results.

Complex 2 (Figure 2.16) places H87 on PC in very close proximity (within 4.5 Å) of Y1 on cyt f, in theory, this would provide for efficient electron transfer between ligands. The primary difficulty in this complex was argued in Pearson et al. (1996) to be poor electrostatic interactions, particularly between positively charged groups surrounding the heme and the positive potential from the copper center. Given such repulsion between key residues important for electron transfer, the logical conclusion would seem to be that this kind of complex would form after rearrangement from a predocking electrostatic complex (i.e., Complex 1) through favorable hydrophobic interactions. In Pearson et al. (1996), we calculated hydrophobic interactions between -6.4 to -12.0 kcal/mol for this complex.

Both of these complexes share the interaction between D44 on PC and K187 on cyt f, the two residues cross-linked by Morand et al. (1989). The activity of such cross-linked complexes has been a matter of some debate; it is possible that those cross-links that are not electron-transfer active (Qin and Kostic, 1993) are of the type of Complex 1, and those that are electron-transfer active (Takabe and Ishikawa, 1989; D.J. Davis, unpublished results) are of the type of Complex 2.

These two manually docked complexes (along with Complex 3, given in Figure 2.17) drove some debate in the literature regarding the nature of the complex formation.
between PC and cyt f. Most prominently, Soriano et al. (1997) assembled complexes deliberately similar to the three above manually docked complexes to enquire about redox potential changes and reorganization energies.

Quantitative results of BD. The purpose of Brownian dynamics simulations was the "less biased" search for possible complexes between PC and cyt f, modeling in some respects the motion of PC in solution under the influence of electrostatic attraction to cyt f. The method does not consider hydrophobic interactions, which (we thought) would provide us with complexes similar to manually-docked Complex 1 (Figure 2.15) and not those resembling Complex 2 (Figure 2.16).

The experimentally observed rate constant is on the order of 10^6 M^-1 s^-1 (Gross et al., 1990; Qin and Kostic, 1992; Kannt et al., 1996). While rates obtained from BD simulations should be taken with a grain of salt because of the limitation of the simulations to only electrostatic interactions, simulations like those involving poplar PC and turnip cyt f give an idea of the probability of complex formation necessary to give physiological rates - between 3 and 10 "successes" per 1000. This seems like an awfully small number, but in these simulations PC and cyt f from whatever species regularly get this fraction of successes when distance between metal centers is within 16-17 Å (at worst) at ionic strengths approaching physiological; control cases get this many only at distances of 20 Å or higher. Studies involving spinach and Chlamydomonas PC reinforce these numbers.

(Immediately on studying these distances, it was obvious that these complexes
were not similar to Complex 1, which had a Cu-Fe distance of 31.4 Å. It would soon become obvious why this was so.

Note also that this definition of "success" does not take into account any events after the electron transfer; the process being modeled reflects only the events leading up to electron transfer, perhaps the interaction between two proteins where electron transfer is prohibited from occurring. This consideration was made to maintain the model's simplicity; the work of Moser et al. (1992) in observing that electron transfer rate depends in large part of the distance between electron transfer partners allows us to make this simplification. In manually-docked Complex 2, the metal centers are within 15.6 Å and the distance between H87 on PC and Y1 on cyt f, ligands to the respective metal centers, was as small as 4.5 Å. Such a complex would certainly be viable for electron transfer.

Qualitative results of BD. Figure 6.1 is a representation of three characteristic complexes from each of the three major PC/cyt f Brownian dynamics experiments - the original poplar PC/turnip cyt f study described in Chapter 3, the control (wild type) spinach PC/turnip cyt f study described in Chapter 4, and the Chlamydomonas PC/cyt f study described in Chapter 5. Table 6.1 describes close contacts between each of those three complexes, in the same format as in the Pearson et al. (1996) paper.

The stunning result of the original study - repeated in the two following studies - was the similarity of the closest and most frequently occurring (as evaluated by "triplet contacts") complexes to Complex 2, the complex thought to require hydrophobic
Figure 6.1. Alpha-carbon traces of characteristic complexes between poplar PC and turnip cyt f, spinach PC and turnip cyt f, and *Chlamydomonas* PC and cyt f. The key point in this image is the coincidence of PC's orientation regardless of the species of PC and cyt f interacting.
<table>
<thead>
<tr>
<th>cyt f residue</th>
<th>poplar PC residues</th>
<th>spinach PC residues</th>
<th>Chlamydomonas PC residues</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y1</td>
<td>L12,P86,H87</td>
<td>P86,H87</td>
<td>P37,P86,H87</td>
</tr>
<tr>
<td>I3</td>
<td>D9,G10</td>
<td>H87</td>
<td></td>
</tr>
<tr>
<td>F4</td>
<td>L12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q7</td>
<td>D9,G10,S11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A62</td>
<td>S85,P86,G89</td>
<td>Q88</td>
<td>Y62,P86</td>
</tr>
<tr>
<td>N63</td>
<td>E59,Q88</td>
<td>Y83,Q88</td>
<td>D61,Y62,E85</td>
</tr>
<tr>
<td>K65</td>
<td>E59</td>
<td></td>
<td>D61</td>
</tr>
<tr>
<td>N99</td>
<td>N64,A65</td>
<td>N64</td>
<td>P66</td>
</tr>
<tr>
<td>S101(trnp)</td>
<td>F35,N64</td>
<td>P66</td>
<td></td>
</tr>
<tr>
<td>Y101(Chls)</td>
<td>F35</td>
<td>P36</td>
<td>F36,P37</td>
</tr>
<tr>
<td>N103</td>
<td>F35</td>
<td>P36</td>
<td>N64</td>
</tr>
<tr>
<td>P117</td>
<td>F35,P36</td>
<td>P36</td>
<td>F36,P37</td>
</tr>
<tr>
<td>P119</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>K187(trnp)</td>
<td>D44</td>
<td></td>
<td></td>
</tr>
<tr>
<td>K188(Chls)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>K189(Chls)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>heme</td>
<td>P86</td>
<td>Q88</td>
<td>Y62,P86</td>
</tr>
</tbody>
</table>

Table 6.1. Table of contact residues in characteristic complexes between PC and cyt f from different species as shown graphically in Figure 6.1. Contact residues must be within 6 Å of one another center-to-center. "trnp" = turnip cyt f; Chls = Chlamydomonas cyt f.
interactions to form. In each complex, H87 on PC and Y1 on cyt f are within 6 Å of one another, similar to Complex 2, and (as said previously) metal centers on these complexes were less than 17 Å from one another. The fact that (as in Figure 3.5, among others) so many of these complexes occur in exactly the same orientation caused us to conclude that "electrostatic interactions tend to steer PC toward a single dominant orientation with respect to cyt f" (Pearson et al., 1998) on the basis of complexes between poplar PC and turnip cyt f alone. Once again, studies between spinach PC and turnip cyt f, and between Chlamydomonas PC and cyt f, reinforced conclusions built on earlier experiments.

Two studies conducted simultaneously to the work that resulted in Pearson et al. (1998) further reinforced the idea that PC and cyt f formed a complex that resulted in H87 and surrounding residues on PC facing Y1, the heme and surrounding residues on cyt f. Ullmann et al. (1997a) also computationally simulated the interaction between PC and cyt f, but instead of using BD sampled the possible complexes between PC and cyt f using a Monte Carlo method, then used molecular dynamics to minimize the energy of those complexes. This considered molecular flexibility in the PC/cyt f complex, but not the diffusion process itself.

Ullmann obtained six ensembles of complexes that were energetically favorable, four of which placed the metal centers too far apart for electron transfer to be feasible. The remaining two complexes placed H87 on PC in close proximity to Y1 on cyt f and were calculated by the Pathways method of Beratan et al. (1991) to be significantly more electron-transfer active than the other four. While Ullmann et al. acknowledged
the likelihood of an electron transfer pathway involving residues surrounding H87/PC and Y1/cyt f, they also pursued a cation-pi pathway for which very little support exists. In addition to Ullmann’s study, Ubbink et al. (1998) were able to observe through NMR experiments chemical shifts in PC nuclei on the binding of cyt f to PC, which they correlated to contacts between the two proteins in the complex. Using these contacts and electrostatic restraints, they performed rigid-body molecular dynamics experiments to model the structure of the complex between the two proteins. The structure of that complex is given in Figure 6.2, with the position of the aforementioned characteristic spinach PC/tumip cyt f complex superimposed. The criteria for the formation of complexes in Ubbink’s modeling were very different than those in the BD studies, particularly in the treatment of electrostatic interactions - while the BD methodology treated electrostatic interactions more rigorously, Ubbink’s rigid-body molecular dynamics required specific charged residues (D42, E43, D44, E59, E60) to interact with "the large positive patch on cytochrome f". This highly specific demand turned out to force D44 on PC closer to K187 on cyt f than observed in the BD simulations.

However, these differences do not take away from the primary similarity between all of these studies and our BD studies: an interaction between H87 on PC and Y1 on cyt f. There is now quite strong evidence that this is the electron-transfer complex between PC and cyt f. The BD work done on PC and cyt f shows that such a complex can form directly, through electrostatic interactions alone, and does not require a pre-docking complex to form before rearrangement can take place.
Figure 6.2. Alpha-carbon trace comparison of the positioning of the characteristic BD complex between spinach PC and turnip cyt f (light lines) and a characteristic complex between the same proteins from the NMR study and rigid-body MD simulation of Ubbink et al. (1998; PDB ID 2PCF) (dark lines). The NMR/MD complex appears to favor electrostatic interactions, particularly between residues 42-45 on PC and residues K187 and R209 on cyt f, far more than the BD complex does.
Modeling of mutants. In a bimolecular interaction where attraction between positive charges on one molecule and negative charges on another are dominant, it would stand to reason that removal of charges would diminish the probability of the interaction. The question here is how much removal of charges, and the position of those charges, impact the interaction between PC and cyt f.

Mutants were modeled in two cases to emulate two different experiments. Spinach PC was "mutated" computationally primarily to mirror experiments performed by Kannt et al. (1996), which produced inconclusive results regarding the effects of the two negatively charged patches on PC. *Chlamydomonas* cyt f was similarly mutated for purposes of simulating studies of Soriano et al. which produced conflicting results regarding the effects of mutating residues in the positive patch *in vivo* (Soriano et al., 1996) against *in vitro* (Soriano et al., 1998). In both cases (when common data is compared, in the case of 100 mM ionic strength and 17 Å distance between metal centers) the general impact of changes are similar. Single mutants of PC negative charges typically cut observed "reactivity" between PC and cyt f in half, with subtle but distinct differences between the 42-45 negative patch and the 59-61 negative patch. Single mutants of the cyt f positive patch typically decreased the probability of success by two-thirds, although variance between K58 (near cyt f's midsection) and K189 (well onto cyt f's small domain) exists. As a systematic study of mutants, both of these studies are far more complete than any mutagenesis study on the PC/cyt f interaction, both as regards PC (where Lee et al. (1995) and Kannt et al. (1996) did not make comparable single mutants on the two patches) and cyt f (where Soriano et al. (1996,
1998) did not make a single double mutant).

With respect to spinach plastocyanin, we continue to argue that the question concerning the difference in importance of the two negative patches is still open. Kannt et al. (1996) argued on the basis of some single mutants in the 42-45 patch and no single mutants in the 59-61 patch that the two patches were equally important. Especially in light of the observation of Lee et al. (1995) concerning severe differences in reactivity between the two negative patches in *Silene* PC, this conclusion seems to be in doubt. However, observations of Sugawara et al. (1999) concerning structural differences between the main chains of *Silene* and spinach plastocyanin crystal structures as well as the structural differences within spinach NMR structures obtained from A. Bergqvist (personal communication) and differences in reactivities of those structures imply that the structure of particular parts of the molecule, most notably the loop containing the 59-61 patch, might have a role in determining the strength and specificity of the electrostatic interactions between the two proteins.

Especially in the *Chlamydomonas* cyt f case, BD simulation of mutagenesis tends to mimic *in vitro* experiments (Soriano et al., 1998). The meaning of Soriano’s *in vivo* experiments (1996), showing no effect of electrostatic mutants and suggesting immediately that diffusional forces alone are capable of promoting docking of the two proteins within the thylakoid space, remain in question. Fernandez-Velasco et al. (1997) was able to observe inhibition of cyt f oxidation by mutating positive charges in permeabilized *Chlamydomonas* cells under low ionic strength conditions. In addition, Haehnel et al. (1989) have shown PC to move over long distances in the chloroplasts of
spinach and pea plants, and Allred and Staehelin (1986) have shown that the distribution of cyt f, in part, throughout the grana stacks and separate from photosystem I in the unstacked lamellae, requires transport over long ranges that would be difficult for solely diffusion to handle.

Comparison of $c_0$ cyt f interaction with PC cyt f interaction. On surface, it appears that the interaction between Chlamydomonas $c_0$ and cyt f is poorer than the interaction between PC and cyt f. It is difficult to compare BD simulations between the two different proteins, because of differing burial of the (deeper) iron center in $c_0$ and the (closer to the surface) copper center of PC. However, no characteristic complex at 100 mM presents itself in the same way that such complexes present themselves in the PC/cyt f interaction. Such a complex does appear at 10 mM, however, implying that electrostatic interactions do have some role in the interaction.

The studies of $c_0$ and cyt f - both computationally and experimentally - are very incomplete, and many questions remain about the interaction between the two proteins. Clearly the mode of interaction is different than that in the PC/cyt f interaction, but how different? Interactions between hydrophobic residues in the $c_0$/cyt f association have not been studied at all, and the lone mutagenesis study performed to date on $c_0$ (de la Cerda et al., 1999) only considers interactions between Synechocystis $c_0$ and Photosystem I.

Overall conclusions. The goal of this work was to describe with computational models the electrostatic interaction between cyt f and its physiological electron transfer
partners, PC and c6. Initially, simple electrostatic analysis was utilized to create hypotheses about the nature of the electrostatic interaction, and then those hypotheses were tested through Brownian dynamics simulation.

Electrostatic interactions have a clear role in complex formation between PC and cyt f in partially all plant and algal species; the importance of electrostatic interactions in the interaction of mobile electron transfer proteins and cyt f in cyanobacteria is still clearly a matter of debate. Based on simple observation of these electrostatic interactions and manual docking, it was hypothesized that the most likely initial conformation of a PC/cyt f complex would place the negative patch of PC and the positive patch of cyt f in a pre-docking situation, with no substantial interaction between any electron-transfer-active face on PC with a complementary face on cyt f; PC and cyt f would then rearrange into a complex competent for electron transfer, likely involving interactions beyond simple electrostatics (i.e., hydrophobic interactions).

Brownian dynamics simulations provided a serious challenge to this hypothesis: in no event did an a complex similar to the postulated pre-docking confirmation materialize, and a single dominant orientation of PC with respect to cyt f tended to form in simulations for PCs and cyt f’s from several different species, in a conformation similar to the electron-transfer-competent complex previously postulated. Such a complex could form through electrostatic interactions alone. The rate of formation of such a complex with copper-iron distances of approximately 16 Å was consistent with experimentally observed bimolecular rate constants.

Modeling of mutant PCs and cyt f’s to remove residues apparently participating
in such electrostatic interactions have predictable effects in diminishing the rate of complex formation. It is necessary at this point for more complete investigations on the role of single and specific mutants to get a reliable read for the role of patches of residues; however, it is clear that all negatively charged residues considered to be part of PC's negative patch (in spinach PC: 42-45, 51, and 59-61) and cyt f's positive patch (in Chlamydomonas cyt f: 58, 65, 188, and 189) have a significant role to play in the interaction in the model and in vitro. Concerning the interaction between Chlamydomonas PC and cyt f, the in vivo experiments of Soriano et al. (1996) are still a matter of debate.

While the electrostatic potential fields of higher plant and algal PCs and higher plant and algal c₆'s are measurably similar (see also Ullmann et al., 1997b) the electrostatic interaction between c₆ and cyt f only approaches the corresponding interaction between PC and cyt f. Hydrophobic interactions might play a greater role in the interaction between c₆ and cyt f, but more study on the interaction between c₆ and cyt f is desperately needed.
BIBLIOGRAPHY


Northrup, S. H., J. O. Boles, and J. C. L. Reynolds. 1988. Brownian dynamics of


Interaction. Tennessee Technological University, Cookeville, TN.

Pearson, D. C., and E. L. Gross. 1995. The docking of cytochrome f with plastocyanin:
three possible complexes. *In* Photosynthesis: from Light to Biosphere, Vol. II. P. Mathis,

Pearson, D. C., E. L. Gross, and E. S. David. 1996. The electrostatic properties of


