Structure and Dynamics of Monoclonal Antibody Domains Using Spins, Scattering, and Simulations.

Veronika A. Szalai[a],*, Christina Bergonzo[b],[c],*, Rachel B. Lyon[b], Zvi Kelman[b],[c], Thomas Schmidt,[d] Alexander Grishaev[b],[c],*

[a] Dr. V. Szalai
Physical Measurement Laboratory,
National Institute of Standards & Technology
Gaithersburg, MD, 20899, United States
E-mail: veronika.szalai@nist.gov
[b] Dr. C. Bergonzo, Dr. R. Lyon, Dr. Z. Kelman, Dr. A. Grishaev
Institute for Bioscience and Biotechnology Research
University of Maryland
9600 Gudelsky Drive, Rockville, Maryland 20850, United States
[c] Dr. C. Bergonzo, Dr. Z. Kelman Dr. A. Grishaev
Material Measurement Laboratory
National Institute of Standards and Technology
100 Bureau Drive, Gaithersburg, Maryland 20899, United States
E-mail: christina.begonzo@nist.gov, alexander.grishaev@nist.gov
[d] Dr. T. Schmidt
Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases
National Institutes of Health
Bethesda, Maryland 20892-0520, United States

Supporting information for this article is given via a link at the end of the document.

Abstract: Antibody-based pharmaceuticals are the leading biologic drug platform (> $75B/year).[1] Despite a wealth of information collected on them, there is still a lack of knowledge on their inter-domain structural distributions, which impedes innovation and development. To address this measurement gap, we have developed a new methodology to derive biomolecular structure ensembles from distance distribution measurements via a library of tagged proteins bound to an unlabeled and otherwise unmodified target biologic. We have employed the NIST monoclonal antibody (NISTmAb) reference material as our development platform for use with spin-labeled affinity protein (SLAP) reagents. Using double electron-electron resonance (DEER) spectroscopy, we have determined inter-spin distance distributions in SLAP complexes of both the isolated Fc domain and the intact NISTmAb. Our SLAP reagents offer a general and extendable technology, compatible with any non-isotopically labeled immunoglobulin G class mAb. Integrating molecular simulations with the DEER and solution X-ray scattering measurements, we enable simultaneous determination of structural distributions and dynamics of mAb-based biologics.

Introduction

Biomolecular flexibility is central to biological function.[2] Flexible connections between protein domains control physiological events such as polyubiquitin-mediated protein degradation,[3],[4] pre-mRNA splicing,[5] and transcription.[6] However, experimental characterization of flexibility continues to be one of structural biology's major unsolved challenges. Our current understanding of both protein and nucleic acid structure emphasizes discrete conformations, sidelining their dynamics because of the limitations of both the simulations and the experimental tools. AI-based approaches for protein structure prediction,[7] trained on largely single-conformation crystal structures in the Protein Data Bank (PDB), are also not parameterized to reproduce dynamic ensembles. The lack of a diverse set of experimentally-determined dynamic models for biomolecules hinders extension of these promising AI techniques to flexible architectures.[8],[9] Single-particle reconstructions via cryo-electron microscopy assume an underlying discrete distribution of the electron density, although partial information on dynamic systems can be extracted by electron tomography.[10] Solution nuclear magnetic resonance (NMR) spectroscopy, while reporting on conformational ensemble averages, comes with caveats such as (i) lower density of restraints and enhanced spin coherence relaxation for biomolecules larger than ≈ 30 kDa in mass, and (ii) the need for isotopic (^{2}H/^{15}N/^{13}C) labeling and unambiguous spectral resonance assignments. These issues greatly complicate NMR structure characterization of flexible systems of pharmaceutical significance such as monoclonal antibody (mAb) based biomedicines with molecular masses of approximately 150 kDa or even their individual domains with masses of approximately 50 kDa. Solution scattering holds promise for structural characterization of flexibly-linked multidomain proteins under native-like conditions. However, when used in isolation, it too suffers from limited information content relative to the underlying multidimensional distributions. In this study, we present a hybrid approach in which structural and dynamic information are extracted via double electron-electron resonance (DEER) spectroscopy, combined with solution small- and wide-angle X-ray scattering (SWAXS), and processed via a judicious computational framework to generate experimentally-verified motional models.

The choice of DEER is motivated by its ability to report on distance distributions up to approximately 16 nm,[11] irrespective of the overall biosystem size, and covering the typical inter-domain spatial separations in biologics such as mAbs. However, characterization of distances beyond 8 nm necessitate high levels (> 95 %) of deuteration for both the buffer and the biomolecule of interest.[11],[12] In addition, as DEER measurements require the presence of two

paramagnetic centers, spin attachment sites need to be introduced via mutations of the target biomolecule.[13] These aspects make application of DEER to biopharmaceuticals problematic, due to challenges of expressing and properly folding sufficient quantities of these mammalian proteins in perdeuterated form and the need for a separate construct for each inter-site distance measurement. Our technology bypasses these issues by creating a library of small bacterially expressed deuterated spin-labeled affinity proteins (SLAPs) that bind tightly ($K_D \approx 30$ nmol L^{-1})[14] and in a known arrangement to the non-deuterated biopharmaceutical. Deuteration of the SLAPs removes ^1H spins near (< 1.5 nm) the unpaired electrons in the tags, extending the accessible distance measurement range beyond 10 nm without any modifications of the target biopharmaceutical antibody platforms. We supplement the inter-probe distances obtained via DEER with SWAXS data as an independent global restraint on the antibody's conformational space, and molecular dynamics (MD) simulations linking the DEER and X-ray scattering distance correlations to the conformational ensembles and providing atomic level insight into antibody dynamics. Combinations of DEER and solution scattering have been used to characterize protein structure in the past,[15–17] including instances when spin labels were introduced via affinity binding.[18,19] These applications probed inter-spin distances of 2 nm to 8 nm and used either single-structure or discrete-set representations of protein conformations. Compared to methods of ensemble optimization that vary both structures and ensemble member populations, we have utilized a continuous representation of the probability distribution function to minimize the number of adjustables in our model. Our generalizable approach allows, via selective deuteration near the spin sites, distance measurements exceeding 10 nm, that probe the inter-domain arrangements on unlabeled and unmodified protein targets, with a data fitting procedure designed to handle continuous conformational distributions.

We demonstrate our approach using the Fc domain of the recombinant, humanized IgG1κ NISTmAb monoclonal antibody as the target biomolecule[20] and individual domains of the bacterial protein A (pA) as SLAPs. C_2-symmetric Fc includes two CH2 domains linked to the CH3 domain dimer. Based on the extensive set of the available IgG Fc crystal structures, the positioning of the CH2 domains relative to the CH3 is expected to be flexible (**Supporting Figure 1**). The extent of this flexibility, which has not been characterized experimentally, is the subject of our study. pA is a well-studied, readily expressed small protein that binds the Fc domain, making it a general reagent for characterization of IgG biotherapeutic platforms. Here, we attach electron spin tags to variants of protein A, which bind at the junction of the CH2 and CH3 domains (**Figure 1**).[21] Binding of two pA SLAPs to the C_2-symmetric Fc dimer creates a complex with spin labels extending into deuterated solvent, away from the protonated sites, and with relative spin/spin separations reporting, along with the solution scattering data, on both the average geometry and the relative motions of the CH2 and CH3 domains. Our technology allows simultaneous characterization of structure and dynamics in addition to the *de novo* derivation of the target/SLAP arrangement.

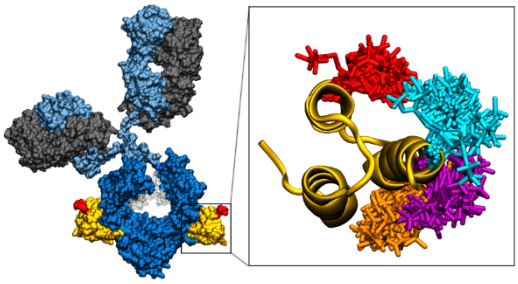

Figure 1. Left: NISTmAb with the Fc domain highlighted in dark blue and the SLAP A25C in yellow, bound at the elbow region between CH2 and CH3 domains. The spin label tag on SLAP A25C is shown in red. Fab domains colored by light blue (heavy chain) and grey (light chain) surface representations. Right: Positions of the spin label on Z domain of protein A used to create SLAPs A25C (red), A38C (orange), A42C (purple), and A50C (cyan). Four label positions are shown in the figure, but only a single label is present on each individual SLAP. Licorice represents the spatial distribution of the spin label for 1000 MD-sampled structures.

Results and Discussion

Measurement of inter-domain distance distributions in a mammalian cell-expressed biologic

Based on the crystal structures of pA bound to the Fc domain (PDB ID 5U4Y) and the molecular dynamics simulations, inter-spin distances in the $Fc(pA)_2$ complex are expected to range from 7 nm to 11 nm depending on the spin label position on the SLAP. Extraction of distances in that range requires collection of DEER echo decays with evolution times of 5 µs to 22 µs to accurately determine the mean inter-spin distance[22] and even longer evolution times of 10 µs to 45 µs for accurate determination of the distribution widths. A total of five SLAPs were constructed for differential placement of the spin label, four of them (A25C, A38C, A42C and A50C) based on the Z domain of protein A and one, K54C based on the closely related B domain of protein A. **Supporting Figure 2** shows individual DEER data traces (from biological and/or technical replicates), fitted to obtain a distance distribution P(r) using either model-free Tikhonov regularization[23,24] or a single Gaussian distance model that includes a best-fit exponential decay background correction. Subsequently, DEER data for each $Fc(SLAP)_2$ sample were fit globally[25] with either the nonparametric Tikhonov regularization approach or a single Gaussian distance model (**Table 1**). Because of challenges in globally fitting data using a nonparametric distance model for one of the variants (see **Supporting Figure 2** legend), mean distances <r> and standard deviations σ(r) shown in **Figure 2** are from global fits using a single Gaussian distance model.[26]

Table 1. Average inter-spin distances <r> and standard deviations σ(r) in nm derived from global analysis of DEER data, using either DeerLab or a Single Gaussian, and molecular dynamics simulations.

Sample	DEER Lab (Global Analysis) ⟨r⟩ ± σ(r)	Single Gaussian (Global Fit) ⟨r⟩ ± σ(r)	MD Average and Standard Deviation (4 µs per Tag)
A25C	8.76 ± 0.21	8.85 ± 0.13	9.01 ± 0.27
A38C	N/A[i]	7.29 ± 0.36	7.61 ± 0.71
A42C	9.23 ± 0.42	9.23 ± 0.53	8.45 ± 0.47
A50C	10.83 ± 0.98	10.74 ± 1.30	10.44 ± 0.25
K54C	10.54 ± 0.50	10.47 ± 0.59	10.38 ± 0.60

[I]See Supporting Figure 2.

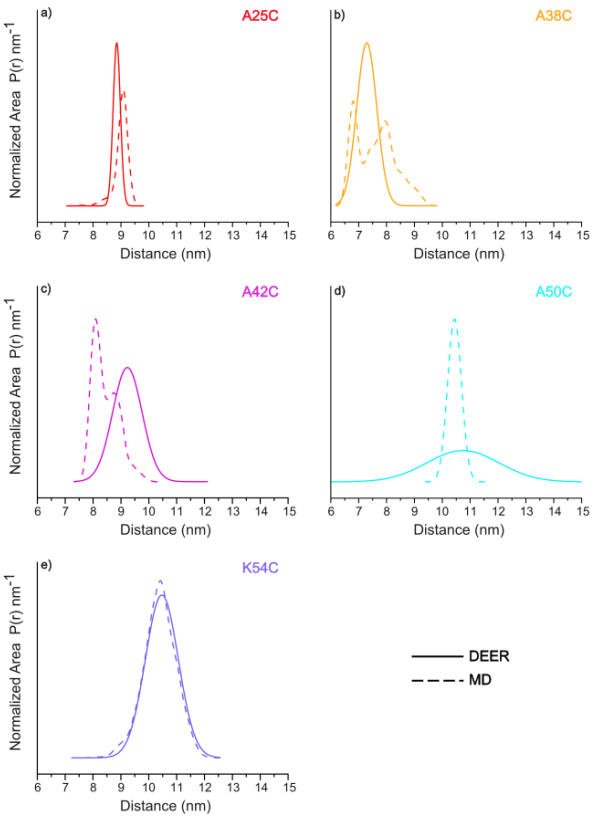

Figure 2. (a-e) Distance distributions between the nitroxide spin labels in $Fc(pA)_2$ complexes, normalized to total area. MD simulations (dashed lines) are shown overlapped with DEER experimental distance distributions (solid lines) for all five tags. MD Data are histogrammed distances calculated over all frames of all simulations, totaling 400000 frames representing 4 μs of simulation.

The inter-spin distance distribution width is a convolution of the distribution of the spin label positions relative to the carrier protein (including rotamer distributions),[27] and the dynamics of the protein(s) to which the carrier proteins are bound. The conventional spin tagging reagent MTSL we used offers simple attachment chemistry but also introduces flexibility in its linkage to pA.[28],[29] Aside from the expected effect of the spin label linker flexibility, molecular dynamics simulations indicate that internal dynamics of pA has a negligible effect on the inter-spin distributions. Within the Fc protein, the "ball and socket" movements of the CH2 domain relative to the CH3 dimer are reflected, via pA's positioning at the "elbow" between the CH2 and CH3 domains, in both changes of the overall SWAXS profile, and the variation of the specific spin/spin distance distributions.

Three of the complexes, $Fc(A25C)_2$, $Fc(A38C)_2$, and $Fc(K54C)_2$, exhibit experimental spin/spin distances and distribution widths from DEER that overlap well with the results of the MD. The P(r) curve for $Fc(A25C)_2$ exhibits a very narrow width, suggesting that the spin label is in a rigid environment and located at a position exhibiting minimal sensitivity to the relative motions of the CH2 and CH3 domains. Except for A50C, DEER-extracted inter-spin distance widths are consistent with unimodal distributions.[30] The half-maximum full-width for the A50C distance distribution is very broad (≈ 3 nm) and is likely a result of the low spin-labeling efficiency of this construct, which decreases the modulation depth in the DEER data, and hinders precise determination of the distribution width. The P(r) curves from DEER of NISTmAb$(SLAP)_2$ complexes are extremely similar to those obtained for the corresponding $Fc(SLAP)_2$ complexes (**Supporting Figure 2**), indicating that the presence of the Fab domains does not disrupt SLAP binding to Fc or change the Fc dynamics appreciably. The Hahn echo decay signals of the $Fc(SLAP)_2$ complexes were fit with stretched exponential functions to determine the apparent phase memory times (T_m) (**Supporting Figure 3** and **Supporting Table 1)**. The phase memory times are sensitive to the proton concentration in the volume surrounding each spin label. Canarie *et al.*[31] have shown that protons within 0.4 nm to 1.2 nm of a nitroxide spin label such as MTSL exhibit the strongest effects on electron spin decoherence. Correspondingly, T_m values for $Fc(SLAP)_2$ complexes inversely correlate with proximity to non-exchangeable ^1H atoms in the complexes (**Supporting Figure 4**). As both Fc and NISTmAb are protonated, our phase memory times for their complexes with SLAPs are not as long as those observed for biomacromolecular systems in which deuterium incorporation is nominally 100 %.[11,32] As our spin labeling levels are lower than 100 %, the modulation of our DEER traces is also suppressed.

Deconvolution of Protein A, spin label, and Fc dynamics

All-atom MD simulations of Fc bound to pairs of each SLAP docked at the Fc elbows were performed and results were aggregated over 4 μs of simulation. **Figure 2** and **Table 1** report the comparison of DEER-derived MTSL label inter-spin distance distributions to those calculated from the MD simulations. Most of the tags exhibit calculated distance averages overlapping with the experimental averages, except for A42C, which deviates by 0.8 nm, although distribution widths are similar. While the MD and DEER average distances are generally in agreement, additional features are evident in the MD distributions (e.g. A38C and A42C), which suggests differences in sampled label rotamers between the simulations and the experiment, either due to the sampling limitations of the MD or the temperature difference between the MD simulation and the DEER measurements.

Independent principal component (PC) analysis was performed on the common set of backbone atoms for simulations of $Fc(pA)_2$ with MTSL spin labels (MD (+pA)), simulations of Fc alone (MD (-pA))[33], and the 40 IgG1 Fc depositions in the PDB (Xtal, as listed in the **Supporting Information**), in order to compare the dynamics exhibited by these systems. **Figure 3** shows the projections along the first two PCs for each system as directional arrows (so-called porcupine plot) on top of the Fc structure. The dominant low frequency motions are the same for each system as described by the first two PCs. However, the crystal structure set has PC1 and PC2 (the variance along the first two orthogonal eigenmodes) swapped from the MD simulation sets. This difference is most likely associated with the relatively small number of crystal structures in the set. The MD simulations of Fc with (+pA) and without (-pA) the SLAPs show similarity between the order of the top two PCs as well as their magnitude (the percentage of overall variance described by each PC). The movements along each PC correlate well with the center of mass opening distance between CH2 domains of each chain of the Fc dimer, as illustrated in **Figure 3** (and quantified in **Supporting Figure 5**). **Table 2** quantifies the similarity between PCs of each

system via the root mean square inner product (RMSIP) between the top two eigenvectors (above the diagonal) and all eigenvectors (below the diagonal) of each system. These results indicate that the motions in all ensembles share a high degree of similarity (average of 0.774 for all eigenvectors and average of 0.961 for the top two eigenvectors).

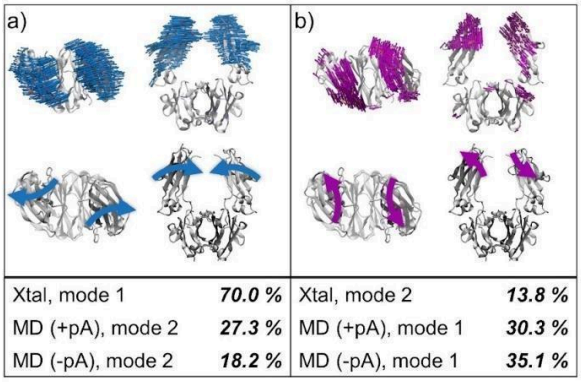

Figure 3. Fc CH2 domain motions projected along the first two principal component eigenmodes (a, b) of 40 crystal structures of Fc domains (Xtal); 40,000 frames from 4 simulations of Fc with pA and MTSL labeled at A25C (MD (+pA)); and 40,000 frames from 4 simulations of Fc truncated from PDB ID: 5VGP to residue 238 for each heavy chain (MD (-pA)). These projections were calculated using Cpptraj and NMWiz and visualized in VMD. The percentages denote the weight of each eigenmode. PCA based on the set of 40 crystal structures is underdetermined due to the limited ensemble size.

Table 2. Root mean square inner product (RMSIP) between eigenvectors 1 and 2 (above the diagonal) and all 15 calculated eigenvectors (below the diagonal) for each set of structures (40 Crystal Structures (Xtal), 40,000 frames from Fc+pA with A25C MTSL label (MD (+pA)), and 40,000 frames from Fc (MD (-pA))). Values of 1 indicate that eigenvectors are identical.

RMSIP	Xtal	MD (+pA)	MD (-pA)
Xtal	--	0.962	0.948
MD (+pA)	0.710	--	0.974
MD (-pA)	0.714	0.897	--

Generally good agreement between the MD-predicted and the experimental DEER and SWAXS observables (**Figure 4b**, bottom curve) indicates that both the experimental data and the computational procedures are reliable and suggests that the inter-domain motional model may be determined directly from the experimental data. Therefore, we set out to investigate whether DEER and SWAXS data are sufficient to uniquely define both the structure and the conformational dynamics of a multi-domain protein such as Fc. We have performed two distinct analyses, one that relies on prior knowledge of the Fc/pA binding interface, and the other that does not. The former examines the possibility of uniquely defining both the structure and dynamics at the flexible inter-domain interface (i.e. between the CH2 and CH3 domains), while the latter evaluates the fidelity with which the biopharmaceutical target/SLAP interface can be determined *de novo*, based on experimental data such as ours.

Fc motional model determined via DEER, SWAXS, and molecular simulation

Crystal structures reported for the IgG1 Fc domains exhibit root-mean-square (RMS) variations of 0.28 nm for the centers of mass of the CH2 domains relative to the CH3 domain dimers, and 4° to 10° RMS variations of the (Θ, φ, ψ) angles for the reactive orientations of their moments of inertia tensor frames, with variables defined in **Supporting Figure 6**. The corresponding RMS values for the Fc MD trajectory are similar: 0.27 nm and 7° to 8°, respectively. Such variability of the CH2 domain positioning relative to the CH3 dimer is consistent with the small surface area of the interdomain interface (4.6 nm^2 for 5VGP PDB ID) and its lack of specific hydrogen bonding or salt bridge contacts. The distributions of both the rotational and the translational parameters of the CH2 domain with respect to the CH3 dimer appear unimodal (**Supporting Figure 7**). A normal distribution representing the minimal six-dimensional joint translational and rotational movement model corresponds to 27 variables, assuming both the equivalence and the lack of correlation of the two CH2 domains (**Supporting Figure 8**). Six of those 27 variables define the vector of means and the remaining 21 the elements of the variance-covariance matrix. Although our experimental data are insufficient to define this number of parameters, the model can be reduced to rotational ball-and-socket type motions considering the short 6-residue linker connecting the CH3 and CH2 domains. To decrease the number of variables in the motional model from 27 to 9, we remove the translational degrees of freedom and project into the 3-dimensional space of (α, β, γ) rotations via pivot points (Materials and Methods Section). Assuming that both the correlations between the rotation angles for a single CH2 domain, and the inter-CH2 domain angular correlations are driven by the avoidance of inter-domain steric clashes, the off-diagonal elements of the variance-covariance matrix can be neglected as well, further reducing the number of fitted model parameters from 9 to 6.

Using this rotational-only model, CH2/CH3 domain pivot point optimizations were carried out for the set of Fc MD frames and the set of the IgG1 Fc crystal structures (**Supporting Figure 9**). The crystal structures set encompasses motions that are closer to the rotations-only model compared to the MD set (pivot point distance of 0.4 nm and 1.0 nm, respectively). The MD-fitted and the crystal structures-fitted pivot points are 0.4 nm apart when projected on the 5VGP structure, and both are located at the interface between the CH2 and CH3 domains, indicating general consistency of the results of this analysis for these two distinct ensembles. Best-fitted pivot points yield the residual r.m.s. of 0.04 nm between the coordinates of the pivot-predicted and crystal structures-based frames representing the CH2 and CH3 domains. Based on these observations, we estimate that our rotational-only approximation corresponds to translational errors ranging between 0.06 nm and 0.2 nm.

Determinations of the average structure of the Fc and its variation relative to the average were carried against the SWAXS data for Fc and the DEER data for the Fc(pA)$_2$ with A25C, A38C, A42C, and A50C SLAPs, assuming either a single structure or a dynamic model, as described in the Materials and Methods Section, via the (α, β, γ) rotations as defined in the **Supporting Figure 10**. We did not exclude DEER data that were challenging to fit such as those for the Fc(A50C)$_2$ construct and we did not optimize rotamer distributions of the spin labels, in order to avoid interpretation biases. CH2 domain orientations for the best-fitting C_2-symmetric single-structure models correspond to the pivot-based (α, β, γ) parameters of (31°, -54°, -146°), and the moments of inertia tensor based (Θ, φ, ψ) parameters of (30°, -43°, -124°). The distances between the CH2 center of mass for the best-fitting single-structure models and the corresponding points for the MD and crystal structure ensembles are 0.13 nm and 0.21 nm, respectively. The

translational and rotational parameters of the single-structure best-fitting models fall within both the MD and the crystal structure ensembles. Dynamic model optimization results in CH2 domain orientational parameters close to the single-structure optimized parameters while including RMS motions of several degrees, with ($\alpha \pm \sigma_\alpha$, $\beta \pm \sigma_\beta$, $\gamma \pm \sigma_\gamma$) values of (32° ± 6°, -56° ± 6°, -145° ± 5°) (**Figure 4a**). The average locations of the CH2 centers of mass for the best-fitting ensembles remain very close to the parameters for the single-structure best-fitting models, as well as both the crystal structure and the MD ensembles (0.11 nm, 0.11 nm, and 0.14 nm, respectively). The best-fitting ensembles exhibit RMS variation of the centers of mass of 0.33 nm relative to their means, also comparable to both the MD and the crystal structure ensembles' parameters. The angular parameters for the CH2/CH3 moments of inertia tensor orientational distributions ($\Theta \pm \sigma_\Theta$, $\varphi \pm \sigma_\varphi$, $\psi \pm \sigma_\psi$) for the best-fitted dynamic ensemble (31° ± 7°, -45° ± 7°, -121° ± 7 °) are also comparable to the respective parameters for the MD (38° ± 8°, -46° ± 8°, -117° ± 7°) and the crystal structure (32° ± 4°, -55° ± 10°, -113° ± 9°) ensembles, while being distinct from either of those. Our best-fitting ensemble models therefore provide an independent assessment of the native-state motions in the Fc, exhibiting both similarities and differences with respect to the MD and the crystal structure ensembles. The best-fitting ensemble fits of both SWAXS and DEER data are within the experimental uncertainties (**Figure 4b, c**). As we have not used the K54C variant DEER distance data for model generation, it was employed for validation of the best-fitting dynamic models. Close agreement between the experimental spin-spin distance distribution parameters (10.5 nm ± 0.6 nm) and those predicted from the best-fitting ensembles (10.3 nm ± 0.6 nm) indicates robustness of the derived dynamic model. Our results suggest that flexible positioning of the CH2 domains relative to the CH3 dimer is an essential native-state feature of the Fc architecture, contributing to its interactions with other proteins.[34],[35] Our approach allows capture of complex intra-molecular dynamics with readily accessible scattering data combined with long-range DEER measurements on an unmodified target. Moderate 5° to 7° RMS orientational variation of the CH2 domains relative to the CH3 dimer is expected to have pronounced effects on the geometry of the proline-rich linkers connecting the Fab and Fc domains, and in turn on their relative positioning within the antibody. Our best-fitting solution corresponds to the average and RMS distances of 1.93 nm ± 0.30 nm between the Cα atoms of residues P238 at the N-termini of the two CH2 domains, with a total range of 1.3 nm to 2.5 nm. An even broader distribution of 2.23 nm ± 0.58 nm is sampled by the crystal structures of unbound Fc, with the P238/P238 distances ranging from 1.2 nm to 3.6 nm. Based on the corresponding crystal structures in the PDB (**Supporting Information, Section 7**), binding of the Fc domain to the Fc γ receptor dramatically sharpens this distribution to 2.50 nm ± 0.04 nm. The changes in the distance between the N-terminal tips of two CH2 domains in the Fc, transmitted via the Fab-Fc linkers, are likely responsible for coupling of the interactions of the Fc and Fab domains with their binding partners.

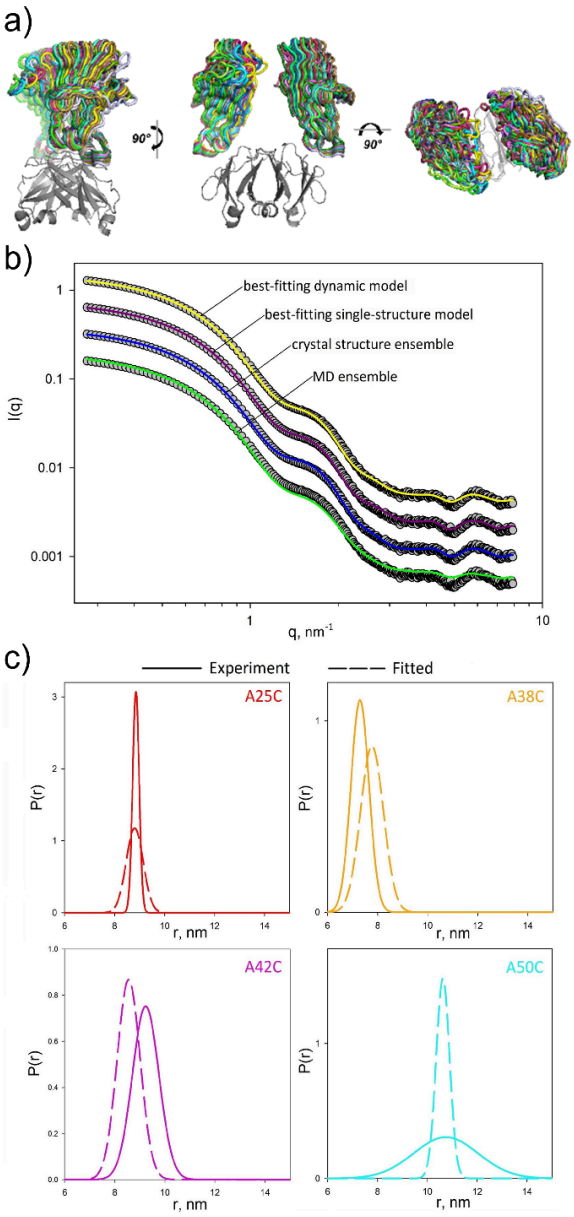

Figure 4. (a) A best-fitting dynamic ensemble model containing 20 overlapping CH2 domain structures, colored individually. (b) Fits of the experimental SWAXS data (black circles), offset for clarity. Bottom to top: MD trajectory (green line, χ^2=3.38), crystal structure ensemble (blue line, χ^2=0.55), best-fitting single structure model (dark pink line, χ^2=0.48), and best-fitting dynamic ensemble model (yellow line, χ^2=0.41). (c) Fits of the DEER distance distributions (solid lines) with the dynamic ensemble model (dashed lines).

De novo determination of the Fc/pA complex structure

The structure and dynamic distribution analysis of the Fc(pA)$_2$ complex relied on *a priori* knowledge of their binding interface. To assess the value of this information, we investigated the possibility of recovering the correct structure of the Fc(pA)$_2$ complex from random trial relative positions relying only on the DEER and SWAXS data supplemented by a cutoff based on the surface area of the binding interface. DEER distance restraints provided a highly effective docking

filter, reducing the initial candidate model pool by 99.99 %. The SWAXS data filter further shrunk the set of models by 95 %, and the interfacial surface area (ISA) cutoff reduced it by another 80 % (**Figure 5a, Supporting Information**), resulting in a tightly defined cluster of solutions. Evolution of the pA positioning accuracy with each applied filter indicates strong convergence towards the correct geometry, exemplified by the 5U4Y crystal structure of the Fc(pA)$_2$ complex. The DEER-filtered structure pool exhibits pA coordinate RMS deviations (RMSDs) from the correct model ranging between 0.2 nm and 7.0 nm, with 1.4 % of those falling at or below 0.7 nm. The 3D distribution of the centers of mass of the filtered pA positions (**Figure 5b**) indicates that despite this filter's dramatic reduction in the number of candidate models, the limited DEER data sets we have alone are insufficient to define the Fc/pA interface, as filtered pA positions cover most of the Fc surface. This finding is consistent with prior studies of protein complex rigid-body docking against limited sets of DEER restraints.[36] Passing the DEER-filtered model set through the SWAXS data filter increases the lowest-RMSD fraction to 28 %, with the three-dimensional distribution of the pA centers of mass now localized to a semi-ellipse passing through the lowest-RMSD cluster, and indicating that our DEER and SWAXS data are still insufficient for unique definition of the contact interface. Approximately one half of the DEER and SWAXS-filtered pool corresponds to RMSD accuracies ranging from 0.8 nm to 2.4 nm, with the majority of those exhibiting inaccurate *orientational* rather than *translational* positioning of pA. This result highlights the facts that (i) the two pA units account for only 19 % of the electron density in the complex, and thus of the total SWAXS signal, and (ii) that the globular shape of the pA units makes model-predicted SWAXS data relatively insensitive to their rotations around their centers of mass. Application of the ISA cutoff of 2 nm^2 (**Supporting Figure S11**) results in an even smaller pool of models, 89 % of which exhibit pA positioning accuracy of 0.7 nm or better and the three-dimensional distribution of the centers of mass nearly completely converging on this main cluster of solutions **(Figure 5b,c)**. The mode of the final cluster of solutions corresponds to 0.25 nm backbone coordinate RMSD to the correct 5U4Y structure. Therefore, our results show the possibility of determining the relative positioning of the target and SLAP in a *de novo* fashion based on DEER and SWAXS data as restraints.

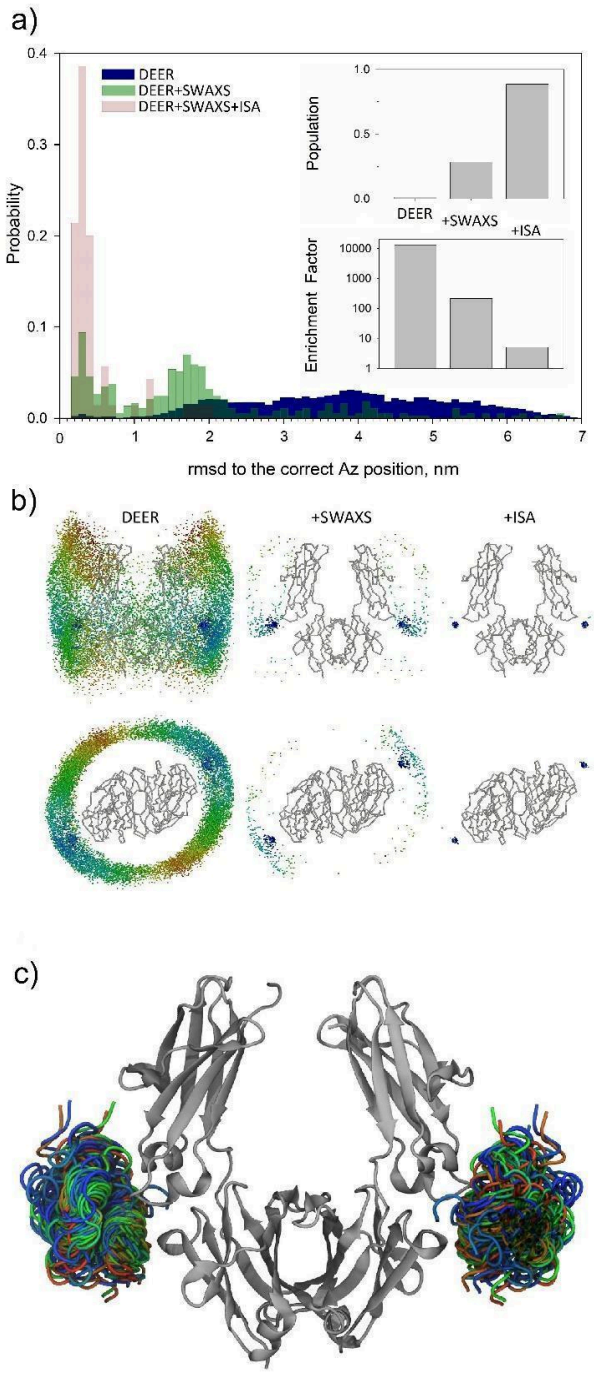

Figure 5. (a) Distributions of the backbone coordinate accuracy for the sets of structures selected against the sequentially applied DEER distance restraints, SWAXS data, and the interfacial surface area (ISA) cutoff. Top insert: population of the correct solution (PDB ID: 5U4Y) with RMSD ≤ 0.7 nm following sequentially applied DEER & SWAXS filters and ISA cutoff. Bottom insert: enrichment factors for DEER, SWAXS, and ISA restraints calculated as the ratios of the number of structures prior to and following the application of each filter. (b) Positions of the centers of mass of the pA SLAPs (colored dots) relative to the Fc backbone with each of the successively applied DEER & SWAXS filters and ISA cutoff. Each dot is colored according to the corresponding protein's backbone RMSD to the correct structure (PDB ID 5U4Y), with the colors from blue corresponding to 0.2 nm to red corresponding to 7.0 nm. The top and bottom views are related by 90° rotations. (c) Superimposition of the pA backbones passing all the DEER & SWAXS filters and ISA cutoff (colored by RMSD to the

correct docked orientation, where blue is the lowest RMSD of 0.2 nm and red is the highest RMSD of 1.1 nm), relative to the Fc backbone (gray).

Conclusion

The range of structures accessible through a protein's dynamic motions determines interactions with its binding partners and impacts factors such as immune response and toxicity when it is employed as a biopharmaceutical. Characterization of the higher order structure of biologic drugs such as mAbs is therefore incomplete without accounting for their inherent flexibility. We devise a novel platform for determination of interdomain arrangements in flexibly-linked proteins by combining spin-labeled affinity protein (SLAP)-based DEER measurements, SWAXS data, and atomistic simulations. Our technology allows quantitative determination of a multi-domain protein's dynamics and structure without the need to label the biologic target isotopically or covalently, while retaining access to long-range (> 10 nm) distance measurements. Moreover, our method is agnostic to the biologic's post-translational modifications (e.g. glycosylation) that are important for drug efficacy. For our model Fc system, we observe 5° to 7° RMS re-orientational motions of CH2 domains relative to the CH3 domain dimer. This degree of flexibility likely contributes to Fc's selectivity and binding to cellular receptor proteins,[34],[35] consistent with the noted allostery[37] connecting the Fab and Fc domains. We also achieve sub-residue accuracy when determining the protein complex structure from the DEER and SWAXS data supplemented by a simple interfacial area restraint, allowing applications of our technology to cases with limited prior structural information. Our work paves the way to determining Fab/Fab and Fab/Fc distributions in full-length mAbs, through use of Fab-binding SLAPs such as those derived from proteins G, L, or A. Applications of approaches such as ours will aid production of much-needed datasets of experimentally determined flexible bio-structures. These training datasets are expected to enable next-generation AI-based flexible structure prediction tools, thereby enhancing our ability to both understand and influence the physicochemical properties and behaviors of these diverse and abundant classes of biomolecular architectures.

Supporting Information

The data that support the findings of this study are available in the Supplementary Information of this article. The authors have cited additional references within the Supporting Information.[38-76]

Acknowledgements

We thank Dr. Brad O'Dell for coining the term SLAP. Funding was through Science & Technical Research Services, National Institute of Standards and Technology.

Notes

This article was funded by the National Institute of Standards and Technology and is not subject to U.S. Copyright. Certain commercial firms and trade names are identified to specify the usage procedures adequately for reproducibility. Such identification is not intended to imply recommendation or endorsement by the National Institute of Standards and Technology, nor is it intended to imply that related products are necessarily the best available for the purpose.

Keywords: Double electron-electron resonance spectroscopy, molecular dynamics, solution X-ray scattering, monoclonal antibody biologics, biopharmaceutical development

Abbreviations: DEER, double electron-electron resonance; SWAXS, small- and wide-angle X-ray scattering; mAb, monoclonal antibody; MD, molecular dynamics; MIT, moments of inertia tensor; MTSL, S-(1-oxyl-2,2,5,5-tetramethyl-2,5-dihydro-1H-pyrrol-3-yl)methyl methanesulfonothioate; pA, protein A; RMS, root mean square; RMSD, root mean square deviation; SLAP, spin-labeled affinity protein.

[1] A. Mullard, *Nat. Rev. Drug Discov.* **2021**, *20*, 491–495.
[2] F. E. Thomasen, K. Lindorff-Larsen, *Biochem. Soc. Trans.* **2022**, *50*, 541–554.
[3] T. Schneider, K. Sawade, F. Berner, C. Peter, M. Kovermann, *Structure* **2023**, *31*, 1259–1274.e10.
[4] R. Yau, M. Rape, *Nat. Cell Biol.* **2016**, *18*, 579–586.
[5] M. T. Bedford, R. Reed, P. Leder, *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 10602–10607.
[6] T. Gomes, P. Martin-Malpartida, L. Ruiz, E. Aragón, T. N. Cordeiro, M. J. Macias, *Comput. Struct. Biotechnol. J.* **2021**, *19*, 5210–5224.
[7] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, A. Bridgland, C. Meyer, S. A. A. Kohl, A. J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A. W. Senior, K. Kavukcuoglu, P. Kohli, D. Hassabis, *Nature* **2021**, *596*, 583–589.
[8] T. Saldaño, N. Escobedo, J. Marchetti, D. J. Zea, J. Mac Donagh, A. J. Velez Rueda, E. Gonik, A. García Melani, J. Novomisky Nechcoff, M. N. Salas, T. Peters, N. Demitroff, S. Fernandez Alberti, N. Palopoli, M. S. Fornasari, G. Parisi, *Bioinformatics* **2022**, *38*, 2742–2748.
[9] J. Abramson, J. Adler, J. Dunger, R. Evans, T. Green, A. Pritzel, O. Ronneberger, L. Willmore, A. J. Ballard, J. Bambrick, S. W. Bodenstein, D. A. Evans, C.-C. Hung, M. O'Neill, D. Reiman, K. Tunyasuvunakool, Z. Wu, A. Žemgulytė, E. Arvaniti, C. Beattie, O. Bertolli, A. Bridgland, A. Cherepanov, M. Congreve, A. I. Cowen-Rivers, A. Cowie, M. Figurnov, F. B. Fuchs, H. Gladman, R. Jain, Y. A. Khan, C. M. R. Low, K. Perlin, A. Potapenko, P. Savy, S. Singh, A. Stecula, A. Thillaisundaram, C. Tong, S. Yakneen, E. D. Zhong, M. Zielinski, A. Žídek, V. Bapst, P. Kohli, M. Jaderberg, D. Hassabis, J. M. Jumper, *Nature* **2024**, 1–3.
[10] M. A. Dunstone, A. de Marco, *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **2017**, *372*, DOI 10.1098/rstb.2016.0210.
[11] T. Schmidt, M. A. Wälti, J. L. Baber, E. J. Hustedt, G. M. Clore, *Angew. Chem. Int. Ed Engl.* **2016**, *55*, 15905–15909.
[12] R. Ward, A. Bowman, H. El-Mkami, T. Owen-Hughes, D. G. Norman, *J. Am. Chem. Soc.* **2009**, *131*, 1348–1349.
[13] W. L. Hubbell, D. S. Cafiso, C. Altenbach, *Nat. Struct. Biol.* **2000**, *7*, 735–739.
[14] M. A. Starovasnik, N. J. Skelton, M. P. O'Connell, R. F. Kelley, W. J. F. Dorothea Reilly, *Biochemistry* **1996**, *35*, 15558–15569.
[15] P. L. Ramachandran, J. E. Lovett, P. J. Carl, M. Cammarata, J. H. Lee, Y. O. Jung, H. Ihee, C. R. Timmel, J. J. van Thor, *J. Am. Chem. Soc.* **2011**, *133*, 9395-9404.

[16] E. Boura, B. Różycki, H.S. Chung, D. Z. Herrick, B. Canagarajah, D. S. Cafiso, W. A. Eaton, G. Hummer, J. H. Hurley,*Structure* **2012**, *20*, 874–886.
[17] N. Alonso-García, I. García-Rubio, J. A. Manso, R. M. Buey, H. Urien, A. Sonnenberg, G. Jeschke, J. M. de Pereda, *Acta Crystallogr. D Biol. Crystallogr.* **2015**, *71*, 969–985.
[18] L. Galazzo, G. Meier, D. Januliene, K. Parey, D. De Vecchis, B. Striednig, H. Hilbi, L. V. Schäfer, I. Kuprov, A. Moeller, E. Bordignon, M. A. Seeger, *Science Advances* **2022**, *8*, eabn6845.
[19] G. Dorn, C. Gmeiner, T. de Vries, E. Dedic, M. Novakovic, F. F. Damberger, C. Maris, E. Finol, C. P. Sarnowski, J. Kohlbrecher, T. J. Welsh, S. Bolisetty, R. Mezzenga, R. Aebersold, A. Leitner, M. Yulikov, G. Jeschke, F. H.-T. Allain, *Nat Commun* **2023**, *14*, 6429.
[20] J. E. Schiel, A. Turner, T. Mouchahoir, K. Yandrofski, S. Telikepalli, J. King, P. DeRose, D. Ripple, K. Phinney, *Anal. Bioanal. Chem.* **2018**, *410*, 2127–2139.
[21] B. Nilsson, T. Moks, B. Jansson, L. Abrahmsén, A. Elmblad, E. Holmgren, C. Henrichson, T. A. Jones, M. Uhlén, *Protein Eng.* **1987**, *1*, 107–113.
[22] G. Jeschke, *Annu. Rev. Phys. Chem.* **2012**, *63*, 419–446.
[23] L. Fábregas Ibáñez, G. Jeschke, S. Stoll, *Magn Reson (Gott)* **2020**, *1*, 209–224.
[24] L. Fábregas-Ibáñez, G. Jeschke, S. Stoll, *J. Magn. Reson.* **2022**, *339*, 107218.
[25] S. Brandon, A. H. Beth, E. J. Hustedt, *J. Magn. Reson.* **2012**, *218*, 93–104.
[26] T. Schmidt, J. Jeon, W.-M. Yau, C. D. Schwieters, R. Tycko, G. M. Clore, *Proc. Natl. Acad. Sci. U. S. A.* **2022**, *119*, e2122308119.
[27] R. Langen, K. J. Oh, W. L. Hubbell, *Biochemistry* **2000**, *39*, 8396–8405.
[28] K. Sale, L. Song, Y.-S. Liu, E. Perozo, P. Fajer, *J. Am. Chem. Soc.* **2005**, *127*, 9334–9335.
[29] G. Hagelueken, R. Ward, J. H. Naismith, O. Schiemann, *Appl. Magn. Reson.* **2012**, *42*, 377–391.
[30] O. Schiemann, C. A. Heubach, D. Abdullin, K. Ackermann, M. Azarkh, E. G. Bagryanskaya, M. Drescher, B. Endeward, J. H. Freed, L. Galazzo, D. Goldfarb, T. Hett, L. E. Hofer, L. F. Ibáñez, E. J. Hustedt, S. Kucher, I. Kuprov, J. E. Lovett, A. Meyer, S. Ruthstein, S. Saxena, S. Stoll, C. R. Timmel, M. Di Valentin, H. S. Mchaourab, T. F. Prisner, B. E. Bode, E. Bordignon, M. Bennati, G. Jeschke, *J. Am. Chem. Soc.* **2021**, *143*, 17875–17890.
[31] E. R. Canarie, S. M. Jahn, S. Stoll, *J. Phys. Chem. Lett.* **2020**, *11*, 3396–3400.
[32] R. Ward, A. Bowman, E. Sozudogru, H. El-Mkami, T. Owen-Hughes, D. G. Norman, *J. Magn. Reson.* **2010**, *207*, 164–167.
[33] C. Bergonzo, J. T. Hoopes, Z. Kelman, D. T. Gallagher, *J. Biomol. Struct. Dyn.* **2023**, 1–9.
[34] C. Yang, B. He, H. Zhang, X. Wang, Q. Zhang, W. Dai, *Pharmaceutics* **2023**, *15*, 187.
[35] G. Delidakis, J. E. Kim, K. George, G. Georgiou, *Annu. Rev. Biomed. Eng.* **2022**, *24*, 249–274.
[36] S. J. Edwards, C. W. Moth, S. Kim, S. Brandon, Z. Zhou, C. E. Cobb, E. J. Hustedt, A. H. Beth, J. A. Smith, T. P. Lybrand, *J. Phys. Chem. B* **2014**, *118*, 4717–4726.
[37] C. Orlandi, D. Deredge, K. Ray, N. Gohain, W. Tolbert, A. L. DeVico, P. Wintrode, M. Pazgier, G. K. Lewis, *Structure* **2020**, *28*, 516–527.
[38] M. Cai, Y. Huang, R. Yang, R. Craigie, G. M. Clore, *J. Biomol. NMR* **2016**, *66*, 85–91.
[39] G. E., H. C., G. A., D. S., M. R. Wilkins, R. D. Appel, B. A., in *The Proteomics Protocols Handbook, Humana Press (2005). Pp. 571-607* (Ed.: J.M. Walker), Humana Press, **2005**, pp. 571–607.
[40] D. T. Gallagher, C. V. Galvin, I. Karageorgos, *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **2018**, *74*, 524–529.
[41] J. E. Schiel, D. L. Davis, O. V. Borisov, Eds. , *State-of-the-Art and Emerging Technologies for Therapeutic Monoclonal Antibody Characterization Volume 2. Biopharmaceutical Characterization: The NISTmAb Case Study*, American Chemical Society, **2015**.
[42] A. Y. Xu, M. A. Blanco, M. M. Castellanos, C. W. Meuse, K. Mattison, I. Karageorgos, H. W. Hatch, V. K. Shen, J. E. Curtis, *J. Phys. Chem. B* **2023**, *127*, 8344–8357.
[43] M. R. Green, J. Sambrook, *Molecular Cloning: A Laboratory Manual, 4th Ed*, Cold Spring Harbor Laboratory Press, **2012**.
[44] B. Duggan, "Biomolecular NMR tools," can be found under http://sopnmr.ucsd.edu/biomol-tools.htm, **2022**.
[45] Anthis, N.J. & Clore, G.M., *Protein Science* **2013**, *22*, 851–858.
[46] H. S. Mchaourab, M. A. Lietzow, K. Hideg, W. L. Hubbell, *Biochemistry* **1996**, *35*, 7692–7704.
[47] P. Z. Qin, K. Warncke, Eds., in *Methods in Enzymology*, Academic Press, **2015**.
[48] A. Band, M. P. Donohue, B. Epel, S. Madhu, V. A. Szalai, *J. Magn. Reson.* **2018**, *288*, 28–36.
[49] Pannier M, Veit S, Godt A, Jeschke G, Spiess HW, *J. Magn. Reson.* **2000**, *142*, 331–340.
[50] C. E. Tait, S. Stoll, *Phys. Chem. Chem. Phys.* **2016**, *18*, 18470–18485.
[51] Z. Hasanbasri, N. A. Moriglioni, S. Saxena, *Phys. Chem. Chem. Phys.* **2023**, *25*, 13275–13288.
[52] L. Fábregas Ibáñez, Advanced Data Analysis and Modeling in Dipolar EPR Spectroscopy, Doctor of Sciences, ETH Zurich, **2022**.
[53] E. J. Hustedt, F. Marinelli, R. A. Stein, J. D. Faraldo-Gómez, H. S. Mchaourab, *Biophys. J.* **2018**, *115*, 1200–1216.
[54] R. A. Stein, A. H. Beth, E. J. Hustedt, *Methods Enzymol.* **2015**, *563*, 531–567.
[55] H. El Mkami, D. G. Norman, *Methods Enzymol.* **2015**, *564*, 125–152.
[56] M. Huber, M. Lindgren, P. Hammarström, L. G. Mårtensson, U. Carlsson, G. R. Eaton, S. S. Eaton, *Biophys. Chem.* **2001**, *94*, 245–256.
[57] M. Ultsch, A. Braisted, H. R. Maun, C. Eigenbrot, *Protein Eng. Des. Sel.* **2017**, *30*, 619–625.
[58] D. R. Roe, B. R. Brooks, *J. Chem. Phys.* **2020**, *153*, 054123.
[59] D. A. Case, H. Metin Aktulga, K. Belfon, I. Ben-Shalom, J. T. Berryman, S. R. Brozell, D. S. Cerutti, T. E. Cheatham III, G. Andrés Cisneros, V. W. D. Cruzeiro, T. A. Darden, R. E. Duke, G. Giambasu, M. K. Gilson, H. Gohlke, A. W. Goetz, R. Harris, S. Izadi, S. A. Izmailov, K. Kasavajhala, M. C. Kaymak, E. King, A. Kovalenko, T. Kurtzman, T. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, M. Machado, V. Man, M. Manathunga, K. M. Merz, Y. Miao, O. Mikhailovskii, G. Monard, H. Nguyen, K. A. O'Hearn, A. Onufriev, F. Pan, S. Pantano, R. Qi, A. Rahnamoun, D. R. Roe, A. Roitberg, C. Sagui, S. Schott-Verdugo, A. Shajan, J. Shen, C. L. Simmerling, N. R. Skrynnikov, J. Smith, J. Swails, R. C. Walker, J. Wang, J. Wang, H. Wei, R. M. Wolf, X. Wu, Y. Xiong, Y. Xue, D. M. York, S. Zhao, P. A. Kollman, *Amber 2022*, University Of California, San Francisco, **2022**.
[60] D. R. Roe, T. E. Cheatham 3rd, *J. Chem. Theory Comput.* **2013**, *9*, 3084–3095.
[61] C. Tian, K. Kasavajhala, K. A. A. Belfon, L. Raguette, H. Huang, A. N. Migues, J. Bickel, Y. Wang, J. Pincay, Q. Wu, C. Simmerling, *J. Chem. Theory Comput.* **2020**, *16*, 528–552.
[62] K. N. Kirschner, A. B. Yongye, S. M. Tschampel, J. González-Outeiriño, C. R. Daniels, B. L. Foley, R. J. Woods, *J. Comput. Chem.* **2008**, *29*, 622–655.
[63] S. Izadi, R. Anandakrishnan, A. V. Onufriev, *J. Phys. Chem. Lett.* **2014**, *5*, 3863–3871.
[64] Arkajyoti Sengupta, Zhen Li, Lin Frank Song, Pengfei Li, and Kenneth M. Merz Jr., *J. Chem. Inf. Model* **2021**, *61*, 3474–3475.
[65] C. W. Hopkins, S. Le Grand, R. C. Walker, A. E. Roitberg, *J. Chem. Theory Comput.* **2015**, *11*, 1864–1874.

[66] Jean-Paul Ryckaert, Giovanni Ciccotti, Herman J.C Berendsen, *J. Comput. Phys.* **1977**, *23*, 327–341.
[67] R. Salomon-Ferrer, A. W. Götz, D. Poole, S. Le Grand, R. C. Walker, *J. Chem. Theory Comput.* **2013**, *9*, 3878–3888.
[68] R. J. Loncharich, B. R. Brooks, R. W. Pastor, *Biopolymers* **1992**, *32*, 523–535.
[69] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, L. G. Pedersen, *J. Chem. Phys.* **1995**, *103*, 8577–8593.
[70] A. Grishaev, L. Guo, T. Irving, A. Bax, *J. Am. Chem. Soc.* **2010**, *132*, 15484–15486.
[71] M. H. Tessmer, S. Stoll, *PLoS Comput. Biol.* **2023**, *19*, e1010834.
[72] C. D. Schwieters, J. J. Kuszewski, N. Tjandra, G. Marius Clore, *J. Magn. Reson.* **2003**, *160*, 65–73.
[73] M. Teucher, E. Bordignon, *J. Magn. Reson.* **2018**, *296*, 103–111.
[74] J. Sambrook, D. W. Russell, *Molecular Cloning: A Laboratory Manual*, **2001**.
[75] J. T. Hoopes, M. A. Elberson, R. J. Preston, P. T. Reddy, Z. Kelman, *Methods Enzymol.* **2015**, *565*, 27–44.
[76] M. R. Green, J. Sambrook, *Molecular Cloning: A Laboratory Manual*, **2012**.