[4] Rigid Body Refinement of Protein Complexes with Long-Range Distance Restraints from Pulsed Dipolar ESR

By JAYA BHATNAGAR, JACK H. FREED, and BRIAN R. CRANE

Abstract

The modeling of protein–protein complexes greatly benefits from the incorporation of experimental distance restraints. Pulsed dipolar electron spin resonance spectroscopy is one such powerful technique for obtaining long-range distance restraints in protein complexes. Measurements of the dipolar interaction between two spins placed specifically within a protein complex give information about the spin–spin separation distance. We have developed a convenient method to incorporate such long-range distance information in the modeling of protein–protein complexes that is based on rigid body refinement of the protein components with the software Crystallography and NMR System (CNS). Factors affecting convergence such as number of restraints, error allocation scheme, and number and position of spin labeling sites were investigated with real and simulated data. The use of 4 to 5 different labeling sites on each protein component was found to provide sufficient coverage for producing accuracies limited by the uncertainty in the spin-label conformation within the complex. With an asymmetric scheme of allocating this uncertainty, addition of simulated restraints revealed the importance of longer distances within a limited set of total restraints. We present two case studies: (1) refinement of the complex formed between the histidine kinase CheA and its coupling protein CheW, and (2) refinement of intra-helical separations in the protein a-synuclein bound to micelles.

Introduction

Elucidation of the structures of protein complexes is often critical for understanding molecular mechanism and function. This is no more evident than for two-component signaling systems where transient associations of proteins mediate the propagation of information. Despite numerous successes, the structure determination of complexes remains a challenge because of the difficulty in growing crystals for X-ray crystallography or in obtaining enough suitable small distance and orientation restraints by NMR. Techniques such as electron microscopy (Frank, 1996), small-angle X-ray scattering (Glatter and Kratky, 1982), and small-angle neutron scattering (Chen and

METHODS IN ENZYMOLOGY, VOL. 423

0076-6879/07 $35.00
DOI: 10.1016/S0076-6879(07)23004-6

Bendedouch, 1986) can provide molecular envelopes for complexes but the results suffer from lack of contrast and resolution. In addition, a number of useful approaches map molecular interfaces by measuring perturbations to interfacial residues, such as changes in cross-linking reactivity, accessibility, or NMR chemical shifts. While these methods provide points of contact between partners, relative orientations can be difficult to discern. Finally, distance measurements between specifically labeled positions on associating molecules are possible with FRET and ESR. The former relies on resonance energy transfer between a donor excited state and an accepter ground state; the latter relies on the direct dipolar coupling between two spins. In each case, probe positioning is often achieved with site-directed cysteine substitution, but whereas FRET usually employs two different types of labels, ESR requires only one, usually a nitroxide derivative. Also, ESR provides the distance directly, since it does not require calibrations nor does it have uncertain parameters. In addition, the distribution in distance, P(r), can readily be obtained. Pulsed ESR techniques, such as Double Electron Electron Resonance (DEER) and Double Quantum Resonance (DQC), are capable of measuring biologically relevant distances in the range of 1 to 8 nm between spin labels (Borbat and Freed, 2007; Chiang *et al.*, 2005).

Such long-range pairwise distance restraints can, in principle, be processed to formulate precise structures. Related methodologies have already been applied to FRET-derived distances (Knight *et al.*, 2005; Mukhopadhyay *et al.*, 2004). In the present study, we have developed a simple and convenient method for modeling the structure of a binary complex by rigid body refinement of known substructures, using as restraints the intermolecular distances derived from pulsed ESR. Also, by testing simulated restraints, we produced a set of guidelines to optimize spin label location, number of labels, and measurement error schemes for achieving reasonable model accuracies. Our method is general enough to be applied to any type of distance restraints provided a reasonable estimate of uncertainty associated with the particular measurement is known.

Method

Rigid Body Minimization with CNS

The software package Crystallography and NMR System (CNS) developed by Brunger *et al.* (1998) is primarily designed for structure determination using data from X-ray crystallography or nuclear magnetic resonance (NMR) spectroscopy. However, the optimization algorithms can readily be applied to other types of structural restraints. Our incorporation of ESR-measured distances into CNS is quite similar to that for distances

derived from Nuclear Overhauser Effect (NOE) data in structure refinement. Distances (d) were input in the form of a table, each line of which specifies the pair of atoms between which the d has been measured and the error limits, d_{minus} and d_{plus}, which represent the minimum and maximum allowed distances associated with that measurement. CNS provides six possible restraining functions associated with NOE-derived distances: biharmonic function, square-well function, soft-square function, symmetry function, 3D NOE-NOE function, and high dimensional function. We have applied the soft square potential (Brunger *et al.*, 1998) to generate the energy term (E_{ESR}), which is then minimized by conjugate gradient refinement based on the agreement between measured distances "d" and model distance "R." Taking default values for most of the constants[1], a simplified version of the function has the form:

$$E_{ESR} = S \times \begin{cases} a + b/\Delta + \Delta; & R > d + d_{plus} + r_{sw} \\ \Delta^2; & R < d + d_{plus} + r_{sw} \end{cases} \quad (1)$$

$$\text{where } \Delta = \begin{cases} R - (d + d_{plus}); & d + d_{plus} < R \\ 0; & d - d_{minus} < R < d + d_{plus} \\ d - d_{minus} - R; & R < d - d_{minus} \end{cases}$$

We assign "d" as the ESR-measured distance between C_β atoms of the corresponding amino acid residues at which the spin label is attached, R is the corresponding distance in the model, d_{plus} and d_{minus} are the positive and negative errors associated with each distance, r_{sw} is a constant with default value 0.5 Å, a and b are determined by the program such that E_{ESR} is a smooth function at point $R = d + d_{plus} + r_{sw}$. S is a scale factor that weights the ESR energy relative to the van der Waals energy. A similar soft square potential has also been used to model constraints for a system of transmembrane helices (Sale *et al.*, 2004) and is analogous to a global penalty function developed by Knight *et al.* (2005) for modeling FRET-derived restraints. However, an additional property of the restraining function in CNS is that it becomes linear for large deviations between experimental and model restraints. This allowance maintains numerical stabilities (Brunger *et al.*, 1999).

Initial Conformation of the Complex

Gradient-descent optimization methods such as conjugate gradient minimization converge to a global minimum of the system if the starting conformation is not very different from the correct structure. Various computational

[1] Softexp = 1; Exp = 2,C = 1; c = 1; $d_{off} = 0$.

procedures have been developed to model initial conformations for refinement, provided distance restraints are available (Faulon *et al.*, 2003; Sale *et al.*, 2004). For our first case study of the complex between chemotaxis proteins CheA and CheW, we test initial conformations determined randomly with those generated with matrix distance geometry from both X-ray crystallography and pulsed ESR, as discussed in our previous work (Borbat and Freed, 2007; Park *et al.*, 2006). With our second case, the protein alpha-synuclein (aS), we compare refinements beginning with either the NMR-determined structure or random orientations of the two synuclein helices.

Evaluation Criterion

In case 1, the separate structures of CheW and the P5 domain of CheA were taken from the crystal structure of CheW and the CheA domains P4-P5, where CheW predominantly binds to the CheA domain, P5 (Park *et al.*, 2006). The final conformation of the complex after rigid body minimization was evaluated by comparing the ESR-refined complex to the coordinates of the P4:P5:CheW crystal structure. The tight binding between CheW and CheAΔ289 (domains P3,P4,P5 collectively called CheAΔ289; $K_b = 100$ nM) makes it unlikely that crystal packing forces significantly alter the association mode of the complex (Park *et al.*, 2004). The measure of agreement was assigned as the root-mean-square deviation (RMSD) in the position of C_α atoms of CheW in the final refined structure with respect to the crystal structure after least square fitting of the P5 domains from both structures (McRee, 1999).

In case 2, we aimed to reproduce the orientation of two anti parallel helices of α-synuclein when bound to micelles, for which an NMR structure has been determined (Bussell *et al.*, 2005; Ulmer *et al.*, 2005). Interhelical distances measured by ESR give information about relative orientation of the helices, which cannot be determined with certainty from NMR data alone. For comparison, the quality of the ESR-refined structure was evaluated by superimposing one of the two α-synuclein helices with the NMR structure of the molecule bound to micelles (Ulmer *et al.*, 2005) and then calculating the RMSD between the ESR refined second helix and that from the NMR structure.

Results

Case Study 1

CheA:CheW complex. CheW forms a complex with the histidine kinase CheA that is necessary for assembly with chemorecepters. To construct the structure of the CheA:CheW complex, 12 intermolecular distances were

measured between nitroxide spin labels on four residues (N553C, S568C, E646C, and D579C) of the P5 domain of *T. maritima* CheAΔ289 (which contains domains P3-P4-P5) and three residues (S15C, S72C and S80C) on *T. maritima* CheW (Fig. 1). Positioning of the labels was achieved by site-directed cysteine mutagenesis followed by reaction with (1-oxyl-2,2,5,5-tetramethylpyrolinyl-3-methyl)-methanethiosulfonate (MTSSL). An initial conformation of the complex was predicted by using a matrix distance geometry method, and it was found to agree with a root mean square deviation (RMSD) in C_α atom positions of about 16 Å when compared with the crystal structure of the complex. Rigid-body refinement using CNS reduced the RMSD to 11 Å. The total energy function in the refinement of CNS is a sum of $E_{EMPIRICAL}$ and $E_{EFFECTIVE}$ terms (Brunger *et al.*, 1998). This force field is similar to the Bundler penalty function used to model transmembrane helices against sparse distance constraints (Sale *et al.*, 2004). $E_{EMPIRICAL}$ describes the energy of the molecule as a function of atomic coordinates (energy associated with bonds, angles, dihedral angles, etc.), whereas $E_{EFFECTIVE}$ refers to restraining energy terms associated with agreement of the model to the ESR data, that is, it equals E_{ESR} given by Equation (1). In rigid body refinement, only energy terms that reflect van der Waals contacts contribute to $E_{EMPIRICAL}$.

The convergence was tested by randomly orienting CheW in various positions and evaluating the refined complex. Within rigid body displacements of 15 Å and rotations of 30°, the same final conformation was found (within an RMSD difference of ~3 Å). In the following sections, we

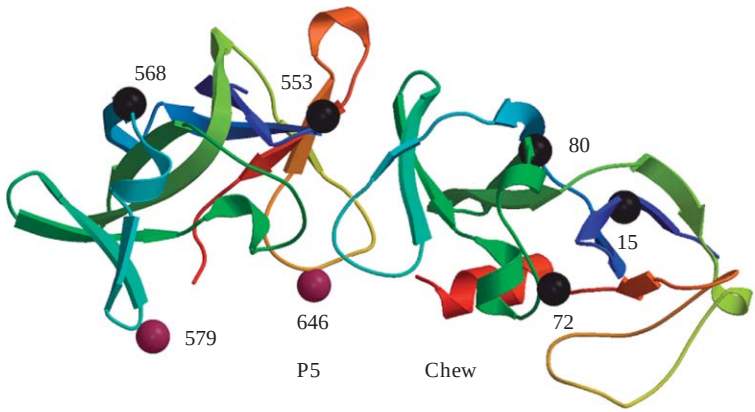

FIG. 1. Crystal structure of CheW-P5 complex showing positions of spin label sites (balls) along the polypeptide. Both proteins shown as ribbon representations colored blue to red from N to C terminus. Sites producing the most aberrant ESR restraints compared to the crystal structure shown in red.

investigate how parameterization of the refinement affects the quality of the final solution. By adding restraints comprising distances taken from the crystal structure with errors derived from the standard deviations observed in the ESR measurements, we also explore how the number and nature of distance restraints affect the modeling results. In particular, we present guidelines to aid selection of potential spin labeling sites on the protein components within a general complex.

Error Allocation Scheme

A pulsed ESR experiment with a pair of nitroxide spin labels measures the separation between the nitroxyl groups of the spin labels, which can have considerable orientational freedom with respect to the protein backbone and with respect to each other because of their flexible tethers. In the absence of information about the spin label orientation, we have assigned the ESR experimental distance to coincide with the C_β position of the native amino acid residue. If the spin label tethers point away from each other in the complex, the model distances will underestimate the nitroxide separations. In fact, the ESR measured distances are almost always larger than those predicted by C_β separations (Table I). In contrast, if the spin label tethers project toward each other in the complex, then the spin–spin separation will be overestimated by C_β separations. However, when globular domains associate, there is a bias against facing labels because they tend to reside on protein surfaces that participate in the interface.

TABLE I
COMPARISON OF ESR-MEASURED DISTANCES TO $C_{\beta\beta}$ SEPARATIONS BETWEEN CORRESPONDING RESIDUES IN THE P5-CHEW CRYSTAL STRUCTURE

Residue P5-CheW	Distance between C_β atoms in crystal structure(R_{crys}) (Å)	ESR measured distances (R_{esr}) (Å)	$R_{esr} - R_{Crys}$ (Å)
553–15	34.9	37	2.1
646–15	31.8	43.7	11.9
568–15	55.4	54.5	−0.9
579–15	52	61	9
553–72	28.3	27	−1.3
646–72	27.5	32.5	5
568–72	47.9	49	1.1
579–72	41	46	5
553–80	23.6	26	2.4
646–80	26.8	39.5	12.7
568–80	44.5	47	2.5
579–80	44.5	54.5	10

To compensate for overall longer experimental distances, on average, we have found that an asymmetric uncertainty model is effective. In previous work (Park *et al.*, 2006), we presented a distance-dependent error allocation scheme. However, better results are obtained by setting $d_{minus} = 5$ Å and $d_{plus} = 1$ Å for all restraints which are the boundaries within which most of the experimental distances are over- or underestimated by C_β separations (Table I). Similar magnitudes in error are consistent with other spin labeling studies (Faulon *et al.*, 2003; Rabenstein and Shin, 1995) which may also benefit from asymmetric error boundaries. However, 4 out of 12 distances do not meet these criteria due to reasons related to the location of spin label site on the protein surface. The reasons for such inaccuracies in distance measurements are discussed in a later section.

Weighting Scheme for Contact Parameters

For ESR restraints to determine the final configuration, $E_{EFFECTIVE}$ must account for a considerable percentage of the total energy. This can be achieved by simply increasing the scale factor (S) in the input file associated with E_{NOE}, or in our case, E_{ESR}. With the ceiling constant assigned to 10^5, the scale factor was increased from 75 to 75,000 in steps of 100 and the RMSD in C_α positions were evaluated. Predictably, the convergence improves progressively as the scale factor increases ($S = 75$ yields an RMSD = 16.38 Å; $S = 75{,}000$, an RMSD = 11.06 Å). Above $S = 75{,}000$, there is no further improvement.

Type and Number of Restraints

Applying all 12 experimental intermolecular distance restraints between CheW and P5 domain, while setting $d_{minus} = 5$ Å and $d_{plus} = 1$ Å, the best structure that could be achieved has an RMSD on C_a positions of 11.06 Å compared to the crystal structure. To evaluate the effect of additional arbitrarily chosen distance restraints, four new label sites on CheW were successively added to the refinement. Each new site generated four new distances to the P5 labels. The standard deviation of the parameter (R_{esr}–R_{crys}) as defined in Table I is 5 Å. In order for the new distances to mimic the experimental ones, the standard deviation obtained previously was added to the C_β separations. Then, for each successive addition of a label site on CheW, the RMSD in positions of the C_α atoms in the final structure was calculated, and the results were plotted against total number of restraints. Two error schemes, $d_{minus} = 5$ Å, $d_{plus} = 1$ Å (Fig. 2A) and $d_{minus} = 5$ Å, $d_{plus} = 5$ Å (Fig. 2B), were used for comparison. The procedure was also repeated for five different initial conformations of the complex. The results indicate that irrespective of the initial conformation

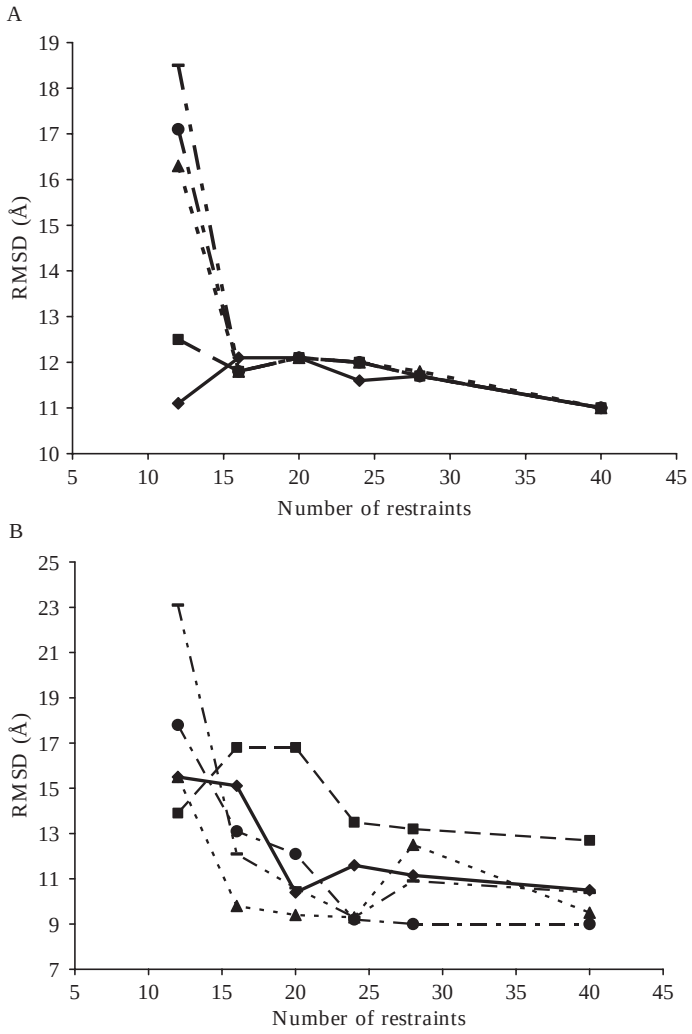

FIG. 2. The effects of different error schemes and simulated restraints on refinement accuracy. RMSDs for the refined CheW/P5 complex are shown for five different initial conformations of the complex ■, ◆, ▲, ●, _ . Two different error schemes: (A) $d_{minus} = 5$ Å, $d_{plus} = 1$ Å and (B) $d_{minus} = 5$ Å, $d_{plus} = 5$ Å.

prior to refinement, addition of random distance restraints leads to an improved RMSD of 8 to 12 Å but beyond 20 and 24 restraints with $d_{minus} = 5$ Å $d_{plus} = 1$ Å and $d_{minus} = 5$ Å $d_{plus} = 5$ Å, respectively, there is no improvement. It is interesting to note that Knight *et al.* (2005) also

reported that model accuracies of only 10 Å RMSD can be obtained with 20 or more FRET restraints.

Addition of More Accurate Restraints

As has been illustrated, about 20 distance restraints with standard deviation of 5 Å from the crystal-structure derived distances are sufficient to produce model accuracies of about 10 Å. The inability of additional restraints to obtain better results suggested that convergence is limited by inaccuracies in the experimental distances. With the initial configuration taken from distance geometry, even the addition of 28 accurate crystal-derived distances (setting $d_{minus} = 1$ Å and $d_{plus} = 1$ Å) to 12 experimental distances (setting $d_{minus} = 5$ Å and $d_{plus} = 5$ Å) only improved the final agreement to a limited degree (from RMSD 15.4 to 10.5 Å); thus, a few distances with large inconsistencies appear to dominate the more accurate restraints. Comparison of 12 experimental distances with crystal separations revealed that 2 of the distances were highly skewed, with average deviations up to 12.7 Å (Table I). If we take the same set of 28 crystal-structure derived distances and 12 experimental distances, and the observed ESR distances are deleted two at a time, beginning with the most deviant ones, the RMSD drastically reduces from 10.5 to 6.2 Å and then becomes constant at 2.6 Å (Fig. 3). Adding only the two highly skewed measurements

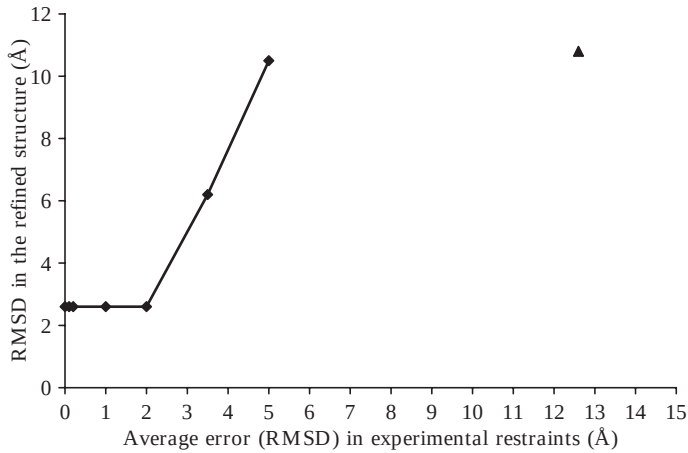

FIG. 3. The effect of aberrant measurements on refinement accuracy in the presence of additional restraints derived from the crystal structure. From a set of 28 crystal distances with ±1 Å and experimental distances with ±5 Å, the most deviant ESR distances were deleted two at a time (◆). Addition of only the two most deviant experimental distances to the defined crystal distances give a slightly higher RMSD compared to the overall set of experimental distances (▲).

to the crystal-structure derived restraints produces a worse RMSD than the entire set of experimental restraints, emphasizing the deleterious effects of these aberrant measurements.

However, in the absence of the simulated restraints, the deletion of the 2 most deviant distances from the set of 12 experimental distances increased the RMSD from 11 to 16 Å. This is probably because the refinement now suffers from underdetermination. Alternatively, since the total energy associated with distance restraints is the sum of individual contributions, improved convergence may result from a weighting scheme based on experimental-to-model agreement that adjusts on successive iterations to reduce the weight of the contribution of aberrant measurements. Simply, if the difference between a measurement and its predicted distance by the refined complex deviates by more than two standard deviations, as given by the distribution of residuals from all the measurements, then the measurement should be removed and the refinement repeated. As we will discuss, due to surface site mobility, interference of labels with complex formation, and other conformational effects, it is reasonable to encounter some outliers in these experiments.

If experimental restraints are deleted successively in the absence of simulated restraints, the RMSD increases as expected. However, the additional increase in RMSD is more sensitive to removal of the shortest, rather than the longest, distance (Fig. 4). This suggests that longer experimental distances in the CheA:CheW system are more inaccurate than shorter ones.

Effect of Spin Label Position

Site-directed spin labeling (SDSL) is a convenient method to attach ESR probes to cysteine residues on proteins (Hubbell and Altenbach, 1994); however, it is unclear how the pattern of sites affects the refinement, apart from the considerations that a solvent-exposed residue is more likely to react with the spin label, and that spin labels in the interfacial region may disrupt complex formation.

To test the effect of label position on predicting the CheA:CheW solution complex, CheW was broadly divided into three sections relative to the CheA interface—front, middle, and back—and from each of these sections, one amino acid residue was randomly selected as a label site. Additional distance restraints from these sites to P5 were measured as before while setting $d_{minus} = 5$ Å and $d_{plus} = 1$ Å. This procedure was repeated seven more times, selecting a random site in each section each time, and finally the RMSD on C_α positions after refinement was averaged for all eight cases per section. The trend in RMSD values

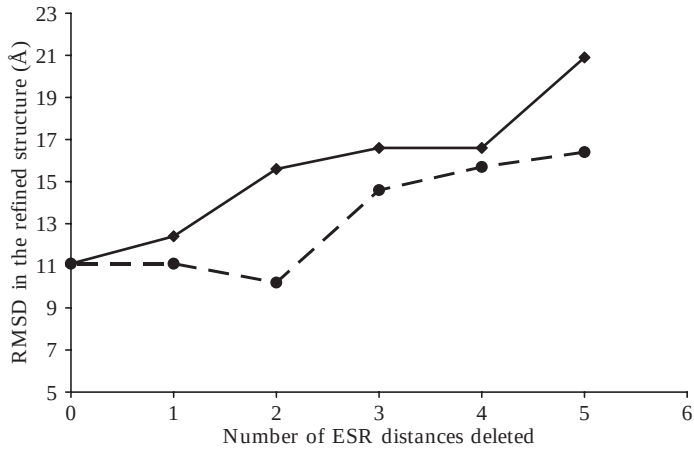

FIG. 4. Long vs short restraints in refinement of CheA/CheW complex. Successive deletions of distances beginning with the shortest (◆) or the longest (●) restraints.

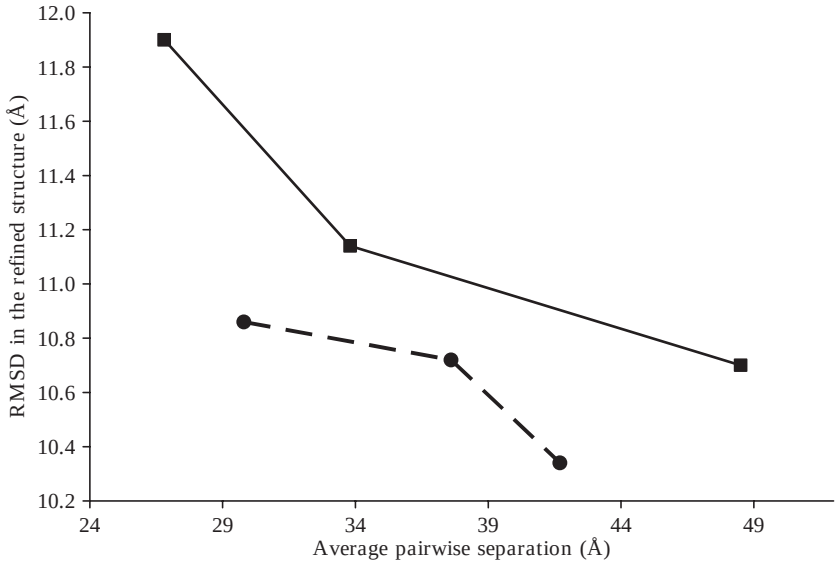

FIG. 5. Variation of RMSD with average pairwise separation of the label sites with addition of one new spin label (■) and four spin labels on CheW (●).

(Fig. 5) showed a slight preference for locating the new sites in the middle and back sections of the protein (from 11.99 to 11.14 to 10.7 Å, respectively).

We also considered the effect of adding four more CheW sites (16 new restraints). The selection of sites was organized the following six ways:

1. All from front section (residues I60, S45, N54, S37)
2. All from middle section (E90, K67, D139, I34)
3. All from back section (V101, K123, N107, N113)
4. Two from front and two from middle section
5. Two from middle and two from back section
6. Two from back and two from front section

For the first three cases, the RMSD from the refined structure shows slightly better agreement with the crystal structure when sites in the distal end of CheW are selected compared to sites closer to the P5 interface. For scenarios 4 through 6, the selection of two residues from each section of CheW was done in eight different ways and the final RMSD was averaged for all eight cases. Plotting the final RMSD in the final structure versus the average pairwise separation of each new label from sites on P5 demonstrates that only minimal improvement in RMSD is seen, no matter how the sites are chosen (Fig. 5).

However, we can conclude that more restraints result in a lower RMSD, and longer restraints play a crucial role only when the total number of restraints is fewer than 16. As more restraints are added, the locations of sites on the surface of CheW have little effect on the refined complex.

Additional considerations affecting the choice of labeling sites are discussed in Chapter 3 of this volume.

Case Study 2

Helix Orientations of α-Synuclein Bound to Micelles. NMR studies on the protein alpha-synuclein (aS) have shown that when bound to sodium dodecyl sulphate (SDS) micelles, the protein adopts a conformation of two separate anti-parallel helices (helix 1 residues: 3–37; helix 2 residues: 45–92) connected by an ordered linker (Bussell *et al.*, 2005; Ulmer *et al.*, 2005). Pulsed ESR has been used to determine interhelix distances between spin labels at various positions on the two helices when the protein is bound to both SDS and lyso-1-palmitoylphosphotidylglycerol (LPPG) micelles (Borbat *et al.*, 2006). In total, 13 interhelical dipolar couplings were measured and from them the average distance (R_{avg}) and its root mean square deviation (RMSD) were evaluated. We tested the ability of our refinement procedure to orient the two helices relative to each other under the assumption that each helix behaves as a rigid body. To generate two rigid bodies, the helices were separated between residues 40 and 41 in the linker. To account for the covalent bonding between residues 40 to 41,

additional restraints were added between residues 40 to 41, 39 to 41, and 40 to 42 (C_i-N_k, $C_{\alpha i}$-$C_{\alpha k}$, C_i-C_k, $C_{\alpha i}$-N_k, N_i-N_k and C_i-$C_{\alpha k}$, for $i = 40$; $k = i + 1$ and $k = i + 2$, for $i = 39$; $k = i + 2$). In this scheme, the restraints were calculated by summing the bond lengths connecting the two atoms of interest and d_{plus} was set to 0 because any distance measured through space is shorter than that measured along the summed bond lengths. For a hypothetical case where no information is available regarding the conformation of the turn residues, the d_{minus} error was given more flexibility by assigning $d_{minus} = 1$ Å for distances between adjacent residues and $d_{minus} = 8$ Å for distances between non-adjacent residues.

For SDS bound α-synuclein (with the exception of two distances, between V3C/E61C and E13C/H50C), 11 interhelical ESR distances, taken as their reported R_{max} values, were incorporated into the refinement. The RMSDs in label position obtained from P(r) measurements were taken as estimates for the d_{minus} error. As the ESR measurements likely overestimate R, as in the CheA/CheW case, d_{plus} was set to a smaller value, but was increased to reflect changes in d_{min} ($d_{plus} = 1$ for 5 Å $< d_{minus} = 8$ Å, $d_{plus} = 2$ for 9 Å $< d_{minus} = 15$ Å and $d_{plus} = 5$ Å for $d_{minus} = 15$ Å). Combining all the restraints, and starting with what was available from the NMR structure, the refinement places the ends of the two rigid helices close to each other (the length of amide bond $C_{40} - N_{41}$ is 2.2 Å compared to ideal bond length 1.3 Å). When the helical fragment from 1 to 40 is superimposed on its position in the NMR structure, the anti-parallel partner helix (residues 41–103) is rotated by an angle $\Phi \sim 30^\circ$ with respect to its position in the NMR structure (Fig. 6). However, the angle separating the two helical axes (θ) is better determined. Thus, the ESR refinement is unable to distinguish which sides of the helices face each other, and this generates inaccuracy in Φ. This is not surprising, since the errors in the spin label position are larger than the width of a helix. We noted that the absolute orientation of the two helices can be determined if precise restraints on the conformation of linker residues are known by other means. If rigid restraints are added for the conformation of residues within the loops, the agreement with the NMR structure is excellent.

Discussion

In this chapter, we have described a simple and readily implemented method for refining association modes of protein complexes from ESR restraints. Agreement with crystal data improves with number of ESR restraints until approximately 20 restraints are available; additional restraints beyond this number result in little further improvement due to errors associated with the knowledge of the label position. A 2005 study

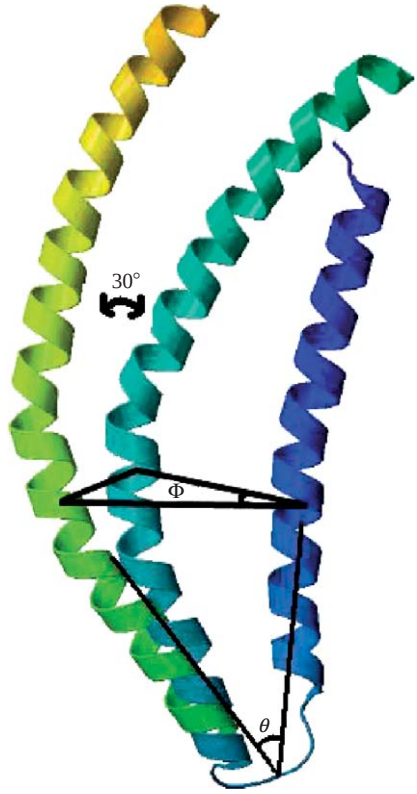

FIG. 6. A comparison of helix orientations in α-synuclein from ESR-refinement and NMR. The orientation of the two anti-parallel α-synuclein helixes (residues 3–34 and 44–94) as derived from NMR is shown in blue. Superposition of N terminal of helix from the rigid body refined structure, places the second helix rotated by angle of 30° with respect to the NMR structure.

reported agreement between C_β to C_β distances and ESR distance restraints with mean errors up to 6 Å (Sale *et al.*, 2005). These errors are similar in magnitude to those accounted for by our asymmetric error scheme in CheA/CheW case study. Molecular modeling approaches such as Monte Carlo simulations and molecular dynamics have been found to be useful in lowering the uncertainty associated with spin-label positions (Borbat *et al.*, 2002; Sale *et al.*, 2005; Schiemann *et al.*, 2004).

Type of Restraints

We investigated how positioning of the spin labels influences convergence of varying accuracies. Addition of longer simulated restraints appeared most effective in driving convergence to the target model,

provided the total number of restraints was fewer than 16. However, removal of the shortest experimental distances has more deleterious effect in the absence of simulated restraints. This apparent contradiction may derive from the longer experimental restraints being unusually aberrant due to conformational properties of these sites in the CheA:CheW system. In addition, any real differences between the solution and crystal complex would be expected to be greatest at sites farthest from the high-affinity interface. Nonetheless, our combination of studies suggests that spin labeling 4 or 5 sites on each protein at positions distributed as far apart as possible on the structures of the individual components is a reasonable strategy for covering the distance space.

Inaccuracies in Distance Measurement

Apart from the technical limitations of the experimental method in measuring accurate distances (cf. Borat and Freed, 2007), local conformational changes in the protein structure, backbone dynamics, and the flexibility of the spin label lead to ambiguity in measurements. The two most deviant intermolecular distances in the CheW/P5 complex were those measured from site 646 on P5 domain (P5/CheW: 646-15, 646-80). In the crystal structure of the complex, P5-646 is very close to the binding interface with CheW, and thus the label conformation may be unusually perturbed in the complex. In addition, the 646 site resides in a loop, which may impart more than usual flexibility (Fig. 1). Aberrant distances involving P5 site 579 may also be caused because this residue resides in a loop with few neighbor contacts and, hence, may be more mobile.

Spatial Resolution of ESR-Derived Structures

Case 2 demonstrates that this method as implemented is less effective at orienting secondary structure elements within a protein than at defining association modes within the complex. It follows that, even with a large number of measurements, it may be difficult to precisely define conformational changes involving small to medium amplitude shifts in secondary structure positions. This limitation could be overcome by more rigid spin labels whose positions on the protein surface are fixed and well defined. In this regard, metal complexes may be an attractive alternative to nitroxide-based labels (Rodriguez-Castaneda *et al.*, 2006).

In conclusion, pulsed dipolar ESR, combined with site-directed spin labeling, can reconstitute structures of protein–protein complexes with reasonable accuracies provided structures of the individual components are well defined. CNS-based rigid-body refinement is a straightforward and accessible method for generating complexes from the distance

restraints. Further improvements may be possible with a weighting scheme that identifies and adjusts the contribution of outliers during the course of refinement.

Acknowledgments

This work has been supported by grants from the National Institutes of Health GM: R01066775 (to B.R.C), NCRR: P41-RR016292 and NIBIB: R01-EB03150 (to J.H.F).

References

Borbat, P., Ramlall, T. F., Freed, J. H., and Eliezer, D. (2006). Inter-helix distances in lysophospholipid micelle-bound alpha-synuclein from pulsed ESR measurements. *J. Am. Chem. Soc.* **128,** 10004–10005.

Borbat, P. P., and Freed, J. H. (2007). Measuring distances by pulsed dipolar ESR spectroscopy: Spin-labeled histidine kinases. *Methods Enzymol.* **423,** 52–116.

Borbat, P. P., Mchaourab, H. S., and Freed, J. H. (2002). Protein structure determination using long-distance constraints from double-quantum coherence ESR: Study of T4 lysozyme. *J. Am. Chem. Soc.* **124,** 5304–5314.

Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998). Crystallography and NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. , Sect. D: Biol. Crystallogr.* **54,** 905–921.

Brunger, A. T., Adams, P. D., and Rice, L. M. (1999). Annealing in crystallography: A powerful optimization tool. *Prog. Biophys. Mol. Biol.* **72,** 135–155.

Bussell, R., Ramlall, T. F., and Eliezer, D. (2005). Helix periodicity, topology, and dynamics of membrane-associated alpha-Synuclein. *Protein Sci.* **14,** 862–872.

Chen, S. H., and Bendedouch, D. (1986). Structure and interactions of proteins in solution studied by small-angle neutron scattering. *Methods Enzymol.* **130,** 79–116.

Chiang, Y. W., Borbat, P. P., and Freed, J. H. (2005). The determination of pair distance distributions by pulsed ESR using Tikhonov regularization. *J. Magn. Reson.* **172,** 279–295.

Faulon, J. L., Sale, K., and Young, M. (2003). Exploring the conformational space of membrane protein folds matching distance constraints. *Protein Sci.* **12,** 1750–1761.

Frank, J. (1996). ''Three-Dimensional Electron Microscopy of Macromolecular Assemblies.'' Academic Press, San Diego, CA.

Glatter, D., and Kratky, O. (1982). ''Small Angle X-ray Scattering.'' Academic Press, London, UK.

Hubbell, W. L., and Altenbach, C. (1994). Investigation of structure and dynamics in membrane-proteins using site-directed spin-labeling. *Curr. Opin. Struct. Biol.* **4,** 566–573.

Knight, J. L., Mekler, V., Mukhopadhyay, J., Ebright, R. H., and Levy, R. M. (2005). Distance-restrained docking of rifampicin and rifamycin SV to RNA polymerase using systematic FRET measurements: Developing benchmarks of model quality and reliability. *Biophys. J.* **88,** 925–938.

McRee, D. E. (1999). XtalView Xfit—A versatile program for manipulating atomic coordinates and electron density. *J. Struct. Biol.* **125,** 156–165.

Mukhopadhyay, J., Sineva, E., Knight, J., Levy, R. M., and Ebright, R. H. (2004). Antibacterial peptide microcin J25 inhibits transcription by binding within and obstructing the RNA polymerase secondary channel. *Mol. Cell* **14,** 739–751.

Park, S. Y., Quezada, C. M., Bilwes, A. M., and Crane, B. R. (2004). Subunit exchange by CheA histidine kinases from the mesophile *Escherichia coli* and the thermophile *Thermotoga maritima*. *Biochemistry* **43,** 2228–2240.

Park, S. Y., Borbat, P. P., Gonzalez-Bonet, G., Bhatnagar, J., Pollard, A. M., Freed, J. H., Bilwes, A. M., and Crane, B. R. (2006). Reconstruction of the chemotaxis receptor–kinase assembly. *Nat. Struct. Mol. Biol.* **13,** 400–407.

Rabenstein, M. D., and Shin, Y. K. (1995). Determination of the distance between 2 spin labels attached to a macromolecule. *Proc. Natl. Acad. Sci. USA* **92,** 8239–8243.

Rodriguez-Castaneda, F., Haberz, P., Leonov, A., and Griesinger, C. (2006). Paramagnetic tagging of diamagnetic proteins for solution NMR. *Magn. Reson. Chem.* **44,** S10–S16.

Sale, K., Faulon, J. L., Gray, G. A., Schoeniger, J. S., and Young, M. M. (2004). Optimal bundling of transmembrane helices using sparse distance constraints. *Protein Sci.* **13,** 2613–2627.

Sale, K., Song, L. K., Liu, Y. S., Perozo, E., and Fajer, P. (2005). Explicit treatment of spin labels in modeling of distance constraints from dipolar EPR and DEER. *J. Am. Chem. Soc.* **127,** 9334–9335.

Schiemann, O., Piton, N., Mu, Y. G., Stock, G., Engels, J. W., and Prisner, T. F. (2004). A PELDOR-based nanometer distance ruler for oligonucleotides. *J. Am. Chem. Soc.* **126,** 5722–5729.

Ulmer, T. S., Bax, A., Cole, N. B., and Nussbaum, R. L. (2005). Structure and dynamics of micelle-bound human alpha-synuclein. *J. Biol. Chem.* **280,** 9595–9603.