# Structural Refinement of Membrane Proteins by Restrained Molecular Dynamics and Solvent Accessibility Data

Pornthep Sompornpisut, Benoît Roux, and Eduardo Perozo
Department of Biochemistry and Molecular Biology and Institute for Biophysical Dynamics, University of Chicago, Chicago, Illinois

ABSTRACT We present an approach for incorporating solvent accessibility data from electron paramagnetic resonance experiments in the structural refinement of membrane proteins through restrained molecular dynamics simulations. The restraints have been parameterized from oxygen ($\Pi O_2$) and nickel-ethylenediaminediacetic acid ($\Pi$NiEdda) collision frequencies, as indicators of lipid or aqueous exposed spin-label sites. These are enforced through interactions between a pseudoatom representation of the covalently attached Nitroxide spin-label and virtual "solvent" particles corresponding to $O_2$ and NiEdda in the surrounding environment. Interactions were computed using an empirical potential function, where the parameters have been optimized to account for the different accessibilities of the spin-label pseudoatoms to the surrounding environment. This approach, "pseudoatom-driven solvent accessibility refinement", was validated by refolding distorted conformations of the *Streptomyces lividans* potassium channel (KcsA), corresponding to a range of 2–30 Å root mean-square deviations away from the native structure. Molecular dynamics simulations based on up to 58 electron paramagnetic resonance restraints derived from spin-label mutants were able to converge toward the native structure within 1–3 Å root mean-square deviations with minimal computational cost. The use of energy-based ranking and structure similarity clustering as selection criteria helped in the convergence and identification of correctly folded structures from a large number of simulations. This approach can be applied to a variety of integral membrane protein systems, regardless of oligomeric state, and should be particularly useful in calculating conformational changes from a known reference crystal structure.

## INTRODUCTION

The three-dimensional (3D) structure of proteins contains crucial information to understand their biological function. Despite their biological and pharmacological importance, relatively few high-resolution membrane protein structures are currently available, corresponding to $<1\%$ of the known 3D protein structures in the Protein Data Bank (1). This is due largely to well-known problems of protein expression, crystallization, and stability, which together hamper ongoing efforts toward high-resolution structure determination by x-ray crystallography or NMR spectroscopy. These are typically carried out in detergent micelles, which can alter their conformation to a variety of nonnative states (2–4). Furthermore, it is very difficult to determine membrane protein structures in more than one conformation, limiting efforts to establish the structural basis of functional mechanisms.

When conventional structure determinations are not possible, biophysical techniques such as solid-state NMR, electron paramagnetic resonance (EPR), and fluorescence spectroscopy can offer complementary structural information under physiological conditions and in a membrane environment (5–7). However, although these approaches provide a limited set of high resolution structural data, when combined with computational analysis they offer a viable alternative to the determination of detailed tertiary folds and conformational rearrangements of membrane proteins in the context of a lipid bilayer (8–11).

Site-directed spin-labeling (SDSL, Fig. 1) and EPR spectroscopy have shown to be versatile tools for the study of the structure and conformational dynamics in membrane protein (12,13–16). Typically, EPR-based structural information from spin-label proteins include i), direct estimation of nitroxide dynamics from line-shape analysis, ii), solvent accessibility of the labeled site through power saturation paramagnetic relaxation experiments, and iii), distances from dipolar coupling between nitroxide spins via spectral broadening (17–19). Analysis of these data can reveal secondary structural elements, membrane protein topology, and in some cases overall architecture (20–22). With some exceptions, however, the correlation between EPR information and protein structure relies heavily on qualitative interpretations of the data. This is due mostly to the intrinsic limitations of the technique (the effect of the spin-label ''tether'') but also is the result of a dearth of quantitative structural descriptions of spin-label proteins (23,24) and the lack of a theoretical framework to use this structural information in structural refinement.

In contrast to NMR methods, where multiple constraints per residue are typically available, a key challenge in the use of EPR structural data is the structural interpretation of significantly fewer constraints in a computationally efficient and

*Submitted July 23, 2008, and accepted for publication July 24, 2008.*

Address reprint requests to Benoît Roux, Dept. of Biochemistry and Molecular Biology and Institute for Biophysical Dynamics, University of Chicago, Chicago, IL. Tel.: 773-834-3557; E-mail: roux@uchicago.edu, or Eduardo Perozo, Dept. of Biochemistry and Molecular Biology and Institute for Biophysical Dynamics, University of Chicago, Chicago, IL. Tel.: 773-834-4747, E-mail: eperozo@uchicago.edu.

Pornthep Sompornpisut is on leave from the Dept. of Chemistry, Faculty of Science, Chulalongkorn University, Bangkok, Thailand.

Editor: Klaus Schulten.


0006-3495/08/12/5349/13 $2.00

doi: 10.1529/biophysj.108.142984

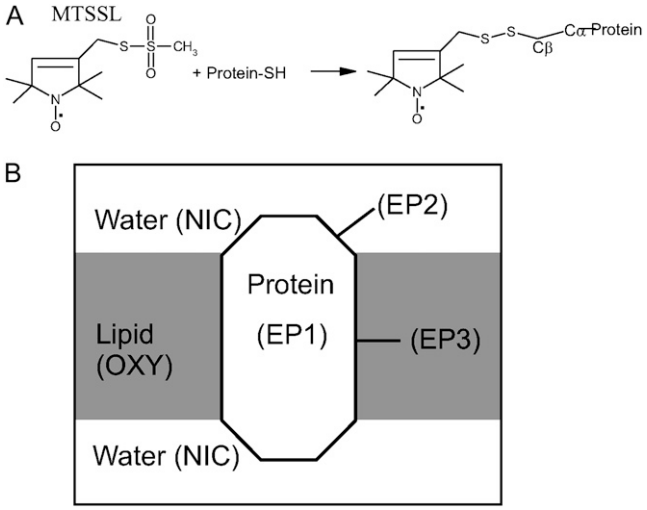


FIGURE 1 Overview of SDSL and the three classes of spin-label pseudoatoms according to environment. (*A*) Methathiosulfonate spin-label (MTSSL) structure, the nitroxide attached to cysteine residue (Protein-SH) via disulfide linkage. (*B*) Simple representation of the pseudospin (EP1, EP2, and EP3) exposure to different environments in a membrane protein.

meaningful way. A number of semiquantitative molecular modeling approaches have been proposed that utilize SDSL/EPR data as restraints to model low-to-moderate resolution structure or conformational changes of membrane proteins (25–27), with recent advances that allow model building in the 1–2.6 Å resolution range (28). There are two general computational approaches when dealing with a limited data set: restraint-based structure generation and model evaluation. Conformational search algorithms such as molecular dynamics (MD), Monte Carlo, or purely geometric methods have been used to generate representative sets of global fold structures, which were subsequently evaluated based on experimentally derived energy or scoring functions. A fundamental aim of these approaches is to produce an ensemble of final conformations that are both energetically favorable and in agreement with experimental restraints.

Previously, we demonstrated the use of SDSL/EPR structure parameters such as spin-spin distances and spin-label solvent accessibility as restraints in calculating conformational changes in ion channels (29,30). These were based on the use of EPR-based constraints to drive defined rigid-body secondary structural elements in a Cartesian space. Using this approach (restraint-driven Cartesian transformations), we estimated conformational changes at the intracellular gate of KcsA by driving rigid-body transmembrane (TM) segment conformation to fit restraints derived from interspin distance changes (29,31). The method was subsequently modified with the addition of a hard-sphere solvent-accessible surface area calculation in the penalty function and applied to the intramembrane conformational changes in the *Escherichia coli* mechanosensitive channel of large conductance (MscL) in the open state (30). Since this approach is purely geometric, energy-based structure refinements were performed separately in the final step of the computation.

Although it has been relatively easy to rank and score structural models based on EPR data, it is much more difficult to drive structure refinement using this information, in particular, ''soft'' constraints like solvent accessibility. Here, we introduce an approach to refine membrane protein structures dynamically, using parameterized constraints derived from EPR solvent accessibility data. The method, pseudoatom-driven solvent accessibility refinement (PaDSAR) translates molecular oxygen ($\Pi O_2$) and nickel-ethylenediaminediacetic acid ($\Pi$NiEdda) collision frequencies into computationally tractable structural restraints, imposed in the form of an empirical potential function. As a test case, we have used the *Streptomyces lividans* potassium channel (KcsA) because of the availability of a high-resolution 3D structure and the experimental abundance of EPR data (32,33). In our hands, the method was able to refold a number of distorted KcsA structures back to the native conformation within a 3 Å or better root mean-square deviation (RMSD) and with high computational efficiency. Ongoing applications of this approach include structural analyses of the *E. coli* mechanosensitive channel of small conductance (MscS) on its closed conformation (34), the overall architecture of the membrane-embedded voltage sensing domain in KvAP (35).

## METHODS

### Generation of starting conformations

We used the 2-Å-resolution x-ray structure of the KcsA potassium channel (accession code: 1K4C (33)) to calibrate and test the method. KcsA is a homotetrameric protein and each subunit is composed of two transmembrane helices encompassing residues 22–51 (TM1) and 86–124 (TM2) (Fig. 2 *A*).

To develop and validate the methodology used here, a set of randomized starting conformations (or ''decoys'') were generated by rigid-body movements of both TM segments. From the 1K4C crystal structure, the antibody FAB fragment and water molecules were removed; and to minimize the energetic contributions of the native side-chain conformations, these were rebuilt using SCWRL 3.0 (36)—a protein side-chain prediction program. All ionizable residues were assigned to their common charge, i.e., $Glu^{-1}$, $Asp^{-1}$, $Lys^{+1}$, and $Arg^{+1}$, except for E71 (37) and the truncated N-terminal S22 and the C-terminal H124. Protein topology and parameters were taken from the CHARMM 19 united-atom force field (38,39). Fourfold rotational symmetry was imposed using the IMAGE facility (38,39) to compute the final tetrameric assembly. Simulations in vacuum were performed by moving TM1 and TM2 helices from the reference KcsA crystal coordinates as rigid bodies as a way to generate different degrees of conformational deformation.

Residues 55–83 corresponding to the pore helix and the selectivity filter were kept fixed in their native conformation. The simulation protocol was repeated to obtain a wide range of randomly distorted conformations—after which energy minimization was subsequently performed. It should be noted that ~90% of the decoy structures were obtained using this procedure. Additional decoys with larger distortions were made by a manual rotation to obtain a specific distortion of conformations, for instance, those by rotating either TM1 or TM2 along its helical axis or by a large degree of TMs movement (~30 Å RMSD; see below). The generated structures were subsequently screened to eliminate those containing either steric clashes or bad covalent geometry. The final set of decoys consisted of distorted conformations or partially unfolded KcsA structures used for the next simulations of refolding with the solvent accessibility restraints.

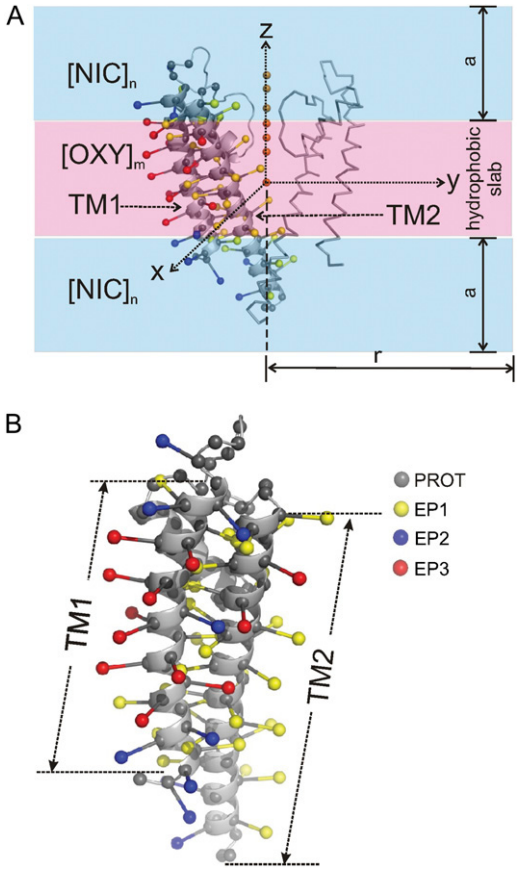


FIGURE 2 System used for MD simulations and pseudoatom representations. (*A*) The KcsA x-ray structure (1k4c), an illustration of two subunits and a definition of restraining boundary for molecular oxygens (OXY) and Ni-EDDA complexes (NIC) for the simulations. Potassium ions are shown as orange spheres (*B*) the attachment of spin-label pseudoatom to protein backbone. The type of spin-labeled pseudoatom is color coded as EP1, yellow, EP2, blue, and EP3, red.

## Solvent accessibility as structural restraint

We developed a simple form of a molecular mechanical potential for the EPR-derived restraints based on the typical accessibility environments found in a spin-label membrane protein (Fig. 1 *B*). The nitroxide spin-label (denoted EP1, EP2, and EP3) can exist in three different environments: buried within the protein itself (EP1), exposed to the aqueous environment (EP2), or exposed to the membrane lipids (EP3). Surrounding the protein and spin-label pseudoatoms, we have included contrasting particles representing molecular oxygen (OXY) and NiEdda, a Ni-chelated complex collisional agent. The spin-label (at EP1, EP2, and EP3), contrasting particles (OXY, NIC), and membrane protein C$\alpha$ particles (PROT) were virtually introduced in the system together with a complete atomistic representation of the protein during MD refinements.

## Pseudoatom descriptions and assignment based on solvent accessibility

Nitroxide spin-labels and contrast agents were present in the system as pseudoatoms called ''spin'' and ''environment'', respectively. Within the two categories, there are six different types of pseudoatoms used in the calculations. The first category is the pseudospin atoms including EP1, EP2, and EP3, classified according to the nitroxide moiety of the spin-label side chain. In the second category of pseudoatoms, PROT, OXY, and NIC represent amino acid residues $O_2$ and NiEdda, respectively, and serve as environmental probe particles with which the pseudospins interact. A description of the pseudoatom types is summarized in Table 1.

The approach requires the spin-label site to be assigned as EP1, EP2, or EP3 according to environmental collision frequency experimental data ($\Pi O_2$ and $\Pi$NiEdda). Thus, EP1 corresponds to sites with both low $\Pi O_2$ and $\Pi$NiEdda values. EP2 represents sites with high $\Pi$NiEdda but low $\Pi O_2$, whereas EP3 sites are defined by high $\Pi O_2$ and low $\Pi$NiEdda values. We assigned EP1 or EP2 or EP3 throughout KcsA TM1 and TM2 residues (Fig. 2 *A*) using available $\Pi O_2$ and $\Pi$NiEdda data from a total of 58 spin-label KcsA mutants (32). But for a few exceptions, all the assigned pseudospin types show excellent agreement with residue positions according to the x-ray structure (1K4C), suggesting that we have chosen a suitable model system for our study. Table 2 and the Supplementary Material (Data S1) summarize the environmental assignment of the 58 spin-label residues.

## Restraint energy

An empirical method to compute EPR restraint energy was employed in the form of a molecular mechanics (MM) potential function, which considered bond length, bond angle, improper torsion, and Lennard-Jones (LJ) potentials:

$$V_{(r)} = \varepsilon[(R_{\min}/r)^{12} - 2(R_{\min}/r)^{6}]$$

in the force field parameters (though not the Coulombic term).

In the PaDSAR approach, nonbonding LJ interactions between the spin and environment particles represent the main driving force of the restraint. A total of nine possible interacting pairs, three spin-label classes versus three environment types (shown in Table 3), were introduced in the MD simulations for structure refinement. They were considered to be matching (EP1-PROT, EP2-NIC, and EP3-OXY) and mismatching (EP1-OXY, EP1-NIC, EP2-OXY, EP2-PROT, EP3-NIC, and EP3-PROT) pairs according to the pseudoatom descriptions defined above. Next, we applied a van der Waals ''repulsion'' to account for interactions of those virtual particles (Table 3) of which the nonbonding energy was described by three sets of LJ parameters, $\varepsilon$, and $R_{\min}$ (Fig. 3). The LJ types I and II were used for the matching pairs, and LJ type III applies to the mismatch between particle types, for instance, buried EP1-OXY or buried EP1-NIC pairs etc. (Table 4). This strategy provides a powerful restraint driving force so that a given spin-label (EP1 or EP2 or EP3) is repelled by the observed environmental mismatch unless it resides in the correct environment with a proper distance defined by their LJ potential function.

In particular, the LJ type II potential describing the EP1-PROT interaction is modified using a switching function (Fig. 3), given as

$$E_{\text{TYPE II}} = \begin{cases} E_{\text{LJ}} & \text{if } R > R_{\min}, \\ E_{\min} & \text{if } R \le R_{\min}. \end{cases} \quad (1)$$

We applied this function to the LJ type II potential and implemented it in CHARMM version c32a2, modified to allow the buried EP1 to overlap with

**TABLE 1 Pseudoatom descriptions**

| Name | Type | Descriptions |
|---|---|---|
| EP1* | Spin-label | Buried |
| EP2* | Spin-label | Water exposed |
| EP3* | Spin-label | Lipid exposed |
| PROT† | Environment | Amino acid residue |
| OXY | Environment | Molecular oxygen ($O_2$) |
| NIC | Environment | NiEDDA complex |

*Nitroxide pseudoatom is covalently attached to the PROT of the labeled residue.
†Residing at C$\alpha$ and applying to all residues.

**TABLE 2 KcsA residues with the assignments of the pseudospin types**

| Residue number | Pseudospin |
|---|---|
| 25, 26, 29, 30, 33, 36, 40, 44, 47, 48, 86, 87, 89, 91, 92, 95, 96, 97, 98, 99, 100, 101, 102, 103, 105, 106, 107, 108, 109, 110, 111, 113, 114, 115, 118, 119 | EP1 |
| 23, 24, 27, 28, 39, 49, 50, 52, 117, 120 | EP2 |
| 31, 32, 34, 35, 37, 38, 41, 42, 45, 46, 90, 94 | EP3 |

PROT atoms without having an effect on the LJ energy. It should be noted that pseudoatoms see neither the protein atoms nor the pseudoatom of the same category. For instance, an EP1 does not see other EP1, protein, and EP2 and EP3 atoms, but interacts only with PROT, OXY, and NIC pseudoatoms. In other words, there are no interactions of the following pairs: protein-pseudoatoms, spin-spin, or environment-environment pseudoatoms. The LJ energy of the restraints is derived from the spin-environment interactions only.

LJ parameters for the spin-environment interacting pairs were calibrated and optimized from MD simulations of the KcsA x-ray structure; 4 ns MD simulations were performed on the x-ray structure to test the effect of various LJ parameters of the EPR restraints on any structural deviations from the native conformation and contrasted with an MD run without the restraints as a control. The bonding and nonbonding parameters of the restraints were incorporated into the united-atom CHARMM19 force field parameters (39). The force field parameters of the restraints are shown in Table 4.

## Pseudoatom patch and system construction

Each system consisted of a decoy protein onto which pseudoatoms were generated for the MD simulations with the EPR restraints. We used the CHARMM patch to construct a starting configuration of the system with the virtual particles by attaching the PROT, EP1, EP2, and EP3 on the backbone of the decoy structure and introduce OXY and NIC pseudoatoms as lipid and aqueous environments, respectively. The pseudospins (EP1, EP2, or EP3) were covalently attached to the C$\alpha$ position (CA) of the assigned residue (Fig. 2 *B*). Their structural parameters including pseudobond length, pseudobond angle, and improper torsion angle were defined to allow a movement of CA-EP1 (or EP2 or EP3) reflecting the dynamics of a real spin-label (40,41). Our testing simulations suggested that a fairly stiff CA-EP bond (Table 4) was more compatible to native structures than a flexible bond. For the environmental contrast particles, PROT atoms were placed at the CA of each amino acid throughout the structure, whereas OXY and NIC coordinates were precalculated from an MD simulation. The OXY and NIC particles were subjected to distribution in the inner and outer membrane compartments, respectively. This was achieved through the miscellaneous mean field potential (MMFP) in CHARMM with restraint distance boundary (Fig. 4).

## Structure refinement

Each starting configuration consisted of a decoy monomer attached to 36 EP1, 10 EP2, 12 EP3, and 103 PROT atoms, plus 300 OXY and 300 NIC

**TABLE 3 Repulsion of pseudospin with the environment pseudoatoms**

| | Repulsion | | |
|---|---|---|---|
| Nitroxide spin pseudoatom | PROT (protein) | NIC (water) | OXY (lipid) |
| EP1 (buried) | No | Yes | Yes |
| EP2 (water) | Yes | No | Yes |
| EP3 (lipid) | Yes | Yes | No |

**TABLE 4 Force field parameters for the pseudoatoms**

| Bonds | $k_b$ (kcal mol$^{-1}$ Å$^{-2}$) | $l_0$ (Å) |
|---|---|---|
| CA-EP1 (EP2, EP3) | 100.0 | 6.00 |
| Angles | $k_\theta$ (kcal mol$^{-1}$ rad$^{-2}$) | $\theta_0$(degree) |
| N-CA-EP1 (EP2, EP3) | 50.0 | 120.0 |
| Improper torsions | $k_\omega$ (kcal mol$^{-1}$ rad$^{-2}$) | $\omega_0$(degree) |
| CA-X-X-EP1 (EP2, EP3) | 55.0 | 0.0 |
| van der Waals | $\varepsilon$ (kcal mol$^{-1}$) | $R_{min}$(Å) |
| EP2-NIC, EP3-OXY | 2.00 (attraction type I)* | 2.00 |
| EP1-PROT | 2.00 (attraction type II)* | 7.00† |
| EP1-OXY, EP1-NIC, EP2-OXY, EP2-PROT, EP3-NIC, EP3-PROT | 0.05 (repulsion type III)* | 6.00 |

*See Fig. 3 for the potential curve.
†Means overlap.

atoms. A fourfold rotational symmetry using the IMAGE facility was imposed throughout the simulations to provide a model of tetrameric channel. In all simulations, the pore helix and the selectivity filter residues were fixed due to the lack of experimental data coming from that region. Key to this decision is that the selectivity filter plays a critical role in K channel assembly and stability and that this region is not expected to undergo major structure changes (at least at the level of resolution afforded by the method) without affecting channel folding and stability. Furthermore, the backbone atoms in TM1 and TM2 residues were constrained by their secondary structure helices; so we have limited all calculations to $\alpha$-helical regions of the channel.

An MMFP with a force constant of 1.0 was employed for OXY particles to partition in a hydrophobic slab of 24 Å thickness and for NIC particles to distribute in the aqueous phases on each side of the membrane (Fig. 2 *A*). NIC atoms move in the aqueous bath within 100 Å of the membrane interface. Additionally, both NIC and OXY were distributed in the cylinder-shape system with a radius $r$ of 50 Å. The simulations were performed without periodic boundary conditions. The system was first relaxed by energy minimization to remove unfavorable steric contacts, and the pseudoatoms were free to move, whereas the positions of the decoy atoms were kept fixed. All positional restraints, except for those in residues 55–83, were removed during the MD simulations.

PaDSAR simulations were carried out using Langevin dynamics with a friction coefficient of 5 ps$^{-1}$ for all the atoms, except hydrogens. MD simulations were performed in a range of 300–600 K using the united-atom model PARAM19, and the dielectric constant was set to 1. The initial velocity of particles was assigned using the Gaussian distribution with a time step of 2 fs. The group-based method for the nonbond list was used with a

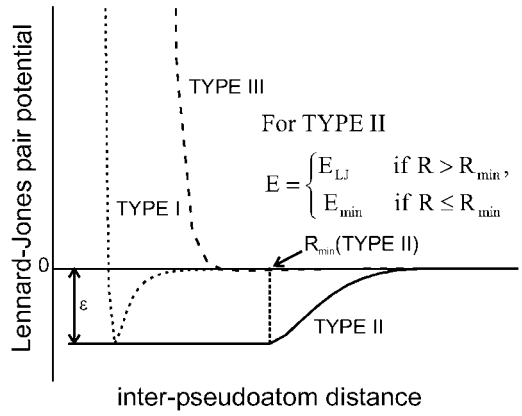


FIGURE 3 LJ potential curves for the different pseudoatoms. Energy profiles for the three types of interactions are given as Type I, dotted line, Type II, solid line, and Type III, dashed line.

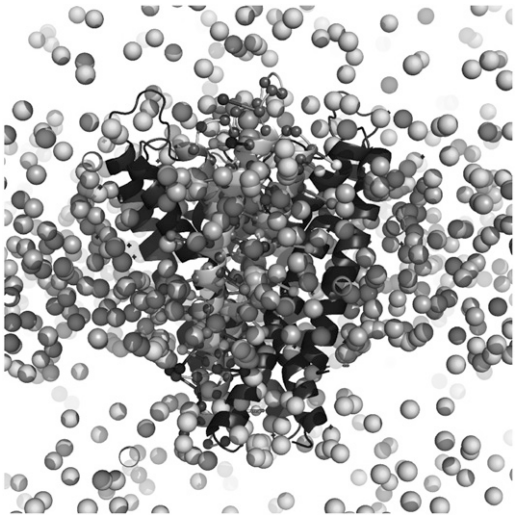

FIGURE 4 Pseudoatom model system of the KcsA tetrameric assembly and contrast agents used in the simulations. The protein is shown in a cartoon representation, pseudoatoms EP1, EP2, and EP3 as ball and stick, and pseudoatoms NIC and OXY as van der Waals spheres.

distance cutoff of 16 Å of actual and image atom lists. The switching function of LJ and electrostatic potentials was applied for an interacting pair distance between 8 Å and 12 Å. We carried out 2 ns MD simulations for the decoy refinement and 4 ns simulations for testing the LJ parameters. The restraint energy and structure fluctuations were used to verify a convergence of the simulations, and we found that all the simulations converged within 1 ns MD runs.

## Analysis and evaluation

We used energy and RMSD relative to the reference structure (1K4C) as the two selection criteria for selecting final structures. For each MD run with PaDSAR, structures of 250–500 snapshots were taken from the trajectory during the last 0.5–1 ns for analysis and evaluation. Unless otherwise specified, we used ''RMSD to native'' or ''RMSD distortion'' to denote a comparison between the obtained structure(s) and the x-ray KcsA structure. $\langle\ \rangle$ indicates the trajectory-average quantity. For instance, $\langle$RMSD$\rangle$ is the trajectory-average RMSD. Each snapshot was used to compute the LJ interaction energy associated with EP1, EP2, and EP3 to identify structures with the correct fold. We found no significant difference between a representative structure selected from the K-mean cluster centroid and that chosen from the lowest interaction energy. Therefore, a representative PaDSAR structure of each run was selected and ranked based on the computed energy. The top 25 structures were chosen, and we measured their structural differences (RMSD). An in-house FORTRAN program for hierarchical clustering algorithm (42) was used to measure dissimilarities among selected structures.

# RESULTS

As defined above, we have devised a strategy to parametrize EPR-based solvent accessibility data to drive the conformation of a preexisting membrane protein crystal structure. We tested and validated our approach by driving structurally distorted decoy structures back to its native conformation based on EPR environmental data (32) from the closed conformation of the potassium channel KcsA (33). A set of 100 decoys with varying degrees of distortion (between 2 and 30 Å RMSD) from the native closed conformation were generated for statistical analysis. In each case, MD simulations with and without EPR restraints were performed to evaluate the efficiency of the approach. The resulting structures from the simulations were validated based on a comparison of RMSD to the native structure.

## Optimizing restraint parameters

Based on $\Pi O_2$ and $\Pi$NiEdda data (32), a total of 58 pseudospins were used in the calculations with the assignment of 36 EP1, 10 EP2, and 12 EP3 (Table 2) along TM1 and TM2 in the transmembrane regions KcsA in its closed conformation. Several MD runs of the x-ray KcsA structure were performed by varying LJ parameters, $\varepsilon$ and $R_{\min}$, of the pseudoatom interacting pairs. Because most of the experimental spin-label accessibility data were consistent with the x-ray KcsA structure, these restraints should satisfy or at least not severely violate the current conformation. Thus, RMSD is an ideal parameter for an initial validation of the results of each run to obtain an optimal well depth for the $\varepsilon$ and $R_{\min}$ parameter set selected from the run giving the structure, on average, closest to the reference structure. Here, we have chosen the parameters from the run that gave the best fitting $\langle$RMSD$\rangle$, relative to the native KcsA conformation, of $1.18 \pm 0.07$ Å. The run that did not include the EPR restraints gave an ensemble of conformations with an RMSD distortion of $0.65 \pm 0.10$ Å. These changes of 0.53 Å RMSD are insignificant compared with the dynamics of the system.

Plots of the radial distribution function, $g(r)$, for the different interacting pairs (Fig. 5) were used to further refine the quality of the pair potential LJ parameters for all pseudoatoms. It is expected that the match pair particles defined with an attractive potential (EP1-PROT, EP2-NIC, and EP3-OXY) should get close to each other with respect to those mismatch pairs. A sharp peak at 6 Å for $g(r)$ of EP1-PROT, EP2-PROT, and EP3-PROT corresponds to the distance from the pseudospin to its own residue, since PROT is located precisely at the C$\alpha$ position (CA-EP1, EP2, or EP3 bond length 6 Å). PROT particles can get much closer to EP1 ($r <$ 2 Å) than can EP2 and EP3 (the closest distances are $\sim$4 Å and $\sim$5 Å, respectively), and EP1 has a greater number of surrounding PROT particles, as indicated by the integration number (Fig. 5, *red line*). This LJ type II function allows EP1 to favor the buried spin type. Furthermore, the simulation showed the first strong peak at 2.0 Å for EP3-OXY and for EP2-NIC, suggesting that OXY particles provide favorable interactions with the lipid-exposure spin EP3 but not with EP1 and EP2. Finally, NIC particles get sufficiently close to give attractive interactions with the water-exposed spin EP2. This is a clear indication that the current set of the LJ functions and parameters adequately describes the pseudoatom features designed for this study. The force field parameters of pseudoatoms are summarized in Table 4. A series of trial parameter sets is provided as Supplementary Material (Data S1).

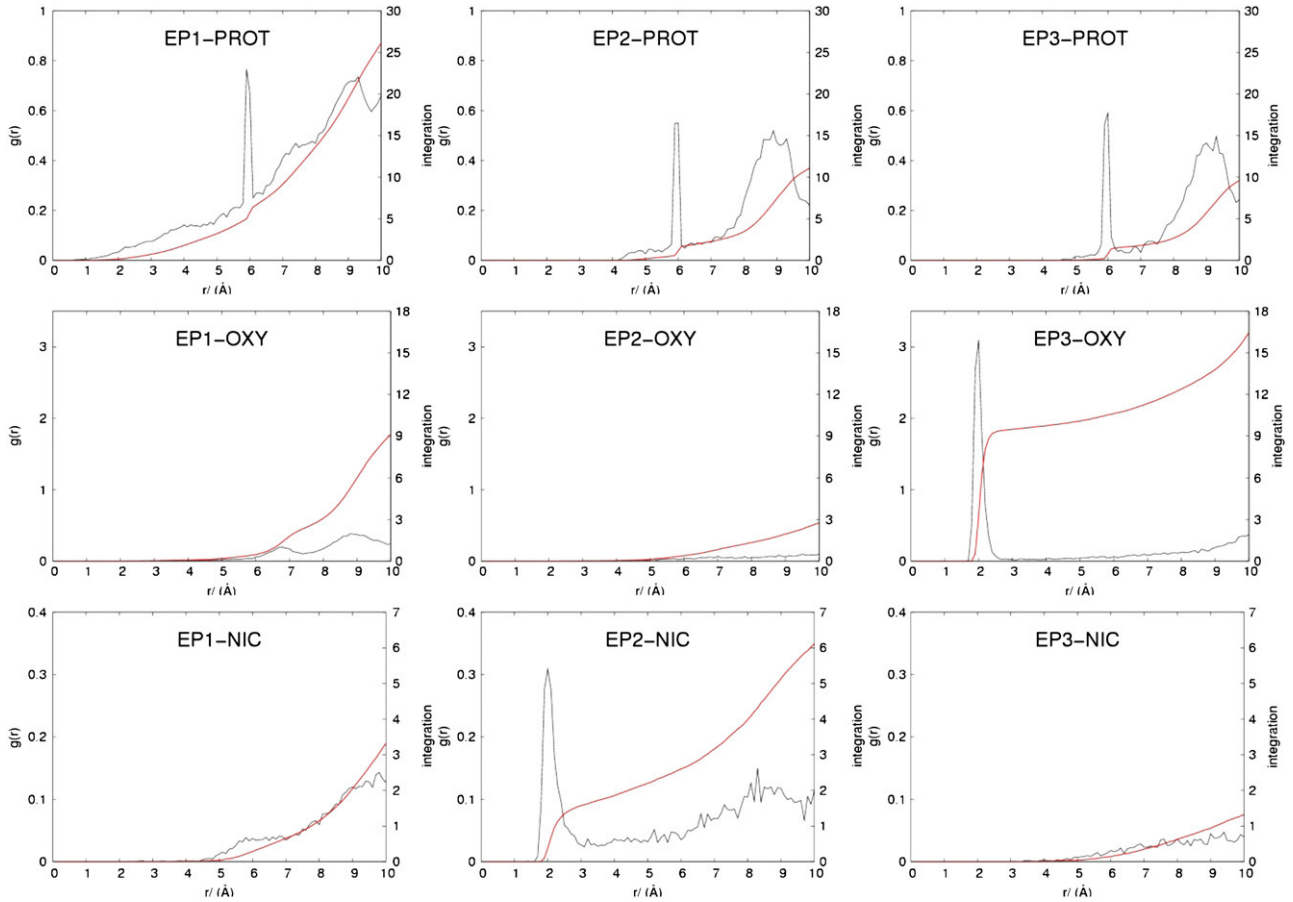


FIGURE 5 Radial distribution function plots, $g(r)$, and its integration (*red line*) of pseudoatom interacting pairs taken from the MD simulation of the SCWRL model.

A major concern during the development of our method here was that the well-defined atomic positions of the x-ray structure may unduly bias the contributions of the EPR restraints in the refolding simulation (43). This influence might have a considerable impact on the selection of conformations close to the native structure. Therefore, we constructed a starting model having the x-ray backbone structure with rebuilt side-chain conformations as the basic template for decoy generation (see Methods). Simulations of the rebuilt side-chain model were performed to compare with that of the x-ray (1K4C), using SCWRL (36) and the Swiss-PdbViewer (44) to predict side-chain conformations; 4 ns MD runs were conducted starting with side-chain-rebuilt models, SCWRL, SWISS, and poly-Alanine and then compared with those from the x-ray model. It should be noted that these four models (1K4C, SCWRL, SWISS, and poly-Alanine) share the same context of the backbone x-ray conformation but differ in their side-chain torsions. In addition, the simulations were performed in the presence and absence of the EPR restraints.

MD results ranked according to ⟨RMSD⟩ (to native) were as follows: 1K4C < SCWRL < SWISS < Alanine (Fig. 6). The order appears to be the same for the simulations in the presence and absence of the restraints. From these results, ⟨RMSD⟩ ranges from 0.65 ± 0.10 Å to 2.19 ± 0.11 Å for the simulations with the absence of the EPR restraints and from 1.18 ± 0.07 Å to 3.70 ± 0.08 Å for the simulations with the EPR restraints (Table 5). As expected, the runs that include the EPR restraints had larger ⟨RMSD⟩to native values compared to those without the restraints. However, the increase in ⟨RMSD⟩ distortion is not dramatic, except for the poly-Alanine model (from 2.19 Å without restraints to 3.70 Å with restraints). These results indicate the importance of side-chain contributions and packing to maintain the structure close to the reference native conformation. The x-ray-based MD runs appear to have a minimal effect on conformation distortion because the side chains have been well defined. Consequently, we decided to use the SCWRL model as a base for generating all the decoy conformations to minimize the strong bias that side-chain packing may have in all refolding simulations.

## Structure refinement

In the simulations here, each set of decoys contains 100 sampling conformations covering a broad range of randomly distorted structures, from 2 Å to 30 Å RMSD relative to the

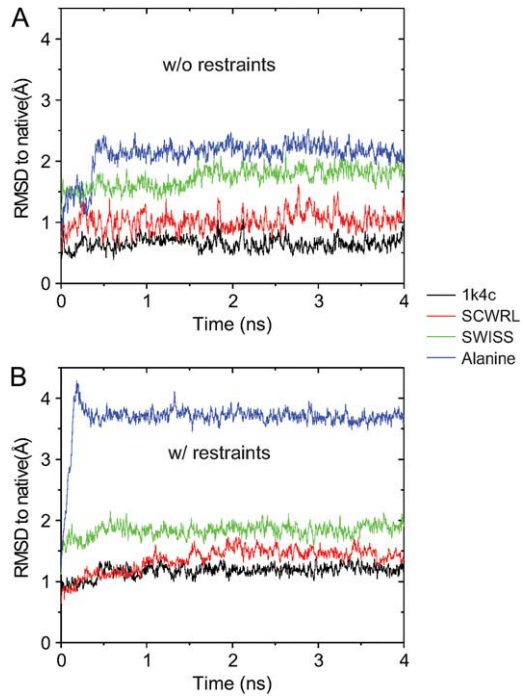


FIGURE 6 RMSD fluctuations from the MD simulations of the backbone-based x-ray structures: 1k4c (*black*), SCWRL (*red*), SWISS-PdbViewer (*green*), and Alanine (*blue*), (*A*) without and (*B*) with the restraints.

native KcsA structure. In our sets, there is an ~20:20:50 relation for structures with RMSD ranges of 2–5 Å, 5–10 Å, and 10–20 Å, respectively. The remaining decoys, with very large RMSDs (~30 Å) were manually generated from the native structure. An analysis of the resulting structural deviations after PaDSAR runs was performed on snapshot structures obtained from all 100 simulations. Fig. 7 shows the RMSD to native before (*solid squares*) and after (*gray bars*) PaDSAR simulations. From the RMSD plot, the majority of the refolding MD simulations successfully drove the starting distorted decoy toward the native conformation (Fig. 7 *A*). Indeed, most of the obtained PaDSAR structures converged to RMSD in a range of 1–4 Å, even those that have starting conformations with RMSD as large as 30 Å. Although convergence proceeds at different rates, depending on initial decoy RMSD, most convergences occurred within the first 1 ns simulation (Fig. 7 *C*).

**TABLE 5 ⟨RMSD⟩ results of the simulations**

| | ⟨RMSD⟩ to the native (Å) | |
|---|---|---|
| Model | With EPR restraints | Without EPR restraints |
| 1K4C | 1.18 ± 0.07 | 0.65 ± 0.10 |
| SCWRL | 1.48 ± 0.09 | 1.06 ± 0.15 |
| SWISS | 1.87 ± 0.09 | 1.82 ± 0.11 |
| Alanine | 3.70 ± 0.08 | 2.19 ± 0.11 |

Data with and without the restraints for the x-ray-based backbone models.

## Structure selection criteria

After considering RMSD to native <3 Å as an acceptable cutoff for the identification of the correct folds, 54 out of the 100 MD runs produced PaDSAR structures that were in good agreement with the native KcsA structure. For the rest, 30 and 16 of the PaDSAR structures exhibited RMSD to native within 3–5 Å and larger than 5 Å, respectively. Among those PaDSAR structures with near-native KcsA conformation (RMSD < 3 Å), 80% had RMSDs to native conformations < 2 Å. In the case of unsuccessful refolding runs, the profiles of RMSDs to native conformations behave in a similar fashion as shown in Fig. 7 *C*. The RMSDs decrease at the beginning of the refolding simulation and then remain unchanged at the value larger than the cutoff (not shown). This suggested that some decoys were trapped by a local minimum of the potential energy surface and were unable to move to a configuration in which the restraints were fulfilled.

As a control, refolding simulations in the absence of the EPR restraints were performed. For a comparison, 50 decoys that adopt starting conformations with RMSDs to native conformations in a range of 2–17 Å were selected from the same decoy set used in the simulations with the restraints. It is clear that simulations including the restraints were on average more successful (Fig. 8 *A*) than those driven by the MD force field alone. The advantage of using the restraints is specially highlighted when the RMSD distortion of the decoys becomes larger (8–17 Å). In this range, MD runs without the restraints failed in the majority of cases (Fig. 8 *B*).

## Restraint energy as the primary selection

As shown in Fig. 7 *A*, the 100 refolding simulations produced a range of PaDSAR structures covering native and distorted fold representations. The individual configurations of the systems in the MD trajectories can be used to extract the energy from the EPR restraints associated with the PaDSAR-generated structures. The idea is to use energy information from the restraint as a criterion in the final structure selection. The LJ interaction energy in particular can be used as a structural indicator, as it is the most relevant parameter to determine the best fit between the calculation and the experiment data. It is, however, useful to quantify the correlation between restraint energy and correct fold structure.

This issue was evaluated by looking at the correlation between the LJ restraint energy components computed from MD snapshots and RMSDs to native from the resulting structures. Each snapshot of the system was taken from the trajectory and used to compute $E_{LJ}$(EP1), $E_{LJ}$(EP2), and $E_{LJ}$(EP3), the LJ energy associated with pseudoatom interactions. $E_{LJ}$(EP1) reflects interactions of EP1 with all the environment pseudoatoms (PROT, OXY, and NIC). The same analogy was applied to $E_{LJ}$(EP2) and $E_{LJ}$(EP3) to define interactions involving EP2 and EP3, respectively. Furthermore, $E_{LJ}$(total) was computed from the sum of each

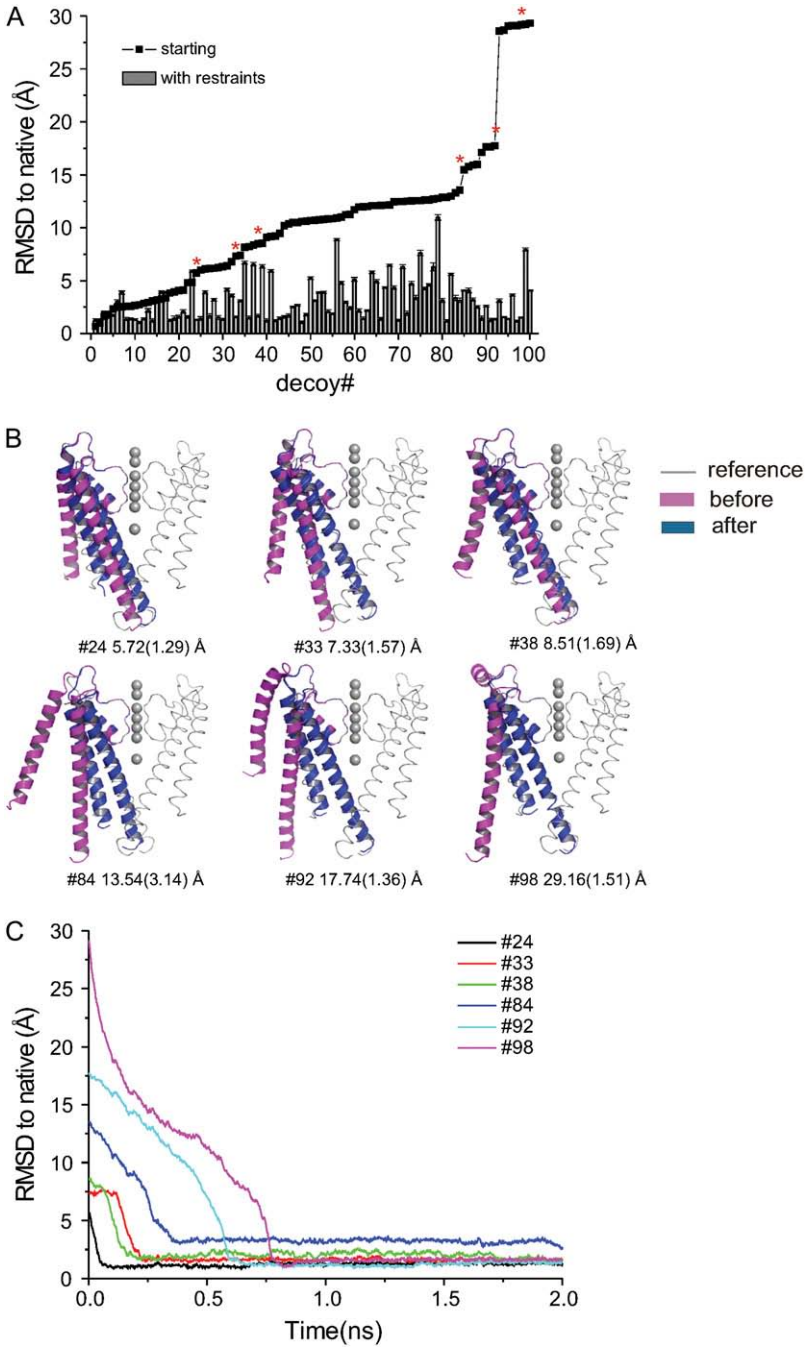


FIGURE 7 (*A*) ⟨RMSD⟩to native before (*solid squares*) and after (*bar charts*) the refolding simulation of 100 decoys. Red stars mark on decoys that their structures and RMSD profiles are shown in *B* and *C*, respectively. In *B*, red and blue ribbons are structures of decoy before and after the simulations, respectively. Decoy number and RMSD are also indicated. The C$\alpha$-trace with light gray is the reference. Potassium ions are shown as CPK (Corey, Pauling, and Koltun atom coloring) mode.

energy component. Energy and RMSD computations were carried out throughout the trajectory to obtain an average value. Fig. 9 *A* shows the $E_{LJ}$(total)-RMSD plot over the 100 runs. From this analysis, we found that both $E_{LJ}$(total) and $E_{LJ}$(EP1) give the best correlation with RMSD with respect to $E_{LJ}$(EP2) and $E_{LJ}$(EP3) (not shown). The $E_{LJ}$(total)-RMSD correlation is essentially similar in pattern to the $E_{LJ}$(EP1)-RMSD correlation, suggesting EP1 interactions provide the major contribution in the studied system.

This is largely a result of the majority of EPR restraints being assigned to the buried EP1, in addition to the defined LJ type II potential function allowing its dense set of interactions with PROT atoms. This can be demonstrated from the $E_{LJ}$(EP1-PROT)-RMSD correlation. Fig. 9 *B* shows that the $E_{LJ}$(EP1-PROT)-RMSD plot is similar to the $E_{LJ}$(total)-RMSD but with a better correlation. Structures obtained from all the PaDSAR simulations were ranked from the lowest (most favorable) to the highest $E_{LJ}$(EP1-PROT) values, and the top 25 runs were selected. Fig. 10 shows a plot of RMSD to native computed from the structures before ($RMSD_{start}$) and after ($RMSD_{refine}$) the refolding simulation. As an illustration, the top 25 runs had the starting decoy conformations with RMSD

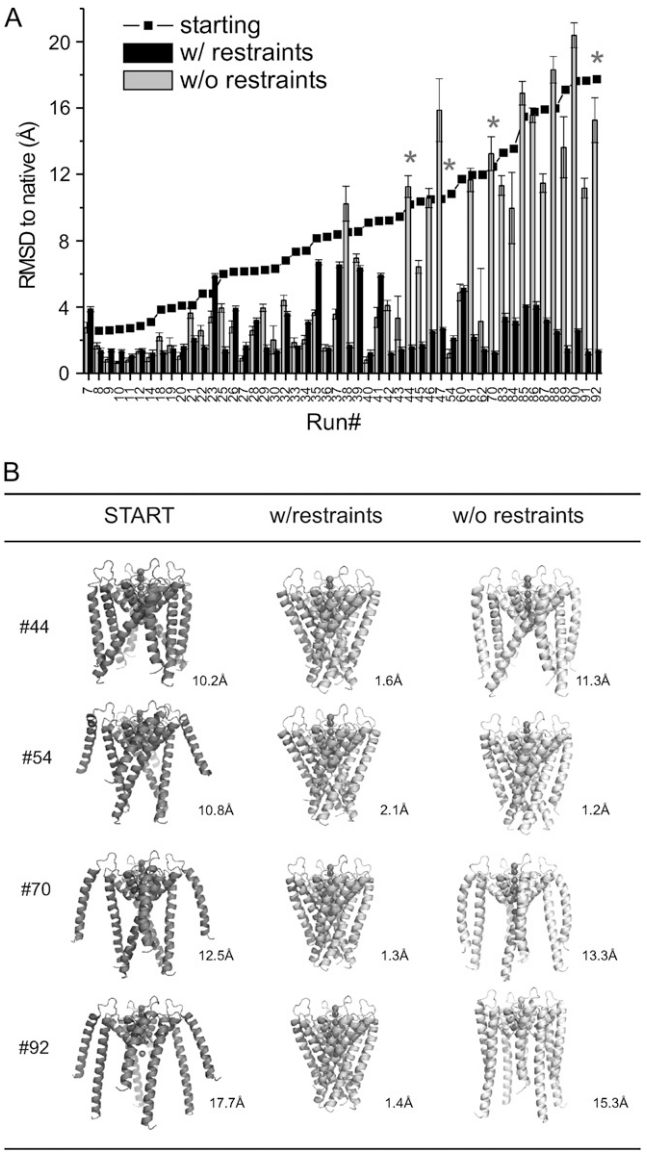


FIGURE 8 ⟨RMSD⟩ to native from MD simulations with and without the restraints. RMSDs of the starting decoy structures are also shown as solid squares. (*B*) Ribbons representation of some selected decoys (marked with stars in *A*, *dark gray*) shown after PaDSAR runs with EPR constraints (in *gray*) and without constraints (*light gray*).

in a range of 2–30 Å and produced PaDSAR structures with RMSD deceased to 1–4.5 Å after structure refinement. Applying a 3 Å cutoff for the RMSD distortion to the current set of PaDSAR structures showed that 20 out of 25 had the native-like conformation. The five remaining structures exhibit RMSD from 3 to 4.5 Å. The $E_{LJ}$(EP1-PROT) of these 25 selected PaDSAR structures is comparable to that calculated from the simulation of the x-ray structure. Thus, we concluded that correlation between the restraint energy and structural quality provides additional useful information for the initial screening of candidate structures from a large number of trial structures.

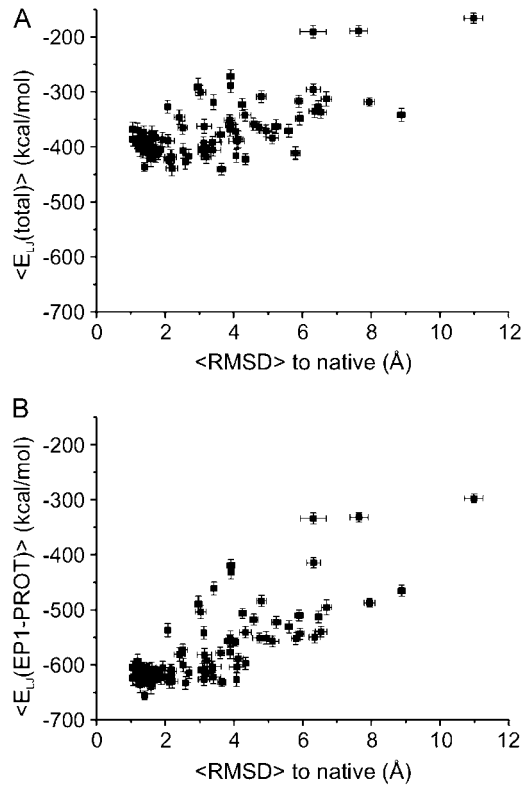


FIGURE 9 Structure discrimination based on the correlation between the LJ restraint energy components and RMSD to the native of the resulting structure. (*A*) ⟨$E_{LJ}$(total)⟩ versus ⟨RMSD⟩ and (*B*) ⟨$E_{LJ}$(EP1-PROT)⟩ versus ⟨RMSD⟩.

## Structure similarity as the secondary selection criterion

Because the above energy components are computed empirically, the analysis of the low energy alone may not be enough to produce robust selection criteria. A number of

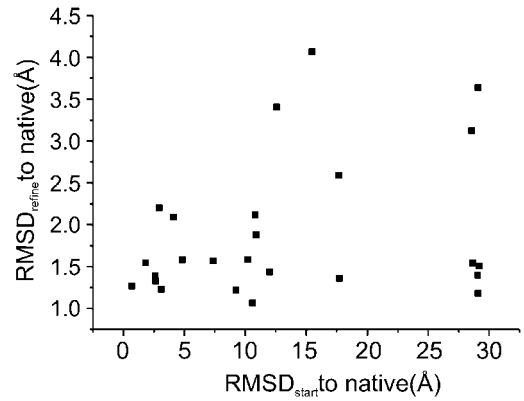


FIGURE 10 Correlation between the restraint energy and structural quality provides additional useful information for the initial screening of candidate structures from a large number of trial structures. RMSD to native computed from the structures before (RMSDstart) and after (RMSDrefine) the refolding simulation. A 3 Å cutoff for the RMSD distortion to the current set of PaDSAR structures shows that 20 out of 25 had the native-like conformation.

clustering algorithms designed to discriminate among a set of fold models have been successfully used in combination with the energy information for protein structure prediction (45,46). We further developed our discrimination criteria by computing pairwise RMSD within the 25 PaDSAR structures, giving rise to a 25 × 25 matrix of RMSD values. The RMSD matrix was then clustered based on a dissimilarity measure using hierarchical agglomerative clustering algorithm (42). In other words, we employed this method to rearrange RMSD values in the matrix from the lowest (*lower left corner*) to the highest (*upper right corner*) values as shown in Fig. 11 *A*. The smaller the RMSD, the lower the difference between the two compared structures. In the RMSD matrix, 25 PaDSAR structures were grouped into three clusters based on the RMSD range (Fig. 11 *A*). Eighteen out of the 25 PaDSAR structures belong to the cluster with the smallest RMSD range <2 Å (Fig. 11 *B*).

## DISCUSSION

Analyses of structure-function relationships, particularly in the case of ion channels and membrane transport proteins, require the determinations and analyses of multiple conformational states, which are difficult to obtain with conventional structural methods. The use of solution NMR with isotope-enriched samples has been successful in a number of cases, though further development is still needed (47–49). Here, we have developed a restrained MD protocol, PaDSAR, to take advantage of probe/reporter group data based on EPR power saturation experiments to drive 3D structural conformation in proteins. An identical treatment is expected for similar types of data obtained through complementary techniques (50) and is complementary to very successful recent protocols to calculate folds based on sparse spin-labeling EPR distance data (28).

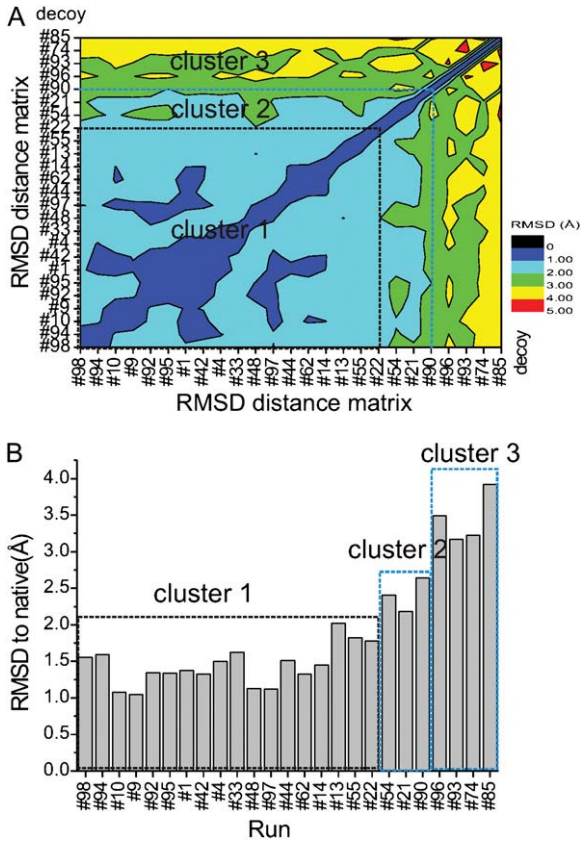


FIGURE 11 (*A*) Pairwise RMSD contour plot of the 25 selected PaDSAR structures rearranged from the lowest (*lower left corner*) to the highest (*upper right corner*) RMSD values. A color legend bar indicates RMSD values. (*B*) RMSD to native versus the 25 PaDSAR structures. The 18 most similar structures are in the dashed square.

One advantage of the approach here is that the EPR accessibility data for each spin-label position holds local structure information that suggests a specific arrangement of the residue backbone with respect to its global fold and the solvent environments. The results demonstrate that novel restraints derived from a pseudoatom-based potential method associated with the experimental accessibility data and used as a ''soft restraint'' represent an effective approach for structural refinement or conformational modification in membrane proteins. The use of a single effective pseudoatom per residue to represent the relatively complex nitroxyde moiety is justified by the low-resolution character of the experimental EPR accessibility data. There is no loss of information, and the combined effect of multiple pseudoatom sites in the restrained MD allows the simulations to refold distorted structures back to the native conformation with ease and computational efficiency.

We have shown that an empirical method where an MD conformational search was guided by the protein force field and the restraints was efficient in computing structural assembly of transmembrane helical proteins. The main driving force of the EPR restraints incorporated in MD simulation is described by nonbonding LJ interactions. In this regard, the definition for the nonbonding term is key to maximize the importance of the restraints. In the bonding term, we used a fairly stiff bond between the pseudospin and the C$\alpha$ backbone. Although this assumption may be only an approximation due to the conformational degree of freedom of the nitroxide side chain (40,41,51), the simulations with fairly rigid CA-EP1 (EP2 or EP3) bonds gave better results than those having flexible bonds.

The approach here is most useful in three key applications for membrane protein structural studies. First, it could be used to refine x-ray structures away from the conformation determined in detergent micelles or in a crystalline lattice to physiologically relevant states in a lipid bilayer. Second, the approach is useful in the case of refining 3D structures with low-resolution or with missing regions of the structure. Third, the method should perform equally well calculating conformational changes for systems for which state-dependent data sets exist.

There are some limitations to the approach and its implementation. One important problem is that the assignment of the spin type was done by relative comparison of acces-

sibility data where ambiguous values were not considered in the calculations. Moreover, an incorrect pseudospin type could be easily assigned due to ambiguous values of solvent accessibility in some positions that cannot be identified as belonging to a specific environment. Unfortunately, at the moment there is no automated method to assign the pseudospin type. As a result, structure refinement needs to be repeated from time to time by changing the type of ambiguous assigned spins until an agreement between the experimental EPR data and the obtained structure is optimized. Other issues to consider include the fact that MD runs with a large number of decoy conformations are time consuming, whereas single MD runs can limit the conformational space to be sampled. Structure refinement was carried out under an assumption of known secondary structure elements. This makes the method less useful when refining the structure of an unstructured flexible region or when calculating structural rearrangements of TM segments associated with a change in secondary structure.

This approach can be further applied to other biophysical data that are derived based on accessibility experiments, for instance, cysteine or tryptophan accessibility scanning analysis. In calculating conformational changes, the method could be further developed to take into account relevant accessibility changes by applying different weights in the LJ parameters. Furthermore, conformational searches combined with scoring methods could be introduced before the MD run for a rapid screening of candidate structures. A recent application of the method has introduced an additional pseudoatom EP4 for the nitroxide spin at the water-lipid interface, as indicated by DOGS-NTA[Ni(II)] (the 1,2-dioleoyl-*sn*-glycero-3-[(*N*-(5-amino-1-carboxypentyl)imidodiacetic acid) succinyl] nickel(II)) (34). A strategy to improve the approach described above is in progress with an effort to predict the structural architecture of the nontransmembrane region of ion channels. These include the closed conformation and N-terminus of a mechanosensitive channel with small conductance (MscS) (34) and the voltage-sensing domain of the potassium KvAP channel in membrane (35).

## CONCLUSION

We report the development of PaDSAR, a novel (to our knowledge) approach to modify the conformation of membrane proteins of known structure, using EPR-determined accessibility data as restraints in MD calculations. The novel restraint is based on different exposure of a spin-label to which $O_2$ or NiEdda molecules collide at a given frequency. We employed a pseudoatom method with empirical force field potentials to treat such collisions as a structural parameter. This study illustrated that structural refinement with the restraints combined with an energy-based screening and clustering method provides a promising tool in refolding a wide range of distorted conformations as much as 30 Å RMSD into the native-like KcsA conformation. Applications of PaDSAR to a number of membrane protein systems are underway in a variety of membrane protein systems.

## APPENDIX

Pseudoatom parameter and topology in CHARMM.
Parameter file PARAM19

```
BOND
PRO     CH1E    100.0    0.000
PRO     CH2E    100.0    0.000

EP1     CH1E    100.0    6.000
EP1     CH2E    100.0    6.000
EP1     PRO     100.0    6.000

EP2     CH1E    100.0    6.000
EP2     CH2E    100.0    6.000
EP2     PRO     100.0    6.000

EP3     CH1E    100.0    6.000
EP3     CH2E    100.0    6.000
EP3     PRO     100.0    6.000

THETAS
NH1     CH1E    EP0     50.0    120.0
NH1     CH2E    EP0     50.0    120.0

NH1     CH1E    EP1     50.0    120.0
NH1     CH2E    EP1     50.0    120.0

NH1     CH1E    EP2     50.0    120.0
NH1     CH2E    EP2     50.0    120.0

NH1     CH1E    EP3     50.0    120.0
NH1     CH2E    EP3     50.0    120.0

IMPHI
CHE1 X  X   EP1   55.0  0  0.0  ! 35.26439
CHE2 X  X   EP1   55.0  0  0.0  ! 35.26439

CHE1 X  X   EP2   55.0  0  0.0  ! 35.26439
CHE2 X  X   EP2   55.0  0  0.0  ! 35.26439

CHE1 X  X   EP3   55.0  0  0.0  ! 35.26439
CHE2 X  X   EP3   55.0  0  0.0  ! 35.26439

NONBONDED NBXMOD 5 ATOM SHIFT VATOM VDISTANCE
  VSHIFT – CUTNB 11.0 CTOFNB 9.0 CTONNB 8.0 EPS 10.0
  E14FAC 0.4 WMIN 1.5
PRO     0.00000    0.000000    0.10000000
EP1     0.00000    0.000000    0.10000000
EP2     0.00000    0.000000    0.10000000
EP3     0.00000    0.000000    0.10000000
XOY     0.00000   −0.100000    2.50000000
NIC     0.00000   −0.100000    4.00000000

NBFIX
XOY     XOY     0.000000    0.10000000
NIC     NIC     0.000000    0.10000000
NIC     OXY     0.000000    0.10000000

EP1     OXY    −0.050000    6.00000000
EP1     NIC    −0.050000    6.00000000
EP1     PRO    −2.000000   −7.00000000

EP2     OXY    −0.050000    6.00000000
EP2     NIC    −2.000000    2.00000000
EP2     PRO    −0.050000    6.00000000

EP3     NIC    −0.050000    6.00000000
EP3     OXY    −2.000000    2.00000000
EP3     PRO    −0.050000    6.00000000
```

Topology file

```
* Patch with pseudo atoms for EPR data
*
read rtf card append
* PATCH for EPR
*
27              1

PRESS           EPR                             0.0
GROU
ATOM            PORT            PRO             0.0
ATOM            EPR             EPO             0.0
BOND             CA              PROT
BOND             CA              EPR
BOND             PROT            EPR
ANGLE            N               CA              EPR
IMPH             CA              N               C           EPR
IC                N               C                * CA       PROT        1.4500      111.60      90.00       51.60       0.1000
IC                N               C                * CA       EPR         1.4500      111.60      180.00      128.40      6.0000

RESI      OXY                           0.00
GROUP
ATOM    OXY    OXY                      0.00
PATCH    FIRST    NONE                  LAST    NONE

RESI      NIC                           0.00
GROUP
ATOM    NIC    NIC                      0.00
PATCH    FIRST    NONE                  LAST    NONE

END
```

## SUPPLEMENTARY MATERIAL

To view all of the supplemental files associated with this article, visit www.biophysj.org.

We thank L. Cuello, Jie P. Wu, V. Vasquez, O. Dalmas, V. Jogini, and A. Sompornpisut for useful comments.

P.S. thanks Chulalongkorn University for supporting research training abroad. This work was supported by National Institutes of Health grants to E.P. and B.R.

## REFERENCES

1. Raman, P., V. Cherezov, and M. Caffrey. 2006. The Membrane Protein Data Bank. *Cell. Mol. Life Sci.* 63:36–51.
2. Klare, J. P., E. Bordignon, M. Doebber, J. Fitter, J. Kriegsmann, I. Chizhov, H. J. Steinhoff, and M. Engelhard. 2006. Effects of solubilization on the structure and function of the sensory rhodopsin II/transducer complex. *J. Mol. Biol.* 356:1207–1221.
3. Tate, C. G. 2006. Comparison of three structures of the multidrug transporter EmrE. *Curr. Opin. Struct. Biol.* 16:457–464.
4. Bezanilla, F. 2005. The voltage-sensor structure in a voltage-gated channel. *Trends Biochem. Sci.* 30:166–168.
5. Vigano, C., L. Manciu, and J. M. Ruysschaert. 2005. Structure, orientation, and conformational changes in transmembrane domains of multidrug transporters. *Acc. Chem. Res.* 38:117–126.
6. Lange, A., S. Becker, K. Seidel, K. Giller, O. Pongs, and M. Baldus. 2005. A concept for rapid protein-structure determination by solid-state NMR spectroscopy. *Angew. Chem. Int. Ed. Engl.* 44:2089–2092.
7. Hubbell, W. L., D. S. Cafiso, and C. Altenbach. 2000. Identifying conformational changes with site-directed spin labeling. *Nat. Struct. Biol.* 7:735–739.
8. Chanda, B., O. K. Asamoah, R. Blunck, B. Roux, and F. Bezanilla. 2005. Gating charge displacement in voltage-gated ion channels involves limited transmembrane movement. *Nature.* 436:852–856.
9. Koteiche, H. A., S. Chiu, R. L. Majdoch, P. L. Stewart, and H. S. McHaourab. 2005. Atomic models by cryo-EM and site-directed spin labeling: application to the N-terminal region of Hsp16.5. *Structure.* 13:1165–1171.
10. Kovacs, J. A., M. Yeager, and R. Abagyan. 2007. Computational prediction of atomic structures of helical membrane proteins aided by EM maps. *Biophys. J.* 93:1950–1959.
11. Mukherjee, P., I. Kass, I. T. Arkin, and M. T. Zanni. 2006. Picosecond dynamics of a membrane protein revealed by 2D IR. *Proc. Natl. Acad. Sci. USA.* 103:3528–3533.
12. Fanucci, G. E., and D. S. Cafiso. 2006. Recent advances and applications of site-directed spin labeling. *Curr. Opin. Struct. Biol.* 16:644–653.
13. Hvorup, R. N., B. A. Goetz, M. Niederer, K. Hollenstein, E. Perozo, and K. P. Locher. 2007. Asymmetry in the structure of the ABC transporter binding protein complex BtuCD-BtuF. *Science.* 317:1387–1390.
14. Li, H. C., L. Song, B. Salzameda, C. R. Cremo, and P. G. Fajer. 2006. Regulatory and catalytic domain dynamics of smooth muscle myosin filaments. *Biochemistry.* 45:6212–6221.
15. Nelson, W. D., S. E. Blakely, Y. E. Nesmelov, and D. D. Thomas. 2005. Site-directed spin labeling reveals a conformational switch in the phosphorylation domain of smooth muscle myosin. *Proc. Natl. Acad. Sci. USA.* 102:4000–4005.
16. Perozo, E., D. M. Cortes, and L. G. Cuello. 1999. Structural rearrangements underlying $K^+$-channel activation gating. *Science.* 285:73–78.
17. Altenbach, C., K. J. Oh, R. J. Trabanino, K. Hideg, and W. L. Hubbell. 2001. Estimation of inter-residue distances in spin labeled proteins at physiological temperatures: experimental strategies and practical limitations. *Biochemistry.* 40:15471–15482.
18. Chiang, Y. W., P. P. Borbat, and J. H. Freed. 2005. The determination of pair distance distributions by pulsed ESR using Tikhonov regularization. *J. Magn. Reson.* 172:279–295.

19. Jeschke, G., and Y. Polyhach. 2007. Distance measurements on spin-labelled biomacromolecules by pulsed electron paramagnetic resonance. *Phys. Chem. Chem. Phys.* 9:1895–1910.

20. Borbat, P. P., A. J. Costa-Filho, K. A. Earle, J. K. Moscicki, and J. H. Freed. 2001. Electron spin resonance in studies of membranes and proteins. *Science.* 291:266–269.

21. Dong, J., G. Yang, and H. S. McHaourab. 2005. Structural basis of energy transduction in the transport cycle of MsbA. *Science.* 308: 1023–1028.

22. Kweon, D. H., C. S. Kim, and Y. K. Shin. 2003. Regulation of neuronal SNARE assembly by the membrane. *Nat. Struct. Biol.* 10:440–447.

23. Hustedt, E. J., A. I. Smirnov, C. F. Laub, C. E. Cobb, and A. H. Beth. 1997. Molecular distances from dipolar coupled spin-labels: the global analysis of multifrequency continuous wave electron paramagnetic resonance data. *Biophys. J.* 72:1861–1877.

24. Steinhoff, H. J., and W. L. Hubbell. 1996. Calculation of electron paramagnetic resonance spectra from Brownian dynamics trajectories: application to nitroxide side chains in proteins. *Biophys. J.* 71:2201–2212.

25. Frazier, A. A., M. A. Wisner, N. J. Malmberg, K. G. Victor, G. E. Fanucci, E. A. Nalefski, J. J. Falke, and D. S. Cafiso. 2002. Membrane orientation and position of the C2 domain from cPLA2 by site-directed spin labeling. *Biochemistry.* 41:6282–6292.

26. Herzyk, P., and R. E. Hubbard. 1998. Combined biophysical and biochemical information confirms arrangement of transmembrane helices visible from the three-dimensional map of frog rhodopsin. *J. Mol. Biol.* 281:741–754.

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

28. Alexander, N., M. Bortolus, A. Al-Mestarihi, H. McHaourab, and J. Meiler. 2008. De novo high-resolution protein structure determination from sparse spin-labeling EPR data. *Structure.* 16:181–195.

29. Liu, Y. S., P. Sompornpisut, and E. Perozo. 2001. Structure of the KcsA channel intracellular gate in the open state. *Nat. Struct. Biol.* 8:883–887.

30. Perozo, E., D. M. Cortes, P. Sompornpisut, A. Kloda, and B. Martinac. 2002. Open channel structure of MscL and the gating mechanism of mechanosensitive channels. *Nature.* 418:942–948.

31. Sompornpisut, P., Y. S. Liu, and E. Perozo. 2001. Calculation of rigid-body conformational changes using restraint-driven Cartesian transformations. *Biophys. J.* 81:2530–2546.

32. Perozo, E., D. M. Cortes, and L. G. Cuello. 1998. Three-dimensional architecture and gating mechanism of a $K^+$ channel studied by EPR spectroscopy. *Nat. Struct. Biol.* 5:459–469.

33. Zhou, Y., J. H. Morais-Cabral, A. Kaufman, and R. MacKinnon. 2001. Chemistry of ion coordination and hydration revealed by a $K^+$ channel-Fab complex at 2.0 A resolution. *Nature.* 414:43–48.

34. Vásquez, V., M. Sotomayor, J. Cordero-Morales, K. Schulten, E. Perozo. 2008. A structural mechanism for MscS gating in lipid bilayers. *Science.* 321:1210–4.

35. Chakrapani, S., L. Cuello, M. D. Cortes, and E. Perozo. 2005. Structural dynamics of an isolated voltage-sensor domain in a lipid bilayer. *Structure.* 16:398–409.

36. Canutescu, A. A., A. A. Shelenkov, and R. L. Dunbrack. 2003. A graph-theory algorithm for rapid protein side-chain prediction. *Protein Sci.* 12:2001–2014.

37. Berneche, S., and B. Roux. 2002. The ionization state and the conformation of Glu-71 in the KcsA $K^+$ channel. *Biophys. J.* 82: 772–780.

38. Brooks, B. R., R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus. 1983. CHARMM—a program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* 4:187–217.

39. Neria, E., S. Fischer, and M. Karplus. 1996. Simulation of activation free energies in molecular systems. *J. Chem. Phys.* 105:1902–1921.

40. Sezer, D., J. H. Freed, and B. Roux. 2008. Simulating electron spin resonance spectra of nitroxide spin labels from molecular dynamics and stochastic trajectories. *J. Chem. Phys.* 128:165106.

41. Sezer, D., J. H. Freed, and B. Roux. 2008. Parametrization, molecular dynamics simulation, and calculation of electron spin resonance spectra of a nitroxide spin label on a polyalanine $\alpha$-helix. *J. Phys. Chem. B.* 112:5755–5767.

42. Orengo, C. A., D. T. Jones, and J. M. Thornton. 2003. Bioinformatics: Genes, Proteins and Computers. BIOS Scientific Publishers, Oxford, England.

43. Zhang, Y., I. A. Hubner, A. K. Arakaki, E. Shakhnovich, and J. Skolnick. 2006. On the origin and highly likely completeness of single-domain protein structures. *Proc. Natl. Acad. Sci. USA.* 103:2605–2610.

44. Guex, N., and M. C. Peitsch. 1997. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis.* 18:2714–2723.

45. Goldsmith-Fischman, S., and B. Honig. 2003. Structural genomics: computational methods for structure analysis. *Protein Sci.* 12:1813–1821.

46. Skolnick, J. 2006. In quest of an empirical potential for protein structure prediction. *Curr. Opin. Struct. Biol.* 16:166–171.

47. Tamm, L. K., and B. Y. Liang. 2006. NMR of membrane proteins in solution. *Prog. Nucl. Magn. Reson. Spectrosc.* 48:201–210.

48. Chill, J. H., J. M. Louis, C. Miller, and A. Bax. 2006. NMR study of the tetrameric KcsA potassium channel in detergent micelles. *Protein Sci.* 15:684–698.

49. Lange, A., K. Giller, S. Hornig, M. F. Martin-Eauclaire, O. Pongs, S. Becker, and M. Baldus. 2006. Toxin-induced conformational changes in a potassium channel revealed by solid-state NMR. *Nature.* 440:959–962.

50. Mansoor, S. E., and D. L. Farrens. 2004. High-throughput protein structural analysis using site-directed fluorescence labeling and the bimane derivative (2-pyridyl)dithiobimane. *Biochemistry.* 43:9426–9438.

51. Langen, R., K. J. Oh, D. Cascio, and W. L. Hubbell. 2000. Crystal structures of spin labeled T4 lysozyme mutants: implications for the interpretation of EPR spectra in terms of structure. *Biochemistry.* 39:8396–8405.