Research Article

Journal of Molecular Recognition

Received: 18 January 2014, Revised: 21 March 2014, Accepted: 18 April 2014, Published online in Wiley Online Library

(wileyonlinelibrary.com) DOI: 10.1002/jmr.2384

Molecular interactions between photosystem I and ferredoxin: an integrated energy frustration and experimental model

Derek J. Cashman[a,b†], Tuo Zhu[a], Richard F. Simmerman[a], Cathy Scott[a], Barry D. Bruce[a]* and Jerome Baudry[a,b]*

The stromal domain (PsaC, PsaD, and PsaE) of photosystem I (PSI) reduces transiently bound ferredoxin (Fd) or flavodoxin. Experimental structures exist for all of these protein partners individually, but no experimental structure of the PSI/Fd or PSI/flavodoxin complexes is presently available. Molecular models of Fd docked onto the stromal domain of the cyanobacterial PSI site are constructed here utilizing X-ray and NMR structures of PSI and Fd, respectively. Predictions of potential protein-protein interaction regions are based on experimental site-directed mutagenesis and cross-linking studies to guide rigid body docking calculations of Fd into PSI, complemented by energy landscape theory to bring together regions of high energetic frustration on each of the interacting proteins. The results identify two regions of high localized frustration on the surface of Fd that contain negatively charged Asp and Glu residues. This study predicts that these regions interact predominantly with regions of high localized frustration on the PsaC, PsaD, and PsaE chains of PSI, which include several residues predicted by previous experimental studies. Copyright © 2014 John Wiley & Sons, Ltd.

Additional supporting information may be found in the online version of this article at the publisher's web site.

Keywords: cross-linking; electron transfer; energy landscape theory; frustration; protein–protein docking

INTRODUCTION

Photosynthesis converts light energy into chemical energy using two large multi-subunit proteins located in the thylakoid membranes of cyanobacteria, and in the chloroplasts of algae and plants of photosystem II (PSII) and photosystem I (PSI). Following photo-excitation, PSII functions to extract electrons from water. These electrons are shuttled from PSII to the B6/F complex and finally arrive at PSI via soluble electron carrier proteins, such as cytochrome C_6 (Cyt C_6) or plastocyanin. Upon a second excitation in PSI, a charge separation occurs followed by a cascade of electron transfer steps through redox co-factors within PSI. This electron ends up in one of the two terminal electron acceptors, the 4Fe-4S centers, F_A or F_B, coordinated by the PsaC subunit. The electron is then transferred out of the PSI membrane complex to either the 2Fe-2S cluster of ferredoxin (Fd) (Zanetti & Merati, 1987; Zilber & Malkin, 1988) or the flavin mononucleotide cofactor of flavodoxin (Muhlenhoff *et al*., 1996; Setif, 2001). The reduced Fd transfers two electrons to the Fd-NADP oxidoreductase (FNR) to produce nicotinamide adenine dinucleotide phosphate-oxidase, which functions as a reductant in reactions of the Calvin cycle and several other metabolic processes (Berg *et al*., 2007). Under iron-limiting conditions, flavodoxin will be expressed and can functionally replace Fd as the PSI electron acceptor and donor to FNR in cyanobacteria (Leonhardt & Straus, 1992). Understanding the binding and affinity of Fd with PSI is a key element of the structure/function relationship of the electron transfer mechanisms during the process of photosynthesis.

NMR structures of cyanobacterial Fd (Baumann *et al*., 1996; Hatanaka *et al*., 1997) and an X-ray crystal structure of cyanobacterial PSI (Jordan *et al*., 2001) are available. The structure of PSI consists of twelve subunits, each labeled PsaA, PsaB, PsaC, PsaD, PsaE, PsaF, PsaI, PsaJ, PsaK, PsaL, PsaM, and either PsaX (plants) or PsaH (cyanobacteria). A summary of each of these subunits is provided in the Supplementary Information, Table S1. These structures suggest amino acid residues that could be involved in non-bonded interactions between PsaC, PsaD, and PsaE (collectively referred to as the *stromal domain*, because these three subunits are not transmembrane subunits and are located on the stromal side of the chloroplast membrane) with Fd or flavodoxin. In addition, these protein–protein interactions have been investigated using chemical cross-linking experiments (Lelong *et al*., 1994; Lelong *et al*., 1996), site-directed mutagenesis, examination of rate constants for electron transfer

* *Correspondence to: J. Baudry, M407 Walters Life Sciences, 1414 Cumberland Ave., Knoxville, TN 37996, USA.*
E-mail: jbaudry@utk.edu
Barry D. Bruce, 226 Hesler Biology, 1414 Cumberland Ave., Knoxville, TN 37996, USA.
E-mail: bbruce@utk.edu

a *D. J. Cashman, T. Zhu, R. F. Simmerman, C. Scott, B. D. Bruce, J. Baudry*
Department of Biochemistry and Cellular and Molecular Biology, University of Tennessee, Knoxville, TN 37996, USA

b *D. J. Cashman, J. Baudry*
UT/ORNL Center for Molecular Biophysics, Oak Ridge National Laboratory, University of Tennessee, Oak Ridge, TN 37831, USA

† *Department of Chemistry, Tennessee Technological University, Box 5055, Cookeville, TN, 38505-0001, USA*

steps using flash photolysis (Rousseau *et al*., 1993; Fischer *et al*., 1998; Fischer *et al*., 1999; Bottin *et al*., 2001; Setif *et al*., 2002; Goni *et al*., 2009; Setif *et al*., 2010), and detection of the PSI:Fd complex by backscattering interferometry (Setif *et al*., 2010). Collectively, these experiments have suggested several residues located on the surface of the stromal domain of PSI that are believed to be important for binding with Fd.

Understanding the molecular basis of the PSI:Fd interaction is important to the overall process of photosynthesis, but it will also increase the understanding of biological electron transfer as well as allowing advances in biotechnology and bioengineering. For example, there have been several attempts to replace the electron acceptor of Fd with a non-biological component such as an organic semiconductor (Das *et al*., 2004), or with an inorganic semiconductor such as TiO_2 to develop biologically based, or *biophotovoltaic*, cells for generating electrical power (Mershin *et al*., 2012; Yehezkeli *et al*., 2014). In addition, it has been demonstrated that the acceptor side of PSI can transfer electrons into catalytic nanoclusters (Millsaps *et al*., 2001; Evans *et al*., 2004; Iwuchukwu *et al*., 2010). Central to this concept of incorporating PSI into any bio-hybrid device to create energy is the need to optimize electron transfer from the site of charge separation within PSI to either the catalyst for hydrogen production or the electrode. Because it may be difficult to decrease the distance from the 4Fe-4S centers F_A/F_B found within PsaC and the non-biological materials, a strategy using the natural electron acceptor, Fd, to function as an intermediary to bridge the distance between F_A/F_B and the hybrid material may increase electron transfer. By directly linking Fd to the PsaC, PsaD, and PsaE subunits, it should be possible to guide the flow of electrons directly from PSI to either a conductive material or a catalytic material depending on the application.

In fact, it has been shown through *in vivo* studies of higher plant chloroplasts that the electron acceptor Fd binds specifically to one or more subunits of PSI, although such results were obtained before genomic, structural, or contemporary mass spectrometry data became available (Zanetti & Merati, 1987; Zilber & Malkin, 1988). However, early cross-linking work in spinach specifically identified a subunit with an apparent molecular weight of either a 22 KDa subunit (Zilber & Malkin, 1988) or a 20 kDa subunit (Zanetti & Merati, 1987) that can cross-link to Fd. It would appear that this early work identified the same subunit that is now called PsaD. Many years later, similar chemical cross-linking results were observed using PSI and Fd from the cyanobacteria *Synechocystis sp*. PCC 6803 (*Syn* 6803), demonstrated that Fd has a close contact with PSI subunits PsaC, PsaD, and PsaE (Zanetti & Merati, 1987; Zilber & Malkin, 1988; Andersen *et al*., 1992; Lelong *et al*., 1994). This suggested that the docking site would allow close proximity between the 2Fe-2S center of Fd and the terminal 4Fe-4S (F_B) of PSI (Lelong *et al*., 1996). This model of interaction was supported by mutagenesis studies in *Syn* 6803 that identified several residues in PsaD that contribute to interactions with Fd (Hanley *et al*., 1996). In addition, research on *Syn* 6803 shows that after a short cross-linking time, Fd only cross-links to PsaD, and the proteolysis of the purified cross-linked product identified a covalent bond between Glu93 of Fd and Lys106 of the PsaD subunit (Lelong *et al*., 1994). Backscattering interferometry data also suggests that a single amino acid of PsaE is crucial for Fd binding to PSI (Setif *et al*., 2010). However, there has been no such work investigating this interaction in organisms that have high-resolution crystal structures, such as pea or *Thermosynechococcus elongatus*. In particular, the availability of both an X-ray structure for PSI (Jordan *et al*., 2001) and NMR structures for Fd (Baumann *et al*., 1996; Hatanaka *et al*., 1997), plus the ability to genetically transform *T. elongatus* (Onai *et al*., 2004), make it an ideal organism to investigate potential interactions of Fd with residues in PsaE using computational methods. Furthermore, the thermophilic nature of this organism may also result in a different interaction modality than what is observed in mesophilic organisms (Kannan & Vishveshwara, 2000; Zeldovich *et al*., 2007) and is of considerable interest to biotechnology (Nguyen & Silberg, 2010). Because these experiments were conducted at a higher growth temperature (55 °C), it may also be important to consider the possibility that interactions should be stronger at biologically relevant temperatures, and even when cross-linked, should last longer than at the operating temperatures of the devices used.

Although there have been many approaches applied to elucidate the mode of interaction between Fd and PSI, the lack of a high-resolution crystal structure of the binary complex has prevented a detailed model. Additionally, the intrinsically transient nature of this interaction, as well as potential structural changes associated with the redox change of the two components, complicates the formation of the complex. In the present work, computational approaches enable the integration of experimental data into predictive models describing how Fd can dock to PSI. The energy landscape of the numerous different configurations that proteins adopt via folding, as well as in binding with other proteins, is described using a statistical mechanical description of the potential energy surface (Onuchic *et al*., 1997). When folding or binding, the tendency of a protein is to adopt a minimally frustrated state, such that the free energy decreases as configurations become more stable and presumably approach the native state (Jenik *et al*., 2012). While the global configuration of folded proteins tends to drive toward this minimally frustrated state, individual regions of the protein can exhibit more highly frustrated regions. The presence of these highly frustrated residues suggests that the corresponding regions are not in their lowest energy configuration, which could potentially be alleviated by binding to another protein. The "frustration concept" implementation of energy landscape theory (Jenik *et al*., 2012) provides a computational algorithm that identifies these regions of high localized energetic frustration in protein structures. The use of this algorithm in identifying protein interaction sites was employed recently to identify key intra-protein binding regions between two helices and the C-terminal region of human frataxin (Roman *et al*., 2012), as well as in the identification of the copper and zinc binding sites in the superoxide dismutase-1 enzyme (Das & Plotkin, 2013). In the present manuscript, we likewise address protein–protein interactions, characterizing the energetic frustration in PSI and Fd proteins to predict potential binding sites on the surface of the proteins and to guide rigid body docking calculations. The goal of the present work is not to investigate every possible protein–protein complex structure through high throughput protein docking approaches but to integrate specific mutagenesis experimental results with protein frustration calculations to limit the number of protein–protein conformations to be investigated, as well as to rank these limited possible complex structures. In this approach, experimental considerations are used not to post-process a large number of computational scenarios but rather to generate an ensemble in conjunction with computational methods that can then be further refined and evaluated using energetic criteria. Essentially, we are simplifying the computational

docking methods necessary by utilizing a more directed-docking, knowledge-based approach based on known biochemical data that Fd binds only to the three stromal subunits (PsaC, PsaD, and PsaE). These results can in the future be used to prioritize and design mutations that may enhance the affinity and stability of these binary complexes to enable more efficient PSI-based biohydrid devices for, such as, direct solar energy conversion.

RESULTS

Mapping ferredoxin-photosystem I interaction using chemical cross-linking

The cross-linking product was analyzed by sodium dodecyl sulfate polyacrylamide gel electrophoresis and immunoblotting with several different antibodies, including anti-Fd, anti-PsaD, and anti-PsaE (Figure 1). These antibodies were raised by our laboratory in rabbits using recombinant forms of the proteins expressed and purified from *Escherichia coli* using the impact system (NEB, Ipswitch, MA). The specificity of these antibodies via western blotting against the isolated PSI complex and Fd was confirmed (see Supplementary Information, Figure S1). As shown in Supplementary Figure S1, A–C, the three different antibodies (anti-Fd, anti-PsaD, and anti-PsaE) do not cross-react with the other subunits. The estimated molecular weight of each subunit or protein is listed—PsaC: 8.8 kDa, PsaD: 15.4 kDa, PsaE: 8.4 kDa, and Fd: 11 kDa (data from Protein Data Bank, PDB accession coordinates *1JB0* and *2CJN*). Several bands are shown in the cross-linking product using anti-Fd antibody. Fd runs higher than its estimated molecular weight in gel electrophoresis. The resulting bands are identified by comparing to the western blot result between different antibodies (anti-Fd, anti-PsaD, and anti-PsaE) and shows that the PsaD is the major subunit, while PsaC is the minor one when Fd is cross-linked to PSI. If PSI and Fd are cross-linked for a longer duration (5 s), a band representing PsaE cross-linked with Fd results. A lighter band representing PsaC cross-linked with Fd also develops as well. Thus, we assume that the docking site of Fd on PSI is primarily located on PsaD and with contributions from PsaC and PsaE.

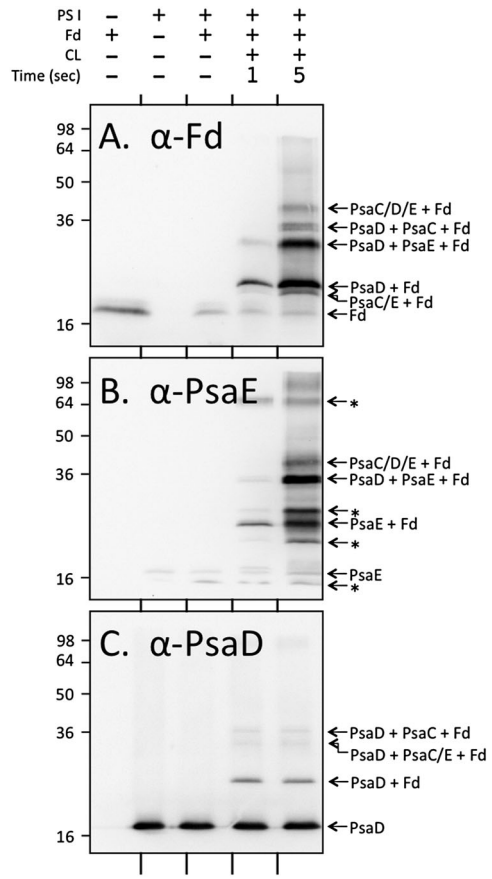

Figure 1. Immunodetection of cross-linked proteins. Three different antibodies were used to detect the cross-linked species. On the basis of the immunodetection and apparent MW of the cross-linked adduct, the composition of the reactive band is shown on the right of each blot. For each panel (a, b, and c), the first lane is ferredoxin (Fd) alone (20% of the input to the cross-linking), the second lane is photosystem I (PSI) alone, the third lane is PSI and Fd without the EDC/NHS cross-linking agent, the fourth lane is the cross-link result of PSI-Fd with a 1 s reaction time, and the fifth lane is the cross-link result of PSI-Fd with a 5 s reaction time. As described in the methods, the cross-linked PSI was re-isolated over a sucrose gradient to remove the free Fd. The bands labeled with an asterisk (*) represent an uncharacterized cross-reactive species in the starting material.

Characterization of the docking surface of the stromal domain of PSI

All residues from the stromal domain of PSI (PsaC, PsaD, and PsaE) with a potential energy surface characterized by a local frustration density of 20% or greater are highlighted in red in Figure 2. For the most part, these highly frustrated residues are contained in five sections (labeled RI through RV in Figure 2) in the PsaC, PsaD, and PsaE complex, and two major sections in Fd. Of particular interest is RIII, that includes residues from Ile94 through Lys107 in PsaD, which also includes His95, Asp98, Glu103, Lys104, and Lys107—all of which were suggested by experimental site-directed mutagenesis and chemical cross-linking studies as potential binding sites for Fd (Hanley *et al*., 1996; Bottin *et al*., 2001), as was Arg109 (Bottin *et al*., 2001), which is located just upstream in the sequence of PsaD from the region of high local frustration. Region II also exhibits a high degree of local frustration in a short segment from residues Glu70 to Asn78. These regions are shown on the right side of Figure 3 in the X-ray crystal structure of PSI (PsaC, PsaD, and PsaE)—indicated by a pink ribbon. These sections are located adjacent to the exposed iron-sulfur (FeS) centers of PsaC. The other section of high localized frustration in PsaD includes the residues Pro123 through Pro138, which interact with chains A and B of the PSI complex (Supplementary Information, Figure S2). Because many of these residues present at the interfaces between the stromal domain (PsaC, PsaD, and PsaE) with PsaA and PsaB were calculated to display a high degree of frustration, this suggests that frustrated residues in proteins may be found at protein–protein interfaces and that the minimal frustration concept could be used to identify binding sites for Fd on PsaC, PsaD, and PsaE.

Configurational energetic frustration of surface residues suggests docking sites of photosystem I

The region of high local frustration predicted for PsaC, from residues Tyr67 through Leu78, is shown at the bottom of the PsaC structure in Figure 3. This section is also in contact with PSI, as are residues Thr46 through Val55 of PsaE, located on the

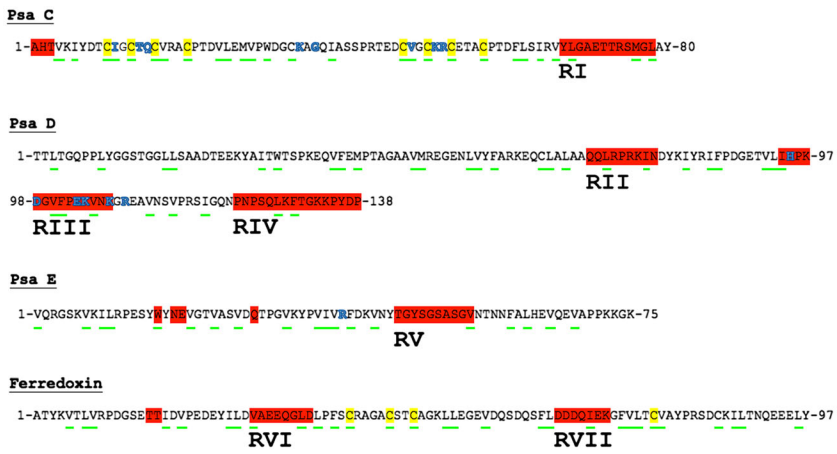

Figure 2. Sequence of PsaC, PsaD, and PsaE and Fd proteins. Cysteine residues are highlighted in yellow. Residues with a frustration density of 20% or greater are highlighted in red. Interaction residues predicted by site-directed mutagenesis experiments are colored in blue. Hydrophobic residues are underlined with green lines.

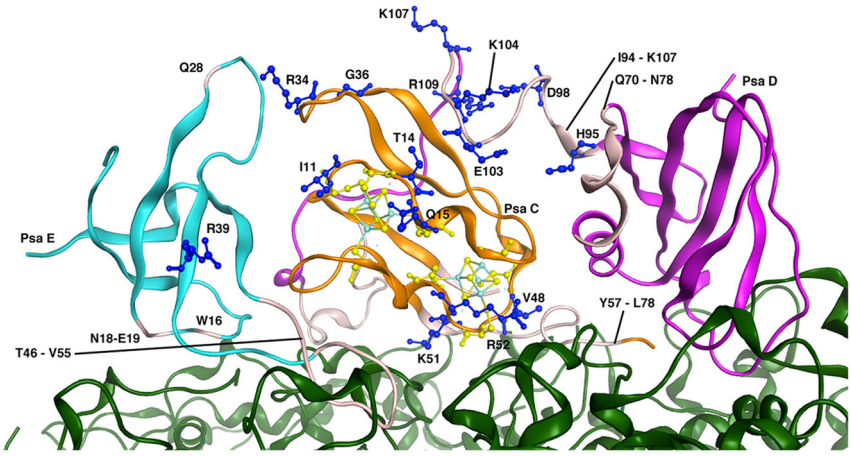

Figure 3. X-ray crystal structure of the photosystem I (PSI), PsaC, PsaD, and PsaE chains used in the present model. The transmembrane domains of PSI (removed from the model) would be located at the bottom of the figure. Cysteine residues 10, 13, 16, 20, 47, 50, 53, and 57 of PsaC are drawn with yellow lines, and the iron-sulfur complexes are drawn with yellow and cyan ball-and-stick models. Interaction residues predicted by previous site-directed mutagenesis experiments are drawn as ball-and-stick models in blue. Residues with a frustration density of 20% or greater are indicated by light pink ribbons. Strands of PsaA and PsaB are drawn in green ribbons for perspective.

other side of the exposed FeS centers of PsaC (Supplementary Information, Figure S2). These domains are interacting with the core subunits PsaA and PsaB of PSI that were deleted from the model in the present calculations. The fact that this region of the isolated stromal domain (PsaC, PsaD, and PsaE) is highly frustrated is a good control here of the relevance of the frustration principle in this system, because it demonstrates that it is capable of identifying regions that are experimentally known to interact with the remainder of the PSI structure.

The Lys34 residue of PsaC and the Arg39 residue of PsaE are not predicted to be as highly frustrated but have been suggested to be potentially involved in PSI/Fd interactions based on previous experimental studies. Site-directed mutagenesis and chemical cross-linking studies have shown that the Lys34 residue of PsaC is potentially involved in docking Fd (Fischer *et al*., 1998). Backscattering interferometry data also indicate that the Arg39 residue of PsaE is also possibly implicated in binding Fd (Setif *et al*., 2010). These residues are both labeled in Figure 3.

Residues on the exposed surface of PsaC in the vicinity of the 4Fe-4S centers (Cys10 through Cys20 and Cys47 through Cys57) do not appear to exhibit a high degree of localized frustration. However, site-directed mutagenesis data suggest several residues of PsaC that could be involved in binding Fd, including Ile11, Thr14, Gln15, Lys34, Gly36, Val48, Lys51, and Arg52 (Fischer *et al*., 1999; Setif *et al*., 2002). These residues are indicated in blue (cf. Figures 2 and 3). If the map of frustrated residues in PSI is merged onto the map of the position of residues shown to be important in experimental studies, a clear region defining a potentially reactive surface can be identified that is centered around the FeS centers of PsaC and bounded by the two regions of high local frustration in PsaD (Figure 3). It is this region of PsaD, where the frustrated residues overlap with experimental residues of interest that was used as the primary target for our rigid body docking calculations. Therefore, this model predicting the docked configuration of Fd and PSI is a knowledge-based model that combines previous experimental data with

Journal of Molecular Recognition

computational predictions of energetic frustration to provide a clearer picture of the overall characteristics of the binding surface to direct future studies.

In the same manner that configurational energetic frustration was used to define the docking surface on the stromal domain of PSI, two areas on the surface of Fd can be identified as potential interaction sites. The residues from Fd calculated to have a configurational frustration density of 20% or greater are also highlighted in red in Figure 2. Both the *2CJN* and *1ROE* NMR structures of Fd were analyzed, and high frustration density was calculated in similar regions of both models (Supplementary Information, Figures S3 and S4). However, the strands in these potential docking regions were observed to be more unordered in *1ROE*, and the *2CJN* structure was chosen for docking due to its organization of these crucial stands into alpha-helices, resulting in a better placement of the Asp and Glu residues for potential interactions with PSI.

The configurational energetically frustrated residues of Fd are concentrated in two sections positioned on opposite sides of the Cys residues constituting the active site (Cys40, Cys45, Cys48, and Cys78) (Figure 4). These highly frustrated sections contain several negatively charged residues—Glu30, Glu31, Asp35, Asp66, Asp67, Asp68, and Glu71. These negatively charged residues can be targeted to positively charged residues in the PsaC, PsaD, and PsaE complexes, such as Lys34 (PsaC), Arg39 (PsaE), Arg73, Lys76, Lys104, and Lys107 (PsaD). In addition to these two sites, the location of hydrophobic residues in the vicinity of the four cysteine residues involved in electron transfer is also shown (green ribbons), as hydrophobic residues on the surface of proteins may play a role in protein–protein binding sites (Figure 4).

Correlation between energetic frustration and conservation of residues

In the study of protein–protein interactions, it is generally expected that the residues at the interface be highly conserved across multiple species. To explore this hypothesis in PSI proteins, the sequences of PsaC, PsaD, and PsaE and Fd in 12 species were obtained from the UniProt Database (http://www.uniprot.org/) and aligned with T-coffee software, freely available on the Internet (EMBL-EBI, Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SD, UK. http://www.ebi.ac.uk/Tools/msa/tcoffee/). While there is evidence of sequence conservation for these proteins, there was no strong correlation between sequence conservation and energetic frustration calculations (Supplementary Information, Figure S5). This suggests that the basis for recognition between PSI and Fd involves complex, non-linear motifs that vary in their details between different species. The use of energetic frustration should help as an additional method in the identification of those interactions independent of sequence conservation.

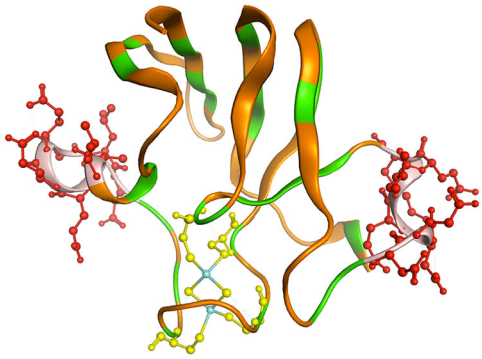

Figure 4. Modeled structure of ferredoxin (based on two NMR structures) used for docking calculations. Cysteine residues 40, 45, 48, and 78 are drawn in yellow lines, connected to the iron-sulfur center drawn with yellow and cyan ball-and-stick models. Residues with a frustration density of 20% or greater are drawn as red ball-and-stick models with light pink ribbons. Hydrophobic residues are indicated by green ribbons.

Rigid body directed-docking

Twelve docking pairs were obtained from rigid body docking calculations (Table 1), as described in Materials and Methods. All twelve docking pairs can be downloaded as Supplementary Information in PDB format. The non-bonded potential interaction energy, (i.e., the sum of van der Waal's and electrostatic potential interaction energies calculated with the CHARMM27 force field as described in Materials and Methods), was used to rank each pose; the docked structures with the most favorable (most negative) potential interaction energies between PSI and Fd are thermodynamically more likely to exist. Each fully minimized complex was structurally superimposed, and the root-mean-square deviation (RMSD) between each of the 12 docking pairs was calculated using all heavy atoms of PsaC, PsaD, and PsaE and Fd (Figure 5). The 12 docking pairs do not correspond to the exact same configurations but are nonetheless clustered into three sets by superimposition of the main chain heavy atoms of each respective protein pair and comparison by RMSD (Figure 6). Cluster I consists of docking pairs **2**, **4**, **5**, and **7** (cluster I), which are oriented with the Val28 to Asp35 residues of Fd in proximity to the Gln70 to Asn78 residues of PsaD. Cluster II consists of docking pairs **1**, **3**, **6**, **10**, **11**, and **12** (cluster II), and they are oriented with the Asp66 to Lys72 residues of Fd in proximity to the Gln70 to Asn78 residues of PsaD. A third cluster, III, consists of docking pairs **8** and **9** (cluster III), which are oriented in the same direction as those of cluster II but centered over residues lysine-34 of PsaC and Lys104 of PsaD instead of directly over the FeS site of PsaC and interacting with Arg39 of PsaE. The RMSD values shown in Figure 5 appear to be clustered similarly to the orientation of Fd docked with PSI.

The two docking pairs (**2** and **3**) yield the most favorable PSI/Fd non-bonded interaction energies, with both differing by ~96 kcal/mol, yet corresponding to entirely different interacting structures, with an RMSD between the two of 9.9 Å. Interestingly, Fd is docked into PSI in opposite directions in each of these two docking pairs—with the frustrated protein sequence from residues Asp66 through Lys72 in proximity to the same Ile94 to Lys107 section of PsaD as in the docking pair **3**, and the frustrated protein sequence from residues Val28 through Asp35 in proximity to the frustrated protein sequence of PsaD from residues Gln70 through Asn78 in docking pair **2**. There are a total of 17 inter-residue protein contacts in **3** as identified by the program MOE. Inter-residue protein contacts were defined as hydrogen bonds of 3 Å or less, non-bonded polar interactions of 4.5 Å or less, or non-bonded hydrophobic contacts of 4.5 Å or less. Seven contacts involve residues calculated to possess high local frustration, and eight contacts involve residues predicted as interaction sites suggested experimentally. Docking

Table 1. Non-bonded interaction energies calculated for photosystem I/ferredoxin docked complexes, sorted in order of lowest to highest total interaction energy

Docked complex	van der Waal's interaction energy	Electrostatic interaction energy	Total interaction energy	Interaction energy difference from 3	F_B–F_{Fd} distance	Cluster
3	2.2	−502.9	−500.6	0.0	9.7	II
2	22.4	−426.9	−404.5	−96.2	11.0	I
6	17.5	−413.2	−395.7	−105.0	11.0	II
10	11.3	−406.8	−395.5	−105.2	10.6	II
5	21.1	−391.3	−370.2	−130.5	11.7	I
4	14.3	−351.9	−337.6	−163.0	11.3	I
12	− 7.6	−312.2	−319.8	−180.8	11.1	II
1	3.5	−318.9	−315.5	−185.2	11.4	II
7	22.4	−337.3	−314.9	−185.8	13.5	I
9	6.0	−296.0	−290.0	−210.7	14.9	III
8	2.5	−270.5	−268.0	−232.6	14.2	III
11	6.1	−247.2	−241.1	−259.6	12.3	II

All energies are in kcal/mol. The fifth column shows the difference in total interaction energy between the highest complex (**3**) and the docked complex indicated. The sixth column shows the distance, in Angstroms, between the closest iron atoms of the Fe_2S_2 complexes in photosystem I and ferredoxin. The seventh column shows the root-mean-square deviation cluster, based on structural superimposition of all heavy atoms of each docked complex with the other docked complexes.

	1	2	3	4	5	6	7	8	9	10	11	12
1		9.9	2.2	10.0	9.9	1.7	10.3	4.3	4.3	1.4	3.1	1.6
2	9.9		9.9	2.5	2.1	9.8	2.9	9.6	9.6	9.9	10.1	10.3
3	2.2	9.9		10.1	9.8	1.3	10.3	3.3	3.3	1.6	3.1	2.6
4	10.0	2.5	10.1		2.9	10.0	2.9	10.1	10.1	10.1	10.1	10.4
5	9.9	2.1	9.8	2.9		9.8	2.9	9.9	9.8	9.8	10.1	10.1
6	1.7	9.8	1.3	10.0	9.8		10.2	3.3	3.4	1.1	3.0	2.4
7	10.3	2.9	10.3	2.9	2.9	10.2		9.9	9.9	10.4	10.3	10.8
8	4.3	9.6	3.3	10.1	9.9	3.3	9.9		0.8	3.5	3.4	5.1
9	4.3	9.6	3.3	10.1	9.8	3.4	9.9	0.8		3.5	3.4	5.1
10	1.4	9.9	1.6	10.1	9.8	1.1	10.4	3.5	3.5		2.7	2.2
11	3.1	10.1	3.1	10.1	10.1	3.0	10.3	3.4	3.4	2.7		3.7
12	1.6	10.3	2.6	10.4	10.1	2.4	10.8	5.1	5.1	2.2	3.7	

Figure 5. Root-mean-square deviations (RMSDs) of each docked PsaC, PsaD, and PsaE/ferredoxin complex superimposed against each other. The colors range from the lowest RMSD (dark green) to the highest RMSD (dark red).

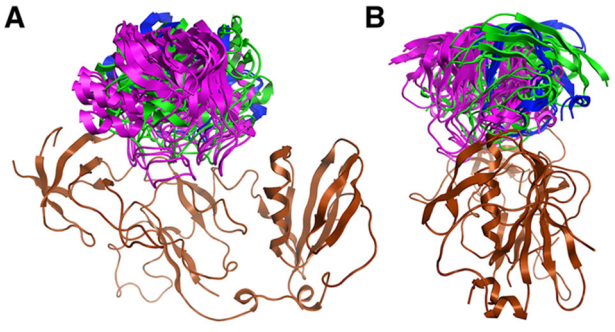

Figure 6. Structural superimposition of PsaC, PsaD, and PsaE showing the different docked configurations of ferredoxin. Cluster I (docking pairs 2, 4, 5, and 7) is drawn in green ribbons, cluster II (docking pairs 1, 3, 6, 10, 11, and 12) is in magenta, and cluster III (docking pairs 8 and 9) is in blue ribbons. A is the same view of PsaC, PsaD, and PsaE as Figure 3; B is rotated 90°.

pair **2** has 12 inter-residue protein contacts, with six of these involving highly frustrated residues and six involving experimentally suggested interacting residues.

In docking pair **3** (Figure 7), the majority of contacts involve salt bridges and electrostatic interactions, with relatively few hydrophobic interactions. This is also clear from looking at the energy in Table 1. This arrangement places one of Fd's iron atoms 9.7 Å from the F_B FeS center of PsaC in **3** (Table 1). This closely compares with the Fe–Fe distances of 8.9–13.1 Å between the two FeS centers F_A/F_B and F_X/F_A) of PsaC observed in six available crystal structures of PSI from *T. elongatus*, *Pisum sativum*, and *Arabidopsis thaliana* (Supplementary Information, Table S2) (Jordan *et al*., 2001; Amunts *et al*., 2007; Amunts *et al*., 2010; Chapman *et al*., 2011; Brunger *et al*., 2012). This close placement suggests that the docking configuration for this pair may provide an adequate proximity for rapid electron transfer (Page *et al*., 1999).

The Lys76 residue of PsaD is engaged in several strong salt-bridge interactions with both the Asp66 (2.5 Å) and Asp68 (bifurcated, 2.6 Å and 2.9 Å). Asp68 is also engaged in a hydrogen bonding interaction with the Gln69 residue (3.0 Å), internal to Fd. This appears to contribute to the further stabilization of this region of Fd. Two other residues in the region of high localized frustration of PsaD are also involved in interactions with Fd. They are Lys104, which is involved in a bifurcated salt bridge with the terminal carboxyl of Tyr97 (2.6 and 2.6 Å), and Lys107, which

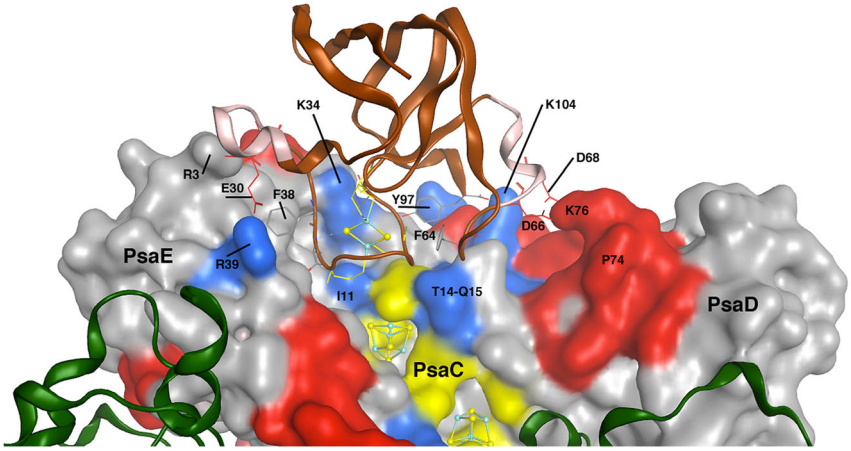

Figure 7. The docking interface between PsaC, PsaD, and PsaE and Fd for docked pair 3. PsaC, PsaD, and PsaE are drawn as a gray surface with blue patches representing residues predicted by experimental data and red patches representing residues with a frustration density of 20% or greater. Ferredoxin is drawn as a brown ribbon representation with light pink ribbons and red atoms representing residues with a frustration density of 20% or greater. The FeS centers of PsaC are represented as ball-and-stick models. Strands of PsaA and PsaB are drawn in green ribbons for perspective.

forms a salt bridge with Glu94 (2.6 Å). The Arg18 residue of PsaC also forms a bifurcated hydrogen bond (2.6 and 2.9 Å) with the carboxyl oxygen of Phe64, which is located two residues downstream in Fd from one of the regions of high localized frustration. Arg18 is also hydrogen-bonded to the Glu15 residue (2.6 Å) internal to PsaC. This residue interacts with Ala44 of Fd (2.9 Å), and is adjacent to its FeS center. The two Cys residues on the surface engage in weak van der Waal's interactions with the PsaC Cys13 residue (4.2 Å) and PsaC Pro58 (4.6 Å), on either side of one of the FeS centers of PsaC. The Lys34 residue of PsaC is involved in several bridging interactions between Fd and the neighboring PsaE chain of PSI. It is involved in a hydrogen bond with Thr47 (2.7 Å) as well as an electrostatic interaction with Glu93 (3.4 Å). It is also involved in a salt bridge with Asp27 (2.5 Å) of PsaE. This residue is adjacent to the highly frustrated Gln28 residue.

Most of the interactions between PsaE and Fd involve the other region of high localized frustration on Fd, from Val28 to Asp35. Asn39 of PsaE forms a hydrogen bond with Glu30 (2.6 Å). Glu30 also engages in a hydrogen bond with Arg3 (2.6 Å), which in turn is hydrogen-bonded to Asp35 of Fd (2.5 Å), which is hydrogen-bonded with Ser25 of PsaE (2.5 Å). Phe38 of Fd is also involved in weak van der Waal's interactions with Ile37 (3.7 Å) and the hydrophobic chain of Arg39 (4.8 Å). Adjacent to this Phe, the Ser39 residue forms a hydrogen bond with Thr57 of PsaE (2.6 Å).

Docking pair **2** (Figure 8) is oriented in the opposite direction as docking pair **3**, with the Val28/Asp35 frustrated residues of Fd in proximity to the frustrated residues of PsaD. This orientation puts the distance between the nearest iron atoms of the FeS complexes of Fd and PsaC at 11.0 Å. Asp35 of Fd engages in a strong, bifurcated salt bridge with Lys76 of PsaD (2.6 Å, 2.8 Å). Glu30 also interacts with the Lys104 residue of PsaD (2.9 Å). This

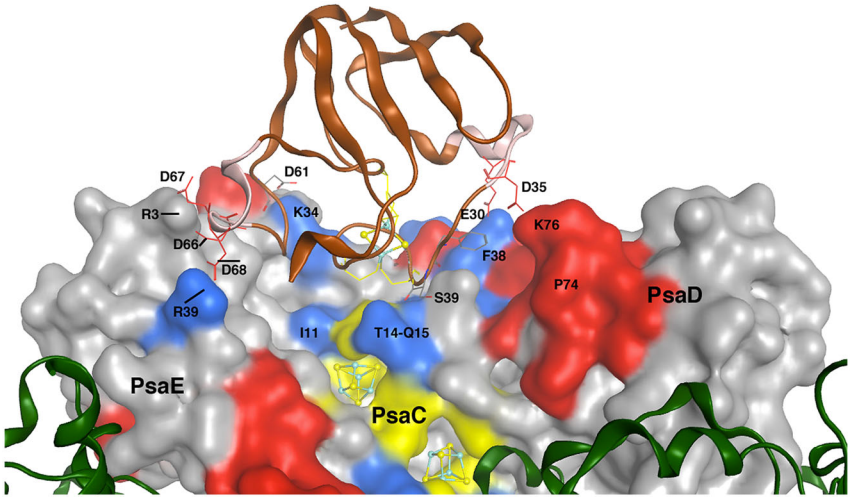

Figure 8. The docking interface between PsaC, PsaD, and PsaE and Fd for docked pair 2. PsaC, PsaD, and PsaE is drawn as a gray surface with blue patches representing residues predicted by experimental data and red patches representing residues with a frustration density of 20% or greater. Ferredoxin is drawn as a brown ribbon representation with light pink ribbons and red atoms representing residues with a frustration density of 20% or greater. The FeS centers of PsaC are represented as ball-and-stick models. Strands of PsaA and PsaB are drawn in green ribbons for perspective.

interaction is further stabilized by several interactions that are internal to Fd, including a bifurcated salt bridge between Glu30 and Arg41 (2.6 Å), and a hydrogen bond between Asp27 and Arg41 (2.5 Å). The Phe101 residue of PsaD interacts in a weak van der Waal's interaction with Phe38 (5.8 Å), and the carboxyl oxygen of this residue hydrogen bonds with Arg18 (2.7 and 2.8 Å).

There are also several interactions in the vicinity of the FeS centers in docking pair **2**. Cys40 of Fd is involved in hydrophobic interaction with Cys13 (4.6 Å) and Thr14 (3.6 Å) of PsaC. Ser39 engages in a hydrogen bond with Glu15 (2.8 Å), while Fd engages in several strong interactions with residues in PsaE. The Lys34 residue of PsaC also appears to bridge interactions between PsaC and PsaE with Fd, forming a hydrogen bond with Asp61 (2.6 Å) as well as with the PsaE Asp27 (2.6 Å and 3.2 Å) and Thr29 (2.7 Å) residues. In the region of high localized frustration, Asp66 engages in a bifurcated salt bridge with Arg3 (2.5 and 2.6 Å). The adjacent Asp67 is also involved in a hydrogen bond with this Arg as well (2.5 Å). And Asp68 forms a strong, bifurcated salt bridge with Arg39 (2.5 and 2.5 Å).

Docking pairs **6** and **10** were in a similar configuration to the strongest binding docked pair **3**, although with −105.0 kcal/mol and −105.2 kcal/mol less favorable PSI/Fd interaction energies and with distances between the FeS centers of 11.0 and 10.6 Å, respectively. Docking pair **5** displayed similar configuration to **2**, with non-bonded interaction energy of 130.5 kcal/mol less than **3** and an FeS center distance of 11.7 Å. The non-bonded interaction energy (Table 1) of the remaining docking pairs between PSI and Fd was considerably less favorable than the first docking pairs and apparently exhibit interaction energies correlated with the distance between the nearest iron atoms in the FeS complexes of PSI and Fd.

DISCUSSION

The rigid body directed-docking calculations yield several observations with regard to possible conformations of Fd docked into PSI. Docking calculations focused predominantly on two regions of high localized frustration in PsaD as potential interaction sites for Fd. These two regions include residues from Glu70 through Asn78 and Ile94 through Lys107 (shown in sequence in Figure 2 and visualized structurally in the upper right corner of Figure 3). Interestingly, several residues in the sequence from Ile94 through Lys107 have also been determined experimentally to be involved in binding Fd—His95, Asp98, Glu103, Lys104, Lys107, and Arg109 (Hanley *et al*., 1996; Bottin *et al*., 2001). This corroboration is strengthened by the aforementioned observation that other regions of high localized frustration in PsaC and PsaE—as well as the third frustrated region in PsaD (Pro123 through Pro138)—are also located in areas that interact with PsaA and PsaB in the PSI complex. This provides a control within the energetic frustration calculations. Because Regions I, IV, and V are known to be present at the protein–protein interface between PsaC and PsaE with PsaA, PsaB, and PsaF (Supplementary Information, Figure S2), the remaining regions of PsaC, PsaD, and PsaE displaying high localized frustration may constitute the interface with Fd.

The presence of six Arg and Lys residues in these two frustrated regions of PsaD is particularly interesting as these residues have been found in the interface region of other protein–protein binding sites (Li *et al*., 2004). As shown in Figures 1 and 3, there are also several Asp and Glu residues concentrated in the two regions of high localized frustration on Fd. These are the residues that our PsaC, PsaD, and PsaE cross-linking studies will target, yet that is still a nonspecific process and it will not yield a specific result. The rigid body directed-docking calculations attempt to position these two regions in contact to fulfill the potential formation of high affinity salt bridges between the Arg or Lys residues of PSI with the Asp or Glu residues of Fd. This is illustrated in the two docking pairs with the lowest non-bonded interaction energy (**3** and **2**) and may also represent a region on both proteins that could be used in an attempt to engineer an increased affinity for *in vitro* docking.

The previous data accounts for the docking of only one of the highly frustrated regions of Fd; however, there are two highly frustrated regions on each side of the electron transfer site on Fd (Figure 4). While there is another region of high localized frustration present on PsaE shown in the bottom of Figure 3, these residues also interact with the unmodeled regions of PSI: PsaA, PsaB, or PsaF. Furthermore, a close analysis of the types of residues present in this region (Figure 2) indicates that there are no strong hydrogen bond donors present, which are capable of forming potential salt bridges (Arg or Lys residues). There are three serine residues present in this region, which could form hydrogen bonds, but they are not as strong as a salt bridge from an Arg or Lys. Previous site-directed mutagenesis studies have shown that removal of the Arg39 residue of PsaE, located adjacent to the region containing the FeS centers of PsaC, significantly decreases the interaction of Fd with PSI (Barth *et al*., 2000). Utilizing this Arg39 residue as a restraint in the rigid body docking calculations also enables Fd to be positioned in proximity to the Lys34 residue of PsaC for which there is also experimental evidence of its importance in binding Fd (Fischer *et al*., 1998). The docking calculations indicate that Lys34 is positioned at the junction between PsaC and PsaE and Fd. The Arg3 residue of PsaE is also located near Arg39. While there is currently no experimental evidence available that suggests this residue is involved or important in binding Fd, it is predicted to be involved in salt-bridge interactions with highly frustrated Asp and Glu residues in several docked complexes. This suggests that Arg3 may contribute to stabilizing Fd binding and contribute to higher affinity with PSI.

Docking these two regions of high frustration in Fd with the frustrated regions of PsaD and the Arg39/Lys34/Arg3 residues of PsaC and PsaE enables the 2Fe-2S center of Fd to be placed within 14 Å of the F_B FeS center present in PsaC, as necessary for electron transfer (based on analysis of structures in the PDB) (Page *et al*., 1999). For each of the docking pairs of PSI/Fd with the lowest non-bonded interaction energy, the distance between the FeS center of Fd and the closest FeS site in PSI is less than 14.0 Å (Table 1). While there are some favorable van der Waal's interactions in each of the docking pairs, the major component of the non-bonded interaction energy is electrostatic interactions, as shown by comparing the van der Waal's and electrostatic energy in Table 1. This is somewhat contradictory to the general expectation that hydrophobic interactions are predominantly involved in protein–protein interactions. However, given the extremely fast time-scale of electron transfer reactions (femtoseconds), the lack of a strong hydrophobic fit should enable Fd to bind, collect electrons, and be dissociated from the solvent quickly. Additionally, many of the residues

identified by experiments as being important to Fd binding are predominantly electrostatic in nature as well (Ullmann *et al*., 2000; Setif, 2001). In addition, the treatment of electrostatic interactions in the present calculations was done using a distance-dependent dielectric that represent an average implicit solvent effect that may overemphasize Coulombic interactions on a short distance scale.

There is some disagreement about the orientation of Fd in the binding site between each of these three pairs. In docking pair **3** (Figure 7), Fd is oriented with the loop containing its four Cys residues facing away from the gap between Lys34 of PsaC and Lys104 of PsaD, while this loop is facing the opposite direction in docking pair **2** (Figure 8). The interaction energy (Table 1) favors docking pair **3** by 96.2 kcal/mol, and the distance between the catalytic FeS sites is 1.3 Å closer in pair **3**. Docking pair 3 is a member of cluster II, and three of the docking pairs with the highest interaction energy are also part of this cluster. Cluster III does not appear to display a favorable orientation, with docking pairs **8** and **9** at the low end of the interaction energy out of the **12** docking runs (Table 1). Docking pairs **2**, **5**, and **4** in Cluster I also show favorable non-bonded interaction energy and yield very plausible configurations of Fd docked with PSI. It is beyond the scope of this paper to designate which conformation is the "correct" one, if any (as several conformations may exist in equilibrium)—this would require molecular dynamics or Monte Carlo simulations to investigate the stability of the models and use of explicit hydration models that were beyond the scope of the present calculations. Experimental cross-linking or covalent attachment of Fd to different parts of the stromal domain involving residues in the highly frustrated regions of Fd may be able to elucidate this further (Wittenberg *et al*., 2013).

Overall, this study provides a new, detailed insight into visualizing possible binding mechanisms between PSI and Fd. These results, if experimentally verified, demonstrate that the concept of protein frustration embodied in energy landscape theory is useful in predicting potential binding sites between two or more proteins. The results suggest several potential structures for Fd bound to PSI. While in the present calculations one orientation exhibits significantly more favorable potential interaction energy than the other Fd orientations investigated, entropic contributions to the binding free energy, not investigated here, as well as hydration effects and redox changes in the FeS complexes, could lower the free energy differences between several orientations of Fd. It is also possible that these effects, in addition to internal structural changes of Fd and PSI, would affect the predicted binding site and orientation of Fd on PSI. For instance, the two NMR structures of Fd (*2CJN* and *1ROE*) indicate significant configurational differences in regions of the protein calculated to exhibit high degrees of localized frustration, as well as some structural variations that render binding of the *1ROE* structure difficult to address with the methods described here. It is possible that Fd may bind in two or more different orientations with PSI. Further analysis of these complexes and their nanosecond time-scale equilibrium could be obtained from molecular dynamics analysis.

Finally, this directed-docking computational approach, which combines site-directed mutagenesis, chemical cross-linking, and electron transfer measurements, provides a unique view of the protein–protein interactions involved in the interaction of Fd with PSI. The combination of computational and experimental models provides valuable information that enables researchers to develop methods that yield significantly improved rates of electron transfer and quantum yields in applied photosynthesis applications (Wittenberg *et al*., 2013). The biological role of Fd requires it to shuttle electrons from PSI to FNR. The need to interact sequentially with two different binding partners limits the affinity for each interaction. However, for cell-free, applied photosynthesis applications, it should be possible to significantly increase the affinity of Fd for PSI. This bioengineering may yield tighter binding, resulting in significantly faster electron transfer.

MATERIALS AND METHODS

Building molecular models

Protein structures from *T. elongatus*, a thermophilic cyanobacterium, were used to model interaction site predictions between the X-ray crystal structure of PSI and the NMR structure of Fd. The structural files were downloaded from the PDB (Bernstein *et al*., 1977) using the PDB accession coordinates *1JB0* for photosystem I and *2CJN* for Fd. Using the program MOE, version 2012 (Chemical Computing Group, Inc., Montreal, Quebec, Canada), the PsaC, PsaD, and PsaE chains, including the FeS centers, were extracted from the PDB file, and hydrogen atoms were added to the structure according to standard residue protonation. The remainder of the system, constituting the membrane regions of PSI, was omitted, because it is not involved in binding with Fd. Because the *2CJN* structure did not include the FeS center of Fd, these atoms were obtained from the *1ROE* Fd NMR structure (Baumann *et al*., 1996). To insert the FeS center into *2CJN*, residues 37 through 48 of *1ROE* were superimposed in MOE with the same residues of *1ROE*, and the FeS cluster was copied from the *1ROE* structure into the *2CJN* structure. A total of 500 cycles of conjugate gradient energy minimization was applied to the full Fd protein, including the FeS cluster. The CHARMM27 force field, as implemented in MOE, was used for all energy minimization and calculation of atomic charges (MacKerell *et al*., 2001), including the addition of angle bending parameters for the FeS bonds based on the angles observed in the *1ROE* NMR structure of Fd (Supplementary Information, Table S3).

Purification of ferredoxin and photosystem I

Photosystem I was isolated from *T. elongatus* following the previously published protocol (Iwuchukwu *et al*., 2010). The gene for Fd (Gene ID: 1011313) was isolated from genomic DNA using PCR. This gene was then inserted in pTYBP2 vector and the protein was expressed and purified from *E. coli* using the IMPACT system (New England Biolabs). The PsaD and PsaE were also cloned from the genomic DNA and inserted into the pTYB2 vector, expressed and purified from *E. coli* using the IMPACT system (New England Biolabs).

Production of anti-Fd, anti-PsaD, and PsaE polyclonal antisera

These polyclonal antibodies were made in rabbits (Pocono Rabbit Farm, Canadensis, PA) using recombinant forms of Fd, PsaD, and PsaE expressed and isolated from *E. coli*. The characterization and specificity of these are shown in Supplemental Information, Figure S1. It is clear that there is no cross-reactivity between these antisera against the other subunits, although

there may be some detection of the CBD fusion protein that was used to purify the proteins.

Chemical cross-linking and immunodetection of cross-linked species

A total of 10 μl of 1 mg Chl/ml *T. elongatus* PSI was used, and *T. elongatus* Fd was added to make the molar ratio 1:1 in 0.1 × PBS (pH 7.2) buffer with 0.03% DDM. The mixture of PSI and Fd was then incubated for 30 min on ice before adding chemical cross-linking agent. EDC (Thermo Scientific) and NHS (Thermo Scientific) were selected to chemically cross-link PSI and Fd. The cross-linker was added to the mixture to make the final concentration 5 mM NHS and 2.5 mM EDC in a final volume of 50 μl. Equal volume (50 μl) of 20 mM glycine in 0.03% DDM was added to quench the cross-linking reaction. Following this, the cross-linking product was centrifuged in a sucrose gradient to separate free Fd from PSI and PSI-Fd cross-linked product. After sucrose gradient separation, 100 μl of the green layer (PSI) was isolated and mixed with sample solubilization buffer and incubated in 37 °C for 30 min to denature the protein. The protein sample was analyzed using 15% tris-glycine sodium dodecyl sulfate polyacrylamide gel electrophoresis and immuno-detected by western blotting. The immunoblottings were probed with different antibodies, including anti-Fd, anti-PsaD, and anti-PsaE (Figure 1).

Identification of protein surface characteristics of photosystem I and ferredoxin for directed-docking

To identify potential protein–protein interaction domains between the stromal subunit of PSI (PsaC, PsaD, and PsaE) and Fd, the program MOE was used to map out key residues that have been experimentally determined on the basis of site-directed mutagenesis, cross-linking, and other studies to be involved in interactions with both systems (Fischer *et al*., 1999; Bottin *et al*., 2001; Setif *et al*., 2002; Setif *et al*., 2010). The position of the cysteine residues bonded to the FeS center of Fd used for electron transfer (Hatanaka *et al*., 1997) and the two FeS centers in PsaC were identified in relation to these experimental residues of interest. The presence of hydrophobic surface patches on both PSI (PsaC, PsaD, and PsaE) and Fd was identified for use as a guide for Fd docking.

Regions exhibiting local configurational frustration density of 20% or greater in PSI and Fd were identified using the protein frustratometer (EMBNet, http://www.embnet.qb.fcen.uba.ar/embnet). The 20% threshold in frustration density was selected as the minimum necessary to differentiate actual calculated frustration from background "noise", thus maximizing the signal-to-noise ratio. In addition to the PsaC, PsaD, and PsaE chains of PSI, the *2CJN* and *1ROE* structures of Fd were also investigated using the protein frustratometer (Supplementary Information, Figures S3 and S4).

Table 2. Distance restraints used for each docking run

1	PsaE: R39 Cζ Fd: D27 Cγ	PsaE: R39 Cζ Fd: E31 Cγ	PsaE: R39 Cζ Fd: D35 Cγ	PsaC: C13 S Fd: C40 C	PsaD: R73 Cζ Fd: D66 Cγ	PsaD: K104 Nζ Fd: D68 Cγ	PsaD: K76 Nζ Fd: E71 Cδ
2	PsaD: R73 Cζ Fd: R35 Cγ	PsaD: K76 Nζ Fd: E30 Cδ	PsaC: C13 S Fd: C40 C	PsaE: R39 Cζ Fd: D66 Cγ	PsaE: R39 Cζ Fd: D68 Cγ	PsaC: K34 Nζ Fd: E71 Cδ	
3	PsaE: R39 Cζ Fd: E31 Cδ	PsaC: K34 Nζ Fd: D35 Cγ	PsaC: C13 S Fd: C40 C	PsaD: R73 Cζ Fd: D66 Cγ	PsaD: R18 Cζ Fd: D68 Cγ	PsaD: K104 Nζ Fd: D68 Cγ	PsaD: K76 Nζ Fd: E71 Cδ
4	PsaD: R18 Nζ Fd: E30 Cδ	PsaD: K76 Nζ Fd: E31 Cδ	PsaC: C13 S Fd: C40 C	PsaE: R39 Cζ Fd: D66 Cγ	PsaC: K34 Nζ Fd: E71 Cδ		
5	PsaD: R73 Cζ Fd: E30 Cδ	PsaD: K76 Nζ Fd: D35 Cγ	PsaC: C13 S Fd: C40 C	PsaE: R39 Cζ Fd: D66 Cγ	PsaE: R39 Cζ Fd: D67 Cγ	PsaE: I37 Cδ Fd: I70 Cγ	
6	PsaE: R39 Cζ Fd: E31 Cδ	PsaC: C13 S Fd: C40 C	PsaD: K104 Nζ Fd: D66 Cγ	PsaD: R73 Cζ Fd: D67 Cγ	PsaD: K76 Nζ Fd: E71 Cδ		
7	PsaD: R73 Cζ Fd: E30 Cδ	PsaD: K104 Nζ Fd: E30 Cδ	PsaD: K76 Nζ Fd: D35 Cγ	PsaC: C13 S Fd: C45 Cα	PsaC: K34 Nζ Fd: D67 Cγ		
8	PsaC: K34 Nζ Fd: E31 Cδ	PsaC: K34 Nζ Fd: D35 Cγ	PsaC: C13 S Fd: C45 N	PsaD: R73 Cζ Fd: D66 Cγ	PsaD: K76 Nζ Fd: D67 Cγ	PsaD: K104 Nζ Fd: E71 Cδ	
9	PsaC: K34 Nζ Fd: E31 Cδ	PsaC: K34 Nζ Fd: D35 Cγ	PsaC: C13 S Fd: C45 N	PsaD: R73 Cζ Fd: D66 Cγ	PsaD: K76 Nζ Fd: D67 Cγ		
10	PsaE: R39 Cζ Fd: E30 Cδ	PsaE: R39 Cζ Fd: E31 Cδ	PsaC: C13 S Fd: C45 Cα	PsaD: R73 Cζ Fd: D66 Cγ	PsaD: K104 Nζ Fd: D68 Cγ	PsaD: K76 Nζ Fd: E71 Cδ	
11	PsaC: C13 S Fd: C40 Cα	PsaD: K104 Nζ Fd: D66 Cγ	PsaD: K76 Nζ Fd: E71 Cδ	PsaD: R39 Cζ Fd: E31 Cδ			
12	PsaE: R39 Cζ Fd: E30 Cδ	PsaE: V55 Cβ Fd: F38 Cζ	PsaC: C13 S Fd: C40 Cβ	PsaD: R73 Cζ Fd: D66 Cγ	PsaD: K76 Nζ Fd: D67 Cγ	PsaD: K104 Nζ Fd: E71 Cδ	

For photosystem I proteins and ferredoxin, the table lists the residues and atoms used in the distance restraints described in the text.

Journal of Molecular Recognition

Rigid body directed-docking of photosystem I and ferredoxin

Protein–protein docking was performed using the energy minimize function in MOE with the CHARMM27 force field modified as described previously with Fd and the PsaC, PsaD, and PsaE proteins of PSI treated as separate rigid bodies. Both of the PDB files were loaded into MOE and arbitrarily separated by approximately 30 Å, with the predicted surfaces of interaction facing each other. Between four and seven distance restraints between selected amino acids were set for each docking run using a harmonic potential of 1 kcal/mol/Å^2 for atomic distances further away than 8 Å (Table 2). The distance restraints were set between one atom of a residue on PSI and another atom of a residue on Fd to insure that the two residues selected would be in proximity in the docked complex structure. The criteria for selection of distance restraints included a comprehensive selection of residues predicted by the protein frustratometer to have 20% or more of their protein contacts displaying high energetic frustration (e.g., a configurational frustration density of 20% or greater), as well as residues predicted by site-directed mutagenesis experiments as potential binding sites. These distance restraints were used to generate potential models for PSI/Fd complexes that bring together the residues that have been experimentally proposed to be involved in PSI/Fd interactions and/or to bring together regions of high energetic frustration. A total of 12 models were generated with a different set of distance restraints (Table 2), and each of these models was then subjected to energy minimization in MOE, with each protein subunit treated as a rigid body and subjected to distance restraints set to 1 and 8 Å as listed in Table 2 until the energy gradient was less than 0.01 RMS kcal/mol/Å^2. The distance restraints were then released, and all of the PSI/Fd complexes were subjected to fully unrestrained energy minimization using a distance-dependent dielectric constant to approximate solvent effects on electrostatic interactions, until the energy gradient was less than 0.01 RMS kcal/mol/Å^2. The non-bonded interaction energy was calculated as the sum of electrostatic and van der Waals interaction between PSI and Fd and used for comparing each docked complex. The protonation state of residues in both the predocked and docked complexes was also calculated using the Protonate3D facility in MOE based on pKa calculations at T = 300 K, pH 7, and salt concentration 0.1 M. This protonation state was found to be identical for all residues in the isolated species as well as in the protein–protein complexes.

Acknowledgements

B. D. B. acknowledges support from TN-SCORE, a multi-disciplinary research program sponsored by NSF-EPSCOR (EPS-1004083) and support from the Gibson Family Foundation. C. S. and B. D. B. acknowledge support from the UTK Graduate Energy Scholars Program. R. S. was supported as an IGERT fellow from the National Science Foundation IGERT program (DGE-0801470). B. D. B. and T. Z. also acknowledge support from the Directors Strategic Initiative, "Understanding Photosystem I as a Biomolecular Reactor for Energy Conversion" at the Army Research Laboratory, Adelphi, MD, USA (ARL Contract #W911NF-11-2-0029).

REFERENCES

Amunts A, Drory O, Nelson N. 2007. The structure of a plant photosystem I supercomplex at 3.4 a resolution. *Nature* **447**: 58–63.

Amunts A, Toporik H, Borovikova A, Nelson N. 2010. Structure determination and improved model of plant photosystem I. *J. Biol. Chem.* **285**: 3478–86.

Andersen B, Scheller HV, Moller BL. 1992. The Psi-E subunit of photosystem-I binds ferredoxin-NADP + oxidoreductase. *Febs Lett* **311**: 169–173.

Barth P, Guillouard I, Setif P, Lagoutte B. 2000. Essential role of a single arginine of photosystem I in stabilizing the electron transfer complex with ferredoxin. *J. Biol. Chem.* **275**: 7030–6.

Baumann B, Sticht H, Scharpf M, Sutter M, Haehnel W, Rosch P. 1996. Structure of Synechococcus elongatus [Fe2S2] ferredoxin in solution. *Biochemistry* **35**: 12831–41.

Berg JM, Tymoczko JL, Stryer L. 2007. Biochemistry (6th *Edition*), W.H. Freeman & Company, New York.

Bernstein FC, Koetzle TF, Williams GJ, Meyer EF, Jr., Brice MD, Rodgers JR, Kennard O, Shimanouchi T, Tasumi M. 1977. The Protein Data Bank: a computer-based archival file for macromolecular structures. *J. Mol. Biol.* **112**: 535–42.

Bottin H, Hanley J, Lagoutte B. 2001. Role of acidic amino acid residues of PsaD subunit on limiting the affinity of photosystem I for ferredoxin. *Biochem. Biophys. Res. Commun.* **287**: 833–6.

Brunger AT, Adams PD, Fromme P, Fromme R, Levitt M, Schroder GF. 2012. Improving the Accuracy of macromolecular structure refinement at 7 angstrom resolution. *Structure* **20**: 957–966.

Chapman HN, Fromme P, Barty A, White TA, Kirian RA, Aquila A, Hunter MS, Schulz J, DePonte DP, Weierstall U, Doak RB, Maia FRNC, Martin AV, Schlichting I, Lomb L, Coppola N, Shoeman RL, Epp SW, Hartmann R, Rolles D, Rudenko A, Foucar L, Kimmel N, Weidenspointner G, Holl P, Liang MN, Barthelmess M, Caleman C, Boutet S, Bogan MJ, Krzywinski J, Bostedt C, Bajt S, Gumprecht L, Rudek B, Erk B, Schmidt C, Homke A, Reich C, Pietschner D, Struder L, Hauser G, Gorke H, Ullrich J, Herrmann S, Schaller G, Schopper F, Soltau H, Kuhnel KU, Messerschmidt M, Bozek JD, Hau-Riege SP, Frank M, Hampton CY, Sierra RG, Starodub D, Williams GJ, Hajdu J, Timneanu N, Seibert MM, Andreasson J, Rocker A, Jonsson O, Svenda M, Stern S, Nass K, Andritschke R, Schroter CD, Krasniqi F, Bott M, Schmidt KE, Wang XY, Grotjohann I, Holton JM, Barends TRM, Neutze R, Marchesini S, Fromme R, Schorb S, Rupp D, Adolph M, Gorkhover T, Andersson I, Hirsemann H, Potdevin G, Graafsma H, Nilsson B, Spence JCH. 2011. Femtosecond X-ray protein nanocrystallography. *Nature* **470**: 73–U81.

Das A, Plotkin SS. 2013. SOD1 exhibits allosteric frustration to facilitate metal binding affinity. *Proc. Natl. Acad. Sci. U. S. A.* **110**: 3871–6.

Das R, Kiley PJ, Segal M, Norville J, Yu AA, Wang LY, Trammell SA, Reddick LE, Kumar R, Stellacci F, Lebedev N, Schnur J, Bruce BD, Zhang SG, Baldo M. 2004. Integration of photosynthetic protein molecular complexes in solid-state electronic devices. *Nano Lett.* **4**: 1079–1083.

Evans BR, O'Neill HM, Hutchens SA, Bruce BD, Greenbaum E. 2004. Enhanced photocatalytic hydrogen evolution by covalent attachment of plastocyanin to photosystem I. *Nano Lett.* **4**: 1815–1819.

Fischer N, Hippler M, Setif P, Jacquot JP, Rochaix JD. 1998. The PsaC subunit of photosystem I provides an essential lysine residue for fast electron transfer to ferredoxin. *EMBO J.* **17**: 849–58.

Fischer N, Setif P, Rochaix JD. 1999. Site-directed mutagenesis of the PsaC subunit of photosystem I. F(b) is the cluster interacting with soluble ferredoxin. *J. Biol. Chem.* **274**: 23333–40.

Goni G, Herguedas B, Hervas M, Peregrina JR, De la Rosa MA, Gomez-Moreno C, Navarro JA, Hermoso JA, Martinez-Julvez M, Medina M. 2009. Flavodoxin: a compromise between efficiency and versatility in the electron transfer from photosystem I to ferredoxin-NADP(+) reductase. *Biochim Biophys Acta* **1787**: 144–54.

Hanley J, Setif P, Bottin H, Lagoutte B. 1996. Mutagenesis of photosystem I in the region of the ferredoxin cross-linking site: modifications of positively charged amino acids. *Biochemistry* **35**: 8563–8571.

Hatanaka H, Tanimura R, Katoh S, Inagaki F. 1997. Solution structure of ferredoxin from the thermophilic cyanobacterium Synechococcus elongatus and its thermostability. *J. Mol. Biol.* **268**: 922–33.

Iwuchukwu IJ, Vaughn M, Myers N, O'Neill H, Frymier P, Bruce BD. 2010. Self-organized photosynthetic nanoparticle for cell-free hydrogen production. *Nat. Nanotechnol.* **5**: 73–79.

Jenik M, Parra RG, Radusky LG, Turjanski A, Wolynes PG, Ferreiro DU. 2012. Protein frustratometer: a tool to localize energetic frustration in protein molecules. *Nucleic Acids Res.* **40**: W348–51.

Jordan P, Fromme P, Witt HT, Klukas O, Saenger W, Krauss N. 2001. Three-dimensional structure of cyanobacterial photosystem I at 2.5 A resolution. *Nature* **411**: 909–17.

Kannan N, Vishveshwara S. 2000. Aromatic clusters: a determinant of thermal stability of thermophilic proteins. *Protein Eng.* **13**: 753–61.

Lelong C, Setif P, Lagoutte B, Bottin H. 1994. Identification of the amino-acids involved in the functional interaction between photosystem-I and ferredoxin from Synechocystis sp-Pcc-6803 by chemical cross-linking. *J. Biol. Chem.* **269**: 10034–10039.

Lelong C, Boekema EJ, Kruip J, Bottin H, Rogner M, Setif P. 1996. Characterization of a redox active cross-linked complex between cyanobacterial photosystem I and soluble ferredoxin. *Embo J.* **15**: 2160–2168.

Leonhardt K, Straus NA. 1992. An iron stress operon involved in photosynthetic electron-transport in the marine cyanobacterium Synechococcus sp Pcc-7002. *J. Gen. Microbiol.* **138**: 1613–1621.

Li X, Keskin O, Ma B, Nussinov R, Liang J. 2004. Protein-protein interactions: hot spots and structurally conserved residues often locate in complemented pockets that pre-organized in the unbound states: implications for docking. *J. Mol. Biol.* **344**: 781–95.

MacKerell AD, Banavali N, Foloppe N. 2001. Development and current status of the CHARMM force field for nucleic acids. *Biopolymers* **56**: 257–265.

Mershin A, Matsumoto K, Kaiser L, Yu D, Vaughn M, Nazeeruddin MK, Bruce BD, Graetzel M, Zhang S. 2012. Self-assembled photosystem-I biophotovoltaics on nanostructured TiO(2)and ZnO. *Sci. Rep.* **2**: 234.

Millsaps JF, Bruce BD, Lee JW, Greenbaum E. 2001. Nanoscale photosynthesis: photocatalytic production of hydrogen by platinized photosystem I reaction centers. *Photochem. Photobiol.* **73**: 630–635.

Muhlenhoff U, Kruip J, Bryant DA, Rogner M, Setif P, Boekema E. 1996. Characterization of a redox-active cross-linked complex between cyanobacterial photosystem I and its physiological acceptor flavodoxin. *Embo J.* **15**: 488–497.

Nguyen PQ, Silberg JJ. 2010. A selection that reports on protein-protein interactions within a thermophilic bacterium. *Protein Eng. Des. Sel.* **23**: 529–536.

Onai K, Morishita M, Kaneko T, Tabata S, Ishiura M. 2004. Natural transformation of the thermophilic cyanobacterium Thermosynechococcus elongatus BP-1: a simple and efficient method for gene transfer. *Mol. Genet. Genomics: MGG* **271**: 50–9.

Onuchic JN, Luthey-Schulten Z, Wolynes PG. 1997. Theory of protein folding: the energy landscape perspective. *Annu. Rev. Phys. Chem.* **48**: 545–600.

Page CC, Moser CC, Chen XX, Dutton PL. 1999. Natural engineering principles of electron tunnelling in biological oxidation-reduction. *Nature* **402**: 47–52.

Roman EA, Faraj SE, Gallo M, Salvay AG, Ferreiro DU, Santos J. 2012. Protein stability and dynamics modulation: the case of human frataxin. *PLoS One* **7**: e45743.

Rousseau F, Setif P, Lagoutte B. 1993. Evidence for the involvement of PSI-E subunit in the reduction of ferredoxin by photosystem I. *EMBO J.* **12**: 1755–65.

Setif P. 2001. Ferredoxin and flavodoxin reduction by photosystem I. *BBA-Bioenergetics* **1507**: 161–179.

Setif P, Fischer N, Lagoutte B, Bottin H, Rochaix JD. 2002. The ferredoxin docking site of photosystem I. *Biochim Biophys Acta* **1555**: 204–9.

Setif P, Harris N, Lagoutte B, Dotson S, Weinberger SR. 2010. Detection of the photosystem I:ferredoxin complex by backscattering interferometry. *J. Am. Chem. Soc.* **132**: 10620–2.

Ullmann GM, Hauswald M, Jensen A, Knapp EW. 2000. Structural alignment of ferredoxin and flavodoxin based on electrostatic potentials: implications for their interactions with photosystem I and ferredoxin-NADP reductase. *Proteins* **38**: 301–9.

Wittenberg G, Sheffler W, Darchi D, Baker D, Noy D. 2013. Accelerated electron transport from photosystem I to redox partners by covalently linked ferredoxin. *Phys. Chem. Chem. Phys.* **15**: 19608–14.

Yehezkeli O, Tel-Vered R, Michaeli D, Willner I, Nechushtai R. 2014. Photosynthetic reaction center-functionalized electrodes for photo-bioelectrochemical cells. *Phycol. Res.* **120**(1-2): 71–85.

Zanetti G, Merati G. 1987. Interaction between photosystem-I and ferredoxin - identification by chemical cross-linking of the polypeptide which binds ferredoxin. *Eur. J. Biochem.* **169**: 143–146.

Zeldovich KB, Berezovsky IN, Shakhnovich EI. 2007. Protein and DNA sequence determinants of thermophilic adaptation. *PLOS Comput Biol.* **3**: e5.

Zilber AL, Malkin R. 1988. Ferredoxin Cross-Links to a 22-Kd Subunit of Photosystem-I. *Plant Physiol.* **88**: 810–814.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version at the publisher's web site.