Heteronuclear NMR and Soft Docking: An Experimental Approach for a Structural Model of the Cytochrome c_{553}−Ferredoxin Complex

Xavier Morelli,[‡] Alain Dolla,[‡] Myrjam Czjzek,[§] P. Nuno Palma,[||,⊥] Francis Blasco,[⊗] Ludwig Krippahl,[||] Jose J. G. Moura,[||] and Françoise Guerlesquin*[,‡]

Unité de Bioénergétique et Ingénierie des Protéines, IBSM-CNRS, 31 chemin Joseph Aiguier, 13402 Marseille cedex 20, France, Architecture et Fonction des Macromolécules Biologiques, IBSM-CNRS, 31 chemin Joseph Aiguier, 13402 Marseille cedex 20, France, Departamento de Química, Centro Química Fina e Biotecnologia, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, 2825−114 Caparica, Portugal, Instituto Superior de Ciências da Saúde−Sul, 2825 Monte de Caparica, Portugal, and Laboratoire de Chimie Bactérienne, IBSM-CNRS, 31 chemin Joseph Aiguier, 13402 Marseille cedex 20, France

Received October 5, 1999; Revised Manuscript Received December 10, 1999

ABSTRACT: The combination of docking algorithms with NMR data has been developed extensively for the studies of protein−ligand interactions. However, to extend this development for the studies of protein−protein interactions, the intermolecular NOE constraints, which are needed, are more difficult to access. In the present work, we describe a new approach that combines an *ab initio* docking calculation and the mapping of an interaction site using chemical shift variation analysis. The cytochrome c_{553}−ferredoxin complex is used as a model of numerous electron-transfer complexes. The ^{15}N-labeling of both molecules has been obtained, and the mapping of the interacting site on each partner, respectively, has been done using HSQC experiments. ^{1}H and ^{15}N chemical shift analysis defines the area of both molecules involved in the recognition interface. Models of the complex were generated by an *ab initio* docking software, the BiGGER program (bimolecular complex generation with global evaluation and ranking). This program generates a population of protein−protein docked geometries ranked by a scoring function, combining relevant stabilization parameters such as geometric complementarity surfaces, electrostatic interactions, desolvation energy, and pairwise affinities of amino acid side chains. We have implemented a new module that includes experimental input (here, NMR mapping of the interacting site) as a filter to select the accurate models. Final structures were energy minimized using the X-PLOR software and then analyzed. The best solution has an interface area (1037.4 Å^2) falling close to the range of generally observed recognition interfaces, with a distance of 10.0 Å between the redox centers.

Electron-transfer reactions are essential in cell metabolism. The electron flow is directed by a molecular pathway along successive electron-transfer complexes where protein recognition is the driving force and the oxidoreduction potential is the thermodynamic parameter. Until recently, only a few X-ray structures of electron-transfer complexes have been solved; therefore, molecular docking is a powerful alternative method to investigate electron-transfer complexes.

Protein−protein recognition results from an assembly where the surface complementarity of the two molecules plays a crucial role (*1*). The simplest model is the rigid body approximation in which the two sterically and chemically complementary surfaces leave no degree of freedom to the component molecules (*2*). Atomic structure of protein−protein recognition sites has been analyzed for 75 protein−protein complexes of known three-dimensional structures (*3*). Protein-docking algorithms, which take into consideration the observed interactions at the interfaces of known protein complexes, can be a useful tool for the structural characterization of such noncovalent complexes. This is especially true if the simulation techniques are combined with experimental information. Covalent cross-linking, site-directed mutagenesis, and fluorescence have been developed extensively to study protein−protein interactions. Recent progress in heteronuclear multidimensional NMR[1] brings this method to the first rank for studying protein recognition (*4*). On the basis of ^{15}N and ^{1}H chemical shift variations, the mapping of an interaction site may be obtained. In the present work, we have developed a new approach where protein docking simulations, carried out with the program BiGGER (*5*), are associated to NMR mapping of the interaction sites to generate consistent models of electron-transfer complexes.

The presence of a "ferredoxin-like" domain has been reported in various iron−sulfur enzymes such as Fe-

[‡] Unité de Bioénergétique et Ingénierie des Protéines.
[§] Architecture et Fonction des Macromolécules Biologiques.
[||] Universidade Nova de Lisboa.
[⊥] Instituto Superior de Ciências da Saúde−Sul.
[⊗] Laboratoire de Chimie Bactérienne.
* Corresponding author. Telephone: 33-4-91-16-43-79. Fax: 33-4-91-16-45-78. E-mail: F.Guerlesquin@ibsm.cnrs-mrs.fr. Internet address: http://garlaban.cnrs-mrs.fr/nmr.html.

[1] Abbreviations: NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser effect spectroscopy; HSQC, heteronuclear single quantum coherence; BiGGER, bimolecular complex generation with global evaluation and ranking; BoGIE, boolean geometric interaction evaluation; PCR, polymerase chain reaction; *Dv*H, *Desulfovibrio vulgaris* Hildenborough; FdI, ferredoxin I.

10.1021/bi992306s CCC: $19.00
Published on Web 02/15/2000

hydrogenase and formate dehydrogenase, two redox partners of *Desulfovibrio vulgaris* Hildenborough cytochrome c_{553}. *D. norvegicum* ferredoxin I was considered as a model for the electron donor of the low oxidoreduction cytochrome. The gene of *D. norvegicum* ferredoxin I was cloned and expressed in *Escherichia coli* in order to obtain the isotopically labeled protein. Heteronuclear NMR experiments (HSQC) were carried out using unlabeled cytochrome c_{553} to map the interacting site on the ferredoxin using induced chemical shift variations. A model of the ferredoxin was calculated and used for an ab initio docking simulation of the cytochrome c_{553}−ferredoxin complex with the program BiGGER (*5*). NMR mapping of the interacting site on cytochrome c_{553} and on ferredoxin was used as a filter to select the best structures resulting from the docking simulations. Final structures were minimized and analyzed on the basis of additional results from the complex. This new strategy is under development and will be adopted for electron-transfer complexes as well as other complexes.

MATERIAL AND METHODS:

Cloning and Sequence of the Gene Encoding Desulfomicrobium norvegicum ferredoxin I. Desulfovibrio desulfuricans (Norway strain) has been classified as *Desulfomicrobium norvegicum* (*6*). *Desulfomicrobium norvegicum* was grown in lactate-sulfate medium as previously described (*7*). Genomic DNA was isolated from 100 mL of culture, using the procedure as described by Voordouw et al (*8*).

The *fdx*I gene, encoding the *D. norvegicum* ferredoxin I, was cloned from the genomic DNA using the polymerase chain reaction (PCR) and two degenerated oligonucleotide primers. Primers used were: FD1Nterm, (5′GGGAATTCATGAC(AGCT)AT(ACT)GT(AGCT)AT(ACT)GA(CT)CA(CT)GA(AG)GA3′); FD1Cterm, (5′GGGAATTCTTA(TC)TC(TC)TT(AGCT)GA(TGA)AT(AGCT)GC(CT)TC3′). These primers were designed on the basis of the N-terminal and C-terminal sequences of the ferredoxin I, respectively (*9*), and contain an *Eco*RI restriction endonuclease site for subsequent subcloning at their 5′-termini. PCRs were carried out in a total volume of 50 μl, using the *Pwo* DNA Polymerase (Boehringer-Mannheim) according to the manufacturer's protocol. The PCR amplification procedure was run for 25 cycles, and the PCR product was then purified using the High Pure PCR Product Purification Kit from Boehringer-Mannheim. The amplicon was first subcloned into pUCBM21 digested by *Sma*I, and its DNA sequence was checked by DNA sequencing using both universal and −50 reverse primers.

To increase the ferredoxin I production in *E. coli*, the *fdxI* gene was subcloned into the pET22b(+) vector (Novagen). The entire *fdx*I gene was first amplified by PCR using two primers. The FD1PET1 primer (5′GGAATTC*CATATG*ACAATAGTGATAGATCATGAAGAA3′) includes a built-in *Nde*1 site (in italics). The FD1PET2 primer (5′AAGCTT*CTCGAG(TCA)*TTACTCTTTGGATATGGCTTCCAC3′) includes a *Xho*I site (in italics) and a stop codon (in parenthesis). The PCR fragment was purified as described above, restricted by *Nde*I and *Xho*I, and ligated into the *Nde*I-*Xho*I restricted pET22b(+). The resulting plasmid, called pETFDI, contains the *fdx*I gene under the T7 promoter.

Ferredoxin I Overproduction in E. coli. Uniformly ^{15}N-labeled ferredoxin I was obtained by culturing *E. coli* BL21 (pETFDI) for 4 h at 37 °C, using a specific medium containing the M9 medium (*10*) supplemented with 30% (w/vol) of the celtone required for the ^{15}N Martek medium (www.martekbio.com), $^{15}NH_4Cl$ (1 g L^{-1}), ampicilin (0.2 g L^{-1}), and ammonium iron citrate (0.01 g L^{-1}), pH of 7.6. IPTG was then added (0.5 mM, final concentration) to induce ferredoxin I production, and the cells were cultured for an additional 6 h time.

The purification of the ferredoxin I was obtained in four steps. First, the cells were broken using a French press. Acidic proteins were adsorbed onto a DEAE-cellulose column equilibrated with a 10 mM Tris-HCl buffer (pH 7.6). The ferredoxin eluted with 400 mM of Tris-HCl buffer, was then adsorbed on a hydroxyapatite column previously equilibrated with 400 mM Tris-HCl, and the protein was eluted in a 50 mM phosphate buffer (pH 5.9). Presence of apoferredoxin was eliminated through a sephadex G50 column equilibrated in a 10 mM phosphate buffer (pH 5.9). The optical spectra was in agreement with 1.2 mg of pure ferredoxin I containing a (4Fe−4S) cluster. The homogeneity of the ferredoxin sample was evidenced by electrophoresis and 1D-NMR spectrum.

The cytochrome c_{553} from *Dv*H was expressed and purified from *Desulfovibrio desulfuricans* G200 as previously described (*11*).

NMR Samples. The oxidized FdI was concentrated to 0.4 mM on an Amicon microconcentrator in a potassium phosphate buffer of 20 mM, pH 5.9, and 10% D_2O. For the titration of the complex formation, the ferredoxin concentration was 0.1 mM.

NMR Spectra. NMR spectra were recorded on a Bruker Avance DRX 500 spectrometer at 296 K. The NMR spectra were processed with UXNMR and Aurelia software provided by Bruker. Two-dimensional ^{1}H-^{15}N HSQC spectra and modified 2D HSQC−NOESY experiments were recorded using a watergate pulse sequence in the TPPI mode. For HSQC, the spectral widths are 2000 Hz for ^{1}H and 2027 Hz for ^{15}N; 1024 data points in t_2, and 64 transients for each 128 t_1 were used. For 2D HSQC−NOESY, the spectral widths were 7000 Hz for ^{1}H and 2027 Hz for ^{15}N; 1024 data points in t_2, and 768 transients for each 128 t_1 were used. ^{1}H chemical shifts were referenced to the H_2O resonance calibrated at 4.792 ppm at 296 K relative to the sodium trimethylsilyl-2,2,3,3-($2H_4$) propionate. ^{15}N chemical shifts were referenced indirectly by using the above ^{1}H reference and the gyromagnetic ratios (0.101329118) (*12*). The ^{1}H chemical shifts of the oxidized ferredoxin I previously reported (*13*) were used for the ^{15}N resonance assignments.

Chemical Shift Analysis. Chemical shift variations observed in 2D heteronuclear experiments were obtained from a single set of ^{1}H-^{15}N HSQC experiments with 0.1 mM ferredoxin I concentration and 0.5, 1, and 2 molar equivalents of cytochrome c_{553}.

Protein Structures. The protein coordinates of *Desulfovibrio vulgaris* Hildenborough cytochrome c_{553} were obtained from the Brookhaven Protein Data Bank (code 1dvh).

The structural model for the ferredoxin I from *Desulfomicrobium norvegicum* was constructed on the basis of the homology with ferredoxin II from *Desulfovibrio gigas*, the crystal structure of which has been determined at 1.7 Å resolution (*14*). In a first step, all the differing amino acids

Table 1: Comparison of the Eleven Models of *Desulfomicrobium norvegicum* Ferredoxin I

	N°1	N°2	N°3	N°4	N°5	N°6	N°7	N°8	N°9	N°10	N°11
temperature (K)	280	300	280	350/280	280	350	300	300	350	300	350
tolerance	10^{-5}	10^{-4}	10^{-4}	10^{-4}	10^{-4}	10^{-4}	10^{-4}	10^{-4}	10^{-4}	10^{-4}	10^{-4}
H_2O			+	+			+	+	+	+	+
RMSD[a] (Å)	1.76	1.93	1.98	1.70	1.99	1.84	2.19	1.90	1.82	1.71	1.75
RMSD[b] (Å)	1.71	1.90	1.80	1.20	2.02	2.04	1.84	1.95	2.05	1.16	1.54

[a] On 26 atoms 9−18, 25−33, 47−56 for the *D. gigas* ferredoxin and 35−44, 51−59, 72−81 for the *D. desulfuricans* ATCC [Fe]-hydrogenase. [b] On 5 atoms 9, 12, 15, 25, and 51 for the *D. gigas* ferredoxin and 35, 38, 41, 51, and 76 for the *D. desulfuricans* ATCC [Fe]-hydrogenase.

were changed with the graphical program TURBO−FRODO (*15*). No insertion or deletion occurred between the two sequences. In a second step, this model was optimized by an energy minimization and molecular dynamic calculation with X-PLOR 3.851 (*16*). The force field defined by the AMBER parameters was used (*17*). Eleven different structures were generated by variation of the following parameters (Table 1): (1) The presence or absence of structural water molecules are deduced from the homologous crystal structure. (2) The molecular dynamic simulations were performed using a heat bath (*18*) at a constant temperature. Three different temperature values were chosen, 280, 300, and 350 K. In all cases, the initial velocities were taken from a Maxwell−Boltzmann distribution at the same temperature as the bath temperature. (3) The SHAKE method (*19*) was used to constrain the geometry of Fe_4S_4 cluster. Two different tolerance values (0.0004 and 0.0005) were used. The integration of the equations of motion was carried out using a 0.001 ps time step. The eleven structures were then compared structurally with the ferredoxin of *D. gigas* and the "ferredoxin-like" domain of the [Fe]-Hydrogenase (*20*). We chose the closest model (*n*°10) according to the lowest rmsd values.

Protein Docking. Molecular interaction simulations were performed using the docking program BiGGER (*5*). The docking procedure is composed of two modules that work in a sequence. The first module (BoGIE, boolean geometric interaction evaluation) is a grid-like search algorithm, which generates a population of docked geometries with maximal surface matching and favorable intermolecular amino acid contacts. The first step is the generation of a 3D matrix composed of small cubic cells of 1 Å size, which represents the complex shape of each molecules. For every relative orientation of the two proteins, the translational interaction space is searched by systematically shifting the matrixes defining one molecule (the probe) relative to the matrixes representing the other partner (the target). Solutions involving unrealistic interpenetration of the docking partners are discarded. The probe is rotated by 15° relative to the target, and this process (digitization, translation, and surface matching) is repeated until a complete nonredundant search is achieved. Conformational changes that occur during the complex formation are among the most difficult to model by rigid body docking methods. The calculation incorporates side chain flexibility for Arg, Lys, Asp, Glu, and Met; the side chains of these residues can penetrate up to the core of the other protein, without causing forbidden core−core overlaps. The module ends by a filtering procedure. A subpopulation of 1000 binding modes is kept using the criteria of geometric complementarity and favorable propensity of each pair of amino acids to be in close contact across the complex interface.

The second module (INTERACT) evaluates the population of putative solutions, according to four different interaction terms (surface matching, side chain contacts, electrostatic, and solvation energies) combined into a global scoring function and computed for every docked geometry. The program has been tested against a wide set of crystallographic protein−protein complexes (*5*) and it shows that a near-correct complex prediction is most often found within a reasonably small group of the top ranked solutions.

NMR Filtering. An experimental filtering taking into account chemical shift variations induced on both the ferredoxin and the cytochrome has been implemented to the software. Chemical shifts, above 0.01 ppm for ^1H and 0.05 ppm for ^{15}N, are considered to be significant and due to the complex formation. They are translated in distance parameters; we considered as an arbitrary cutoff that the chemical shift of one nucleus is affected when it is at less than 5 Å of any atoms of the other partner. The 1000 solutions obtained from the BiGGER procedure are analyzed using these distance constraints and are ranked according to the number of NMR constraints that are violated. Note, however, that these experimental constraints are not used at all for the ab initio scoring of putative binding modes as predicted by BiGGER. They constitute an independent evaluation of the structural solutions and are used to cross-validate the models predicted.

Finally, an additional "metal−metal distance filtering" was implemented for electron-transfer complexes, assuming that an electron-transfer process takes place when the distance between the two redox centers is less than 20 Å.

Molecular Dynamics and Minimization of the Complex Structures. The structures of the complex, as obtained by the docking algorithm filtered by NMR data, were first minimized in a preparation step, using the conjugate gradient method by Powell (*21*), with X-PLOR 3.851. All parameters were minimized. Then, a molecular dynamics calculation (AMBER force field) was performed on the selected complex structures with a heat bath temperature of 280 K and initial velocities from a Maxwell−Boltzmann distribution at 280 K. The SHAKE constraints (*22*) were applied for both the Fe_4S_4 cluster and the heme group with a tolerance of 0.0004 and integration of the motion equations at 0.001 ps time steps. The coordinates for the three "best structures" (solution 1, 2, and 3) have been deposited at the Protein Data Bank (accession code 1dwl).

RESULTS

Ferredoxin Labeling. Degenerated primers were designed in order to clone the gene encoding *D. norvegicum* ferredoxin I by PCR from genomic DNA. The protein sequence, deduced from the determined nucleotidic sequence (Figure

```
                ATGACAATAGTGATAGATCATGAAGAATGCATCGGCTGCGAAAGTTGTGTGGAA
0               M  T  I  V  I  D  H  E  E  C  I  G  C  E  S  C  V  E

                CTTTGTCCTGAAGTTTTTGCCATGATCGACGGCGAAGAAAAAGCCATGGTCACGGCCCCG
18              L  C  P  E  V  F  A  M  I  D  G  E  E  K  A  M  V  T  A  P

                GACTCCACAGCCGAATGCGCCCAGGACGCCATCGACGCCTGTCCGGTGGAAGCCATATCC
38              D  S  T  A  E  C  A  Q  D  A  I  D  A  C  P  V  E  A  I  S

                AAAGAGTAA
58              K  E  *
```

FIGURE 1: Nucleotidic sequence of *D. norvegicum* ferredoxin I gene (*fdxI*). The translated sequence corresponding to the protein sequence is shown in italics. The initiating methionine has been numbered as Met0.

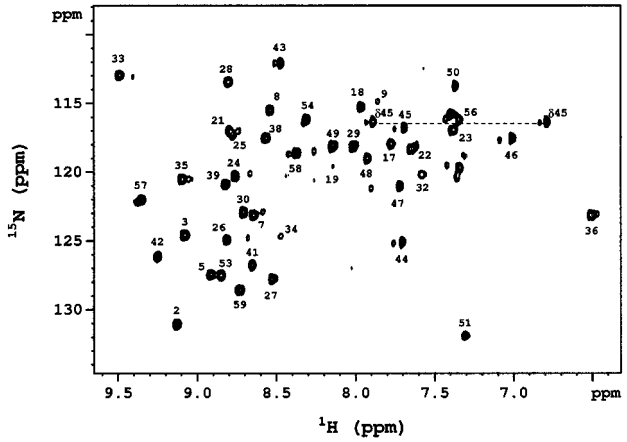

FIGURE 2: HSQC spectra of the *D. norvegicum* ferredoxin I (0.4mM, in phosphate buffer, pH 5.9) at 296 K. The assignment of the ^{1}H/^{15}N cross-peaks of NH groups are indicated by the sequential number of the corresponding amino acid.

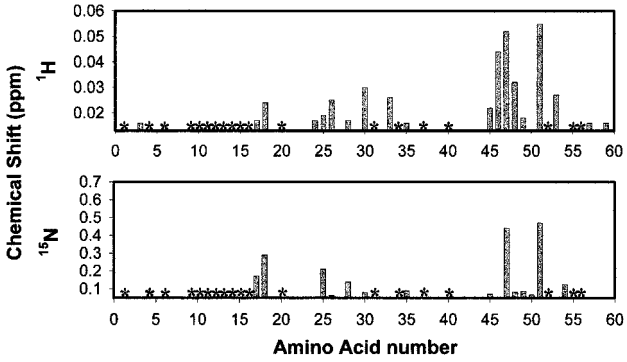

FIGURE 3: ^{1}H and ^{15}N chemical shift variations of NH groups of ferredoxin, induced by the presence of unlabeled cytochrome c_{553}. Stars indicate the unassigned residues.

1) was in complete agreement with the previously published sequence by Bruschi et al. (*9*). When the *fdx*I gene was subcloned into a PUC series vector, the rate of ferredoxin I production was too low for the labeling experiments. Thus, we decided to use another expression vector, based on the T7 promoter. In this case, the protein was produced in *E. coli* with a correct folding as demonstrated by a 1D NMR spectrum, at a high level (1 mg/L) on LB medium, as compared to the expression in the sulfate-reducing bacteria (6×10^{-3}mg/L). The expression of ferredoxin I on M9 medium was too low to obtain a labeled protein. To produce the labeled protein, we grew the cells on a M9 medium where labeled celtone from Martek (3.6 g/L) was added instead of Yeast extract; 1.2 mg of pure labeled ferredoxin were obtained from 4 L culture.

Mapping the Interacting Site by Heteronuclear NMR. Heteronuclear NMR spectra (2D HSQC−NOESY or HSQC−TOCSY) of the 0.4 mM ferredoxin solution were used to obtain the ^{15}N assignment on the basis of the reported proton assignment (*13*). The ^{1}H-^{15}N HSQC spectrum is presented in Figure 2. Out of the 59 amino acids of the protein, 42 NH groups were identified with 2D NMR experiments. The sequence residues from 9 to 16, including the cysteines that ligate the cluster, were not found. The ^{15}N of the N-terminal, and of the three prolines, were not observable and six NH groups (I4, H6, K31, T40, A55, and I56), probably too close to the cluster, were not present in the HSQC. The ^{1}H and ^{15}N chemical shifts at 296 K are included in the Supporting Information.

The mapping of the interaction site was obtained by performing a titration of the complex formation and recording HSQC experiments on a 0.1 mM sample of ferredoxin at various unlabeled cytochrome c_{553} concentrations (0, 0.5, 1, and 2 molar equivalents). Only the NH groups of the ferredoxin are observable here. Seven NH cross-peaks are significantly altered by the complex formation when considering the ^{15}N chemical shift variations and were used for NMR filtering. Fifteen cross-peaks are affected when considering ^{1}H chemical shifts (Figure 3). The residues are found in two parts of the molecule close to the cluster (Figure 4).

Docking of the Complex. Figure 5 summarizes the combined *ab initio*/experimental approach used for protein recognition analysis. Figure 6 represents the results mapped on each of the two proteins (Figure 6: top, ferredoxin; bottom, cytochrome c_{553}). The set of 1000 putative docked geometries initially generated by BiGGER were independently evaluated and ranked according to two approaches: an *ab initio* procedure and an experimental approach.

The *ab initio* procedure uses exclusively the interaction scoring function implemented in BiGGER (Figure 6A). As shown in a series of test cases (*5*), one should expect that this method will necessarily pinpoint the correct structure of the complex as one of the top ranking solutions of the docking. Then, the group of top-ranking solutions should rather be carefully analyzed.

The experimental approach uses the NMR data on the mapping of the interaction site on cytochrome c_{553} (*23*) and on the ferredoxin (the present work) to limit the set of solutions. Nineteen NMR constraints (twelve on the cytochrome and seven on the ferredoxin) were used to score the 1000 docking solutions according to their compatibility with this experimental data. In this case, the scoring is simply

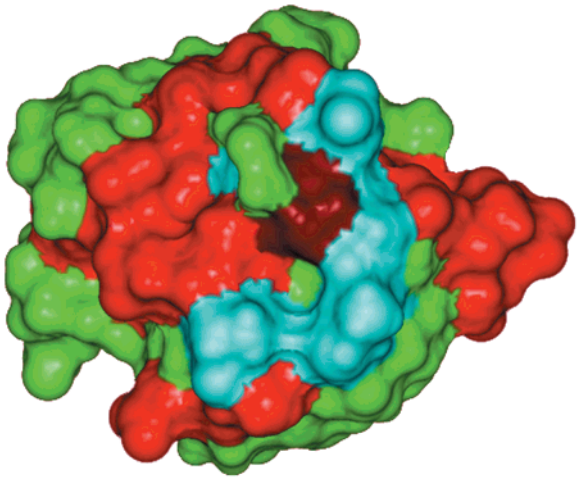

FIGURE 4: The mapping of the interacting site on the ferredoxin, obtained by heteronuclear experiments is represented using InsightII from MSI. In red are NH groups whose resonances undergo chemical shift variations. The Cys12, which is perturbed in the 1D NMR experiments (*23*) is shown in brown. In blue are the unassigned NH groups. In green are unaffected residues.

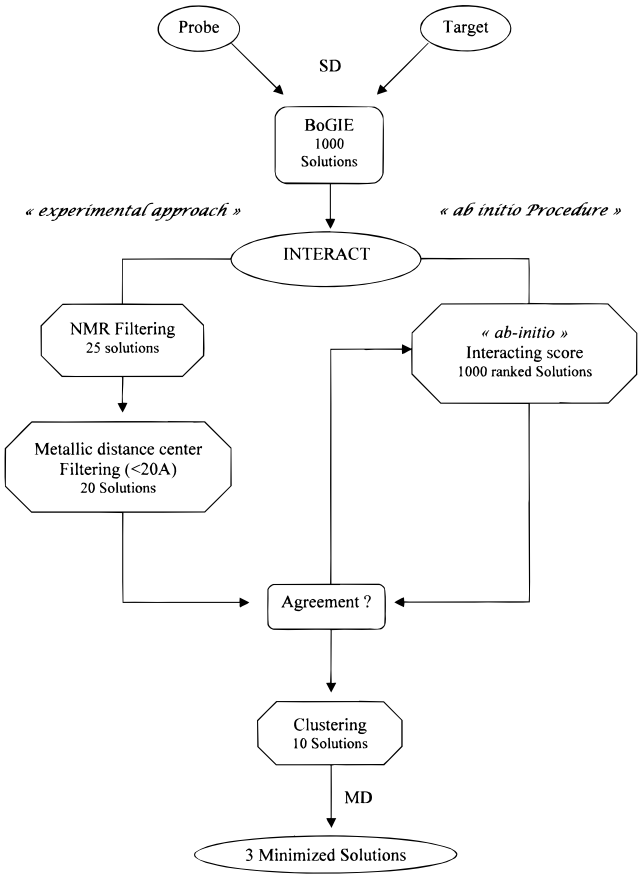

FIGURE 5: Strategy for the obtention of structural models using the *ab initio* docking and the NMR filtering. SD means soft docking and MD means molecular dynamics.

the number of NMR constraints that are fulfilled for each docked geometry and its value is color-coded in Figure 6B. Here, there are several solutions with a similar level of agreement with the experimental constraints. However, if compared, the top ranking solutions given by the two independent scoring criteria clearly indicate that the same areas on the surface of the ferredoxin and the cytochrome should be involved in binding (Figure 6C). The resulting areas involve a surface patch surrounding the heme crevice on the cytochrome and the two parallel alpha helices accommodating the Fe−S cluster on the ferredoxin moiety.

Among the 1000 structures generated, 25 are found to be compatible with the NMR mapping, with a maximum of 5 out of the 19 NMR constraints being simultaneously violated. Note that no single structure is found to fulfill all 19 constraints simultaneously, which may reveal a dynamic adaptation process during binding.

An additional filter was applied considering that a maximum distance of 20 Å between the iron of the heme and the cluster of the ferredoxin (before minimization) is coherent for the electron-transfer phenomenon. This procedure filters out 5 of the 25 solutions previously selected. Additionally, some of the 20 structures, selected in this way, are very similar to each other and may be clustered together into 10 different families, which will be considered as simple solution structures. Three families of solutions contain the best part of the 20 resulting structures: five structures were clustered in solution 1, and four structures in each of solutions 2 and 3 (each group of solutions will be represented by the structure with highest docking score). These three families are retained as the most likely solutions.

In the first round of calculations, the four structures of solution 2 were minimized as described in the section material and methods. The results indicated that after minimization they could still be considered as identical solutions. Therefore, the representative structure of each family in agreement with NMR data, metal center distance, surface area, and interaction score was submitted to energy minimization. The resulting properties of the three complex solutions after minimization are summarized in Table 2.

DISCUSSION

The recent report by Lo Conte and co-workers summarizes the properties of protein−protein recognition interfaces of cocrystallized complexes (*3*). Among all the complexes present in the PDB, only four different structures of electron-transfer complexes have been solved in the past, even though numerous X-ray structures of metalloproteins are available. In the case of electron-transfer complexes such as cytochrome c_{553}−ferredoxin I complex, the cocrystallization is likely to fail, due to its K_a value (10^6 M^{-1}). Furthermore, doubt remains that such crystalline complexes represent only one of a number of productive associations. The mode of interaction of electron-transfer systems makes the analysis by NMR the more adequate technique. However, on the basis of X-ray or NMR structures, docking simulations may be carried out. Flexible docking coupled to NMR mapping is now well-established to study protein−ligand interactions (*24*) and is a promising approach for drug design. In the present work, a similar approach has been developed for protein−protein recognition analysis. The flexible docking generally used for protein−ligand interaction has been replaced by a soft docking algorithm implemented in the BiGGER program and the nOe used to restrain the ligand have been replaced by an NMR filter using chemical shift analysis on both proteins. The method described here, combining NMR mapping experiments with an adapted docking algorithm, is the least time-consuming and most appropriate way to get a clear image of how these proteins interact.

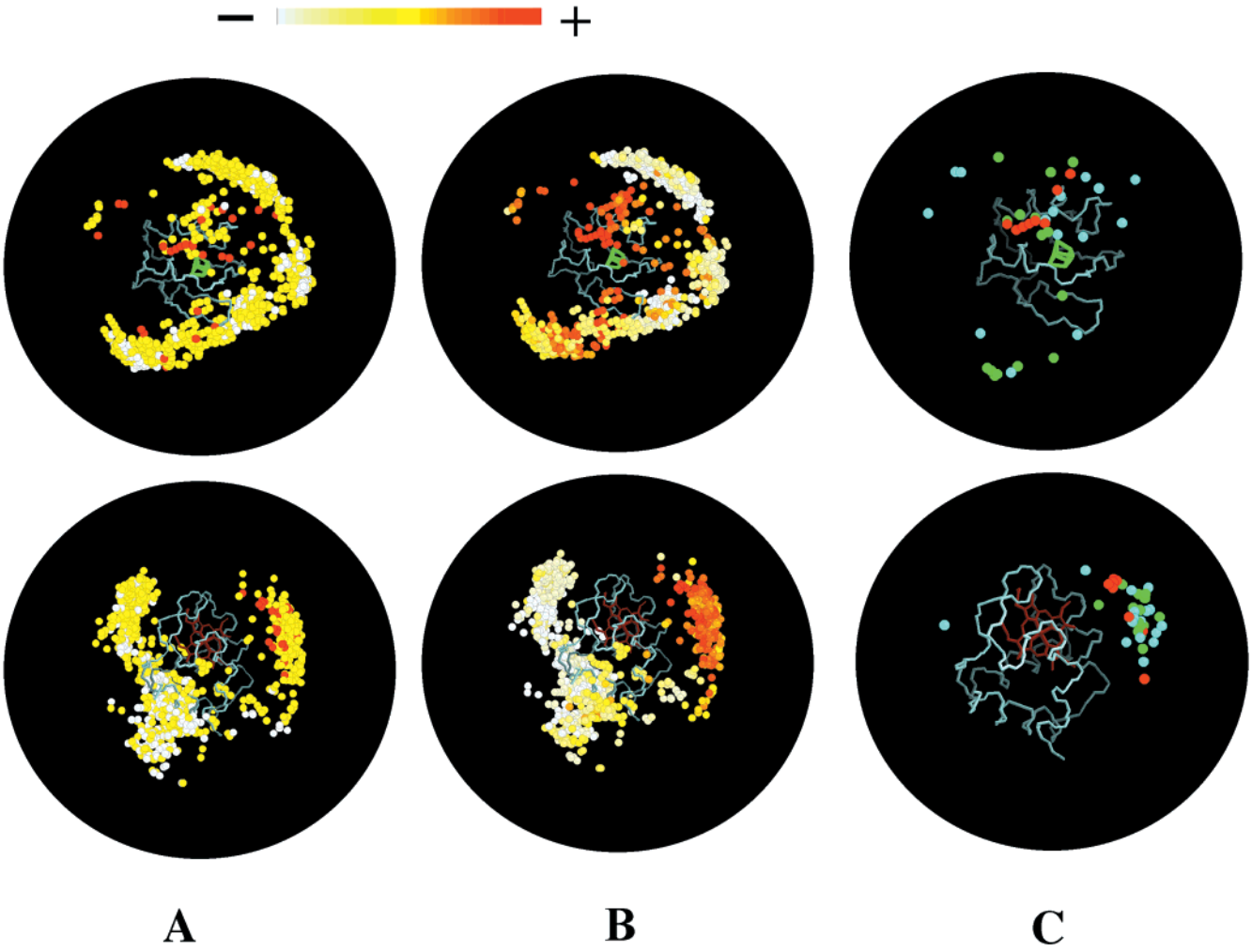

FIGURE 6: Comparison of the *ab initio* docking with the NMR filtering. In the top row, ferredoxin is shown as a backbone trace and the center of mass of the cytochrome, in each of the 1000 putative docking positions, is represented by a small sphere. The bottom row represents, in the same way, the docking positions of the ferredoxin, mapped around the structure of the cytochrome. The spheres are color coded to indicate: (A) the relative *ab initio* interaction score of the corresponding solution or (B) the level of agreement between each solution and the set of NMR mapping. Panel C compares the 25 best docking solutions, according to the *ab initio* (cyan) and the NMR mapping (green) scoring methods. Solutions which are common to both sets are colored in red.

Table 2: Structural Comparison of the Three Final Models of the Cytochrome c_{553}−Ferredoxin I Complex

properties	solution 1	solution 2	solution 3
energy (kcal mole^{-1})	−5738	−5699	−5694
VDW-energy (kcal mole^{-1})	−602	−620	−590
electrostatic energy (kcal mole^{-1})	−5743	−5660	−5713
interface area (Å^2)	1015.9	1037.4	835.4
% of surface covered	13.5	13.8	11.1
atoms on interface	101	130	116
nonpolar contacts	17	32	25
hydrophilic contacts	20	16	21
distance Fe_4S_4−Fe (Å)	12.36	10.0	12.28

In general, the three best "docking NMR-filtered" solutions obtained for the cytochrome c_{553}−ferredoxin I complex seem to have roughly equivalent probabilities of occurrence. When analyzing the three solutions, with respect to the criteria statistically gathered on 75 recognition sites of protein−protein complexes (3), all three solutions have interfaces with properties close to those already described (around 1000 Å^2) (Table 2). Complex solution 1 has the most negative total energy, a reasonable size of interface within the range classified as "small-size" following Lo Conte et al. (*3*), and the distance between the two electron-transfer centers is 12.36 Å. Complex solution 3 has the smallest interface area, clearly at the lower limit of a typical surface area buried in complexes, however, on energetic terms, it seems more favorable than solution 2. It also seems to be less favorable than solution 2 due to the distance of 12.28 Å separating the electron-transfer centers. Complex solution 2 has an interface area of 1037.4 Å^2, falling close to the range of generally observed recognition interfaces, but also has the shortest distance of 10.0 Å between the centers, which are supposed to be brought into catalytic efficient close contact.

Remarkably, the buried area of the large electron-transfer complex cytochrome peroxidase−cytochrome *c* is one of the smallest reported. This seems to indicate that in electron-transfer complexes the contact area is small and structural changes are weak; this has already been pointed out by Lo Conte et al. (*3*).

The three solutions provide the same interacting site involving the residues affected in NMR experiments. The docking simulation give the relative orientation of one molecule to the other, around this interacting site. The heme and Fe−S cluster are thus facing each other with a small variation of the orientations. The heme−FeS distances are

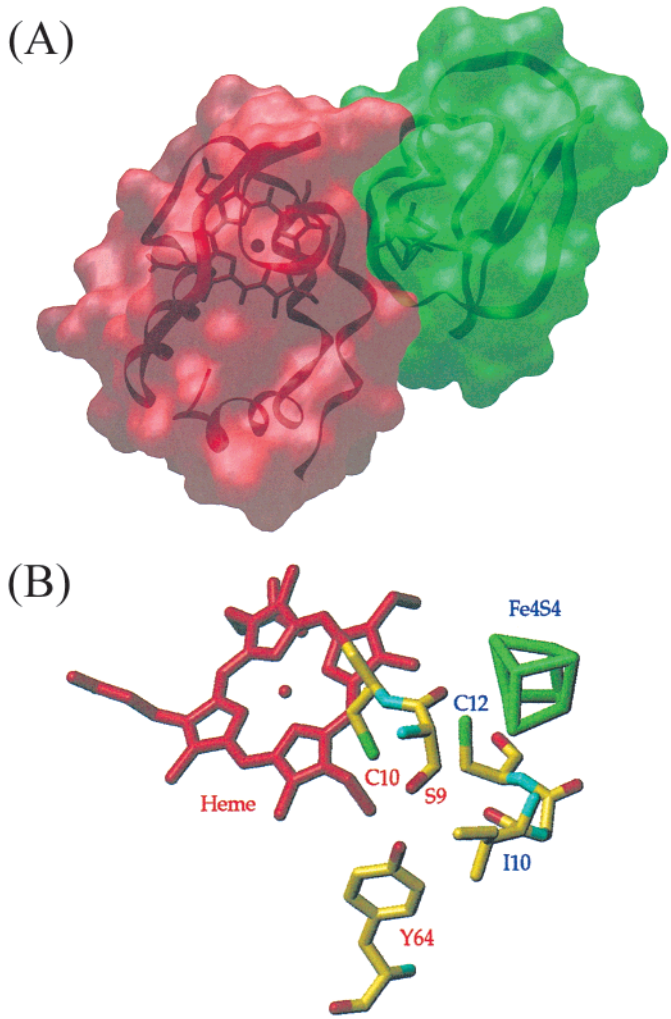

FIGURE 7: Model of solution 2 using the GRASP software (*28*). (A) Ribbon model of the complex with the molecular surface in transparency. The ferredoxin is in green and the cytochrome c_{553} is in red. (B) View of the interface. Both the heme (red) of the cytochrome and the cluster (green) of the ferredoxin are shown. Ile10 and Cys12 of the ferredoxin (blue labels) and Ser9, Cys10, and Tyr64 of the cytochrome (red labels) are represented in cpk color code.

quite similar (from 10 to 12.36 Å). The modification of this orientation is correlated to a variation of the interacting residues, highlighted in Table 2 by a difference of nonpolar contacts (from 17 to 31) and hydrophilic contacts (from 16 to 21). These interactions are essential to definitively conclude on the recognition process. Two alternatives are possible here (1) either to investigate the electron-transfer complex by structural NMR or (2) to study the complex by site-directed mutagenesis. Concerning structural NMR, ^{13}C-labeling would be necessary. However, we have already reported site-directed mutagenesis on the cytochrome c_{553}, and we can analyze the three models in the light of these previous results. We have already reported that Lys62 and Lys63 residues of cytochrome c_{553} are involved in the interaction with the formate dehydrogenase (*25*). Lys63 is found in a salt bridge with glutamate residues in two of the above-described solutions; in solution 2, it is at 3.3 Å from Glu29 of the ferredoxin, whereas it is at 3.2 Å from Glu13 of the ferredoxin in solution 1. In solution 3, however, it makes a hydrogen bond to the main chain carboxyl group of Ile10 in the ferredoxin. Therefore, the three models are in agreement with the idea that Lys63 plays a key role in the recognition process.

We have also reported that Tyr64 in cytochrome c_{553} was required for electron-transfer with formate dehydrogenase (*26*). This tyrosine residue, which is only positioned at the interface area in solution 2, makes a hydrogen bond to Ile10 (3.5 Å), a conserved residue in ferredoxins and in the hydrogenase "ferredoxin-like" domain. In the other two solutions, this tyrosine residue is positioned away from the complex interface and makes intramolecular hydrogen bonds within the cytochrome molecule. Moreover, Ser9, the amino acid in the cytochrome, which is the most perturbed in the heteronuclear NMR experiments (*23*), is at the center of the interaction site in solution 2. For these reasons, we think that solution 2 is the best solution (Figure 7A).

The analysis of the molecular interface of solution 2 (besides Tyr64-OH/Ile10-O) shows another short distance existing between atoms of the ligands of the two redox centers (Figure 7B). A S−S contact of 4.1 Å between Cys12 of the ferredoxin and Cys10, covalently linked to the porphyric ring of the cytochrome, is observed. Short S−S van der Waals contacts between sulfur atoms, making overlaps of the π-orbitals possible, are known to be electron-transferring features; they are responsible for supraconductivity in organo−metallic charge-transfer complexes (*27*). This contact is only present in complex solution 2. It represents a possible short electron pathway to facilitate the electron-transfer from Cys12, one bond is directly leading to the Fe_4S_4 center, and Cys10 is covalently linked to the heme six bonds away from the iron. Ser9 of the cytochrome is also at the interface. The role of Ser9 and Tyr64 in the protein recognition and/or the electron-transfer pathway is not yet very well-understood. The present model of the complex gives us a new hypothesis for site-directed mutagenesis experiments. The structure modelized here is the first complete description of an interface involving a heme and a [4Fe−4S] cluster. The strategy hereby described can be helpful to get an overall view of electron-transfer complexes, such as the ones involving hydrogenase, formate dehydrogenase, or other enzymes of this type.

ACKNOWLEDGMENT

The authors would like to thank Dr. Olivier Bornet for helpful technical assistance and Susan D. Wells for reading the manuscript. We are grateful to the French embassy and to the ICCTI which support this work (project n°316C2). The NMR equipment was partially provided by the Conseil Général des Bouches-du-Rhône (France).

REFERENCES

1. Janin, J., and Chothia, C. (1990) *J. Biol. Chem. 265*, 16027−16030.
2. Janin, J. (1997) *Proteins: Struct., Funct., Genet. 28*, 153−161.
3. Lo Conte, LL., Chothia, C., and Janin, J. (1999) *J. Mol. Biol. 285*, 2177−2198.
4. Pellecchia, M., Sebbel, P., Hermanns, U., Wuthrich, K., and Glockshuber, R. (1999) *Nature Struct. Biol. 6*, 336−339.
5. Palma, P. N., Krippahl, L., Wampler, J. E., and Moura, J. J. G. (1999) *in press.*
6. Genthner, B. R. S, Friedman, S. D., and Devreux, R. (1997) *Int. J. Sys. Bacteriol. 47*, 889−892.
7. Aubert, C., Leroy, G., Bruschi, M., Wall, J., and Dolla, A. (1997) *J. Biol. Chem. 272*, 15128−15134.

8. Voordouw, G., Walker, J. E., and Brenner, S. (1985) *Eur. J. Biochem. 148*, 509−514.
9. Bruschi, M. H., Guerlesquin, F. A., Bovier-Lapierre, G. E., Bonicel, J., and Couchoud, P. M. (1985) *J. Biol. Chem. 260*, 8292−8296.
10. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). Molecular Cloning, a laboratory manual. Second Edition, pp. A-3, Cold Spring Harbor Laboratory Press.
11. Blanchard, L., Marion, D., Pollock, B., Voordouw, G., Wall, J., Bruschi, M., and Guerlesquin, F. (1993) *Eur. J. Biochem. 218*, 293−301.
12. Wishart, D. S., and Sykes, B. D. (1994). *J. Biomol. NMR 4*, 171−180.
13. Lebrun, E. (1998) *Magn. Reson. Chem. 36*, 913−920.
14. Kissinger, C. R, Sieker, L. C, Adman, E. T., and Jensen, L. H. (1991) *J. Mol. Biol. 219*, 693−715.
15. Roussel, A., and Cambillau, C., (1992) *TURBO−FRODO program, the manual*, Biographics, AFMB, Marseille, France.
16. Brünger, A. T. (1996). *X-PLOR program*, Version 3.8, Yale University, New Haven, CT.
17. Jorgensen, W., and Tirado-Rives, J. (1988) *J. Am. Chem. Soc. 110*, 1657−1666.
18. Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, N. F., DiNola, A., and Haak, J. R. (1984) *J. Chem. Phys. 81*, 3684−3690.
19. Ryckaert, J.-P., Ciccotti, G., and Berendsen, H. J. C. (1997) *J. Comput. Phys.* 23, 327−341.
20. Nicolet, Y., Piras, C., Legrand, P., Hatchikian, C. E., and Fontecilla-Camps, J. C. (1999). *Structure Fold. Des. 7*, 13−23.
21. Powell, M. J. D. (1977) *Mathematical Programming 12*, 241−254.
22. Maurer, M. C., Trosset, J. Y., Lester, C. C., DiBella, E. E, and Scheraga, H. A. (1999) *Proteins: Struct., Funct., Genet. 34*, 29−48.
23. Morelli, X., and Guerlesquin, F., (1999) *FEBS Letters 460*, 77−80.
24. Burkhard, P., Hommel, U., Sanner, M., and Walkinshaw, M. D. (1999) *J. Mol. Biol. 287*, 853−858.
25. Sebban-Kreuzer, C., Dolla, A., and Guerlesquin, F. (1998a). *Eur. J. Biochem. 251*, 787−794.
26. Sebban-Kreuzer, C., Blackledge, M., Dolla, A., Marion, D., and Guerlesquin, F. (1998) *Biochemistry* 37, 8331−8340.
27. Bousseau, M., Valade, L., Legros, J. P., Cassoux, P., Garbauska, M. (1986) *J. Am. Chem. Soc. 108*, 1908−1916
28. Nicholls, A., Sharp, K., and Honig, B. (1991) *Proteins: Struct., Funct., Genet. 11*, 281−296.

BI992306S