Structure-Based Experimental Datasets for Benchmarking Protein Simulation Force Fields [Article v1.0]

Chapin E. Cavender[1*], **David A. Case**[2], **Julian C.-H. Chen**[3], **Lillian T. Chong**[4], **Daniel A. Keedy**[5], **Kresten Lindorff-Larsen**[6], **David L. Mobley**[7], **O. H. Samuli Ollila**[8], **Chris Oostenbrink**[9], **Paul Robustelli**[10], **Vincent A. Voelz**[11], **Michael E. Wall**[12], **David C. Wych**[12], **Michael K. Gilson**[1*]

[1]Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California San Diego, La Jolla, CA, USA; [2]Department of Chemistry & Chemical Biology, Rutgers University, Piscataway, NJ, USA; [3]Bioscience Division, Los Alamos National Laboratory, Los Alamos, NM, USA; Department of Chemistry and Biochemistry, The University of Toledo, Toledo, OH, USA; [4]Department of Chemistry, University of Pittsburgh, Pittsburgh, PA, USA; [5]Structural Biology Initiative, CUNY Advanced Science Research Center, New York, NY, USA; Department of Chemistry and Biochemistry, City College of New York, New York, NY, USA; PhD Programs in Biochemistry, Biology, and Chemistry, CUNY Graduate Center, New York, NY, USA; [6]Linderstrøm-Lang Centre for Protein Science, Department of Biology, University of Copenhagen, Copenhagen N, Denmark; [7]Department of Pharmaceutical Sciences, University of California Irvine, Irvine, CA, USA; [8]Institute of Biotechnology, University of Helsinki, Helsinki, Finland; VTT Technical Research Centre of Finland, Espoo, Finland; [9]Institute for Molecular Modeling and Simulation, BOKU University, Vienna, Austria; [10]Department of Chemistry, Dartmouth College, Hanover, NH, USA; [11]Department of Chemistry, Temple University, Philadelphia, PA, USA; [12]Computer, Computational, and Statistical Sciences Division, Los Alamos National Laboratory, Los Alamos, NM, USA; The Center for Nonlinear Studies, Los Alamos National Laboratory, Los Alamos, NM, USA

This LiveCoMS document is maintained online on GitHub at https://github.com/openforcefield/review-protein-benchmark-datasets; to provide feedback, suggestions, or help improve it, please visit the GitHub repository and participate via the issue tracker.

This version dated October 25, 2025

Abstract This review article provides an overview of structurally oriented experimental datasets that can be used to benchmark protein force fields, focusing on data generated by nuclear magnetic resonance (NMR) spectroscopy and room temperature (RT) protein crystallography. We discuss what the observables are, what they tell us about structure and dynamics, what makes them useful for assessing force field accuracy, and how they can be connected to molecular dynamics simulations carried out using the force field one wishes to benchmark. We also touch on best practices for setup and analysis of benchmark simulations. We hope this article will be particularly useful to computational researchers and trainees who develop, benchmark, or use protein force fields or machine learning models that generate protein ensembles.

*For correspondence:
chapin.cavender@openforcefield.org (CEC); mgilson@health.ucsd.edu (MKG)

Contents

It is a truth universally acknowledged that a research group in possession of a good force field must be in want of a benchmark.

1 Introduction

The earliest computer simulations of biomolecules could explore only sub-nanosecond phenomena [1–3], but advances in computing hardware [4, 5] and simulation algorithms [6, 7] now enable the simulation of processes that occur on biologically relevant time scales, enabling informative studies of, for example, protein conformational changes [8, 9], ligand binding [10–12], protein folding [13, 14], and the assembly of multi-protein complexes [15]. As a consequence, simulations have become increasingly useful tools to improve our understanding of protein functions, elucidate molecular mechanisms of human disease, and design small molecule drugs that work by binding targeted proteins. Although knowledge-based heuristic approaches [16] and approaches based on deep learning [17] may be the current methods of choice for predicting the most stable conformation of a protein [18], studying the thermodynamics of phenomena such as conformational changes and ligand binding still requires simulation-based sampling of conformations away from the global energy minimum.

Molecular simulations—such as molecular dynamics and Monte Carlo simulations—are, in effect, importance-weighted sampling methods [19]. They sample conformations from a Boltzmann-weighted ensemble, where the Boltzmann weight is based on a potential function, also known as an energy model. A simulation with a sufficiently accurate potential function will sample protein conformations from a probability distribution function similar to that of the ground truth thermodynamic ensemble. These conformations, or snapshots, make up a simulation trajectory. Physical observables can be computed as Boltzmann averages over a simulation trajectory because most experimental observables are averages over time for many molecules. The more conformations are sampled, the greater the numerical precision of the physical property estimates. Given a sufficiently large number of conformations, the reliability of the predictions is no longer determined by the amount of sampling and depends instead on the accuracy of the energy model and the method used to compute the property of

interest from the trajectory. Not surprisingly, there is a strong tradeoff between precision, which is determined by the amount of sampling done, and accuracy, because more accurate energy models are more computationally costly and thus reduce the amount of sampling that can be achieved. For example, quantum chemical energy models can be highly accurate but remain too slow to provide adequately converged estimates of many properties of interest or to simulate phenomena on biologically relevant time scales. Therefore, the simulation community has a continued interest in far more computationally efficient energy models called force fields. These are parameterized, physics-based energy models that use simple approximations of interatomic interactions and thus can be evaluated quickly, enabling relatively fast conformational sampling.

Force fields specifically for proteins composed of the 20 canonical amino acids date back to the 1980s and 1990s, and their parameters were typically derived from quantum chemical calculations or from the bulk properties of neat liquids chosen as small molecule analogs of protein fragments [20–22]. The quality of these force fields was then evaluated by running simulations of proteins in water and computing the root mean square deviation (RMSD) of the simulated protein's atomic coordinates from the corresponding coordinates of a static crystal structure of the protein. Over time, a growing collection of structural and dynamical data on proteins, generated by X-ray diffraction and nuclear magnetic resonance (NMR) experiments, has enabled more detailed evaluations of the conformational ensembles produced by protein force fields [23–28]. Deficiencies identified in these assessments have motivated further tuning of protein force fields, such as the addition of protein-specific corrections to the force fields' torsional parameters [29–35]. Further information on the history of today's protein force fields is available in prior reviews [36–40].

Although modern protein force fields have been fruitfully applied, they still provide only approximations to the true quantum mechanical energies, and it is not possible to predict *a priori* the consequences of the approximations made for the calculation of quantities of interest. Additionally, the different research groups that develop force fields often prioritize conflicting goals, such as accuracy for the simulation of folded versus disordered proteins or for the prediction of structural versus thermodynamic properties. A given protein force field therefore models some observables better and others less well. Thus, to obtain a well-rounded picture of a force field's strengths and weaknesses, it is essential to benchmark it against a variety of data types. Moreover, publications of new force fields often validate their parameters on different subsets of the available experimental datasets, making it difficult to directly compare a new force field with

established ones [41]. It is therefore important to work toward a common set of accepted benchmarks that will enable force fields to be compared on a common footing.

In this article, we review and describe experimental data that interrogate a range of structural and dynamical features of proteins and therefore are useful to benchmark the accuracy of protein force fields. While not the focus of the article, we also note that many of these datasets could be used to benchmark the increasing number of machine learning models that generate protein conformational ensembles or predictions of protein motions [42, 43]. We focus on experimental observables that provide detailed information about protein conformational ensembles under conditions similar to those of greatest interest for protein simulations, i.e., well-hydrated proteins near room temperature and pressure. We consider only datasets that involve water-soluble (i.e., non-membrane bound) proteins and peptides without ligands or co-factors, because when non-protein molecules (e.g., membrane lipids or drug molecules) are present in the simulation, the results depend on not only the protein force field but also the force field used for these other compounds. That said, since we are interested in proteins in water, the accuracy of the simulations necessarily also depends on the choice of water model. This article is intended for computational researchers who develop, benchmark, or use protein force fields. We thus assume familiarity with molecular dynamics techniques, force field terms, and the basics of protein structure, but not with the experimental techniques used to generate the data.

The remainder of the review is organized into four sections. Section 2 summarizes key points and recommendations. Section 3 describes experiments using nuclear magnetic resonance (NMR) spectroscopy, and Section 4 describes experiments using room-temperature (RT) crystallography. In each of these two sections, we first review the types of observables provided by the experiments, why they are useful for interrogating protein conformational dynamics, and how they can be calculated from simulation trajectories. Then, we describe specific datasets containing measurements of these observables for peptide or protein systems, as summarized in Table 1. Finally, Section 5 discusses best practices for setting up and analyzing benchmark simulations.

2 Overall Recommendations for Choosing Benchmark Datasets

Although early benchmarks of protein force fields compared simulations of solvated proteins to corresponding X-ray structures—i.e., to single structural models from crystal diffraction experiments—more recent benchmarks utilize

experimental observables from solution experiments like NMR. Such experiments match the conditions of dilute aqueous solutions most commonly used for protein simulations, and NMR observables include information about excursions from native structures that contribute to solution-phase ensembles. In contrast, there are fewer studies that evaluate the ability of force fields to reproduce diffraction observables in crystal simulations. Crystal simulations are not applicable to disordered proteins and require larger and longer simulations to converge estimates of observables. In addition, methods of comparing experimental observables with those estimated from simulations are not as well developed. Nonetheless, we have included a section on crystal observables because these provide additional information on the structure, interactions, and dynamics of proteins. As computational throughput continues to increase, we hope that benchmarks targeting crystal observables become more commonplace.

Table 1 summarizes many existing experimental datasets—both NMR and crystallographic—suitable for benchmark studies of protein force fields. We recommend that a minimal benchmark include at least one dataset containing peptides, folded proteins, and disordered proteins—because a generally applicable protein FF should work for all three cases. We further recommend using datasets that target the following observables, as listed in the second column of Table 1:

- chemical shifts
- scalar couplings
- residual dipolar couplings
- electron and nuclear densities from Bragg diffraction

The first three are useful because they have low experimental uncertainty, fast convergence in simulations, and a consensus on best practices for estimating the observables. Although Bragg diffraction lacks such a consensus and converges slowly, it stands out from the other observables considered in this review because it informs on protein tertiary structures while having low experimental uncertainty. See Table 2 for a summary of the salient qualitative features of each observable: whether the experimental uncertainty is high or low, whether the estimate of the observable from simulations converges quickly or slowly, whether the observable primarily provides information about the local structure or tertiary structure of proteins, and whether there is a consensus in the literature on best practices for comparing simulations to experiments.

Convergence always requires sampling the main free energy basins of the protein ensemble, and this can be slow. However, it is still useful to define "quickly" converging properties, which require only a modest amount of sam-

pling within each free energy basin, and "slowly" converging properties, which require extensive sampling within each free energy basin. Additionally, consensus on best practices represents the present authors' opinion about whether studies have converged on a single method for computing observables from simulations. Detailed discussion of convergence behavior and consensus on methodology is provided in the respective sections on each observable below.

The other observables (nuclear Overhauser effect intensities, spin relaxation rates, paramagnetic relaxation enhancements, crystallographic B-factors, crystallographic alternate conformations, and diffuse crystallographic scattering intensities) are less ideal because each exhibits some combination of high experimental uncertainties, slow convergence in simulations, or lack of consensus on best practices for estimating the observable. We therefore recommend that they be deprioritized under circumstances of limited computational resources. Nonetheless, these observables can provide useful alternative characterizations of protein structural ensembles. Note, too, that some can be estimated from the same trajectories used to estimate observables in the first list. For example, simulations used to estimate Bragg diffraction intensities could also be used to estimate B factors. One approach is to perform a two-tier benchmark, in which the first tier uses the first list of observables to assess many candidate force fields, and the second tier uses the second list to assess force fields that perform well on the first tier.

Finally, although one may aspire to compare molecular simulations "directly" with experiment, this is impossible. Experimental observables are raw data like radio frequency signals collected during NMR studies and speckle patterns collected during crystallographic experiments, while simulations give atomic coordinates over time. Thus, models, with their own assumptions and parameters, are needed to bring calculation together with experiment for comparison. For example, computing chemical shifts from molecular simulations requires a model (Section 3.1.1), and a protein crystal structure is a model based on diffraction data (Section 4.1.1). Thus, the agreement between experimental and computed observables should be interpreted with an understanding of the models used in estimating the observable from a simulation and the experimental observable itself from the raw data, as previously emphasized [104, 105].

3 Nuclear magnetic resonance (NMR) spectroscopy

Nuclear magnetic resonance (NMR) spectroscopy measures the responses of nuclear magnetic moments in a strong external magnetic field to perturbations by weak oscillating external magnetic fields tuned to the resonant frequency of

Table 1. Summary of experimental datasets for benchmarking protein force fields. The "Proteins" column counts the number of proteins in the dataset. For crystal datasets, this includes datasets of the same protein crystallized in different space groups.

Dataset	Observables	Description	Proteins
Nuclear magnetic resonance spectroscopy			
Beauchamp short peptides [44]	Chemical shifts, J-couplings	Short, unstructured peptides	31
Designed β-hairpins [45–65] and Trp-cage miniproteins [66–78]	Chemical shifts, NOEs, folding stabilities	Small, folded peptides with mutations	92
Stroet folded proteins [79]	J-couplings, RDCs, NOEs	Folded proteins	13
Mao folded proteins [80]	Chemical shifts, NOEs	Diverse folded proteins	41
Robustelli a99SB-disp [81]	Chemical shifts, J-couplings, RDCs, NOEs, PREs	Folded proteins and disordered proteins	21
Spin relaxation datasets [82–85]	Spin relaxation rates	Folded proteins and disordered proteins	136
Salt bridge stabilities [86]	Salt bridge stabilities derived from chemical shifts	Folded proteins	1
Room-temperature crystallography			
Scorpion toxin [87]	Electron density, B-factors	Folded protein with little secondary structure	1
Hen egg white lysozyme [88–93]	Electron density, B-factors, diffuse scattering	Rigid, folded protein	3
Crambin [94–96]	Electron density, B-factors, neutron scattering	Rigid, folded protein	1
Cyclophilin A [97–99]	Electron density, B-factors, alternate conformations	Folded protein with multiple accessible states and crystal data at different temperatures	1
Ubiquitin [100]	Electron density, B-factors, alternate conformations	Folded protein with mutations	2
PTP1B [101]	Electron density, B-factors, alternate conformations	Folded protein with multiple accessible states and crystal data at different temperatures	1
Endoglucanase [102]	Electron density, B-factors, neutron scattering	Folded protein with measured solvent density	1
Staphylococcal nuclease [103]	Electron density, B-factors, diffuse scattering	Folded protein	1

the nuclei. The observed responses are sensitive to the local magnetic fields at the nuclei, which in turn depend on what other nuclei are nearby. As a consequence, NMR spectroscopy provides information about the local chemical environments of atoms and hence about the conformational distributions and dynamics of the molecules they belong to.

NMR is applicable to nuclei with an odd number of nucleons and hence nonzero nuclear spin—such as ^{1}H, ^{13}C, ^{15}N, ^{19}F, and ^{31}P—because these possess a magnetic dipole moment. In the presence of a strong external magnetic field,

this dipole precesses around the external field, much as a spinning top precesses around the downward gravitational field. The angular frequency of the precession, which is called the Larmor frequency, is proportional to the strength of the magnetic field at the nucleus, with a proportionality constant characteristic of the nuclear isotope. If a second, weak magnetic field that oscillates near the Larmor frequency is applied, the axis of precession will rotate away from the direction of the strong external field. An NMR spectrometer detects the resulting transverse magnetiza-

Table 2. Summary of features of experimental observables.

Observable	Experimental uncertainty	Convergence	Structural information	Consensus on best practices
Nuclear magnetic resonance spectroscopy				
Chemical shifts	Low	Fast	Local	Yes
Scalar couplings	Low	Fast	Local	Yes
Residual dipolar couplings	Low	Fast	Tertiary	Yes
Nuclear Overhauser effect spectroscopy	High	Slow	Tertiary	No
Spin relaxation	High	Slow	Tertiary	No
Paramagnetic relaxation enhancement	High	Slow	Tertiary	No
Room-temperature crystallography				
Bragg diffraction (X-ray and neutron)	Low	Slow	Tertiary	No
B-factors	High	Fast	Local	Yes
Alternate conformations	Low	Slow	Local	No
Diffuse scattering	High	Slow	Tertiary	No

tion by measuring the electrical current it induces in a coil of wire. Although the value of the Larmor frequency is determined chiefly by the strong external magnetic field, the electronic structure near each nucleus, which is influenced by through-bond and through-space interactions, modulates the field felt by the nucleus and hence its Larmor frequency, so NMR probes the local environment of the nucleus. Additionally, magnetization can be transferred to nearby nuclei via through-space interactions, so NMR can report on the relative dispositions in space of pairs of nuclei. These two effects give NMR spectroscopy its sensitivity to the structure and dynamics of proteins. In practice, specific sequences of radiofrequency pulses are crafted to interrogate various aspects of these relaxation phenomena.

NMR observables have several useful features that have led to their adoption as targets for both training and validation of protein force fields [106]. First, NMR experiments are typically performed in laboratory conditions that are similar to the desired setup for most simulation applications, namely dilute aqueous solution. Note, however, that NMR studies are often done at moderately low pH (often around pH 6) to minimize effects from amide proton exchange with solvent. Thus, it is important that simulations meant for comparison against NMR data assign pH-appropriate protonation states of titratable residues and uncapped protein termini. Second, in contrast with typical X-ray or neutron crystallography experiments, NMR spectroscopy can provide useful information about disordered proteins, i.e., proteins that do not fold into a well-defined structure that can crystallize. Furthermore, whereas other methods applicable to disordered proteins, such as small-angle X-ray

scattering, provide only low-resolution structural information, NMR observables can report on structural features that are closely related to specific force field terms, such as the probability distributions of rotatable bonds, which connect closely with the torsional energy terms for a particular dihedral angle. Finally, because NMR observables are averages over ensembles that include deviations from native structures, they can report on the thermodynamic balance among native states, near-native states with local rearrangements, and unfolded states, a perspective not readily available from crystallography. Indeed, while early protein force fields tended to maintain the correct structures of folded proteins, which is favorable, they often underestimated local fluctuations of folded proteins about their mean structures yet overestimated the amount of transient secondary structure in loops or disordered regions of proteins [29–34]. Such deficiencies were identified and largely corrected by comparisons with NMR observables that report on the fluctuations in folded proteins and the amount of residual structure in more disordered regions.

Here, Section 3.1 discusses NMR observables that can be used to assess the accuracy of molecular simulations and hence of the force fields used in simulations. The observables considered are chemical shifts, scalar couplings (also known as J-couplings), residual dipolar couplings (RDCs), the nuclear Overhauser effect (NOE), spin relaxation, and paramagnetic relaxation enhancement (PRE). Section 3.2 then presents available experimental NMR datasets that are well suited for evaluating simulations.

3.1 NMR observables

3.1.1 Chemical shifts

General principles

The chemical shift of a nucleus—the difference of its Larmor frequency from that of the same isotope in a reference compound—probes the degree to which the nuclear spin "feels" the externally imposed magnetic field. This is determined by its electronic environment, which in turn is controlled by the details of the local molecular structure. In a protein, one can consider the chemical shift of a nucleus as having a baseline offset—resulting from its local covalent connectivity and bond hybridization—and an additional shift determined by both its local geometry (bond lengths, bond angles, and dihedral angles) and through-space interactions—resulting from electric fields, hydrogen bonds, and the proximity of chemical groups that contain substantial magnetic anisotropy, such as aromatic rings [107–109]. The sensitivity of a given nucleus to each of these influences depends on its chemical identity (atomic number), covalent structure (bond hybridization), and chemical environment. If ω is the observed Larmor frequency of a nucleus in a particular molecular context, and the reference Larmor frequency for the same isotope is ω_0, the chemical shift, δ, is reported as

$$\delta = \frac{\omega - \omega_0}{\omega_0} \qquad (1)$$

Because ω and ω_0 are similar, δ is typically reported in parts per million (ppm).

Empirical algorithms have been developed to predict the chemical shifts of protein backbone atoms for a given set of three-dimensional coordinates. These algorithms—implemented in software packages that include SHIFTS [110], PROSHIFT [111], Camshift [112], SPARTA+ [113], SHIFTX2/SHIFTX+ [114], PPM/PPM_One [115, 116], UCB-Shift [117], and GraphNMR [118]—are trained on databases of proteins for which both high-resolution X-ray structures and solution backbone NMR chemical shift assignments are available. They take protein coordinates as inputs and output a chemical shift prediction for the backbone nuclei of Cα, Cβ, C′, N, HN, and Hα atoms (Figure1). Each nucleus type has a baseline offset, δ_{random}, determined by its identity, which reflects the chemical shift expected if the protein were in a random coil conformation. This baseline offset is modified by additive terms for specific structural features. The structural features with the largest influence on the backbone nuclei of a given residue are the backbone and side chain dihedral angles of that residue and of neighboring residues within two positions in primary sequence; the distances and orientations of nearby aromatic rings and other chemical groups with substantial magnetic anisotropy; the
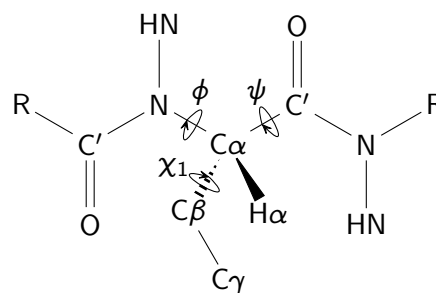

Figure 1. Atom names and named dihedral angles associated with commonly measured chemical shifts and ^3J-couplings in a typical amino acid. Side chain hydrogens such as $H\beta$ are omitted for clarity.

presence and geometry of hydrogen bonds; the proximity of polar and charged nuclei; and the solvent exposure of the residue.

Of these, the main determinants of backbone chemical shifts are local dihedral angles (see Figure 1). The Cα and Cβ shifts of residue i are most sensitive to the ϕ and ψ angles of the same residue, ϕ_i and ψ_i. C′ shifts are most sensitive to the ψ angles of residue i and the following residue i + 1, ψ_i and ψ_{i+1}. N shifts are most sensitive to the χ_1 angle of that residue and the ψ angle of the preceding residue i–1, χ_{1i} and ψ_{i-1}. Proton shifts are more sensitive to non-bonded interactions than are those of carbon and nitrogen atoms. In particular, HN and Hα shifts are very sensitive to the presence of aromatic ring currents, HN shifts are particularly sensitive to hydrogen bond geometries, and Hα shifts are sensitive to electric fields from nearby polar and charged atoms.

The field of empirical protein backbone chemical shift prediction is relatively mature, with a high degree of consensus in the predictions of algorithms published in the last 20 years. Three of the first broadly applicable protein chemical shift prediction algorithms are SHIFTS [110], SHIFTX [109], and PROSHIFT [111]. Of these, SHIFTX generally produces the most accurate predictions on protein structures not contained in its training databases, and it requires only a few seconds to predict all backbone shifts in a protein. Later, the SPARTA algorithm [119] provided a small improvement in prediction accuracy, with somewhat slower calculation times. The program Camshift [112], published in 2009, produced comparable accuracy to SHIFTX and SPARTA. Camshift utilizes polynomials based on interatomic distances that can be evaluated in milliseconds and are differentiable with respect to atomic coordinates, enabling the computationally efficient incorporation of chemical shifts as structural restraints in MD simulations. Taken together, SPARTA, SHIFTX, and Camshift represented important milestones in the field, as their predictions were good enough to enable the prediction of accurate protein structures

using only NMR chemical shifts combined with molecular mechanics force fields or knowledge-based potential energy functions [120–124]. SPARTA, SHIFTX, and Camshift were also found to be sensitive to the conformational fluctuations of proteins observed in MD simulations, and they were utilized to computationally generate protein conformational ensembles that accurately model the dynamics of proteins [125–128] and to guide the optimization of protein force field torsion terms [81, 129]. PPM was parameterized using MD simulations as input (rather than X-ray structures), thus aiming to link conformational ensembles to chemical shifts [115].

A more recent generation of empirical shift predictors—including SPARTA+ [113], SHIFTX2/SHIFTX+ [114], PPM_One [116], UCB-Shift [117, 130], and GraphNMR [118]—were developed using machine learning techniques (although the earlier method PROSHIFT [111] is also based on a neural network). They provide improved accuracy, have very similar accuracy to one another, and typically produce very similar results in practical applications such as the validation of MD ensembles, the fitting of force field corrections, the calculation of protein structures and structural ensembles, and the reweighting of MD trajectories. Despite the accuracy of these methods, their errors are still an order of magnitude larger than the experimental uncertainties in the chemical shifts themselves, so experimental uncertainties can generally be neglected when comparing calculation to experiment.

In addition to empirical prediction algorithms trained on databases of protein structures, other prediction algorithms use quantum mechanical calculations on local fragments of proteins to estimate chemical shifts. Quantum predictors aim to estimate the shielding of the magnetic moment of a nucleus due to its local environment in a particular conformation, and these methods are sensitive to small changes in the coordinates of the input structure. While quantum predictors can achieve good agreement with experimental chemical shifts when averaged over many conformations in a protein's ensemble [131, 132], they typically provide lower accuracy and computational speed than empirical predictors [133–138].

Programs have also been developed to calculate the NMR chemical shifts of protein side chain atoms, but they tend to be less informative than backbone chemical shift predictors [114, 116, 130, 139, 140]. This is because the predictors tend to be less accurate, due to a smaller database of training data, and because the experimental chemical shifts for each atom type tend to have small ranges and can take on similar values for different side chain conformations.

Evaluation of MD simulations using chemical shifts

The chemical shift of a given nucleus in a protein can be estimated from a simulation as a simple average of the predicted chemical shifts over the N_{frame} simulation frames:

$$\langle \delta \rangle = \frac{1}{N_{frame}} \sum_{t=1}^{N_{frame}} \delta_t \tag{2}$$

Here t indexes simulation frames, and δ_t is the chemical shift predicted for simulation frame t using one of the methods discussed above. Deviations between chemical shifts predicted from MD trajectories and experimental shifts are sensitive indicators of errors in MD ensembles. Accordingly, NMR chemical shift predictions are routinely used to assess the accuracy of MD simulations, and torsion terms in protein force fields have been parameterized by optimizing the agreement between experiment and simulation [81, 129].

Backbone chemical shift predictions computed from MD ensembles that have large deviations from experiment—i.e. greater than one- to two-fold higher than the average predictor accuracy for a given backbone atom type (e.g., $C\alpha$) as assessed on a predictor's training databases—indicate substantial inaccuracies in the simulated ensemble. The largest deviations generally reflect incorrect secondary structures and/or side chain rotamers. Smaller deviations can result from inaccurate distributions of tertiary interactions in the MD ensemble, including hydrogen bonds, electrostatic interactions from charged groups, and aromatic ring currents. The latter are particularly important for proton chemical shifts. For disordered proteins, backbone chemical shifts are very sensitive to the presence of partially populated secondary structure elements, and deviations from experiment can indicate that MD ensembles overestimate or underestimate α-helical, PP_{II}, or β-sheet propensities.

When comparing calculated and experimental backbone chemical shifts, one should keep the following considerations in mind.

First, chemical shift prediction algorithms are imperfect, so isolated errors in individual atoms or residues may not imply problems with a simulation. However, a series of residues with several atoms having large deviations between predicted and experimental shifts probably does reflect an error in the conformational distribution of the simulation. The average prediction accuracy on the database of X-ray structures used to train a given predictor provides a reasonable baseline to identify problematic simulations. For example, the standard deviation of $C\alpha$ shift predictions made by SPARTA+ is 0.94 ppm for its training database, so a contiguous stretch of residues with $C\alpha$ prediction errors greater than 1.5 ppm probably indicates a problem with the simulated ensemble, whereas a simulation where the errors

are less than 1.00 ppm is relatively reliable.

Second, as noted above, the chemical shifts of the backbone nuclei (atom types $C\alpha$, $C\beta$, C', N, HN, $H\alpha$) of each residue type have large baseline offsets, δ_{random}, which are determined by chemical identity—i.e., by their covalent structure—rather than by any time-dependent conformational variation that might be sampled in a simulation [141–144]. The random coil shifts of a given backbone atom type can vary more strongly across residue types (e.g., ~50 ppm for $C\beta$ between Ala and Thr [145]) than across residues of a given type (e.g., 3 ppm to 5 ppm for Ala $C\beta$ atoms) due to conformational variations [146, 147]. As a result, it can be highly misleading to visualize correlations and report correlation coefficients between experimental chemical shifts and shifts predicted from MD simulations, as most of the magnitude in deviation between different residues can be explained by differences in their baseline δ_{random} values. Indeed, extremely high correlation coefficients (close to 1.0) can be obtained by comparing experimental chemical shifts to database random coil chemical shift values without utilizing any structural information [112–114, 142]. A more informative visualization of the accuracy of chemical shift predictions from an MD simulation is provided by comparing the deviations between predicted and experimental shifts separately for each residue type. Alternatively, one can compare predicted and experimental "secondary" chemical shifts, which subtract the random coil chemical shift values from both experimental and predicted shifts and thus put all residues on comparable footing.

Third, empirical chemical shift predictions depend on many structural features, so one cannot be sure what conformational error causes a given prediction error. For example, a deviation of 1.5 ppm for an N atom may result from a side chain populating an incorrect rotamer or an aromatic group being incorrectly positioned.

Fourth, chemical shift prediction error is sequence- and conformation-specific and so should not be used to compare the accuracy of simulations of two different proteins or of two different regions of one protein. For example, chemical shift prediction errors for β-sheet proteins are substantially higher than for α-helical proteins [118], so a simulation of a β-sheet protein may yield worse agreement with experimental chemical shifts than an equally accurate simulation of an α-helical protein. In contrast, when one protein is simulated with multiple force fields, the accuracy of the chemical shift predictions is a clear indication of the relative accuracy of the simulations and hence of the force fields. Note that scalar couplings do not have this limitation, as detailed in Section 3.1.2.

Fifth, in principle, chemical shift predictions computed from accurate, thermalized conformational ensembles should be more accurate than predictions computed from static protein structures. However, the most accurate chemical shift prediction algorithms are trained against static X-ray structures (with PPM being one exception [115]). As a result, some amount of conformational averaging is "baked in" to these prediction algorithms; that is, chemical shifts predicted from static conformations already implicitly account for the solution-phase averaging over conformations. This fact may help account for the fact that chemical shifts predicted based on X-ray structures are often (though not always) more accurate than predictions from MD ensembles [125–128]. However, the lower accuracy obtained with simulations does not necessarily point to inaccuracies of simulations. Indeed, MD simulations that yield excellent agreement with other NMR observables (such as NMR scalar couplings or residual dipolar couplings; see below), which are thus presumably quite accurate, may produce less accurate chemical shift predictions than static high-resolution X-ray structures [81, 125, 127].

Sixth, comparisons between simulated and experimental chemical shifts are particularly informative regarding the accuracy of simulations of disordered proteins and peptides. The lack of stable tertiary structure in such systems means that through-space interactions are largely washed out so that chemical shifts are dominated by backbone dihedral angles, which are the most accurately parameterized and least noisy relationships in backbone chemical shift predictors [148]. As a result, chemical shifts computed for disordered proteins can be more accurate than chemical shifts computed for training databases of X-ray structures of folded proteins. For example, C' chemical shift predictions obtained from simulations of disordered proteins and peptides that agree with orthogonal NMR data or with circular dichroism data frequently have RMSDs from experiment less than 0.5 ppm, whereas C' prediction errors for folded proteins average ~1 ppm. Indeed, simulations of disordered proteins with state-of-the-art force fields regularly achieve prediction RMSDs a factor of 2 lower than the average predictor errors observed on databases of folded proteins, with a substantial dynamic range such that agreement with chemical shifts correlates well with the agreement with orthogonal experimental data [81, 128, 148, 149]. Therefore, when several different force fields yield predicted chemical shifts for disordered proteins with RMSDs lower than typically observed for folded proteins, differences across the force fields being tested should not be dismissed as being within prediction error but instead should be regarded as meaningful and informative.

Seventh, there is not a one-to-one mapping between conformational distributions and chemical shifts [150]. As a consequence, many different conformational ensembles can

produce equivalent agreement with experiment [151–153]. For example, the same C′ chemical shift prediction can be obtained from many different ϕ/ψ dihedral distributions; and two ensembles with extremely different ϕ/ψ dihedral distributions for a given residue can produce identical chemical shift predictions for a Cα atom if, for example, it is exposed to an aromatic ring current in one ensemble and not the other. The degeneracy between conformations and chemical shifts can often be resolved by analyzing the chemical shifts of multiple nuclei.

Eighth, different programs for predicting chemical shifts tend to yield very similar results, so the choice of method is not expected to have much effect on conclusions. However, some chemical shift predictors utilize sequence homology for part or all of their predictions, and a prediction based on sequence homology rather than three-dimensional structure *ipso facto* does not report on the accuracy of a structure. For example, the SHIFTX2 predictor is a weighted average between a structure-based predictor, SHIFTX+, and a homology-based predictor, SHIFTY. It is recommended to use only structure-based predictors for validation of MD simulations and to use the same method consistently across the trajectories being compared.

Evaluation of force fields via structural properties derived from chemical shifts

Measured chemical shifts can be used to derive structural information, such as the helicity of a peptide in solution. Simulations can then be evaluated based on their agreement with these experimentally derived structural data. This indirect approach avoids the complications of predicting chemical shifts from simulations, instead relying on the availability of reliable methods to map from chemical shifts to structures. While we generally recommend direct comparison between observables computed from simulations and those measured experimentally, indirect structural comparisons can be more interpretable and can help guide improvements in parameters. Also, as noted in the final paragraph of section 2, even measured observables like chemical shifts rely on interpretation to obtain the observables from raw experimental data, so there may be no real escape from the need to use interpretation, which may be imperfect. Two examples of structural models that can be used in this way, helical propensities and stabilities of salt bridges and salt bridge analogs, are now considered.

Accurate protein force fields should be able to model the preferences for proteins to adopt particular secondary structures. Helical propensity in a particular sequence context is often modeled using the expected fraction of time that a particular non-terminal residue adopts an α-helical backbone conformation, and the fractional helicity of a residue can be measured because α-helical residues form backbone hydrogen bonds that alter the chemical shift of the C′ carbons in ^{13}C-labeled proteins [154]. We note here that this commonly used dataset was measured in D$_2$O as a solvent. Since the helicity will likely be different in H$_2$O and D$_2$O [155], one should not expect perfect agreement in simulations with water models that represent H$_2$O. Assuming a two-state helix-coil transition, the helical fraction can be calculated using

$$f_{\text{helix}} = \frac{\delta_{\text{obs}} - \delta_{\text{coil}}}{\delta_{\text{helix}} - \delta_{\text{coil}}} \qquad (3)$$

where δ_{obs}, δ_{helix}, and δ_{coil}, are, respectively, the C′ chemical shifts observed in the experiment, in the reference helical state, and in the reference coil state. Meanwhile, a simulation can be analyzed to provide the fraction of time each residue is in a helical conformation, based on its backbone dihedral angles or hydrogen bond occupancies, and the results can be compared with the results inferred from chemical shifts [156]. Alternatively, researchers can fit simulated conformations to a helix-coil transition model such as the Lifson-Roig model [157]. A commonly used variant of the Lifson-Roig model expresses the partition function of a two-state helix-coil system in terms of two parameters: a nucleation parameter describing the likelihood for a residue to sample the helical dihedral angles in the coil state and an extension parameter describing the likelihood for a neighboring residue at the end of a helix to transition from the coil to the helix state. These two parameters can be allowed to take different values for each residue type, and the resulting set of parameters can be fit to the fraction of helix from the simulation and then compared to parameters derived from experimental data [156].

Another important characteristic of protein force fields is the ability to accurately model the formation of salt bridges—pairs of amino acids whose oppositely charged side chains are within hydrogen bonding distance [158]. When a salt bridge is formed, the presence of the anionic side chain alters the chemical shift of nitrogen in the cationic side chain, and this perturbation can be measured in ^{15}N-labeled proteins [86] or in small molecule analogs of these side chains. Similar to the helical fraction, the fraction of salt bridge formation can be calculated from the chemical shifts observed in the experiment, along with reference shifts measured in the presence and absence of the salt bridge. Simulations can then be analyzed to provide the fraction of time the salt bridge is present based on geometric criteria for hydrogen bonding between the charged side chains.

3.1.2 Scalar couplings

General principles

NMR scalar couplings, also known as indirect couplings or J-couplings, are electron-mediated, spin-spin couplings that act particularly strongly through covalent bonds and hence are often described as through-bond couplings [159, 160]. Scalar couples occur when the magnetic field of a nuclear spin modifies the behavior of nearby electrons, and this modification propagates to nearby nuclei, resulting in an indirect nucleus-nucleus coupling. Scalar couplings can be observed in an NMR spectrum as the splitting of the resonance peak of one nucleus into two (sometimes overlapping) peaks. The value of the scalar coupling, also called the coupling constant, is the distance between the split peaks in the frequency domain and thus has units of reciprocal time (e.g., Hz). These couplings can propagate through multiple bonds, and three-bond couplings are of particular interest because they provide a readout of the dihedral angle of the central bond. Accordingly, three-bond couplings, termed ^3J-coupling constants, have been utilized extensively in the conformational analysis of small molecules [160] and peptides [161] since the discovery of the Karplus relationship [159], or Karplus equation, which relates the ^3J-coupling constant between two nuclear spins to the intervening dihedral angle θ:

$$^3J(\theta) = A\cos^2(\theta) + B\cos(\theta) + C \qquad (4)$$

Here the Karplus coefficients—A, B, and C—depend on the identities of the nuclei involved in the coupling (e.g., ^{13}C or ^1H) and their local chemical environments, including bond hybridizations, bond lengths, bond angles, and the electronegativities of nearby substituents [162]. There are also one-bond ^1J and two-bond ^2J scalar couplings, which can have more complex Karplus relationships that depend on more than one dihedral angle [163–167].

Karplus equation coefficients used in conformational analyses are generally empirically determined, based on measurements of J-coupling constants in molecules of known structure, and then transferred to analyze the dihedral angles between nuclei in similar chemical environments [159, 160, 162, 168, 169]. For a given pair of coupled nuclei, the experimentally observed J-coupling is the probability-weighted average of their instantaneous coupling over the molecule's thermalized conformational distribution.

In proteins and peptides, the ^3J-couplings between backbone amide protons and alpha protons ($^3J_{HN-H\alpha}$) report on the ϕ dihedral angle of the peptide backbone (Fig.1) and thus distinguish between α and β secondary structure. These couplings were adopted as structural restraints in early NMR protein structure calculations [170]. The values

of the Karplus coefficients for these ^3J-coupling constants have been the subject of frequent reexamination and scrutiny [169, 171–176], including studies that examine the consequences of harmonic motion and conformational dynamics [169, 171, 172, 175, 177]. Although $^3J_{HN-H\alpha}$ are the most frequently measured and reported ^3J-couplings for the protein backbone, five additional coupling constants also report on the ϕ dihedral angle [178]: $^3J_{HN-C'}$, $^3J_{HN-C\beta}$, $^3J_{C'(i-1)-H\alpha}$, $^3J_{C'(i-1)-C'}$, and $^3J_{C'(i-1)-C\beta}$. It has also been shown that one-bond ^1J-couplings, such as $^1J_{C\alpha-H\alpha}$ and $^1J_{C\alpha-C\beta}$, are sensitive to the ψ angle of protein backbones [163, 164, 167]. The χ_1 angles of protein side chains can be analyzed via the following ^3J-couplings [179, 180]: $^3J_{H\alpha-H\beta}$, $^3J_{N-H\beta}$, $^3J_{C'-H\beta}$, $^3J_{H\alpha-C\gamma}$, $^3J_{N-C\gamma}$, and $^3J_{C'-C\gamma}$. Additionally, $^3J_{HN-C'}$ hydrogen bond scalar couplings (scalar couplings between protein backbone nitrogen and carbonyl atoms in different residues that are mediated through hydrogen bonds) provide quantitative information about hydrogen bond geometries [181]. Once Karplus coefficients are known, no specialized software is required to compute scalar couplings for a protein structure. One simply needs to calculate the dihedral angles of interest and use them to predict the corresponding scalar couplings via the Karplus relationship applied with the appropriate Karplus coefficients. Special attention, however, is needed to ensure correct mapping to the correct atoms in cases when stereospecific assignments are available so that, for example, the same definition of the atoms is used in experiments and simulations.

Evaluation of force fields using scalar couplings

Observed J-coupling constants can be estimated from the snapshots of an MD trajectory as

$$\langle J(\theta) \rangle = \frac{1}{N_{\text{frame}}} \sum_{t=1}^{N_{\text{frame}}} J(\theta_t) \qquad (5)$$

where t indexes the N_{frame} simulation frames and $J(\theta)$ is given by a Karplus relationship (Eq. 4). This mean can be compared with the corresponding measured J-coupling. One may therefore use J-coupling data to test [33, 44, 182–184] and parameterize [30, 33, 81, 156, 185, 186] simulation force fields. In such calculations, one must choose between "static" and "dynamics-corrected" Karplus coefficients. Static coefficients are obtained from empirical fits of ensemble-averaged solution data to high-resolution X-ray structures and therefore do not explicitly account for the complexities of conformational distributions during parametrization, analogous to the "baked in" dynamics of many chemical shift predictors (Section 3.1.1). Dynamics-corrected Karplus coefficients are obtained from empirical fits that seek to account explicitly for the distributions of dihedral angles of the molecule in solution using a variety of approaches,

including single- and multiple-well harmonic motion models as well as fitting of coefficients to rotamer population distributions obtained from other types of experimental data [169, 171, 172, 175, 177, 180].

A recent study [81] examined the accuracy of ^3J-couplings computed with long MD simulations using several sets of Karplus coefficients and seven different force fields for a large set of NMR data spanning folded and disordered proteins. The static $^3J_{HN-H\alpha}$ Karplus coefficients from Vogeli et al. [169] and the static $^3J_{C'(i-1)-C'}$ Karplus coefficients from Li et al. [176] produced the lowest RMSD values from experiment on average across force fields and protein systems. The dynamics-corrected Karplus coefficients from Lee et al. [175] gave similar trends in accuracy across force fields but larger average deviations from the experimental couplings. Karplus coefficients could be derived not empirically but from quantum calculations for specific conformations, thus avoiding the question of conformational averaging. Karplus coefficients derived in this way have similar values as dynamics-corrected empirical Karplus coefficients [169, 171, 172, 175, 177, 180].

When interpreting the NMR scalar couplings calculated from MD simulations, it is essential to consider the uncertainties in the Karplus coefficients. This is frequently done using a χ^2 value [81, 186]:

$$\chi^2 = \frac{1}{N_{obs}} \sum_{i=1}^{N_{obs}} \frac{\left(\langle J_i(\theta)\rangle - J_{i,exp}\right)^2}{\sigma_i^2} \qquad (6)$$

where i indexes the N_{obs} observables, $\langle J_i(\theta)\rangle$ and $J_{i,exp}$ are the computed (Eq. 5) and experimental J-coupling constants, respectively, and σ_i is the RMSD between predicted and measured scalar couplings obtained in fitting of the Karplus parameters. A χ^2 value less than 1.0 indicates agreement within the estimated uncertainty. It is also worth recalling the words of Martin Karplus [160]: "The person who attempts to estimate dihedral angles to an accuracy of one or two degrees does so at ...[their] own peril."

3.1.3   Residual dipolar couplings
General principles
The direct interaction of the magnetic dipoles associated with two nuclear spins leads to an experimentally measurable coupling, known as a dipolar coupling. Similar to scalar couplings, the dipolar coupling can be observed as the splitting of a peak in the frequency domain and thus has units of reciprocal time (Hz). The magnitude of the dipolar coupling between the nuclear spins of atoms A and B for an instantaneous configuration and orientation of the molecule is given by

$$D_{AB} = \frac{D_0}{R_{AB}^3} \left(3\cos^2\theta_{AB} - 1\right) = \frac{D_0}{R_{AB}^3} \hat{r}_{AB}^T \mathcal{A} \hat{r}_{AB} \qquad (7)$$

where D_0 is a constant that depends on the identities of the nuclei, R_{AB} is the distance between the nuclei, and θ_{AB} is the angle between the magnetic field imposed by the NMR instrument and the vector connecting atoms A and B [187–189]. The second equality expresses the quantity in parentheses in terms of $\hat{r}_{AB}$, the unit vector joining the two nuclei, defined in the internal coordinates of the molecule, and the alignment tensor $\mathcal{A}$, which relates an internal coordinate system to the lab-frame magnetic field.

In a solution of freely tumbling molecules, D_{AB} takes on all possible values with equal probability, and the observed dipolar couplings—which are averages over molecules and time—are zero. However, if the molecules can be even weakly aligned relative to the instrument's magnetic field, then D_{AB} no longer averages to zero, and dipolar couplings can report on the structure and conformational dynamics of proteins. Such alignment may be achieved by linking the protein to a prosthetic group that tends to align with the field; by placing it in an aqueous liquid crystal formed by, for example, bicelles [190, 191]; or by anisotropic compression of acrylamide gels [192]. The dipolar coupling measured in a weakly aligned sample is called a residual dipolar coupling (RDC). RDCs are often measured between atoms that are directly bonded. In proteins, these are often an amide proton and nitrogen, resulting in a $^1D_{NH}$ RDC, but it is also possible to measure RDCs between nuclei that are not directly bonded. Although the alignment procedure could in principle perturb the protein's conformational ensemble, consistency across different methods for alignment suggests this is not a substantial concern [193].

Evaluation of MD simulations using residual dipolar couplings
To use protein RDCs to benchmark a simulation, one must estimate the alignment tensor, $\mathcal{A}$. For folded proteins, it is often reasonable to assume that the internal motions of the protein and the alignment are mostly decoupled or that any coupling does not contribute substantially to the RDCs [83, 194]. In this case, one may keep the concept of an alignment tensor, noting that this should be fitted over the full ensemble rather than using a single structure [195, 196]. The procedure and equations used in this fitting are the same as for rigid proteins and typically rely on alignment of the MD frames, followed by singular value decomposition (SVD) of the unit vectors of the bonds of interest. With a set of experimental RDCs and a simulation, one can fit the five independent parameters of the alignment tensor. In contrast, for unfolded proteins, flexible peptides, and intrinsically disor-

dered proteins, it is not possible to fit an average alignment from the data because the alignment varies across conformations. Therefore, most analyses use a physical model to predict the alignment tensor for each conformation (rather than to fit it from the experiments) and use this to calculate per-frame RDCs that can then be averaged. These physical models can be tested against data for folded proteins, where the alignment can be determined from experiment using an SVD fitting procedure. Such tests suggest that the predicted alignments are accurate enough to be useful, in particular when alignment is dominated by steric interactions between the macromolecule and the alignment medium. Among the different methods for predicting alignment tensors, the PALES software is probably the most commonly used [197]. While it is possible to predict alignment for the full chain, sometimes local alignment over short stretches of ca. 15 residues is used instead [198].

RDCs have been used extensively to benchmark protein force fields [29, 81, 184, 199], as they offer a number of advantages. In particular, they can be measured very precisely [189], they are averaged over long time scales and thus report on conformational ensembles [193], and they can be measured for multiple bonds in both the backbone and side chains [189]. Additionally, RDCs for atoms that are not directly bonded can be measured at long range, because the RDC falls off with only the third power of the interatomic distance (Eq. 7). By contrast, the nuclear Overhauser effect (Section 3.1.4) falls off with the sixth power of the interatomic distance.

Because the magnitudes of the RDCs depend on the alignment strength [200, 201], the quality of agreement between experiments and simulations is typically evaluated in terms of a normalized RMSD (the quality factor, Q [202]), with high-resolution crystal structures typically having $Q < 25\%$ [188].

3.1.4 Nuclear Overhauser effect spectroscopy
General principles
The nuclear Overhauser effect (NOE) is the change in intensity of the resonance peak of one nucleus that occurs when the resonance of a nearby nucleus is saturated by radio frequency irradiation. Saturation of a nucleus greatly perturbs its magnetization away from its equilibrium value, and the interaction between the magnetic dipoles of two nuclei nearby in space allows the system to relax toward equilibrium via a concerted flip of both nuclear spins. This concerted flip is known as dipolar cross relaxation and requires an exchange of energy between the two-spin system and nearby atoms. In NMR spectroscopy, the environment surrounding a system of spins is called the lattice. (However, this term does not imply a repeating arrangement of atoms related by translational symmetry as in a crystal lattice.)
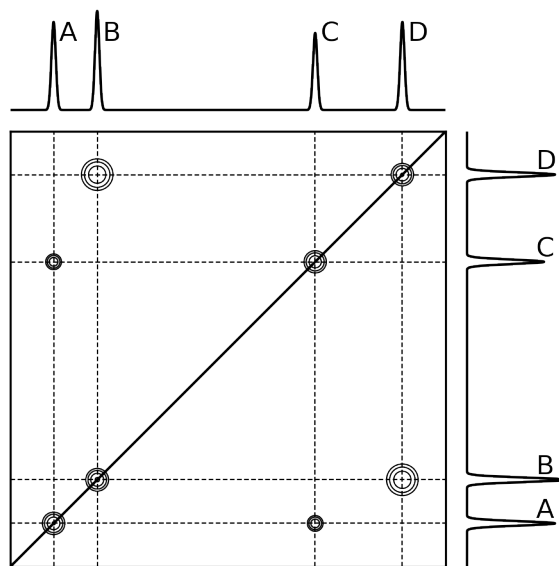

Figure 2. Idealized NOESY spectrum. The one-dimensional spectrum displayed along the top and right contains peaks for four hydrogen nuclei, labeled "A" through "D". Apart from the self-peaks along the diagonal, there are NOE cross peaks between two pairs of nuclei. The volumes of the cross peaks indicate a weak NOE between A and C and a strong NOE between B and D nuclei. The lack of cross peaks between A and D and between B and C suggests no significant NOE between these pairs of nuclei.

Thus, dipolar cross relaxation involves coupling to the motions of the lattice. The NOE is a through-space, rather than through-bond, effect that is typically detectable only when the mean internuclear distance is less than 6 Å.

NOEs for proteins are typically measured using a two-dimensional nuclear Overhauser effect spectroscopy (NOESY) experiment, in which one dimension corresponds to the chemical shift of the saturated nucleus, δ_A, and the second dimension corresponds to the chemical shift of the observed nucleus, δ_B. The intensity of the NOE between nuclei A and B is measured as the integral of the cross peak located at (δ_A, δ_B) in the two-dimensional NOESY spectrum (Fig. 2). This integral is proportional to the time elapsed between the saturation and observation pulses, i.e., the mixing time.

NOESY experiments for proteins report on the average, or effective, distances between pairs of nuclei, where the averages are taken over the mixing time of the NOESY experiment, and the form of the average is determined by the time scales of intramolecular vs. overall motions of the protein [203, 204]. In the limit where intramolecular motions are much slower than molecular tumbling, the effective distance between nuclei A and B, R_{AB}, takes the form

$$R_{AB}^{\text{slow}} = \left\langle \frac{1}{R_{AB}^6} \right\rangle^{-1/6} \tag{8}$$

In the limit where variations in the orientations of internuclear displacements are small and where intramolecular motions are much faster than molecular tumbling, the average becomes

$$R_{AB}^{\text{fast}} = \left\langle \frac{1}{R_{AB}^3} \right\rangle^{-1/3} \tag{9}$$

Appendix A provides additional information about the precise meanings of these assumptions and how they give rise to the corresponding effective distances, and more detailed discussions of this topic are available in the literature [203, 204].

The previous discussion has assumed that the cross peak intensities in a NOESY experiment result purely from direct dipolar cross relaxation between the saturated and observed spins. However, observed NOESY intensities also include contributions from indirect dipolar cross relaxation events mediated by additional spins nearby in space. For example, after saturation of spin A, the cross-peak intensity between spins A and C will include not only a direct contribution but also an indirect contribution due to dipolar cross relaxation between spins A and B followed by dipolar cross relaxation between spins B and C. This indirect relaxation, called spin diffusion, can reduce the internuclear distances inferred from NOEs. Although spin diffusion is non-negligible in most protein systems, it is slower than direct relaxation, so it contributes less to NOESY spectra with shorter mixing times, making such experiments more straightforward to interpret and simpler to model with MD simulations.

Evaluation of MD simulations using NOESY

The literature describes many approaches for assessing the accuracy of a protein ensemble against a NOESY experiment. Most can be grouped into three broad categories, according to the level of detail used in modeling the dependence of the NOEs on intramolecular motions, as now discussed.

The first approach treats the dependence on intramolecular motions implicitly by assigning the experimental NOESY cross peak intensities to categories associated with upper boundaries on the corresponding effective internuclear distances—e.g., 2.5 Å for strong, 3.0 Å for medium, and 4.5 Å for weak NOEs [205]. Appropriate ensemble-averaged effective distances, i.e., R_{AB}^{slow} (Eq. 8) or R_{AB}^{fast} (Eq. 9), can be computed from a simulation and compared to these upper boundaries, and force fields can be assessed by counting violations of these boundaries. The choice of which distance average to use depends on assumptions about the time scales of intramolecular motions relative to overall molecular tumbling. For typical globular proteins, time constants

for molecular tumbling are on the order of 10 ns to 100 ns, and intramolecular motions are assumed to occur on faster time scales, so that R_{AB}^{fast} is an appropriate choice of effective distance. For short, unstructured peptides, time constants for molecular tumbling are much faster, and intramolecular motions are assumed to occur on slower time scales, so that R_{AB}^{slow} is an appropriate choice of effective distance. For intrinsically disordered proteins and for folded proteins that undergo slow, large-scale conformational fluctuations, it is likely that neither assumption is valid, and a different approach for evaluating NOESY intensities should be used (see below). Note that this first approach ignores the effects of spin diffusion and therefore is most appropriate for comparison against NOESY spectra collected at short mixing times. Sometimes experimental measurements are reported only as upper boundaries on distances rather than as direct NOESY intensities, in which case this approach is the only one available. The experimental uncertainty in this approach is mostly due to the discretization of the measured intensities into rough distance restraints, which should be kept in mind when interpreting the number and size of any violations between simulated and experimental effective distances.

In the second approach, effective distances are calculated from the experimental NOESY cross peak intensities by assuming that the cross peak intensities V_{AB} are proportional to the reciprocal sixth power of the effective distance.

$$R_{AB}^{\text{eff}} = \left(\frac{\kappa}{V_{AB}} \right)^{1/6} \tag{10}$$

Here, the value of κ is determined from the cross peak intensities for nuclei with known, rigidly constrained distances, such as vicinal protons in aromatic rings. Experimental effective distances for mobile proton pairs can be computed from the derived value of the κ and the measured cross peak intensities of these pairs. Then, an ensemble-averaged effective distance is computed from an MD trajectory, and force fields can be assessed using an RMSE between the effective distances calculated from the simulations and from the experimental cross peak intensities. The choice of whether to use R_{AB}^{slow} or R_{AB}^{fast} to calculate the effective distances from the simulations relies on the same considerations as those described in the previous approach. This approach, too, is less appropriate for intrinsically disordered proteins and proteins that undergo large conformational changes, as well as for NOESY spectra collected at long mixing times, as these can have non-negligible contributions from spin diffusion.

The third class of approaches uses more detailed models for the dependence of the NOE on intramolecular motions [204, 206] and can be applied to compute the dipolar cross relaxation rates and NOESY cross peak intensities di-

rectly (Eq. 34 and Appendix B). Such models may also attempt to account for spin diffusion or for anisotropic molecular tumbling.

In all three approaches, the experimental study may not distinguish between individual hydrogen atoms, either because they are chemically indistinguishable (e.g., the three protons of a methyl group) or because stereospecific hydrogen atoms could not be assigned (e.g. the β-hydrogens in an amino acid side chain). These cases may either be omitted from the evaluation or addressed by pseudoatom corrections such as the "center average" approach [207, 208].

3.1.5   Spin relaxation
General principles
A radiofrequency pulse at a nucleus's resonant frequency tips the spin away from the axis of the external field around which it is precessing, increasing the amplitude of its precession. It also makes the spins of multiple nuclei precess coherently. Spin relaxation is the decay of a nucleus's spin magnetization back to its equilibrium distribution following such a pulse. This process is typically characterized by two independent exponential decay processes. Spin-lattice relaxation, or T_1 relaxation, is the decay of the component of the nuclear spin magnetization parallel to the external magnetic field, conventionally defined as the z axis. This decay is caused by the interaction of the aligned spins with their surroundings (often called the lattice despite not being ordered), i.e., an exchange of energy between the spin and the lattice. This longitudinal component of the magnetization, $M_z(t)$, decays according to

$$M_z(t) = M_{z,\text{eq}} - [M_{z,\text{eq}} - M_z(0)] \exp\left(-\frac{t}{T_1}\right) \quad (11)$$

where $M_{z,\text{eq}}$ is the longitudinal component of the magnetization at thermal equilibrium, time $t = 0$ corresponds to the end of the pulse, and T_1 is the decay constant for the spin-lattice relaxation. Spin-spin relaxation, or T_2 relaxation, is the decay of the net magnetization transverse to the external magnetic field due to dwindling coherence of the phases of the spins of individual nuclei. The transverse component of the magnetization, $M_{xy}(t)$, decays according to

$$M_{xy}(t) = M_{xy}(0) \exp\left(-\frac{t}{T_2}\right) \quad (12)$$

where T_2 is the decay constant for the spin-spin relaxation. For most systems, spin-spin relaxation is faster than spin-lattice relaxation; i.e., $T_1 > T_2$. These decay constants can be measured for NMR-active isotopes in proteins, typically ^{15}N in labeled backbone amides or ^{13}C- and ^{2}H-labeled side chains. Although protein force fields are usually benchmarked against the spin relaxation times of backbone nuclei, it is also possible to use side chain data [209], though this

is generated less often than backbone data. In addition, heteronuclear NOE (hetNOE) relaxation, which is typically sensitive to fast motions, can be measured.

The Redfield equations [210] (see Appendix B) demonstrate that spin relaxation times are related to molecular dynamics via rotational correlation functions of bonds in a complex and magnetic field-dependent manner whose intuitive interpretation is often not straightforward. To simplify this, spin relaxation analyses [211–215] typically employ the Lipari-Szabo or model-free approach [216, 217], where rotations of bonds are assumed to be dominated by two independent types of motion: overall molecular tumbling and intramolecular motions in the molecular frame. This analysis gives the time scales of protein overall and intramolecular motions, as well as order parameters, S^2, characterizing the orientational freedom of bond vectors in the protein frame of reference. These order parameters range between 0 and 1 [217], where 0 implies free, isotropic rotation of the bond vector within the molecular frame of reference, and 1 implies a complete absence of such freedom. Experimental values for the order parameters can be derived for each residue by fitting the value of the order parameter and the time constant for internal motion, along with one additional time constant for molecular tumbling, to the observed spin relaxation rates and NOE enhancements for all residues. The time constant for molecular tumbling can also be separately fit to the T_1/T_2 ratios for all residues in the protein, which are independent of the residue-specific order parameters and time constants for internal motion as long as the internal motions are fast relative to molecular tumbling, an assumption that is typically valid for globular proteins [152].

Alternatively, instead of estimating both T_1 and T_2 values from order parameters using the Lipari-Szabo model, as just discussed, one may estimate T_2 from a linear transformation of the effective correlation time for molecular tumbling, obtained from a multi-exponential decay describing the multiple time scales for protein backbone rotational motions [218]. This can facilitate rapid interpretation of molecular dynamics from spin relaxation times and their comparison with MD simulations.

Evaluation of MD simulations using spin relaxation
Order parameters and time scales of internal motions computed from MD simulations of folded proteins have often been compared with values extracted from experiments using the Lipari-Szabo model [211–215, 219, 220]. It is worth noting that this approach cannot be used for proteins with disordered regions because their overall motions cannot be uniquely defined, and their rotational dynamics may not be modeled by a single time scale as

15 of 47

https://doi.org/10.33011/livecoms.6.1.3871

Living J. Comp. Mol. Sci. 2025, 6(1), 3871

assumed in the Lipari-Szabo model. Nevertheless, spin relaxation times for such systems can be directly calculated from MD simulations by substituting Fourier transformations of multiple rotational correlation functions to the Redfield equations and then compared with experiments [152, 221–223]. Several codes with slightly different implementation details and varying levels of documentation are available for this: https://github.com/zharmad/SpinRelax [221], https://github.com/ohsOllila/ProteinDynamics [222], and https://github.com/nencini/NMR_FF_tools/blob/master/relaxation_times/RelaxationTimes.ipynb [224]. Additionally, https://github.com/fahoffmann/MethylRelax [223] provides code for the calculation of spin relaxation rates for methyl groups in protein side chains.

Sufficiently long simulations [225] can reproduce experimental spin relaxation times to good accuracy without any further corrections [152], given that the simulations use a water model that gives an accurate viscosity of water. In particular, the TIP3P water model [226] underestimates the viscosity significantly and therefore leads to unrealistically short time constants for overall protein rotation (tumbling) [227]. For simulations that use a water model that gives an inaccurate viscosity, overall motion is usually removed mathematically before computing internal relaxation rates [209, 221, 222, 227–229].

Comparisons of simulated spin relaxation times to experiments can be used to evaluate the accuracy of both protein dynamics and the conformational ensembles of proteins with long disordered regions [152, 230], multi-domain proteins [231], and peptides in micelles [224]. Spin relaxation rates are particularly informative for such systems because they report on dynamics, whereas other observables—such as chemical shifts and scalar couplings—may only report time-averaged properties [152]. For fully folded proteins, comparisons of spin relaxation times with experiment can be used to evaluate the accuracy of the computed rotational diffusion rates [222].

The experimental errors are typically between 0.01 s and 0.1 s for T_1 and T_2 values and, when using the full Redfield equations (Appendix B), between 0.05 and 0.1 for the backbone amide heteronuclear NOE values [232]. These experimental uncertainties are smaller than the usual statistical uncertainty in MD simulations [231].

3.1.6 Paramagnetic relaxation enhancement
General principles
Paramagnetic relaxation enhancement (PRE) is an increase in the NMR relaxation rate of a nuclear spin due to its dipolar interactions with unpaired electrons at a paramagnetic site; i.e., at an atom with an unpaired electron [233, 234]. Although the dipolar cross relaxation that causes PRE is

essentially the same as that which causes NOE, PRE ranges over much longer distances, as far as 25 Å, because the magnetic moment of an unpaired electron is much larger than that of a nucleus. Some metal atoms in metalloproteins are paramagnetic and generate measurable PREs. For other proteins, paramagnetic sites can be artificially introduced by using chemical reactions to attach extrinsic labels, called spin labels. This is usually done by engineering proteins to have just a single, reactive cysteine residue, which can then be reacted with a nitroxide-containing compound—such as MTSL—or a chelating agent carrying a paramagnetic metal—such as an EDTA derivative. Although it is possible to measure the PRE for longitudinal NMR relaxation rates ($R_1 = T_1^{-1}$), most applications focus on transverse relaxation rates ($R_2 = T_2^{-1}$), and these will be the focus of this section. PRE depends on both the distribution of distances and the time scales of motion of the protein, along with the location and dynamics of the spin label. The strength, range, and strong dependence on distance make PRE particularly suitable to detect and quantify transient and low-probability interactions.

In a typical PRE experiment, one measures the transverse relaxation rates of various groups in both the spin-labeled (paramagnetic) protein and in the same protein without the spin label (the diamagnetic protein). The PRE for nuclei across the protein then is obtained from the difference between the spin relaxations of these two measurements. Alternatively, because the PRE leads to line broadening, one may estimate the PRE from the ratio of the intensities (peak heights) in, e.g., heteronuclear single quantum correlation spectra of the paramagnetic and diamagnetic samples. However, although this has commonly been done and may lead to useful insights, the analysis of such data comes with additional uncertainty and assumptions so that one should, if possible, measure the PRE via relaxation rates [234]. When a nitroxide spin label is used, the diamagnetic protein can be generated simply by reducing the nitroxide with ascorbic acid. It is possible to measure PREs of various nuclei and chemical groups, but they are most commonly measured at backbone amides.

Evaluation of MD simulations using PRE
Like NOEs (Section 3.1.4) and spin relaxation rates (Section 3.1.5), PREs depend on the intramolecular motions of proteins in a complex manner (see Appendix C), and approaches used to handle intramolecular motions for NOEs and spin relaxation rates are applicable to PREs as well. Specifically, the PRE can be interpreted as providing a probe of the ensemble-averaged distances between probed spins and the spin label. An effective ensemble-averaged effective distance calculated from an MD trajectory, such as R_{AB}^{slow} (Eq. 8), can be com-

pared to distances calculated from experimental PREs. Alternatively, a model for the intramolecular motion of the protein, such as the Lipari-Szabo approach, can be used to compare order parameters or PREs directly to the corresponding experimental values.

The PRE has been used most extensively to benchmark MD simulations of disordered proteins with weak and/or transient interactions [81, 235], but it can also be used to probe transient interactions between folded proteins and between proteins and nucleic acids, and could thus be used to benchmark simulations of such systems. PREs have also been used to parameterize coarse-grained force fields for disordered proteins [236, 237]. Often, PRE experiments involve measurements on proteins labeled at multiple sites, one at a time, to get a global view of the structural dynamics of a protein. Each experiment probes the distance between the spin label site and all backbone amide protons. Thus, one is faced with the challenge of assessing a simulation of a protein based on all these measurements.

When calculating PREs from MD simulations and comparing them to spin label experiments, one must decide how to model the spin labels. One approach is to simulate each variant of the protein with its covalently linked spin label. This requires generating force field parameters for the covalently modified protein and often also requires mutating the sequence to exclude/include cysteine residues to match the experimental protein constructs. One must then repeat the simulation for each spin label site of interest. An easier alternative is to simulate the unlabeled wild-type protein and calculate distances between protein atoms as proxies for the distances between the spin label and amide protons. For example, one may use an atom in the side chain of the reference (wild-type) sequence as a proxy for the location of the unpaired electron. However, this approach risks missing genuine effects of the spin labels on the protein's dynamics. A compromise between these two extremes is to perform simulations of the wild-type protein and then model the spin label onto this simulation using a rotamer library developed to describe the structural preferences of the label. Such rotamer libraries are available for the commonly used MTSL spin label [238]. For applications to large MD simulations, placing the spin labels and sampling the rotamers may be achieved by tools such as Rotamer-ConvolveMD in the MD-Analysis package [239, 240] and DEER-PREdict [241], with the latter also implementing calculations of the PREs from the simulations. Which of these approaches to choose depends in part on the desired accuracy and whether there is experimental evidence that the spin label itself introduces a change in the conformational ensemble.

While PREs can be measured relatively accurately, a number of issues complicate their use in assessing force field

accuracies. First, unless the spin label is modeled explicitly, there may be uncertainty regarding the degree to which the spin label affects the protein [241, 242]. Second, PREs are often probed indirectly via measurements of intensity ratios, limiting accuracy and interoperability [234]. Finally, unless the time scales of motions are calculated directly from the simulations [243], calculations of PREs require either estimating [241] or fitting [237] the time scales, approaches that introduce uncertainty. Due to these complications, the root mean square error between simulated and experimental values (PREs directly or, indirectly, effective distances or order parameters) should be compared between simulations of the same protein with different force fields but should not be compared between simulations of different proteins.

3.2 NMR datasets

Many NMR datasets for proteins and peptides are available in the literature and in the Biological Magnetic Resonance Bank (BMRB) [146, 147]. We focus here on datasets, listed in Table 1, that appear particularly useful for benchmarking force field parameters because they largely meet the following criteria: 1) the data are available in a machine-readable format; 2) estimates of experimental uncertainty are included; 3) a diversity of structural motifs are present, including different secondary and tertiary structure elements and disordered regions; 4) the proteins are small enough that relatively short simulations suffice to provide well-converged estimates of the NMR observables. Note that many NMR observables for proteins were collected under low pH conditions, so it is essential to assign protonation states that are consistent with the environment of the experimental measurements (see Section 5.1). This can complicate force field benchmarks, because some force fields may not have parameters for protomers of the standard amino acids, especially for a protonated C terminus.

3.2.1 Beauchamp short peptides and ubiquitin

Beauchamp et al. [44] curated from the literature and Biological Magnetic Resonance Bank 524 NMR chemical shifts and scalar J-couplings for 19 capped 1-mers of the form Ace-X-Nme, where X are non-proline amino acids; 11 uncapped 3-mers; tetraalanine; and the protein ubiquitin. The short peptides in this dataset provide an opportunity to assess the backbone preferences of amino acid residues in the absence of a defined secondary structure. It is important for a force field to capture these preferences in order to accurately predict the conformational distributions of flexible loops in folded proteins and of unfolded and intrinsically disordered proteins.

3.2.2 Designed β-hairpins and Trp-cage miniproteins

Many groups have used NMR to characterize the solution structure, stability, and dynamics of designed beta-hairpin sequences [45–65] and miniproteins [66–72] in a variety of solution conditions. One series of studies of Trp-cage miniprotein sequences [73–78] includes NOE restraints for four Trp-cages (PDB entries 1L2Y [73], 2JOF [75], 2M7D [77], 6D37 [78]), folding rates measured by NMR resonance line-broadening due to folded/unfolded-state exchange [60, 244], and temperature-dependent chemical shift deviations (CSDs) for dozens of related sequences measured in the same solution conditions. These measurements offer a high-quality benchmark set that reports on how mutations perturb folding. Molecular simulations are now routinely able to access the microsecond time scales required to make accurate comparisons with these observables in only days of wall time.

3.2.3 Stroet folded proteins

Stroet et al. curated a dataset of 52 high-resolution protein structures, ranging in size from 17 to 326 residues [79]. All of the proteins are monomeric in solution, and none contain ligands or co-factors. 13 protein models, ranging in size from 17 to 129 residues, were derived from NMR experiments. For nine of these, NMR observables—including upper bounds for interatomic distances derived from NOE data, ^3J-couplings, and RDCs—are available from either the PDB or the Biological Magnetic Resonance Data Bank (BMRB) [146, 147, 232]. This dataset, which is available from the Australasian Computational and Simulation Commons (ACSC) Molecular Simulation Data Repository [245] at https://molecular-dynamics.atb.uq.edu.au/collection/protein-force-field-validation-set, was used to validate backbone torsional parameters [34] and to study the impact of nonbonded cutoff schemes on experimental observables [246].

In preparing the dataset, Stroet et al. curated the NOESY studies that provide interatomic distances derived from NOESY intensities but not the intensities themselves. Any pseudoatom corrections included in the experimental data were first removed, and then the pseudoatom corrections proposed by Wüthrich [247] were applied. Pseudoatom corrections were applied in this way to six of the nine proteins with NOE distance restraints. For two of the remaining three proteins, no information on any pseudoatom corrections to the NOE-derived upper bounds could be found, and so the NOE restraints reported for these proteins may be less rigorous. For the third, mercury binding protein from *Shigella flexneri* (PDB ID 1AFI), a large number of violations in NOEs assigned to Phe47 were observed in initial simulations. Inspection of the structure and the violations points to wrongly assigned explicit phenyl ring hydrogens Hε1/2 and

Hδ1/2. Replacing the assignment of these protons on the symmetrical phenyl ring by pseudoatoms HE* and HD*, and adding the appropriate pseudoatom corrections [247], removed all violations above 0.5 Å [79].

3.2.4 Mao folded proteins

This is a collection of 41 folded proteins for which both X-ray structures and NMR data, comprising backbone chemical shifts and NOESY intensities, have been measured by the Northeast Structural Genomic Consortium [80]. These data have been used to assess the accuracy of NMR structures, X-ray structures, and Rosetta refinements [80], and to compare the accuracy of MD simulations run with various force fields [81, 248].

3.2.5 Robustelli folded and disordered proteins

Robustelli, Lindorff-Larsen, and co-workers assembled a dataset (https://github.com/paulrobustelli/Force-Fields) that spans 21 proteins and peptides with over 9,000 experimental data points and probes the ability of a protein force field to simultaneously describe the properties of folded proteins, weakly structured peptides, and disordered proteins with a range of residual secondary structure propensities [29, 81, 184, 235, 248].

The dataset contains four fully folded proteins (ubiquitin, GB3, hen egg white lysozyme (HEWL), and bovine pancreatic trypsin inhibitor) with extensive NMR data, including backbone and side chain J-couplings, RDCs, and backbone and side chain spin relaxation order parameters. It also includes calmodulin and the bZip domain of the GCN4 transcription factor, both of which contain folded and flexible components. Calmodulin, which comprises two folded domains connected by a flexible linker, probes the ability of a force field to simultaneously describe the flexibility of the linker region, the stability of folded domains, and the propensity of the folded domains to associate. The NMR data for calmodulin comprise chemical shifts and RDCs. The bZip domain of the GCN4 transcription factor is a partially disordered dimer with an ordered, helical, coiled-coil dimerization domain, for which NMR chemical shifts and backbone amide spin-relaxation parameters are available.

The dataset also includes nine proteins that are disordered under physiological conditions and for which extensive sets of NMR data are available. These test the ability of a force field to accurately describe the dimensions and secondary structure propensities of intrinsically disordered proteins. The proteins range in size from 40 to 140 amino acids, which was important, as a number of force fields that produced reasonable dimensions for proteins containing <70 amino acids produced conformations that were substantially over-collapsed for longer sequences. The available

NMR data for these disordered proteins include chemical shifts, RDCs, backbone *J*-couplings, and PREs. Scalar couplings of the disordered Ala5 peptide were also included. Code to calculate NMR and SAXS data for five IDPs in this benchmark (Aβ40, drkN SH3, ACTR, PaaA2 and α-synuclein) is included in a recent manuscript describing a maximum entropy reweighting approach to calculate conformational ensembles of IDPs from experimental data [249] (https://github.com/paulrobustelli/Borthakur_MaxEnt_IDPs_2024).

The Robustelli dataset is enriched by a number of non-NMR data, including radii of gyration obtained by various experimental methods and data on the temperature-dependent stability of fast-folding proteins. The dataset has also been expanded [248] to include the free energies of association of 14 protein-protein complexes, the osmotic coefficients of 18 organic and inorganic salts, the position of the first peak of the radial distribution functions of seven ion-water and ion-ion pairs, comparisons of Ramachandran distributions of blocked amino-acids in water, Ramachandran distributions obtained from X-ray coil libraries, Lifson-Roig helix extension parameters for the 20 amino acids (estimated from NMR studies of Ace-(AAXAA)3-Nme peptides), the folding free energies of mutants of 22 mutants of Trp-cage, the folding enthalpies of 10 fast-folding proteins, the Kirkwood-Buff integrals of ethanol water-mixtures, and the melting curves of the Trpzip1 and GB1 β-hairpin-forming peptides.

3.2.6 Spin relaxation datasets

Spin relaxation data have been reported for a large number of proteins; see, for example, a comprehensive review of spin relaxation measurements published before 2006 [82]. Furthermore, the BMRB contains spin relaxation data for approximately 350 proteins [146, 147, 232]. A curated dataset of BMRB spin relaxation measurements for 133 proteins, which were used to train the Dyna-1 deep learning model for predicting slow dynamics in proteins [250], is available from https://github.com/WaymentSteeleLab/Dyna-1#datasets. Although most reported spin relaxation datasets have not so far been analyzed in conjunction with molecular simulations, there are some useful examples of such analyses.

First, the membrane-bound, bacterial TonB proteins possess a long, disordered region that links their C-terminal domain to their transmembrane N-terminus. Spin relaxation times have been measured for backbone N-H bonds in the C-terminal domain with varying linker lengths (HpTonB194-285, HpTonB179-285, and HpTonB36-285) [84], and these data can be used to evaluate the conformational ensembles predicted by MD simulations. In one such study [152], simulations with the Amber ff03ws force field reproduced the experimental spin relaxation times of partially disordered fragments of TonB as well as Engrailed 2 (see below), while CHARMM36m and Amber ff99-ILDN gave less accurate results, apparently because they yielded overly collapsed conformational ensembles. Interestingly, however, the three force fields gave similar accuracy for chemical shifts, presumably because of degeneracy in chemical shifts with respect to the structural features important for describing this ensemble. Second, backbone N-H spin relaxation data at multiple magnetic fields are available for the partially disordered 143–259 region of the Engrailed 2 transcription factor [85]. This region is highly conserved and is involved in the binding of transcriptional regulators. Simulations with the Amber ff03ws force field yielded good agreement with the experimental data, except for serine and aspartate residues [152].

Backbone N-H spin relaxation times have also been used to find the best ensembles among diverse simulation data for four different multi-domain proteins using the quality evaluation based simulation selection (QEBSS) approach [231], and to characterize dynamics of six different peptide-micelle systems in conjunction with MD simulations [224].

In addition, spin relaxation times for side chain methyls, rather than backbone N-H groups, have been reported for T4 lysozyme [251], which exhibits anisotropic tumbling in solution. Modeling the anisotropic tumbling in the calculation of spin relaxation observables from MD simulations was found to improve the agreement with experimental observables. This dataset of side chain relaxation rates was also used to validate the ABSURDer reweighting approach for calculating relaxation rates from MD simulations [252].

3.2.7 NMR-derived salt bridge stabilities

In some cases, NMR data can be used to measure secondary properties, which can in turn be compared with simulations. For example, the thermodynamic stability of weak, solvent-exposed salt bridges can be assessed by monitoring NMR chemical shifts as a function of pH. One such study examined three potential salt bridges in the context of a folded protein, the B1 domain of protein G (GB1) [253]. These involve lysine-carboxylate ionic interactions and were identified from crystal structures. The stabilities of these potential salt bridges were assessed via ^{15}N and ^{1}H chemical shifts and the hydrogen-deuterium exchange rates of the lysine ammonium group during titration of the carboxylates [86]. The NMR data indicate that two of the salt bridges are not formed in solution, while the third is only weakly formed. Interestingly, most force fields tested overestimated the stability of the salt bridges [253], a result also reported by previous studies [185, 254] looking at the association constants of oppositely charged amino acids in water, where the experimental data were obtained by potentiometric titration rather than

NMR. Here, the results were significantly improved by atomic charge derivation strategies that implicitly incorporate solvent polarization [255] and by the use of the more expensive, polarizable CHARMM Drude-2013 [256] and AMOEBA force fields [257].

4 Room-temperature protein crystallography

The earliest protein crystal structures were determined by X-ray diffraction for specimens at or near room temperature [258]. Later, methods of working with protein crystals at low temperature were widely adopted for their practical advantages, such as slowed radiation damage of the crystal specimen [259, 260]. More recently, however, there has been a revival of interest in room temperature protein crystallography because it gives insight into molecular motions that are quenched at cryogenic temperatures [259–262]. At cryogenic temperatures, there is little motion to simulate, and it is also not clear that the conformational variability observed at cryogenic temperatures corresponds to a well-defined thermodynamic ensemble at any particular temperature [262, 263]. In contrast, the conformational variability observed by RT crystallography is more clearly attributable to room temperature thermal motions. Thus, although most crystallographic data available in the Protein Data Bank (PDB) were measured at cryogenic temperatures, the PDB contains a growing number of room temperature structures. Neutron diffraction crystallography (Section 4.1.5), which also can be carried out at room temperature, goes beyond X-ray crystallography because it can resolve the positions of hydrogen atoms in protein structures. However, there are not many neutron diffraction structures in the PDB, as there are not many suitable neutron sources where these studies can be done, and the method requires larger crystals.

Water typically occupies about 50% of a protein crystal's volume [264], so crystallized proteins are usually quite well-solvated. Nonetheless, the conformational distribution of a crystallized protein is likely to differ from that of the same protein in solution due to protein-protein contacts, perturbations of water structure, and the presence of co-solutes added to facilitate crystallization. Therefore, when one uses crystallographic data to benchmark molecular simulations, one should simulate the crystal rather than the protein in solution. Crystal simulations with periodic boundary conditions can in principle be used to simulate a single unit cell, but this does not allow variations between adjacent lattice positions. Therefore, it is preferable to simulate a supercell comprising multiple unit cells, which can sample such lattice variations (Fig. 3). This requirement adds complexity and
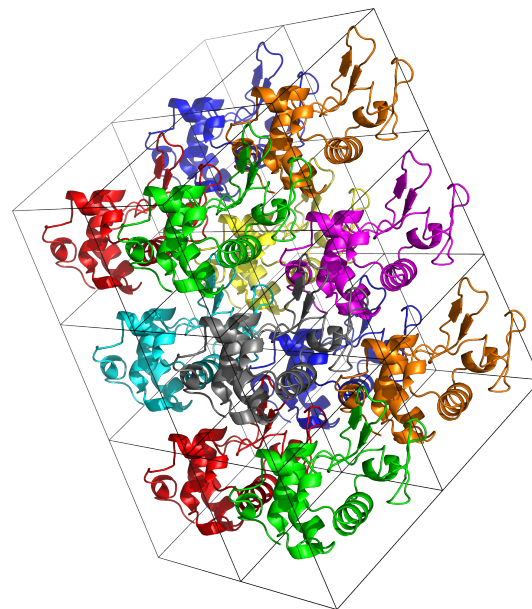

Figure 3. A $3 \times 2 \times 2$ supercell containing 12 independent copies of hen egg white lysozyme (4LZT [90]), each colored differently from its neighbors. Black lines represent cell boundaries of the triclinic unit cell.

can lead to larger and hence slower simulations than those used with NMR benchmark data, depending on whether the size of the supercell exceeds the size of the simulation box for the solvated system. At the same time, the extra copies of the protein in supercell simulations often can be included in computing statistical averages, enabling some analyses to be performed using shorter simulation durations than for conventional solution simulations. Note, however, that supercell simulations include periodic boundaries with exact translational symmetries not present in experimental crystals, and this difference in periodicity can complicate the prediction of some crystallographic observables, such as Bragg diffraction. Different proteins, and even different crystal structures of the same protein, have different levels and characters of both conformational variation and experimental error. Therefore, much as for NMR benchmarking, force fields should be compared against a single crystallographic dataset, rather than attempting to compare force fields based on benchmarks against different crystal structures. Methods of simulating protein crystals have been discussed in recent reviews [265, 266].

This section reviews crystallographic observables that are relatively well developed for use in benchmarking protein simulations, touches on the additional topic of diffuse scattering, and lists crystallographic datasets that are well-suited for benchmarking.

4.1 Crystallographic observables

4.1.1 Bragg diffraction data

General principles

X-ray scattering from an ideal crystal would be focused into sharp spots on the detector, located at positions that are generally called the Bragg positions, and the intensities at these positions are sensitive to the time- and space-average of the electron density. Diffuse scattering, which reports on correlated motions and crystalline disorder [267], is discussed in Section 4.1.4. The Bragg intensities are related to the mean unit-cell electron density through the complex structure factors, i.e., the Fourier transform of the unit-cell electron density distribution. In particular, the Bragg intensities are proportional to the squared amplitudes of the structure factors. The phases are also needed to compute the electron density, and in the absence of an anomalous dispersion signal from which experimental phases can be obtained, the phases must be obtained from an atomic model. Therefore, the electron density is not typically an "experimental observable" in the usual sense. Still, for a well-refined model, with good resolution and an R-factor below 15% or so, it is common to refer to the "observed electron density".

Evaluation of MD simulations using Bragg diffraction data

The simplest way to assess the accuracy of a simulation of a protein crystal is to compare the average structure from the simulation to the crystallographic model structure available in the Protein Data Bank (PDB) [268–270]. This intuitive and quick comparison may be useful when there are large differences between two force fields of interest. In addition, the focus on a single structure may help to identify specific problems in the force field, such as insufficient stability of alpha-helices or side chains preferring uncommon rotamers, and thus can yield feedback to improve it. To assess protein backbone agreement, RMSD of $C\alpha$ atoms or of all backbone heavy atoms can be used. To assess protein side chain agreement, RMSD of side chain atoms, dihedral-angle agreement within an angular threshold, or match between defined rotamers [271, 272], can be used. Note that a rigorous measure of RMSD must account for symmetries; for example, a $180°$ rotation of the χ_2 angle of a phenylalanine in the model should not alter the computed RMSD. The possibility of alternate conformations (Section 4.1.3) should also be factored in. While this general approach is widely used and simple to implement, the structural models used as the basis for the comparison are typically derived by refinement informed by an empirical force field. Therefore, comparisons against the structural model risk assessing, at least in part, the agree-

ment of the force field of interest with the force field used during refinement, rather than agreement with experimental data. For example, an active site proline in ligand-bound triosephosphate isomerase adopts a planar ring conformation, but structural models refined using force fields place the proline ring in a puckered conformation that is inconsistent with the experimental electron density [273].

A second approach is to compute the mean electron density from the simulation [274–277] and compare it with the crystallographic electron density. This avoids the intermediary of a structural model inferred from the crystallographic density. Note that the observed electron density is averaged over time and over many unit cells of the crystal and hence reflects disorder in the experimental system. Because the density is often sharply peaked around atomic positions, standard methods of comparing density, such as the Pearson correlation, can be more sensitive to small, local differences than is desired for assessing the relative accuracy between force fields, although these issues may be addressed by approaches such as local model-based map alignments [278]. Existing crystallographic software, such as EDSTATS in CCP4 [276] and `phenix.real_space_correlation` in PHENIX [279], can be used to compare electron density maps.

Other methods of comparing simulations to crystallographic data are also possible. For example, one could compute the mean electron density from the simulation using mdv2map in AmberTools with CCP4 [276, 277] or xtraj in LUNUS with CCBTX [274, 275], use crystallographic software to refine an atomic model into this density, and then use structural metrics such as atomic RMSD to compare this model with the experimental structure model [266]. Alternatively, one could directly compare computed and measured Bragg intensities. This properly takes care of disorder but can be more difficult to interpret, as modest deviations (by an RMSD less than 0.5 Å) of the protein away from the "correct" structure can greatly reduce the real-space correlations between computed and experimental diffraction intensities [280]. In such cases, other measures of accuracy than correlation may be more useful. One may also focus on the accuracy of protein-protein contacts at interfaces, as these depend on non-covalent and solvent-mediated interactions that are important to get correct. Most crystal contact analysis methods, such as the PISA software [281], use a single average structure as input, so it is difficult to account directly for disorder. As with the comparison of average protein structures discussed above, this sort of comparison may be most useful when there are relatively large differences between the results from different proposed force fields.

4.1.2 Debye-Waller factors (B-factors)

General principles

The Debye-Waller factor of an atom, also known as its B-factor or temperature factor, is closely related to the atom's fluctuations about its modeled position. B-factors are, in principle at least, related to the mean square displacement of the atom, $\langle |\vec{u}|^2 \rangle$, by the equation

$$B = \frac{8}{3}\pi^2 \langle |\vec{u}|^2 \rangle = 8\pi^2 \langle u_s^2 \rangle \tag{13}$$

where $u_s = \frac{\vec{u} \cdot \vec{s}}{|\vec{s}|}$ is the projection of the displacement vector $\vec{u}$ onto the direction of the scattering vector $\vec{s}$. The statistic $\langle u_s^2 \rangle$ is sometimes referred to as the isotropic mean square displacement [282]. However, B-factors can also have contributions from various sources of error and are sometimes viewed as "slop factors" that absorb other errors in the model to improve the fit to the diffraction data. They may result from incorrect assignments of the identities of atoms of similar atomic number, i.e., K vs. Ca, and may also reflect, in a nonspecific manner, incorrect modeling of side chains, alternate conformations, and disordered regions. Debates continue between crystallographers about whether a region that exhibits signs of disorder should be left unmodeled, modeled with zero occupancy, or modeled with B-factors allowed to refine to high values; as a result, their meaning can vary between different crystal structures. However, when derived from well-validated structural models [283] that are modeled correctly and carefully interpreted, they provide a meaningful measure of the mobility of atoms within a protein structure. For structures with resolutions better than 1.5 Å, it is often appropriate to refine anisotropic B-factors, such that each atom is assigned a tensor of six parameters that define a three-dimensional Gaussian distribution of atomic fluctuations. A more detailed hierarchical model—which partitions the total disorder, normally assigned to atomic B-factors, into a series of contributions across length scales—was recently developed [284] and applied to ensemble refinement [285] to produce structural models with lower R-factors.

Evaluation of MD simulations using Debye-Waller factors

Given a simulation of a protein crystal, it is straightforward to compute the mean square displacement $\langle |\vec{u}|^2 \rangle$ of each atom from its mean position and thus obtain computed B-factors; a modest amount of additional complexity is involved in computing anisotropic B-factors. Alternatively, one may, again, refine a structural model into the electron density computed from the simulation and compare the resulting simulated B-factors with the experimental B-factors. Depending on the software used for refinement, constant offsets may be applied to reported B-factors using methods that do not distinguish contributions from molecular motion from other contributions [279]. Constant offsets are often ignored in B-factor profile comparisons by using, e.g., a Pearson correlation to assess the agreement. Nevertheless, it is sometimes possible to account even for the constant offset using MD simulations [280].

4.1.3 Alternate conformations

General principles

In contrast to B-factors, which in principle describe fluctuations within a local energy well, alternate conformations explicitly describe jumps between discrete conformational states in separate local energy wells. For example, alternate conformations—also known as alternative conformations, alt confs, alternate locations, or altlocs—can be used to model amino acid side chains that switch between different rotamer [271, 272] conformations. Alternate conformations are a promising avenue for comparison to and improvement of simulations. In a protein structure file (e.g., a PDB file), an alternate conformation is given as an additional set of coordinates for a group of atoms (often a residue or series of residues) and marked with a single-character identifier (A, B, C, etc.) that is unique within the file. Each alternate conformation is also assigned a partial occupancy (i.e., probability) from 0 to 1 that is determined by the crystallographic refinement. For covalently bound atoms, like those of proteins, the occupancies of all alternate conformations for a given atom typically sum to one. For molecules that are not covalently bound to the protein, such as ligands and water molecules, the occupancies may sum to less than one. A crystal structure model that contains alternate conformations is termed a multiconformer model.

Alternate conformations are often left unmodeled in crystal structure models even when they are evident in the electron density maps [286]. Such missing alternate conformations can be modeled in an automated and unbiased manner with tools such as PanDDA [278] and qFit [287, 288], and MD simulations of protein crystals can be used to find alternate conformations that are missed by such automated methods [266]. Alternate conformations are significantly more prevalent at RT than at cryogenic temperatures (cryo) [289], and some alternate conformations that are observed only at RT and not at cryo are critical to biological function [261] (see Section 4.2.4) and can modulate ligand binding in important ways [263]. Thus, RT crystallography can reveal biologically relevant conformational heterogeneity.

When different parts of a biomolecule have alternate conformations, these conformations may not be mutually independent. For example, a rearrangement of one part of a protein might restrict which conformations its neighbors can adopt. However, the RCSB PDB format for crystal structure

models does not provide a mechanism for specifying which alternate conformations are physically compatible with one another, other than the A, B, C, etc. identifiers themselves. This bookkeeping issue creates ambiguity in situations where, for example, residue X with conformations A and B is near residue Y with conformations A, B, and C. In such a case, it is unclear which conformations of residue X are energetically compatible with conformation C of residue Y. The corresponding alternate conformations might be coupled, uncoupled, or partially coupled.

Evaluation of MD simulations using alternate conformations

The fractional occupancies of alternate conformations inferred from RT crystallography can be readily compared with the corresponding results from a simulation. Conformational probabilities can be extracted from simulations based on the occupancy of rotameric energy wells, e.g., the rotamer name strings [272] that are output by software such as the model validation suite MolProbity [283] or the related tool `phenix.rotalyze` from PHENIX [279]. A figure of merit for the ability to reproduce alternate conformations could be the root-mean-square error between rotamer probabilities from simulations and experimental occupancies of alternate conformers, although future studies could explore more sophisticated approaches. Note that there are datasets where occupancies have been shown to vary with temperature [263]. In addition, the coordinates and occupancies of alternate conformations can be combined with B-factors to calculate so-called crystallographic order parameters, which may be compared with experimental NMR order parameters [290]; a similar framework could be useful for comparing RT crystallographic models to simulations.

When comparing computed and observed populations, it is important to bear in mind that small changes in the free energy difference between two conformations will lead to large changes in occupancy if the free energy difference is near zero, but essentially no change if the free energy difference is far from zero. This is because, given two conformers A and B, the probability of being in state A is

$$p_A = \left(1 + \exp\left(-\frac{\Delta G_{AB}}{RT} \right) \right)^{-1} \tag{14}$$

where ΔG_{AB} is the free energy change upon going from A to B; and $p_B = 1 - p_A$, assuming only two accessible conformations.

4.1.4 Diffuse scattering (non-Bragg reflections)

Diffuse or continuous scattering refers to the cloudy, streaked, speckled, halo-shaped, or otherwise patterned weak scattering that lies between the Bragg peaks. In contrast with the Bragg scattering, which is associated with

correlations in the mean electron density, the diffuse scattering is associated with spatial correlations in the deviation of the density from the mean. Thus, much as the Bragg peaks can be used to model the mean structure of the molecules within the unit cell, the diffuse scattering can, at least in principle, be used to model the coupling of variations in atom positions. There have been several protein crystal MD simulations of diffuse scattering, and it is straightforward to compute diffuse intensities from a collection of simulation snapshots [89, 93, 280, 291–296]. Although a limited study did show that the force field can influence the simulated diffuse scattering [296], it is not yet clear what sort of variation one should expect. Future crystal simulations should help clarify whether diffuse scattering data can be useful for benchmarking force fields.

4.1.5 Hydrogen coordinates from neutron diffraction

Protein X-ray crystallography generally resolves the coordinates only of atoms with atomic number greater than one; hydrogens are visible only in well-ordered regions of ultra-high-resolution structures. This is because X-rays are scattered by electrons, and a hydrogen atom, with only one electron, scatters only weakly. In contrast, neutrons are scattered by atomic nuclei, and hydrogen atoms have high scattering cross-sections. As a consequence, neutron diffraction protein crystallography can provide experimental information on hydrogen atom positions and protonation states [297, 298]. These data are valuable because hydrogens make up nearly half the atoms in proteins and can play critical functional and structural roles [299, 300]. Thus, neutron diffraction studies are uniquely suited to discern hydrogen bonding patterns [102, 301], the orientations of solvent and side chain groups [302, 303], and the protonation states of critical catalytic amino acids—such as the histidine in the catalytic triad of a serine protease [304, 305]. By the same token, they provide distinctive information to benchmark protein force fields.

Neutron diffraction is generally weak and therefore requires large crystals and long data collection times. In addition, there are only a few suitable neutron sources in the world. However, macromolecular neutron crystallography has become more practical in recent years with the commissioning of new, higher flux neutron sources, as well as improved techniques for crystal growth and molecular biology techniques for producing large amounts of purified proteins.

4.2 Protein crystallography datasets

Here, we highlight protein crystallography datasets that are well-suited to benchmark force field accuracy, prioritizing the following features: 1) resolution better than 1.2Å, so

that observable data have low uncertainty; 2) relatively unambiguous assignment of protonation states; 3) absence of ligands or co-factors which might conflate the accuracy of the small molecule force field with that of the protein force field (note that benchmark systems for binding free energies have been reviewed elsewhere [306, 307]); 4) diversity of secondary and tertiary structural motifs; 5) availability of crystal data for the same protein system in multiple symmetry groups to explore the possible impact of different crystallographic contacts; and 6) availability of crystal data for the same protein system at multiple temperatures near room temperature to probe the ability of the force field to capture the temperature dependence of the observables.

4.2.1 Scorpion toxin

Scorpion toxin (PDB ID 1AHO) is a 64-residue globular protein with a room-temperature X-ray data set at 0.96 Å resolution [87]. It has little regular secondary structure—one 9-residue helix and two short β-strands—but is stabilized by four disulfide bonds. It was the subject of an early MD simulation [308] that compared simulations of the crystal conducted with four protein force fields that would now be considered obsolete. The notable finding at the time was how diverse the simulation results were, even for simple metrics like average backbone structure and computed B-factors. This study supported the idea that crystal simulations could be used for testing protein simulations and could report on both the stronger interactions that determine the conformation of an individual chain and also on the weaker (often solvent-mediated) interactions that stabilize the crystal lattice. The small size of the unit cell was a more important consideration in 2010 than it would be today.

4.2.2 Hen egg-white lysozyme

Hen egg-white lysozyme (HEWL) was one of the first proteins whose structure was solved by X-ray crystallography, and it is widely used as an experimental model system, in part because of its ease of crystallization. This protein has 129 residues, eight α-helices, two β-sheets, and four disulfide bonds. Given these constraints, simulations with different force fields are expected to reproduce the structural features of the deposited models with similar fidelity but to exhibit differences in fluctuations that are captured by observables such as electron densities and B-factors.

There are three RT crystal structures of the triclinic form (P1 space group) of HEWL (PDB IDs 2LZT [88], 2.0 Å resolution; 6O2H [89], 1.2 Å; 4LZT [90], 0.95 Å), as well as a high-resolution cryo version (PDB ID 2VB1 [91], 0.65 Å). For all four structures, nitrate ions predominate in neutralizing the protein, and a large number of solvent waters are visible in the experimental electron density. The crystal density

of the solvent has been determined to high precision [309] and is very close to that estimated by the MD simulations discussed below; this suggests that we know the amount of water in the unit cell to within an uncertainty of just a few water molecules.

Two studies of the triclinic HEWL crystal provide guidance for future studies. First, Janowski et al. [310] simulated a triclinic supercell with 12 protein chains and made extensive comparisons to the reflection intensities from 4LZT. Perhaps most insightful was a comparison of two newly refined atomic models, one refined against the experimental data and a second refined against the average electron density from an MD simulation of 3 µs. The backbone of the structure refined against the simulated density was about 0.4 Å away from the experimental structure, which is well outside the expected coordinate uncertainty of 0.1 Å for a static crystal model but small compared to differences between structural models of the same protein derived from crystal data in different space groups [311]. In addition, simulations using four protein force fields showed systematic differences in how close the B-factors refined against simulated densities were to experiment (Fig. 4). Rather strikingly, B-factors computed with the ff14SB simulations are considerably closer to experiment (black) than those computed with three other protein force fields. This result further supports the utility of crystallographic data to benchmark protein force fields.

Second, Meisburger et al. [89] simulated triclinic HEWL with a primary focus on interpreting diffuse X-ray intensities. Since collective motions of many protein chains are important for diffuse scattering, simulations were carried out using 1, 27, 125, and 343 unit cells, but they used only a single protein force field, so this study did not directly address the suitability of diffuse scattering for force field benchmarking.

There are also RT crystal structures for two other crystal forms—orthorhombic (PDB IDs 1AKI [92] and 8DYZ [93]) and tetragonal (PDB ID 8DZ7 [93])—of HEWL, and these may further probe the ability of force fields to capture crystal-packing interactions [93]. Additionally, NMR data for HEWL in solution is available as part of the Stroet (Section 3.2.3) and Robustelli (Section 3.2.5) datasets.

4.2.3 Crambin

The small hydrophobic protein crambin, isolated from the seeds of the Abyssinian cabbage (*Crambe abyssinica*), was found early on to form exceptionally well-ordered crystals [94] and has been used for the development of experimental phasing techniques. Crambin has only 46 residues, contains both α-helical and β-strand secondary structure, and is held together by disulfide bridges. Of note here is a study combining RT X-ray and neutron diffraction to 1.1 Å
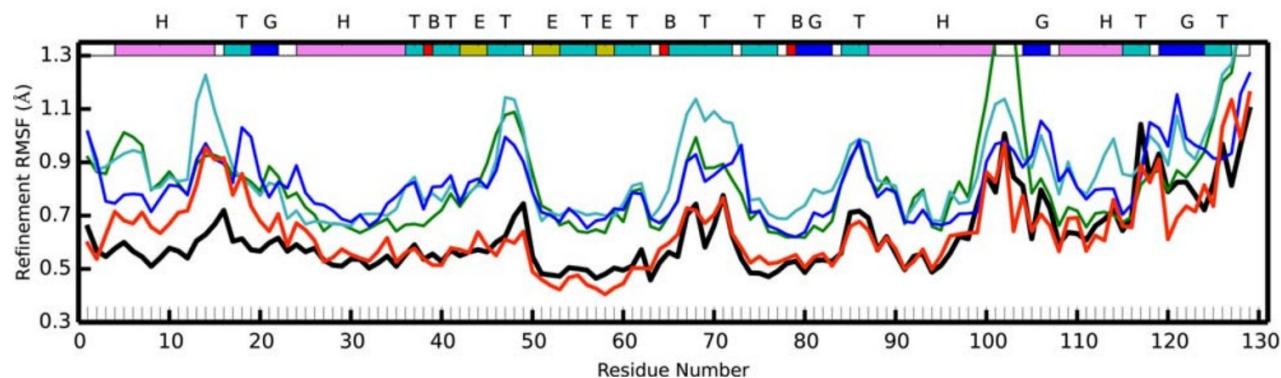
Figure 4. Refined Cα root-mean-square fluctuations for four crystal simulations of HEWL, compared to experiment (black). Red: ff14SB, blue: CHARMM36, cyan: ff14ipq, green: ff99SB. The colored band across the top describes the secondary structure (T: turn, E: β-sheet, H: α-helix, G: 3_{10} helix, B: isolated bridge). Adapted from Janowski et al. [310].

resolution (PDB ID 3U7T) and allowing for the modeling of anisotropic displacement parameters for select hydrogen atoms [95]. More recently, a new dataset of crambin was collected at RT to 0.70 Å (PDB ID 9EWK) [96], and the model was refined to an R-factor of 0.0591 for reflections four standard deviations above background and of 0.0759 for all reflections, with stereochemical restraints released during the refinement process. This model has one of the lowest R-factors in the PDB. The estimated standard deviation for C-C bonds (a measure of the uncertainty in bond lengths for a model refined without restraints on bond lengths) was 1.6 pm overall (0.6 pm for atoms with no alternate conformations). These estimated standard deviations are similar in magnitude to those found in the Cambridge Structural Database (CSD) [312], a database of small molecule crystals containing particularly reliable molecular geometries. NMR data has also been reported on recombinantly expressed crambin [313].

4.2.4 Cyclophilin A

Human cyclophilin A (CypA) is a proline *cis/trans* isomerase that is attractive for assessing force fields due to its well-studied conformational dynamics and high-resolution RT crystallographic data. Several residues (Ser99, Phe113, Met61, and Arg55) in the active site form a network that exhibits correlated conformational dynamics on a similar time scale (millisecond) as catalytic turnover, as measured by NMR relaxation experiments [314]. The alternate conformations in the CypA active site are primarily related to one another by transitions between side chain rotamers, but they also involve backbone shifts that are well-modeled by the back-rub motion [315, 316]. More recently, multi-temperature crystallography across a series of eight temperatures added

nuance to our understanding of CypA's dynamic active-site network, revealing evidence for more complex, hierarchical coupling, in which dynamics for some active-site residues are dependent on the conformation of another key active-site residue (Phe113) [97]. Diffuse scattering [98] and X-ray solution scattering [99] have also been measured for CypA. Thus, CypA has rich dynamics that are well-studied by RT crystallography and other experiments and are ripe for comparison to MD simulations.

An attractive crystallography dataset for validating simulations of CypA is the high-resolution (1.2 Å) RT structure (PDB ID 4YUO) [97], where various residues adopt two (A, B) or three (A, B, C) alternate conformations. This complicates the creation of a single-conformer starting model for a given simulation; we recommend beginning with either state A or state B, which have distinct rotamers for Phe113, which is thought to be the linchpin of the network [97]. One challenge in using CypA to validate simulations is that its catalysis and matching dynamics have been reported to be on the millisecond time scale [314]. However, multi-temperature crystallography also suggests motions at nanosecond time scales [97]. Mutagenesis and solution scattering also point to a separate, independent loop region with faster dynamics [99]. Thus, even simulations at shorter than millisecond time scales are likely to uncover relevant dynamic features in CypA that can be useful for validation and force field optimization/development. Moreover, CypA is moderately sized, with 165 amino acid residues, which enables reasonably long simulations. Enhanced sampling techniques [317] that accelerate sampling of solute conformations, such as replica exchange with solute tempering [318], may also be of use. Time scale issues are considered further in Section 5.2. In addition to the 1.2 Å RT crystal structure, a similarly high-resolution

(1.25 Å) cryo structure of CypA (PDB ID 3K0M) [261] in the same crystal lattice is also available.

4.2.5 Ubiquitin variants

Computational protein design and directed evolution have been combined to generate ubiquitin variants with enhanced affinity to the protein USP7 through stabilization of a specific conformation of the $\beta 1\beta 2$ loop [319]. In a subsequent study, RT X-ray crystallography was used to examine the structural basis for the increasing affinity of ubiquitin variants over the course of the design/selection process [100]. The resulting high-resolution RT crystal structures revealed that the earlier "core" mutant of ubiquitin (with 6 mutations) exhibits multiple discrete alternate conformations of the $\beta 1\beta 2$ loop, whereas the later "affinity-matured" mutant (with 3 additional mutations) adopts a singular conformation of this loop. Also, elsewhere in the protein, both variants exhibit alternate conformations for a peptide flip of residues 52-53.

High-resolution RT crystallography data and models are available for both ubiquitin mutant proteins at resolutions of 1.12 Å for the "core" mutant (PDB ID 5TOF) and 1.08 Å for the "affinity-matured" mutant (PDB ID 5TOG). Simulations based on these datasets should ideally capture the differences in $\beta 1\beta 2$ loop conformational heterogeneity observed in the crystallographic electron density maps for the two variants (more flexible for the core mutant, more rigid for the affinity-matured mutant) as well as the peptide flip shared by both variants. For this system, a number of metrics could be used to quantify the match between simulations and experimental data: root-mean-square fluctuations (RMSF) of backbone atoms in the loop for simulations vs. for multi-conformer crystal structures, local real-space fit to the electron density map in the $\beta 1\beta 2$ loop region, recapitulation of fractional occupancies for the different loop conformations (perhaps after clustering the simulation snapshots), etc. Thus, the high-resolution RT crystallography datasets available for these ubiquitin variants provide an opportunity to benchmark force fields for their ability to provide accurate simulations of protein backbone conformational heterogeneity, as well as accurate predictions of the effects of amino acid substitutions on conformational heterogeneity. Note that NMR data for wild-type ubiquitin in solution is available as part of the Beauchamp (Section 3.2.1), Stroet (Section 3.2.3), and Robustelli (Section 3.2.5) datasets.

4.2.6 Protein tyrosine phosphatase 1B

Human protein tyrosine phosphatase 1B (PTP1B; also known as PTPN1) exhibits structural dynamics across a range of time scales and length scales, as revealed by numerous structural

biophysics experiments. Recently, multi-temperature X-ray crystallography of apo PTP1B at temperatures from cryo to RT [101] provided insights into correlated conformational heterogeneity. The resulting series of crystal structures featured alternative conformations, each modeled with partial occupancy, for the active-site WPD loop (open vs. closed) as well as distal allosteric regions. As temperature increased, the WPD loop shifted crystallographic occupancy from the closed to the open state. Simultaneously, the distal $\alpha 7$ helix, a key component of PTP1B's allosteric network, shifted occupancy from the ordered state to a disordered state, moving into an adjacent solvent channel in the crystal lattice. Several residues between these regions exhibited smaller-scale conformational shifts between alternate conformations of the WPD loop and the $\alpha 7$ helix, mimicking the shifts seen previously for an allosteric inhibitor that displaces $\alpha 7$ [320].

Other studies have used NMR spectroscopy to characterize time scales of motion for various regions of PTP1B. NMR relaxation experiments showed that the active-site WPD loop closes on a time scale corresponding to the rate of catalysis (millisecond) [321]. Beyond the active site, NMR relaxation experiments, mutagenesis, and molecular dynamics simulations restrained by NMR chemical shifts showed that faster dynamics are key to allosteric regulation via $\alpha 7$ [322]. Thus, PTP1B exhibits motions that may be amenable to various different types of simulations, from short, traditional simulations to long, enhanced-sampling simulations.

Both of the major states of the protein are modeled in the 1.74 Å RT (278 K) crystal structure of apo PTP1B (PDB ID 6B8X) [101]: the closed state (alternate conformation A) and the open state (B). Most regions of the structure are modeled either with no alternate conformations or with both A and B conformations. By contrast, because it is disordered in the open state of the protein, the $\alpha 7$ helix is modeled as only the A conformation with partial occupancy, with no coordinates for the B conformation. Note that crystallography was performed with residues 1-321 of PTP1B, but only residues 1-298 are visible in the electron density, even in the closed state with $\alpha 7$ ordered; the remaining residues are always disordered in an open region of bulk solvent within the crystal lattice. The wealth of types and extents of conformational heterogeneity it features make this a promising candidate for force field evaluation. Simulations of PTP1B based on 6B8X should be initiated from either the open state or the closed state of the protein and should be assessed on the basis of their ability to recapitulate the allosteric coupling observed in various experiments: as the WPD loop opens, the allosteric network should shift toward the corresponding open-like state, and $\alpha 7$ should become disordered. In addition to unbiased simulations [323], one could perform biased or

guided simulations to enforce a shift in one region (WPD or α7), then examine whether the other region allosterically responds as expected [324].

4.2.7 Endoglucanase

Combined X-ray and neutron-diffraction data are available for endoglucanase (EG) from *Phanerochaete chrysosporium* at room temperature (PDB ID 3X2P) [102]. As noted in Section 4.1.5, the neutron data make it possible to use experimentally derived protonation states for protein residues [297, 298, 301], removing a potential cause of modeling errors and uncertainties, while also allowing direct determination of the orientation of ordered waters and other factors. Water structure in this case is also much more clearly defined and amenable to careful curation [302, 303]. Note, however, that this structure includes a non-standard amino acid and an oligosaccharide ligand.

In one simulation study [303], the EG-cellopentaose complex from 3X2P was re-refined from 1.5 Å neutron and 1.0 Å X-ray diffraction data, with careful attention to placing H/D atoms based on neutron scattering data and H-bond interactions with the local environment. The re-refined structure featured several differences in protonation states with respect to the deposited model and the replacement of an imidic acid form for an asparagine side chain in the deposited model with canonical asparagine. This particular study focused on the ability of MD simulations to correctly position water molecules defined by the crystallographic studies. This was accomplished by computing electron density maps from crystalline simulations of a $2 \times 2 \times 2$ periodic supercell under several different solvent conditions, with several different restraints of the protein heavy atoms and ligand atoms. Clear indications of force field limitations were identified. For example, although recovery of experimentally observed crystallographic waters was over 90% when the protein was restrained modestly to crystallographic positions, this fraction dropped to 50% without restraints. The drop was traced to small, local protein motions that disrupted entire water networks [303], consistent with prior work [325].

While this study focused primarily on recovery of crystallographic water molecules, it suggests a route forward in terms of benchmarking force fields on this and similar systems [303]. The carefully curated, re-refined structures provide a valuable starting point and should be used as input for simulations. Force fields could be benchmarked by repeating the simulation approach used here while comparing results at different restraint strengths. Presumably, as force fields are improved, this will result in better recovery of crystallographic water structures at lower protein restraint strengths or with no restraints. Additional benchmarking

studies could more closely focus on how well various force fields preserve the structure of the protein.

4.2.8 Staphylococcal nuclease

Staphylococcal nuclease (SNase) has provided a particularly valuable dataset for diffuse scattering studies (PDB ID 4WOR) [103]. Until fairly recently, this was the only complete, high-quality, three-dimensional (i.e., sampled on a 3-D reciprocal space grid using data from many diffraction images obtained at different crystal orientations) diffuse scattering dataset available for a protein crystal [295]. Crystalline MD simulations greater than 1 µs in length allowed calculation of diffuse scattering intensity for direct comparison with experiment [295]. More recent work [280] extended these simulations to cover a $2 \times 2 \times 2$ supercell, with roughly 5 µs of data, and obtained improved agreement with experimental scattering data, perhaps because the simulations used a starting structure that modeled missing terminal residues and a bound ligand that better represented the system in the crystal experiment. Different force fields gave different levels of agreement with experimental diffuse scattering patterns [296], supporting the concept of using such data to benchmark protein force fields.

5 Best practices for setup and analysis of benchmark simulations

5.1 Setup of benchmark simulations

Simulations intended to benchmark protein force fields against experimental observables should strive to replicate the conditions under which those observables were measured as closely as possible. Below we provide specific recommendations to achieve this goal.

5.1.1 Initial coordinates

Benchmark simulations must be started from a set of initial coordinates for the protein. For folded proteins with structural models refined against experimental data, the initial coordinates can simply be those of the appropriate structural model. When both crystal and NMR models are available, one should select the model that corresponds to the desired conditions of the simulation, i.e., NMR models should be selected for simulations of proteins in dilute solution, while crystal models should be selected for simulations of protein crystals. For short peptides, the initial coordinates are often set to an extended conformation in which all backbone dihedrals are set to 180°. For disordered proteins, a structural model for a partially folded conformation can be used when available, and an extended conformation can be used for residues not included in the structural model. An extended conformation can also be used for the entire

disordered protein, although this setup requires a very large simulation box with many solvent molecules. To avoid large boxes, several methods [326–331] can be used to generate one or more initial conformations consistent with an expected radius of gyration.

While it is common practice to remove hydrogen atoms present in structural models and add them back using standard residue libraries, it is essential to assign correct protonation states to titratable side chains and protein termini. High-resolution crystal models, especially those refined against neutron diffraction data, can include the co-ordinates of hydrogen atoms and thus give high confidence in protonation states. Otherwise, the choice of protonation states should generally reflect the expected protonation based on the pH of the experiments used to measure the observables. However, the local environment of a residue—such as hydrogen bonds to nearby residues, long-range electrostatics, and solvation—may indicate a protonation state different from the expected one at the experimental pH. Similarly, the local environment of cysteine residues can indicate whether they should be modeled with a disulfide bond. Tools such as PROPKA [332] and PDB2PQR [333] can assign protonation states and disulfide bonds based on local environment and a specified pH.

5.1.2 Solvation and periodic box construction

Most benchmark simulations model water molecules explicitly, and the water model should be considered part of the force field being benchmarked just as much as the parameters applied to the protein. Indeed, the choice of force field parameters for the water can significantly impact the resulting protein ensembles [35, 235, 334]. In particular, the widely used 3-point water model TIP3P [226] favors compact structures excessively [35, 334] and has low viscosity compared to experimental water [335]. The consequences will be more pronounced for simulations of disordered proteins and for calculations of kinetic observables, such as NMR spin relaxation rates. In such applications, simulations using TIP3P should be expected to exhibit some disagreement with experiment due to the properties of the water model rather than the protein force field.

For crystal simulations, the simplest setup is a periodic box containing a single unit cell from the crystal. A single unit cell might be adequate for some Bragg data applications or for modeling the isotropic component of diffuse scattering data. Other applications might require a supercell in which the periodic simulation box includes multiple copies of the crystal unit cell that can move independently. For example, in diffuse scattering studies, using two copies of the unit cell along each lattice vector avoids artificial correlations between pairs of atoms that are separated by a distance

smaller than the unit cell. This setup can substantially increase the agreement with the anisotropic component of diffuse scattering data [280]. An appropriate choice of the supercell also is required for direct comparison of simulations with the full set of diffuse scattering measurements: in general, a full dataset with N_h, N_k, and N_l points sampled per reciprocal lattice vector along Miller indices hkl can be simulated using a box with N_h, N_k, and N_l copies of the unit cell along lattice vectors $\vec{a}$, $\vec{b}$, and $\vec{c}$. Representative lattices generated by this approach may be, e.g., $2 \times 2 \times 2$ or $7 \times 7 \times 5$.

For dilute solution simulations, the periodic box should contain the protein and a large enough number of solvent molecules to effectively screen interactions between periodic images of the protein (see next paragraph). Many system-building tools provide automated methods to construct periodic boxes filled with pre-equilibrated solvent around a solute, with the box geometry defined by user-provided shape and dimensions. While triclinic simulation boxes are straightforward to build and visualize, other polyhedra that can fill space using only translations can provide the same minimum distances between periodic images with lower volume and thus faster simulation speed. The rhombic dodecahedron [336] has the smallest volume of these space-filling polyhedra and is the recommended box shape for simulations of solutes that are roughly spherical, including most globular proteins.

The size of the periodic simulation box can be specified using either the total box length or the solvent padding, i.e., the shortest acceptable distance between a protein atom and the edge of the simulation box. If the solvent padding is shorter than 10 Å, enough space for two to three shells of water, then interactions between periodic images can perturb observed quantities such as solvation free energies and secondary structure preferences [337, 338]. Based on these considerations, for folded proteins that are expected to sample only compact structures during the simulation, a solvent padding of 10 Å to 15 Å is sufficient to avoid such artifacts. For disordered proteins, a larger solvent padding should be used to accommodate these less compact structures. A common practice for disordered proteins is to use a total box length that is twice the radius of gyration plus 20 Å.

5.1.3 Salt ions and co-solutes

For proteins with a net charge, counterions should be added to the simulation box to neutralize the system. If the observables targeted by the simulation were measured in a solution containing salt, then additional ion pairs should be added to the simulation box to model the experimental ionic strength of the bulk salt. If the protein has no net charge, the desired number of ion pairs can be estimated from the number of water molecules used to solvate the system N_{water}.

$$N_{\text{ions}} = \frac{N_{\text{water}}}{C_{\text{water}}} C_{\text{ions}} \tag{15}$$

where C_{water} is the expected concentration of bulk water in dilute solution, $55.4 \,\text{mol}\,\text{L}^{-1}$, and C_{ions} is the experimental concentration of bulk salt. If the protein has a net charge, then the solvent near the protein will be depleted of ions of like charge and enriched in ions of opposite charge, relative to bulk [339]. In this case, the desired number of ion pairs can be estimated using the SLTCAP method (this expression is equivalent to Eq. 5 in Schmit et al. [339]).

$$N_{\text{ions}} = \frac{N_{\text{water}}}{C_{\text{water}}} C_{\text{ions}} \left(\sqrt{1 + \lambda^2} - \lambda \right)$$
$$\lambda = \frac{|Q| \, C_{\text{water}}}{2 N_{\text{water}} C_{\text{ions}}} \tag{16}$$

In addition to salt, co-solutes such as buffering agents or crowding agents may be present in solutions used to measure observables. Such co-solutes are particularly important for protein crystals, which often require co-solutes in the mother liquor in order to induce crystallization. It is sometimes necessary to model co-solutes explicitly to accurately describe the solvent environment within protein crystals [303]. Modeling co-solutes explicitly can improve the fidelity of benchmark simulations, but care must be taken to ensure that appropriate force field parameters are available for the co-solutes and that sampling is long enough to observe changes in co-solute locations that can influence protein conformations and convergence of the protein ensemble.

5.1.4 Thermodynamic ensembles

Because the observables of interest here are typically measured at constant pressure and temperature, benchmark simulations should sample from the isobaric-isothermal (NPT) ensemble. Some commonly used methods for controlling the pressure and temperature in simulations produce samples from different, unphysical ensembles. In particular, Berendsen or weak-coupling methods fail to preserve the equipartition of energy and can lead to artifacts such as the flying ice cube effect [340, 341]. These methods should be avoided in favor of methods that sample correctly from physical ensembles, such as Langevin dynamics, the canonical velocity rescaling thermostat, and the Monte Carlo barostat. Additional care is needed when computing kinetic quantities, as some thermostats—most notably Langevin dynamics, with its added frictional damping and random forces—may introduce significant kinetic artifacts.

5.2 Analysis of benchmark simulations

5.2.1 Equilibration

It is recommended to discard the initial segment of a molecular dynamics simulation before collecting the frames that will be used to calculate observables. The discarded initial segment, often termed the equilibration or burn-in run, allows the system to relax from its initial configuration, which may be far from the equilibrium distribution dictated by the force field. Several methods are available to set the length of the equilibration run. One common practice is to monitor the volume or density of the barostatted system and to discard frames until this quantity reaches a plateau. However, for novel force fields or complex systems such as protein crystals, a plateau in volume can be reached before strain from the initial configuration can be dissipated. Thus, we recommend monitoring quantities such as potential energies as well to determine when a system is equilibrated. An equilibration time can also be chosen based on the autocorrelation function of a time series of an observable [342].

5.2.2 Convergence

Precise estimation of quantities from molecular dynamics simulations requires a large number of samples drawn from the system's ensemble. It is good practice to assess the convergence of simulations by determining whether estimates of observables are robust to changes in the set of conformational samples used to estimate them. A number of methods have been developed to assess convergence [343]. In one simple approach, the time series of observable estimates is visualized to check for drift over the course of the simulation. Blocking analysis, a more rigorous approach, divides the trajectory into blocks and looks for convergence of the standard error of the mean as the block size is increased [344].

Assessment of convergence will indicate when the populations of states visited during a simulation are not known precisely, but such analysis cannot identify when a simulation fails entirely to visit regions of a protein's configuration space that carry significant weight in the Boltzmann distribution. This failure mode can occur, for example, when high free energy barriers exist between minima that contribute to the protein's ensemble in the experimental measurements. Diagnosing this problem requires some knowledge of the proteins being simulated, either from experimental measurements or from a reference ensemble known to model the system well. If benchmark simulations of a particular protein consistently fail to visit relevant regions of the protein's configuration space across multiple force fields, enhanced sampling methods [317] can accelerate barrier crossing at the cost of additional complexity in obtaining estimates of observables and uncertainties [343].

5.2.3 Statistical analysis

When comparing sets of simulations against experimental data, it is essential to determine whether the differences are statistically significant [105]. Statistical uncertainties in simulated data are perhaps most reliable if obtained by running multiple independent simulations but can also be obtained by analysis of a single simulation [343]. The raw experimental data also have uncertainties, and additional uncertainty may result from model assumptions, such as the values of the Karplus parameters for ^3J-values or the representation of NOESY intensities as upper bounds to interatomic distances. When multiple proteins are simulated and several properties are considered, it may not be straightforward to determine if two sets of simulations are significantly different or even which simulation shows the overall better agreement with the available experimental data [79, 345, 346]. The differences in a particular property between different force fields will be affected by both the variability due to the choice of protein and the variability between independent replicate simulations. A statistical approach that takes mixed effects explicitly into account [79] may be most appropriate.

Author Contributions

All authors contributed to conceptualizing and defining the scope of the review article. Contributions to the initial draft of the manuscript were as follows:

- Section 2 on recommendations by C.E.C. and M.K.G.
- Section 3.1.1 on chemical shifts by L.T.C. and P.J.R.
- Section 3.1.2 on J couplings by P.J.R.
- Section 3.1.3 on RDCs by K.L.-L.
- Section 3.1.4 on NOESY by C.E.C. and C.O.
- Section 3.1.5 on spin relaxation by O.H.S.O.
- Section 3.1.6 on PREs by K.L.-L.
- Section 3.2.1 on the Beauchamp dataset by P.J.R.
- Section 3.2.2 on designed peptides by V.A.V.
- Section 3.2.3 on the Stroet dataset by C.O.
- Section 3.2.4 on the Mao dataset by P.J.R.
- Section 3.2.5 on the Robustelli dataset by P.J.R.
- Section 3.2.6 on spin relaxation datasets by O.H.S.O.
- Section 3.2.7 on the salt bridge stability dataset by L.T.C.
- Section 4.1.1 on Bragg diffraction by D.A.C.
- Section 4.1.2 on B-factors by J.C.-H.C.
- Section 4.1.3 on alternate conformations by D.A.K.
- Section 4.1.4 on diffuse scattering by M.E.W., D.C.W., and D.A.C.
- Section 4.1.5 on neutron diffraction by J.C.-H.C.
- Section 4.2.1 on scorpion toxin by D.A.C.
- Section 4.2.2 on lysozyme by D.A.C. and J.C.-H.C.
- Section 4.2.3 on crambin by J.C.-H.C.
- Section 4.2.4 on cyclophilin A by D.A.K.
- Section 4.2.5 on ubiquitin by D.A.K.
- Section 4.2.6 on PTP1B by D.A.K.
- Section 4.2.7 on endoglucanase by D.L.M.
- Section 4.2.8 on SNase by D.L.M. and M.E.W.
- Section 5 on best practices by C.E.C. and C.O.

All authors contributed to revisions of the initial draft. C.E.C. and M.K.G. coordinated and managed the writing and revision of the manuscript.

For a more detailed description of author contributions, see the GitHub issue tracking and changelog at https://github.com/openforcefield/review-protein-benchmark-datasets.

Other Contributions

The authors acknowledge Bernard Brooks, David Cerutti, Thomas Cheatham III, John Chodera, Gabriel Rocklin, Adrian Roitberg, Benoit Roux, Michael Shirts, and Junmei Wang for early discussions about benchmarking datasets.

For a more detailed description of contributions from the community and others, see the GitHub issue tracking and changelog at https://github.com/openforcefield/review-protein-benchmark-datasets.

Potentially Conflicting Interests

L.T.C. serves on the scientific advisory board of OpenEye Scientific Software. K.L.-L. holds stock options in and is a consultant for Peptone Ltd. D.L.M. serves on the scientific advisory board of OpenEye Scientific Software and is an Open Science Fellow with Psivant Sciences. M.E.W. is a consultant for Eli Lilly and Company. M.K.G. has an equity interest in and is a cofounder and scientific advisor of VeraChem LLC and a member of the Scientific Advisory Boards of InCerebro Co., Ltd., Denovicon Therapeutics, Beren Therapeutics, and Cold Start Therapeutics.

Funding Information

Author Information

ORCID:
Chapin E. Cavender: 0000-0002-5899-7953
David A. Case: 0000-0003-2314-2346
Julian C.-H. Chen: 0000-0003-0341-165X
Lillian T. Chong: 0000-0002-0590-483X
Daniel A. Keedy: 0000-0002-9184-7586
Kresten Lindorff-Larsen: 0000-0002-4750-6039
David L. Mobley: 0000-0002-1083-5533
O.H. Samuli Ollila: 0000-0002-8728-1006
Chris Oostenbrink: 0000-0002-4232-2556
Paul Robustelli: 0000-0002-9282-8993
Vincent A. Voelz: 0000-0002-1054-2124
Michael E. Wall: 0000-0003-1000-688X
David C. Wych: 0000-0001-9209-4371
Michael K. Gilson: 0000-0002-3375-1738

References

[1] **McCammon JA**, Gelin BR, Karplus M. Dynamics of Folded Proteins. Nature. 1977; 267(5612):585–590. https://doi.org/10.1038/267585a0.

[2] **Case DA**, Karplus M. Dynamics of Ligand Binding to Heme Proteins. J Mol Biol. 1979; 132(3):343–368. https://doi.org/10.1016/0022-2836(79)90265-1.

[3] **Levitt M**, Sharon R. Accurate Simulation of Protein Dynamics in Solution. Proc Nat Acad Sci. 1988; 85(20):7557–7561. https://doi.org/10.1073/pnas.85.20.7557.

[4] **Stone JE**, Phillips JC, Freddolino L, Hardy DJ, Trabuco LG, Schulten K. Accelerating Molecular Modeling Applications with Graphics Processors. J Comput Chem. 2007; 28(16):2618–2640. https://doi.org/10.1002/jcc.20829.

[5] **Shaw DE**, Deneroff MM, Dror RO, Kuskin JS, Larson RH, Salmon JK, Young C, Batson B, Bowers KJ, Chao JC, Eastwood MP, Gagliardo J, Grossman JP, Ho CR, Ierardi DJ, Kolossváry I, Klepeis JL, Layman T, McLeavey C, Moraes MA, et al. Anton, a Special-purpose Machine for Molecular Dynamics Simulation. Commun ACM. 2008; 51(7):91–97. https://doi.org/10.1145/1364782.1364802.

[6] **Darden T**, York D, Pedersen L. Particle Mesh Ewald: An N Log(N) Method for Ewald Sums in Large Systems. J Chem Phys. 1993; 98(12):10089–10092. https://doi.org/10.1063/1.464397.

[7] **Wennberg CL**, Murtola T, Páll S, Abraham MJ, Hess B, Lindahl E. Direct-Space Corrections Enable Fast and Accurate Lorentz–Berthelot Combination Rule Lennard-Jones Lattice Summation. J Chem Theory Comput. 2015; 11(12):5737–5746. https://doi.org/10.1021/acs.jctc.5b00726.

[8] **Grant BJ**, Gorfe AA, McCammon JA. Large Conformational Changes in Proteins: Signaling and Other Functions. Curr Opin Struct Biol. 2010; 20(2):142–147. https://doi.org/10.1016/j.sbi.2009.12.004.

[9] **Anandakrishnan R**, Drozdetski A, Walker RC, Onufriev AV. Speed of Conformational Change: Comparing Explicit and Implicit Solvent Molecular Dynamics Simulations. Biophys J. 2015; 108(5):1153–1164. https://doi.org/10.1016/j.bpj.2014.12.047.

[10] **Gilson MK**, Zhou HX. Calculation of Protein-Ligand Binding Affinities. Annu Rev Biophys. 2007; 36(1):21–42. https://doi.org/10.1146/annurev.biophys.36.040306.132550.

[11] **Wang K**, Chodera JD, Yang Y, Shirts MR. Identifying Ligand Binding Sites and Poses Using GPU-accelerated Hamiltonian Replica Exchange Molecular Dynamics. J Comput-Aided Mol Des. 2013; 27(12):989–1007. https://doi.org/10.1007/s10822-013-9689-8.

[12] **Reif MM**, Oostenbrink C. Net Charge Changes in the Calculation of Relative Ligand-binding Free Energies via Classical Atomistic Molecular Dynamics Simulation. J Comput Chem. 2014; 35(3):227–243. https://doi.org/10.1002/jcc.23490.

[13] **Lindorff-Larsen K**, Piana S, Dror RO, Shaw DE. How Fast-Folding Proteins Fold. Science. 2011; 334(6055):517–520. https://doi.org/10.1126/science.1208351.

[14] **Shaw DE**, Maragakis P, Lindorff-Larsen K, Piana S, Dror RO, Eastwood MP, Bank JA, Jumper JM, Salmon JK, Shan Y, Wriggers W. Atomic-Level Characterization of the Structural Dynamics of Proteins. Science. 2010; 330(6002):341–346. https://doi.org/10.1126/science.1187409.

[15] **Saglam AS**, Chong LT. Protein–protein Binding Pathways and Calculations of Rate Constants Using Fully-continuous, Explicit-solvent Simulations. Chem Sci. 2019; 10(8):2360–2372. https://doi.org/10.1039/c8sc04811h.

[16] **Rohl CA**, Strauss CEM, Misura KMS, Baker D. Protein Structure Prediction Using Rosetta. In: *Numerical Computer Methods - Part D*, vol. 383 of Methods in Enzymology Academic Press; 2004.p. 66–93. https://doi.org/10.1016/S0076-6879(04)83004-0.

[17] **Jumper J**, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Žídek A, Potapenko A, Bridgland A, Meyer C, Kohl SAA, Ballard AJ, Cowie A, Romera-Paredes B, Nikolov S, Jain R, Adler J, Back T, et al. Highly Accurate Protein Structure Prediction with AlphaFold. Nature.

2021; 596(7873):583–589. https://doi.org/10.1038/s41586-021-03819-2.

[18] **Kryshtafovych A**, Schwede T, Topf M, Fidelis K, Moult J. Critical Assessment of Methods of Protein Structure Prediction (CASP)—Round XIV. Proteins. 2021; 89(12):1607–1617. https://doi.org/10.1002/prot.26237.

[19] **Frenkel D**, Smit B. Understanding Molecular Simulation: from Algorithms to Applications. Elsevier; 2001. https://doi.org/10.1016/B978-0-12-267351-1.X5000-7.

[20] **Jorgensen WL**, Tirado-Rives J. The OPLS [Optimized Potentials for Liquid Simulations] Potential Functions for Proteins, Energy Minimizations for Crystals of Cyclic Peptides and Crambin. J Am Chem Soc. 1988; 110(6):1657–1666. https://doi.org/10.1021/ja00214a001.

[21] **Cornell WD**, Cieplak P, Bayly CI, Gould IR, Merz KM, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. J Am Chem Soc. 1995; 117(19):5179–5197. https://doi.org/10.1021/ja00124a002.

[22] **MacKerell AD**, Bashford D, Bellott M, Dunbrack RL, Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S, Joseph-McCarthy D, Kuchnir L, Kuczera K, Lau FTK, Mattos C, Michnick S, Ngo T, Nguyen DT, Prodhom B, Reiher WE, et al. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. J Phys Chem B. 1998; 102(18):3586–3616. https://doi.org/10.1021/jp973084f.

[23] **Smith LJ**, Mark AE, Dobson CM, van Gunsteren WF. Comparison of MD Simulations and NMR Experiments for Hen Lysozyme. Analysis of Local Fluctuations, Cooperative Motions, and Global Changes. Biochemistry. 1995; 34(34):10918–10931. https://doi.org/10.1021/bi00034a026, publisher: American Chemical Society.

[24] **van der Spoel D**, van Buuren AR, Tieleman DP, Berendsen HJC. Molecular Dynamics Simulations of Peptides from BPTI: A Closer Look At Amide-Aromatic Interactions. J Biomol NMR. 1996; 8(3):229–238. https://doi.org/10.1007/bf00410322.

[25] **van der Spoel D**. The Solution Conformations of Amino Acids from Molecular Dynamics Simulations of Gly-X-Gly Peptides: Comparison with NMR Parameters. Biochem Cell Biol. 1998; 76(2-3):164–170. https://doi.org/10.1139/o98-025, publisher: NRC Research Press.

[26] **Stocker U**, van Gunsteren WF. Molecular Dynamics Simulation of Hen Egg White Lysozyme: A Test of the GROMOS96 Force Field Against Nuclear Magnetic Resonance Data. Proteins. 2000; 40(1):145–153. https://doi.org/10.1002/(sici)1097-0134(20000701)40:1<145::aid-prot160>3.3.co;2-p.

[27] **van der Spoel D**, Lindahl E. Brute-Force Molecular Dynamics Simulations of Villin Headpiece: Comparison with NMR Parameters. J Phys Chem B. 2003; 107(40):11178–11187. https://doi.org/10.1021/jp034108n, publisher: American Chemical Society.

[28] **Oostenbrink C**, Soares TA, van der Vegt NFA, van Gunsteren WF. Validation of the 53A6 GROMOS Force Field. Eur Biophys J. 2005; 34(4):273–284. https://doi.org/10.1007/s00249-004-0448-6.

[29] **Lindorff-Larsen K**, Piana S, Palmo K, Maragakis P, Klepeis JL, Dror RO, Shaw DE. Improved Side-chain Torsion Potentials for the Amber ff99SB Protein Force Field. Proteins. 2010; 78(8):1950–1958. https://doi.org/10.1002/prot.22711.

[30] **Best RB**, Zhu X, Shim J, Lopes PEM, Mittal J, Feig M, MacKerell AD. Optimization of the Additive CHARMM All-atom Protein Force Field Targeting Improved Sampling of the Backbone φ, ψ and Side-chain χ1 and χ2 Dihedral Angles. J Chem Theory Comput. 2012; 8(9):3257–3273. https://doi.org/10.1021/ct300400x.

[31] **Mackerell AD**, Feig M, Brooks CL. Extending the Treatment of Backbone Energetics in Protein Force Fields: Limitations of Gas-phase Quantum Mechanics in Reproducing Protein Conformational Distributions in Molecular Dynamics Simulations. J Comput Chem. 2004; 25(11):1400–1415. https://doi.org/10.1002/jcc.20065.

[32] **Hornak V**, Abel R, Okur A, Strockbine B, Roitberg A, Simmerling C. Comparison of Multiple Amber Force Fields and Development of Improved Protein Backbone Parameters. Proteins. 2006; 65(3):712–725. https://doi.org/10.1002/prot.21123.

[33] **Maier JA**, Martinez C, Kasavajhala K, Wickstrom L, Hauser KE, Simmerling C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. J Chem Theory Comput. 2015; 11(8):3696–3713. https://doi.org/10.1021/acs.jctc.5b00255.

[34] **Diem M**, Oostenbrink C. Hamiltonian Reweighing To Refine Protein Backbone Dihedral Angle Parameters in the GROMOS Force Field. J Chem Inf Model. 2020; 60(1):279–288. https://doi.org/10.1021/acs.jcim.9b01034.

[35] **Tian C**, Kasavajhala K, Belfon KAA, Raguette L, Huang H, Migues AN, Bickel J, Wang Y, Pincay J, Wu Q, Simmerling C. ff19SB: Amino-Acid-Specific Protein Backbone Parameters Trained Against Quantum Mechanics Energy Surfaces in Solution. J Chem Theory Comput. 2020; 16(1):528–552. https://doi.org/10.1021/acs.jctc.9b00591.

[36] **Lifson S**. Recent Developments in the Consistent Force Field Calculations. In: Sadron C, editor. *Dynamic Aspects of Conformation Changes in Biological Macromolecules* Dordrecht: Springer Netherlands; 1973. p. 421–429. https://doi.org/10.1007/978-94-010-2579-9_28.

[37] **Ponder JW**, Case DA. Force Fields for Protein Simulations. In: *Advances in Protein Chemistry*, vol. 66 of Protein Simulations Academic Press; 2003.p. 27–85. https://doi.org/10.1016/S0065-3233(03)66002-X.

[38] **Guvench O**, MacKerell AD. Comparison of Protein Force Fields for Molecular Dynamics Simulations. In: Kukol A, editor. *Molecular Modeling of Proteins* Methods Molecular Biology, Totowa, NJ: Humana Press; 2008.p. 63–88. https://doi.org/10.1007/978-1-59745-177-2_4.

[39] **Zhu X**, Lopes PEM, MacKerell AD. Recent Developments and Applications of the CHARMM Force Fields. WIREs Comput Mol Sci. 2012; 2(1):167–185. https://doi.org/10.1002/wcms.74.

[40] **Dauber-Osguthorpe P**, Hagler AT. Biomolecular Force Fields: Where Have We Been, Where Are We Now, Where Do We Need to Go and How Do We Get There? J Comput-Aided Mol Des. 2019; 33(2):133–203. https://doi.org/10.1007/s10822-018-0111-4.

[41] **Hagler AT**. Force Field Development Phase II: Relaxation of Physics-based Criteria... or Inclusion of More Rigorous Physics into the Representation of Molecular Energetics. J Comput-Aided Mol Des. 2019; 33(2):205–264. https://doi.org/10.1007/s10822-018-0134-x.

[42] **Janson G**, Feig M. Generation of Protein Dynamics by Machine Learning. Curr Opin Struct Biol. 2025; 93:103115. https://doi.org/10.1016/j.sbi.2025.103115.

[43] **Brownless ALR**, Yehorova D, Welsh CL, Kamerlin SCL. Generative AI Techniques for Conformational Diversity and Evolutionary Adaptation of Proteins. Curr Opin Struct Biol. 2025; 94:103135. https://doi.org/10.1016/j.sbi.2025.103135.

[44] **Beauchamp KA**, Lin YS, Das R, Pande VS. Are Protein Force Fields Getting Better? A Systematic Benchmark on 524 Diverse NMR Measurements. J Chem Theory Comput. 2012; 8(4):1409–1414. https://doi.org/10.1021/ct2007814.

[45] **Blanco FJ**, Rivas G, Serrano L. A Short Linear Peptide That Folds into a Native Stable β-Hairpin in Aqueous Solution. Nat Struct Mol Biol. 1994; 1(9):584–590. https://doi.org/10.1038/nsb0994-584.

[46] **Ramírez-Alvarado M**, Blanco FJ, Serrano L. De Novo Design and Structural Analysis of a Model β-Hairpin Peptide System. Nat Struct Biol. 1996; 3(7):604–612. https://doi.org/10.1038/nsb0796-604.

[47] **de Alba E**, Jiménez MA, Rico M. Turn Residue Sequence Determines β-Hairpin Conformation in Designed Peptides. J Am Chem Soc. 1997; 119(1):175–183. https://doi.org/10.1021/ja962325e.

[48] **Maynard AJ**, Sharman GJ, Searle MS. Origin of β-Hairpin Stability in Solution: Structural and Thermodynamic Analysis of the Folding of a Model Peptide Supports Hydrophobic Stabilization in Water. J Am Chem Soc. 1998; 120(9):1996–2007. https://doi.org/10.1021/ja9726769.

[49] **Stanger HE**, Gellman SH. Rules for Antiparallel β-Sheet Design: D-Pro-Gly Is Superior to L-Asn-Gly for β-Hairpin Nucleation. J Am Chem Soc. 1998; 120(17):4236–4237. https://doi.org/10.1021/ja973704q.

[50] **Cochran AG**, Skelton NJ, Starovasnik MA. Tryptophan Zippers: Stable, Monomeric β-Hairpins. Proc Nat Acad Sci. 2001; 98(10):5578–5583. https://doi.org/10.1073/pnas.091100898.

[51] **Ramírez-Alvarado M**, Blanco FJ, Serrano L. Elongation of the BH8 β-Hairpin Peptide: Electrostatic Interactions in β-Hairpin Formation and Stability. Prot Sci. 2001; 10(7):1381–1392. https://doi.org/10.1110/ps.52901.

[52] **Pastor MT**, López de la Paz M, Lacroix E, Serrano L, Pérez-Payá E. Combinatorial Approaches: A New Tool to Search for Highly Structured β-Hairpin Peptides. Proc Nat Acad Sci. 2002; 99(2):614–619. https://doi.org/10.1073/pnas.012583999.

[53] **Russell SJ**, Blandl T, Skelton NJ, Cochran AG. Stability of Cyclic β-Hairpins: Asymmetric Contributions from Side Chains of a Hydrogen-Bonded Cross-Strand Residue Pair. J Am Chem Soc. 2003; 125(2):388–395. https://doi.org/10.1021/ja028075l.

[54] **Skelton NJ**, Blandl T, Russell SJ, Starovasnik MA, Cochran AG. β-Hairpin Polypeptides by Design and Selection. J Spectrosc. 2003; 17(2-3):213–230. https://doi.org/10.1155/2003/148024.

[55] **Fesinmeyer RM**, Hudson FM, Andersen NH. Enhanced Hairpin Stability Through Loop Design: The Case of the Protein G B1 Domain Hairpin. J Am Chem Soc. 2004; 126(23):7238–7243. https://doi.org/10.1021/ja0379520.

[56] **Espinosa JF**, Syud FA, Gellman SH. An Autonomously Folding β-Hairpin Derived from the Human YAP65 WW Domain: Attempts to Define a Minimum Ligand-binding Motif. Pept Sci. 2005; 80(2-3):303–311. https://doi.org/10.1002/bip.20205.

[57] **Andersen NH**, Olsen KA, Fesinmeyer RM, Tan X, Hudson FM, Eidenschink LA, Farazi SR. Minimization and Optimization of Designed β-Hairpin Folds. J Am Chem Soc. 2006; 128(18):6101–6110. https://doi.org/10.1021/ja054971w.

[58] **Honda S**, Akiba T, Kato YS, Sawada Y, Sekijima M, Ishimura M, Ooishi A, Watanabe H, Odahara T, Harata K. Crystal Structure of a Ten-amino Acid Protein. J Am Chem Soc. 2008; 130(46):15327–15331. https://doi.org/10.1021/ja8030533.

[59] **Kier BL**, Shu I, Eidenschink LA, Andersen NH. Stabilizing Capping Motif for β-Hairpins and Sheets. Proc Nat Acad Sci. 2010; 107(23):10466–10471. https://doi.org/10.1073/pnas.0913534107.

[60] **Scian M**, Shu I, Olsen KA, Hassam K, Andersen NH. Mutational Effects on the Folding Dynamics of a Minimized Hairpin. Biochemistry. 2013; 52(15):2556–2564. https://doi.org/10.1021/bi400146c.

[61] **Jiménez MA**. Design of Monomeric Water-soluble β-Hairpin and β-Sheet Peptides. In: Köhler V, editor. *Protein Design: Methods and Applications* New York, NY: Springer; 2014.p. 15–52. https://doi.org/10.1007/978-1-4939-1486-9_2.

[62] **Anderson JM**, Jurban B, Huggins KNL, Shcherbakov AA, Shu I, Kier B, Andersen NH. Nascent Hairpins in Proteins: Identifying Turn Loci and Quantitating Turn Contributions to Hairpin Stability. Biochemistry. 2016; 55(39):5537–5553. https://doi.org/10.1021/acs.biochem.6b00732.

[63] **Morales P**, Jiménez MA. Design and Structural Characterisation of Monomeric Water-soluble α-Helix and β-Hairpin Peptides: State-of-the-art. Arch Biochem Biophys. 2019; 661:149–167. https://doi.org/10.1016/j.abb.2018.11.014.

[64] **Richaud AD**, Zhao G, Hobloss S, Roche SP. Folding in Place: Design of β-Strap Motifs to Stabilize the Folding of Hairpins with Long Loops. J Org Chem. 2021; 86(19):13535–13547. https://doi.org/10.1021/acs.joc.1c01442.

[65] **Peintner S**, Erdélyi M. Pushing the Limits of Characterising a Weak Halogen Bond in Solution. Chem Eur J. 2022; 28(5):e202103559. https://doi.org/10.1002/chem.202103559.

[66] **Struthers MD**, Cheng RP, Imperiali B. Design of a Monomeric 23-residue Polypeptide with Defined Tertiary Structure. Science. 1996; 271(5247):342–345. https://doi.org/10.1126/science.271.5247.342.

[67] **McKnight CJ**, Matsudaira PT, Kim PS. NMR Structure of the 35-residue Villin Headpiece Subdomain. Nat Struct Biol. 1997; 4(3):180–184. https://doi.org/10.1038/nsb0397-180.

[68] **Struthers M**, Ottesen JJ, Imperiali B. Design and NMR Analyses of Compact, Independently Folded BBA Motifs. Folding Des. 1998; 3(2):95–103. https://doi.org/10.1016/s1359-0278(98)00015-7.

[69] **Imperiali B**, Ottesen Jj. Uniquely Folded Mini-protein Motifs. J Pept Res. 1999; 54(3):177–184. https://doi.org/10.1034/j.1399-3011.1999.00121.x.

[70] **Vermeulen W**, Vanhaesebrouck P, Van Troys M, Verschueren M, Fant F, Goethals M, Ampe C, Martins JC, Borremans FAM. Solution Structures of the C-terminal Headpiece Subdomains of Human Villin and Advillin, Evaluation of Headpiece F-actin-binding Requirements. Prot Sci. 2004; 13(5):1276–1287. https://doi.org/10.1110/ps.03518104.

[71] **Cornilescu G**, Hadley EB, Woll MG, Markley JL, Gellman SH, Cornilescu CC. Solution Structure of a Small Protein Containing a Fluorinated Side Chain in the Core. Prot Sci. 2007; 16(1):14–19. https://doi.org/10.1110/ps.062557707.

[72] **Gronwald W**, Hohm T, Hoffmann D. Evolutionary Pareto-optimization of Stably Folding Peptides. BMC Bioinf. 2008; 9(1):109. https://doi.org/10.1186/1471-2105-9-109.

[73] **Neidigh JW**, Fesinmeyer RM, Andersen NH. Designing a 20-residue Protein. Nat Struct Biol. 2002; 9(6):425–430. https://doi.org/10.1038/nsb798.

[74] **Lin JC**, Barua B, Andersen NH. The Helical Alanine Controversy: An (Ala)6 Insertion Dramatically Increases Helicity. J Am Chem Soc. 2004; 126(42):13679–13684. https://doi.org/10.1021/ja047265o.

[75] **Barua B**, Lin JC, Williams VD, Kummler P, Neidigh JW, Andersen NH. The Trp-cage: Optimizing the Stability of a Globular Miniprotein. Protein Eng Des Sel. 2008; 21(3):171–185. https://doi.org/10.1093/protein/gzm082.

[76] **Scian M**, Lin JC, Le Trong I, Makhatadze GI, Stenkamp RE, Andersen NH. Crystal and NMR Structures of a Trp-cage Mini-protein Benchmark for Computational Fold Prediction. Proc Nat Acad Sci. 2012; 109(31):12521–12525. https://doi.org/10.1073/pnas.1121421109.

[77] **Byrne A**, Kier BL, Williams DV, Scian M, Andersen NH. Circular Permutation of the Trp-cage: Fold Rescue upon Addition of a Hydrophobic Staple. RSC Adv. 2013; 3(43):19824–19829. https://doi.org/10.1039/c3ra43674h.

[78] **Graham KA**, Byrne A, Son R, Andersen NH. Reversing the Typical PH Stability Profile of the Trp-cage. Biopolymers. 2019; 110(3):e23260. https://doi.org/10.1002/bip.23260.

[79] **Stroet M**, Setz M, Lee T, Malde AK, van den Bergen G, Sykacek P, Oostenbrink C, Mark AE. On the Validation of Protein Force Fields Based on Structural Criteria. J Phys Chem B. 2024; 128(19):4602–4620. https://doi.org/10.1021/acs.jpcb.3c08469.

[80] **Mao B**, Tejero R, Baker D, Montelione GT. Protein NMR Structures Refined with Rosetta Have Higher Accuracy Relative to Corresponding X-ray Crystal Structures. J Am Chem Soc. 2014; 136(5):1893–1906. https://doi.org/10.1021/ja409845w.

[81] **Robustelli P**, Piana S, Shaw DE. Developing a Molecular Dynamics Force Field for Both Folded and Disordered Protein States. Proc Nat Acad Sci. 2018; 115(21):E4758–E4766. https://doi.org/10.1073/pnas.1800690115.

[82] **Jarymowycz VA**, Stone MJ. Fast Time Scale Dynamics of Protein Backbones: NMR Relaxation Methods, Applications, and Functional Consequences. Chem Rev. 2006; 106(5):1624–1671. https://doi.org/10.1021/cr040421p.

[83] **Louhivuori M**, Otten R, Lindorff-Larsen K, Annila A. Conformational Fluctuations Affect Protein Alignment in Dilute Liquid Crystal Media. J Am Chem Soc. 2006; 128(13):4371–4376. https://doi.org/10.1021/ja0576334.

[84] **Ciragan A**, Backlund SM, Mikula KM, Beyer HM, Samuli Ollila OH, Iwaï H. NMR Structure and Dynamics of TonB Investigated by Scar-less Segmental Isotopic Labeling Using a Salt-inducible Split Intein. Front Chem. 2020; 8. https://doi.org/10.3389/fchem.2020.00136.

[85] **Khan SN**, Charlier C, Augustyniak R, Salvi N, Déjean V, Bodenhausen G, Lequin O, Pelupessy P, Ferrage F. Distribution of Pico- and Nanosecond Motions in Disordered Proteins from Nuclear Spin Relaxation. Biophys J. 2015; 109(5):988–999. https://doi.org/10.1016/j.bpj.2015.06.069.

[86] **Tomlinson JH**, Ullah S, Hansen PE, Williamson MP. Characterization of Salt Bridges to Lysines in the Protein G B1 Domain. J Am Chem Soc. 2009; 131(13):4674–4684. https://doi.org/10.1021/ja808223p.

[87] **Smith GD**, Blessing RH, Ealick SE, Fontecilla-Camps JC, Hauptman HA, Housset D, Langs DA, Miller R. Ab Initio Structure Determination and Refinement of a Scorpion Protein Toxin. Acta Crystallogr D. 1997; 53(5):551–557. https://doi.org/10.1107/s0907444997005386.

[88] **Ramanadham M**, Sieker LC, Jensen LH. Refinement of Triclinic Lysozyme: II. The Method of Stereochemically Restrained Least Squares. Acta Crystallogr B. 1990; 46(1):63–69. https://doi.org/10.1107/s0108768189009195.

[89] **Meisburger SP**, Case DA, Ando N. Diffuse X-ray Scattering from Correlated Motions in a Protein Crystal. Nat Commun. 2020; 11(1):1271. https://doi.org/10.1038/s41467-020-14933-6.

[90] **Walsh MA**, Schneider TR, Sieker LC, Dauter Z, Lamzin VS, Wilson KS. Refinement of Triclinic Hen Egg-white Lysozyme At Atomic Resolution. Acta Crystallogr D. 1998; 54(4):522–546. https://doi.org/10.1107/s0907444997013656.

[91] **Wang J**, Dauter M, Alkire R, Joachimiak A, Dauter Z. Triclinic Lysozyme At 0.65 Å Resolution. Acta Crystallogr D. 2007; 63(12):1254–1268. https://doi.org/10.1107/s0907444907054224.

[92] **Artymiuk PJ**, Blake CCF, Rice DW, Wilson KS. The Structures of the Monoclinic and Orthorhombic Forms of Hen Egg-white Lysozyme At 6 Å Resolution. Acta Crystallogr B. 1982; 38(3):778–783. https://doi.org/10.1107/s0567740882004075.

[93] **Meisburger SP**, Case DA, Ando N. Robust Total X-ray Scattering Workflow to Study Correlated Motion of Proteins in Crystals. Nat Commun. 2023; 14(1):1228. https://doi.org/10.1038/s41467-023-36734-3.

[94] **Teeter MM**. Water Structure of a Hydrophobic Protein At Atomic Resolution: Pentagon Rings of Water Molecules in Crystals of Crambin. Proc Nat Acad Sci. 1984; 81(19):6014–6018. https://doi.org/10.1073/pnas.81.19.6014.

[95] **Chen JCH**, Fisher Z, Kovalevsky AY, Mustyakimov M, Hanson BL, Zhurov VV, Langan P. Room-temperature Ultrahigh-resolution Time-of-flight Neutron and X-ray Diffraction Studies of H/D-exchanged Crambin. Acta Crystallogr F. 2012; 68(2):119–123. https://doi.org/10.1107/s1744309111051499.

[96] **Chen JCH**, Gilski M, Chang C, Borek D, Rosenbaum G, Lavens A, Otwinowski Z, Kubicki M, Dauter Z, Jaskolski M, Joachimiak A. Solvent Organization in the Ultrahigh-resolution Crystal Structure of Crambin At Room Temperature. IUCrJ. 2024; 11(5):649–663. https://doi.org/10.1107/s2052252524007784.

[97] **Keedy DA**, Kenner LR, Warkentin M, Woldeyes RA, Hopkins JB, Thompson MC, Brewster AS, Van Benschoten AH, Baxter EL, Uervirojnangkoorn M, McPhillips SE, Song J, Alonso-Mori R, Holton JM, Weis WI, Brunger AT, Soltis SM, Lemke H, Gonzalez A, Sauter NK, et al. Mapping the Conformational Landscape of a Dynamic Enzyme by Multitemperature and XFEL Crystallography. eLife. 2015; 4:e07574. https://doi.org/10.7554/elife.07574.

[98] **Van Benschoten AH**, Liu L, Gonzalez A, Brewster AS, Sauter NK, Fraser JS, Wall ME. Measuring and Modeling Diffuse Scattering in Protein X-ray Crystallography. Proc Nat Acad Sci. 2016; 113(15):4069–4074. https://doi.org/10.1073/pnas.1524048113.

[99] **Thompson MC**, Barad BA, Wolff AM, Sun Cho H, Schotte F, Schwarz DMC, Anfinrud P, Fraser JS. Temperature-jump Solution X-ray Scattering Reveals Distinct Motions in a Dynamic Enzyme. Nat Chem. 2019; 11(11):1058–1066. https://doi.org/10.1038/s41557-019-0329-3.

[100] **Biel JT**, Thompson MC, Cunningham CN, Corn JE, Fraser JS. Flexibility and Design: Conformational Heterogeneity Along the Evolutionary Trajectory of a Redesigned Ubiquitin. Structure. 2017; 25(5):739–749.e3. https://doi.org/10.1016/j.str.2017.03.009.

[101] **Keedy DA**, Hill ZB, Biel JT, Kang E, Rettenmaier TJ, Brandão-Neto J, Pearce NM, von Delft F, Wells JA, Fraser JS. An Expanded Allosteric Network in PTP1B by Multitemperature Crystallography, Fragment Screening, and Covalent Tethering. eLife. 2018; 7:e36307. https://doi.org/10.7554/elife.36307.

[102] **Nakamura A**, Ishida T, Kusaka K, Yamada T, Fushinobu S, Tanaka I, Kaneko S, Ohta K, Tanaka H, Inaka K, Higuchi Y, Niimura N, Samejima M, Igarashi K. "Newton's Cradle" Proton Relay with Amide–imidic Acid Tautomerization in Inverting Cellulase Visualized by Neutron Crystallography. Sci Adv. 2015; 1(7):e1500263. https://doi.org/10.1126/sciadv.1500263.

[103] **Wall ME**, Ealick SE, Gruner SM. Three-dimensional Diffuse X-ray Scattering from Crystals of Staphylococcal Nuclease. Proc Nat Acad Sci. 1997; 94(12):6180–6184. https://doi.org/10.1073/pnas.94.12.6180.

[104] **van Gunsteren WF**, Allison JR, Daura X, Dolenc J, Hansen N, Mark AE, Oostenbrink C, Rusu VH, Smith LJ. Deriving Structural Information from Experimentally Measured Data on Biomolecules. Angew Chem. 2016; 55(52):15990–16010. https://doi.org/10.1002/anie.201601828.

[105] **van Gunsteren WF**, Daura X, Hansen N, Mark AE, Oostenbrink C, Riniker S, Smith LJ. Validation of Molecular Simulation: An Overview of Issues. Angew Chem. 2018; 57(4):884–902. https://doi.org/10.1002/anie.201702945.

[106] **Allison JR**. Assessing and Refining Molecular Dynamics Simulations of Proteins with Nuclear Magnetic Resonance Data. Biophys Rev. 2012; 4(3):189–203. https://doi.org/10.1007/s12551-012-0087-6.

[107] **Wishart DS**, Sykes BD. [12] Chemical Shifts As a Tool for Structure Determination. In: *Nuclear Magnetic Resonance - Part C*, vol. 239 of Methods in Enzymology Academic Press; 1994.p. 363–392. https://doi.org/10.1016/S0076-6879(94)39014-2.

[108] **Wishart DS**, Case DA. [1] - Use of Chemical Shifts in Macromolecular Structure Determination. In: James TL, Dötsch V, Schmitz U, editors. *Nuclear Magnetic Resonance of Biological Macromolecules - Part A*, vol. 338 of Methods in Enzymology Academic Press; 2002.p. 3–34. https://doi.org/10.1016/S0076-6879(02)38214-4.

[109] **Neal S**, Nip AM, Zhang H, Wishart DS. Rapid and Accurate Calculation of Protein 1H, 13C and 15N Chemical Shifts. J Biomol NMR. 2003; 26(3):215–240. https://doi.org/10.1023/a:1023812930288.

[110] **Xu XP**, Case DA. Automated Prediction of 15N, 13Cα, 13Cβ and 13C' Chemical Shifts in Proteins Using a Density Functional Database. J Biomol NMR. 2001; 21(4):321–333. https://doi.org/10.1023/a:1013324104681.

[111] **Meiler J**. PROSHIFT: Protein Chemical Shift Prediction Using Artificial Neural Networks. J Biomol NMR. 2003; 26(1):25–37. https://doi.org/10.1023/a:1023060720156.

[112] **Kohlhoff KJ**, Robustelli P, Cavalli A, Salvatella X, Vendruscolo M. Fast and Accurate Predictions of Protein NMR Chemical Shifts from Interatomic Distances. J Am Chem Soc. 2009; 131(39):13894–13895. https://doi.org/10.1021/ja903772t.

[113] **Shen Y**, Bax A. SPARTA+: a Modest Improvement in Empirical NMR Chemical Shift Prediction by Means of an Artificial Neural Network. J Biomol NMR. 2010; 48(1):13–22. https://doi.org/10.1007/s10858-010-9433-9.

[114] **Han B**, Liu Y, Ginzinger SW, Wishart DS. SHIFTX2: Significantly Improved Protein Chemical Shift Prediction. J Biomol NMR. 2011; 50(1):43–57. https://doi.org/10.1007/s10858-011-9478-4.

[115] **Li DW**, Brüschweiler R. PPM: a Side-chain and Backbone Chemical Shift Predictor for the Assessment of Protein Conformational Ensembles. J Biomol NMR. 2012; 54(3):257–265. https://doi.org/10.1007/s10858-012-9668-8.

[116] **Li D**, Brüschweiler R. PPM_One: a Static Protein Structure Based Chemical Shift Predictor. J Biomol NMR. 2015; 62(3):403–409. https://doi.org/10.1007/s10858-015-9958-z.

[117] **Li J**, Bennett KC, Liu Y, Martin MV, Head-Gordon T. Accurate Prediction of Chemical Shifts for Aqueous Protein Structure on "Real World" Data. Chem Sci. 2020; 11(12):3180–3191. https://doi.org/10.1039/c9sc06561j.

[118] **Yang Z**, Chakraborty M, White AD. Predicting Chemical Shifts with Graph Neural Networks. Chem Sci. 2021; 12(32):10802–10809. https://doi.org/10.1039/d1sc01895g.

[119] **Shen Y**, Bax A. Protein Backbone Chemical Shifts Predicted from Searching a Database for Torsion Angle and Sequence Homology. J Biomol NMR. 2007; 38(4):289–302. https://doi.org/10.1007/s10858-007-9166-6.

[120] **Cavalli A**, Salvatella X, Dobson CM, Vendruscolo M. Protein Structure Determination from NMR Chemical Shifts. Proc Nat Acad Sci. 2007; 104(23):9615–9620. https://doi.org/10.1073/pnas.0610313104.

[121] **Shen Y**, Lange O, Delaglio F, Rossi P, Aramini JM, Liu G, Eletsky A, Wu Y, Singarapu KK, Lemak A, Ignatchenko A, Arrowsmith CH, Szyperski T, Montelione GT, Baker D, Bax A. Consistent Blind Protein Structure Generation from NMR Chemical Shift Data. Proc Nat Acad Sci. 2008; 105(12):4685–4690. https://doi.org/10.1073/pnas.0800256105.

[122] **Wishart DS**, Arndt D, Berjanskii M, Tang P, Zhou J, Lin G. CS23D: a Web Server for Rapid Protein Structure Generation Using NMR Chemical Shifts and Sequence Data. Nucleic Acids Res. 2008; 36(Web Server):W496–W502. https://doi.org/10.1093/nar/gkn305.

[123] **Robustelli P**, Cavalli A, Dobson CM, Vendruscolo M, Salvatella X. Folding of Small Proteins by Monte Carlo Simulations with Chemical Shift Restraints Without the Use of Molecular Fragment Replacement or Structural Homology. J Phys Chem B. 2009; 113(22):7890–7896. https://doi.org/10.1021/jp900780b.

[124] **Robustelli P**, Kohlhoff K, Cavalli A, Vendruscolo M. Using NMR Chemical Shifts As Structural Restraints in Molecular Dynamics Simulations of Proteins. Structure. 2010; 18(8):923–933. https://doi.org/10.1016/j.str.2010.04.016.

[125] **Li DW**, Brüschweiler R. Certification of Molecular Dynamics Trajectories with NMR Chemical Shifts. J Phys Chem Lett. 2010; 1(1):246–248. https://doi.org/10.1021/jz9001345.

[126] **Markwick PRL**, Cervantes CF, Abel BL, Komives EA, Blackledge M, McCammon JA. Enhanced Conformational Space Sampling Improves the Prediction of Chemical Shifts in Proteins. J Am Chem Soc. 2010; 132(4):1220–1221. https://doi.org/10.1021/ja9093692.

[127] **Robustelli P**, Stafford KA, Palmer AG. Interpreting Protein Structural Dynamics from NMR Chemical Shifts. J Am Chem Soc. 2012; 134(14):6365–6374. https://doi.org/10.1021/ja300265w.

[128] **Robustelli P**, Trbovic N, Friesner RA, Palmer AG. Conformational Dynamics of the Partially Disordered Yeast Transcription Factor GCN4. J Chem Theory Comput. 2013; 9(11):5190–5200. https://doi.org/10.1021/ct400654r.

[129] **Li DW**, Brüschweiler R. NMR-based Protein Potentials. Angew Chem. 2010; 49(38):6778–6780. https://doi.org/10.1002/anie.201001898.

[130] **Ptaszek AL**, Li J, Konrat R, Platzer G, Head-Gordon T. UCBShift 2.0: Bridging the Gap from Backbone to Side Chain Protein Chemical Shift Prediction for Protein Structures. J Am Chem Soc. 2024; 146(46):31733–31745. https://doi.org/10.1021/jacs.4c10474.

[131] **Yi X**, Fritzsching KJ, Rogawski R, Xu Y, McDermott AE. Contribution of Protein Conformational Heterogeneity to NMR Lineshapes At Cryogenic Temperatures. Proc Nat Acad Sci. 2024; 121(8):e2301053120. https://doi.org/10.1073/pnas.2301053120.

[132] **Yi X**, Zhang L, Friesner RA, McDermott A. Predicted and Experimental NMR Chemical Shifts At Variable Temperatures: The Effect of Protein Conformational Dynamics. J Phys Chem Lett. 2024; 15(8):2270–2278. https://doi.org/10.1021/acs.jpclett.3c02589.

[133] **Mulder FAA**, Filatov M. NMR Chemical Shift Data and Ab Initio Shielding Calculations: Emerging Tools for Protein Structure Determination. Chem Soc Rev. 2010; 39(2):578–590. https://doi.org/10.1039/b811366c.

[134] **Frank A**, Möller HM, Exner TE. Toward the Quantum Chemical Calculation of NMR Chemical Shifts of Proteins. 2. Level of Theory, Basis Set, and Solvents Model Dependence. J Chem Theory Comput. 2012; 8(4):1480–1492. https://doi.org/10.1021/ct200913r.

[135] **Sumowski CV**, Hanni M, Schweizer S, Ochsenfeld C. Sensitivity of Ab Initio Vs Empirical Methods in Computing Structural Effects on NMR Chemical Shifts for the Example of Peptides. J Chem Theory Comput. 2014; 10(1):122–133. https://doi.org/10.1021/ct400713t.

[136] **Fritz M**, Quinn CM, Wang M, Hou G, Lu X, Koharudin LMI, Struppe J, Case DA, Polenova T, Gronenborn AM. Determination of Accurate Backbone Chemical Shift Tensors in Microcrystalline Proteins by Integrating MAS NMR and QM/MM. Phys Chem Chem Phys. 2018; 20(14):9543–9553. https://doi.org/10.1039/c8cp00647d.

[137] **Case DA**. Using Quantum Chemistry to Estimate Chemical Shifts in Biomolecules. Biophys Chem. 2020; 267:106476. https://doi.org/10.1016/j.bpc.2020.106476.

[138] **Chandy SK**, Thapa B, Raghavachari K. Accurate and Cost-effective NMR Chemical Shift Predictions for Proteins Using a Molecules-in-molecules Fragmentation-based Method. Phys Chem Chem Phys. 2020; 22(47):27781–27799. https://doi.org/10.1039/d0cp05064d.

[139] **Sahakyan AB**, Vranken WF, Cavalli A, Vendruscolo M. Structure-based Prediction of Methyl Chemical Shifts in Proteins. J Biomol NMR. 2011; 50(4):331–346. https://doi.org/10.1007/s10858-011-9524-2.

[140] **Sahakyan AB**, Vranken WF, Cavalli A, Vendruscolo M. Using Side-chain Aromatic Proton Chemical Shifts for a Quantitative Analysis of Protein Structures. Angew Chem. 2011; 123(41):9794–9797. https://doi.org/10.1002/ange.201101641.

[141] **Schwarzinger S**, Kroon GJA, Foss TR, Chung J, Wright PE, Dyson HJ. Sequence-dependent Correction of Random Coil NMR Chemical Shifts. J Am Chem Soc. 2001; 123(13):2970–2978. https://doi.org/10.1021/ja003760i.

[142] **De Simone A**, Cavalli A, Hsu STD, Vranken W, Vendruscolo M. Accurate Random Coil Chemical Shifts from an Analysis of Loop Regions in Native States of Proteins. J Am Chem Soc. 2009; 131(45):16332–16333. https://doi.org/10.1021/ja904937a.

[143] **Tamiola K**, Acar B, Mulder FAA. Sequence-specific Random Coil Chemical Shifts of Intrinsically Disordered Proteins. J Am Chem Soc. 2010; 132(51):18000–18003. https://doi.org/10.1021/ja105656t.

[144] **Kjaergaard M**, Poulsen FM. Sequence Correction of Random Coil Chemical Shifts: Correlation Between Neighbor Correction Factors and Changes in the Ramachandran Distribution. J Biomol NMR. 2011; 50(2):157–165. https://doi.org/10.1007/s10858-011-9508-2.

[145] **Kjaergaard M**, Poulsen FM. Disordered Proteins Studied by Chemical Shifts. Prog Nucl Magn Reson Spectrosc. 2012; 60:42–51. https://doi.org/10.1016/j.pnmrs.2011.10.001.

[146] **Ulrich EL**, Akutsu H, Doreleijers JF, Harano Y, Ioannidis YE, Lin J, Livny M, Mading S, Maziuk D, Miller Z, Nakatani E, Schulte CF, Tolmie DE, Kent Wenger R, Yao H, Markley JL. BioMagResBank. Nucleic Acids Res. 2007; 36(Database):D402–D408. https://doi.org/10.1093/nar/gkm957.

[147] **Romero PR**, Kobayashi N, Wedell JR, Baskaran K, Iwata T, Yokochi M, Maziuk D, Yao H, Fujiwara T, Kurusu G, Ulrich EL, Hoch JC, Markley JL. BioMagResBank (BMRB) as a Resource for Structural Biology. In: Gáspári Z, editor. *Structural Bioinformatics: Methods and Protocols* Methods in Molecular Biology, New York, NY: Springer US; 2020.p. 187–218. https://doi.org/10.1007/978-1-0716-0270-6_14.

[148] **Salmon L**, Nodet G, Ozenne V, Yin G, Jensen MR, Zweckstetter M, Blackledge M. NMR Characterization of Long-range Order in Intrinsically Disordered Proteins. J Am Chem Soc. 2010; 132(24):8407–8418. https://doi.org/10.1021/ja101645g.

[149] **Schwalbe M**, Ozenne V, Bibow S, Jaremko M, Jaremko L, Gajda M, Jensen MR, Biernat J, Becker S, Mandelkow E, Zweckstetter M, Blackledge M. Predictive Atomic Resolution Descriptions of Intrinsically Disordered hTau40 and α-Synuclein in Solution from NMR and Small Angle Scattering. Structure. 2014; 22(2):238–249. https://doi.org/10.1016/j.str.2013.10.020.

[150] **Ozenne V**, Schneider R, Yao M, Huang Jr, Salmon L, Zweckstetter M, Jensen MR, Blackledge M. Mapping the Potential Energy Landscape of Intrinsically Disordered Proteins At Amino Acid Resolution. J Am Chem Soc. 2012; 134(36):15138–15148. https://doi.org/10.1021/ja306905s.

[151] **Wood GPF**, Rothlisberger U. Secondary Structure Assignment of Amyloid-β Peptide Using Chemical Shifts. J Chem Theory Comput. 2011; 7(5):1552–1563. https://doi.org/10.1021/ct200156e.

[152] **Virtanen SI**, Kiirikki AM, Mikula KM, Iwaï H, Ollila OHS. Heterogeneous Dynamics in Partially Disordered Proteins. Phys Chem Chem Phys. 2020; 22(37):21185–21196. https://doi.org/10.1039/d0cp03473h.

[153] **Shrestha UR**, Smith JC, Petridis L. Full Structural Ensembles of Intrinsically Disordered Proteins from Unbiased Molecular Dynamics Simulations. Commun Biol. 2021; 4(1):1–8. https://doi.org/10.1038/s42003-021-01759-1.

[154] **Shalongo W**, Dugad L, Stellwagen E. Distribution of Helicity Within the Model Peptide Acetyl(AAQAA)3amide. J Am Chem Soc. 1994; 116(18):8288–8293. https://doi.org/10.1021/ja00097a039.

[155] **Shi Z**, Olson CA, Kallenbach NR, Sosnick TR. D/H Amide Isotope Effect in Model α-Helical Peptides. J Am Chem Soc. 2002; 124(47):13994–13995. https://doi.org/10.1021/ja027740n.

[156] **Best RB**, Hummer G. Optimized Molecular Dynamics Force Fields Applied to the Helix−Coil Transition of Polypeptides. J Phys Chem B. 2009; 113(26):9004–9015. https://doi.org/10.1021/jp901540t.

[157] **Lifson S**, Roig A. On the Theory of Helix—Coil Transition in Polypeptides. J Chem Phys. 1961; 34(6):1963–1974. https://doi.org/10.1063/1.1731802.

[158] **Donald JE**, Kulp DW, DeGrado WF. Salt Bridges: Geometrically Specific, Designable Interactions. Proteins. 2011; 79(3):898–915. https://doi.org/10.1002/prot.22927.

[159] **Karplus M**. Contact Electron-Spin Coupling of Nuclear Magnetic Moments. J Chem Phys. 1959; 30(1):11–15. https://doi.org/10.1063/1.1729860.

[160] **Karplus M**. Vicinal Proton Coupling in Nuclear Magnetic Resonance. J Am Chem Soc. 1963; 85(18):2870–2871. https://doi.org/10.1021/ja00901a059.

[161] **Bystrov VF**. Spin—spin Coupling and the Conformational States of Peptide Systems. Prog Nucl Magn Reson Spectrosc. 1976; 10(2):41–82. https://doi.org/10.1016/0079-6565(76)80001-5.

[162] **Haasnoot CAG**, de Leeuw FAAM, Altona C. The Relationship Between Proton-proton NMR Coupling Constants and Substituent Electronegativities—I: An Empirical Generalization of the Karplus Equation. Tetrahedron. 1980; 36(19):2783–2792. https://doi.org/10.1016/0040-4020(80)80155-4.

[163] **Vuister G**, Delaglio F, Bax A. The Use of 1JCαHα Coupling Constants As a Probe for Protein Backbone Conformation. J Biomol NMR. 1993; 3(1):67–80. https://doi.org/10.1007/bf00242476.

[164] **Cornilescu G**, Bax A, Case DA. Large Variations in One-Bond 13Cα−13Cβ J Couplings in Polypeptides Correlate with Backbone Conformation. J Am Chem Soc. 2000; 122(10):2168–2171. https://doi.org/10.1021/ja993609c.

[165] **Wirmer J**, Schwalbe H. Angular Dependence of 1J(Ni,Cα i) and 2J(Ni,Cα (i − 1)) Coupling Constants Measured in J-modulated HSQCs. J Biomol NMR. 2002; 23(1):47–55. https://doi.org/10.1023/a:1015384805098.

[166] **Ding K**, Gronenborn AM. Protein Backbone 1HN−13Cα and 15N−13Cα Residual Dipolar and J Couplings: New Constraints for NMR Structure Determination. J Am Chem Soc. 2004; 126(20):6232–6233. https://doi.org/10.1021/ja049049l.

[167] **Gapsys V**, Narayanan RL, Xiang S, de Groot BL, Zweckstetter M. Improved Validation of IDP Ensembles by One-bond Cα–Hα Scalar Couplings. J Biomol NMR. 2015; 63(3):299–307. https://doi.org/10.1007/s10858-015-9990-z.

[168] **Wang AC**, Bax A. Determination of the Backbone Dihedral Angles φ in Human Ubiquitin from Reparametrized Empirical Karplus Equations. J Am Chem Soc. 1996; 118(10):2483–2494. https://doi.org/10.1021/ja9535524.

[169] **Vögeli B**, Ying J, Grishaev A, Bax A. Limits on Variations in Protein Backbone Dynamics from Precise Measurements of Scalar Couplings. J Am Chem Soc. 2007; 129(30):9377–9385. https://doi.org/10.1021/ja070324o.

[170] **Pardi A**, Billeter M, Wüthrich K. Calibration of the Angular Dependence of the Amide Proton-Cα Proton Coupling Constants, 3JHNα, in a Globular Protein: Use of 3JHNα for Identification of Helical Secondary Structure. J Mol Biol. 1984; 180(3):741–751. https://doi.org/10.1016/0022-2836(84)90035-4.

[171] **Case DA**, Scheurer C, Brüschweiler R. Static and Dynamic Effects on Vicinal Scalar J Couplings in Proteins and Peptides: A MD/DFT Analysis. J Am Chem Soc. 2000; 122(42):10390–10397. https://doi.org/10.1021/ja001798p.

[172] **Lindorff-Larsen K**, Best RB, Vendruscolo M. Interpreting Dynamically-averaged Scalar Couplings in Proteins. J Biomol NMR. 2005; 32(4):273–280. https://doi.org/10.1007/s10858-005-8873-0.

[173] **Altona C**. Vicinal Coupling Constants and Conformation of Biomolecules. In: *eMagRes* John Wiley & Sons, Ltd; 2007.https://doi.org/10.1002/9780470034590.emrstm0587.

[174] **Wang B**, He X, Merz KM. Quantum Mechanical Study of Vicinal J Spin–Spin Coupling Constants for the Protein Backbone. J Chem Theory Comput. 2013; 9(10):4653–4659. https://doi.org/10.1021/ct400631b.

[175] **Lee JH**, Li F, Grishaev A, Bax A. Quantitative Residue-Specific Protein Backbone Torsion Angle Dynamics from Concerted Measurement of 3J Couplings. J Am Chem Soc. 2015; 137(4):1432–1435. https://doi.org/10.1021/ja512593s.

[176] **Li F**, Lee JH, Grishaev A, Ying J, Bax A. High Accuracy of Karplus Equations for Relating Three-bond J Couplings to Protein Backbone Torsion Angles. ChemPhysChem. 2015; 16(3):572–578. https://doi.org/10.1002/cphc.201402704.

[177] **Brueschweiler R**, Case DA. Adding Harmonic Motion to the Karplus Relation for Spin-spin Coupling. J Am Chem Soc. 1994; 116(24):11199–11200. https://doi.org/10.1021/ja00103a062.

[178] **Schmidt JM**, Blümel M, Löhr F, Rüterjans H. Self-consistent 3J Coupling Analysis for the Joint Calibration of Karplus Coefficients and Evaluation of Torsion Angles. J Biomol NMR. 1999; 14(1):1–12. https://doi.org/10.1023/a:1008345303942.

[179] **Pérez C**, Löhr F, Rüterjans H, Schmidt JM. Self-consistent Karplus Parametrization of 3J Couplings Depending on the Polypeptide Side-chain Torsion χ1. J Am Chem Soc. 2001; 123(29):7081–7093. https://doi.org/10.1021/ja003724j.

[180] **Chou JJ**, Case DA, Bax A. Insights into the Mobility of Methyl-bearing Side Chains in Proteins from 3JCC and 3JCN Couplings. J Am Chem Soc. 2003; 125(29):8959–8966. https://doi.org/10.1021/ja029972s.

[181] **Barfield M**. Structural Dependencies of Interresidue Scalar Coupling H3JNC' and Donor 1H Chemical Shifts in the Hydrogen Bonding Regions of Proteins. J Am Chem Soc. 2002; 124(15):4158–4168. https://doi.org/10.1021/ja012674v.

[182] **Graf J**, Nguyen PH, Stock G, Schwalbe H. Structure and Dynamics of the Homologous Series of Alanine Peptides: A Joint Molecular Dynamics/NMR Study. J Am Chem Soc. 2007; 129(5):1179–1189. https://doi.org/10.1021/ja0660406.

[183] **Best RB**, Buchete NV, Hummer G. Are Current Molecular Dynamics Force Fields Too Helical? Biophys J. 2008; 95(1):L07–L09. https://doi.org/10.1529/biophysj.108.132696.

[184] **Lindorff-Larsen K**, Maragakis P, Piana S, Eastwood MP, Dror RO, Shaw DE. Systematic Validation of Protein Force Fields Against Experimental Data. PLoS One. 2012; 7(2):e32131. https://doi.org/10.1371/journal.pone.0032131.

[185] **Piana S**, Lindorff-Larsen K, Shaw DE. How Robust Are Protein Folding Simulations with Respect to Force Field Parameterization? Biophys J. 2011; 100(9):L47–L49. https://doi.org/10.1016/j.bpj.2011.03.051.

[186] **Best RB**, Zheng W, Mittal J. Balanced Protein–water Interactions Improve Properties of Disordered Proteins and Non-Specific Protein Association. J Chem Theory Comput. 2014; 10(11):5113–5124. https://doi.org/10.1021/ct500569b.

[187] **Bax A**, Kontaxis G, Tjandra N. [8] - Dipolar Couplings in Macromolecular Structure Determination. In: *Nuclear Magnetic Resonance of Biological Macromolecules - Part B*, vol. 339 of Methods in Enzymology Academic Press; 2001.p. 127–174. https://doi.org/10.1016/s0076-6879(01)39313-8.

[188] **Bax A**. Weak Alignment Offers New NMR Opportunities to Study Protein Structure and Dynamics. Prot Sci. 2003; 12(1):1–16. https://doi.org/10.1110/ps.0233303.

[189] **Chiliveri SC**, Robertson AJ, Shen Y, Torchia DA, Bax A. Advances in NMR Spectroscopy of Weakly Aligned Biomolecular Systems. Chem Rev. 2022; 122(10):9307–9330. https://doi.org/10.1021/acs.chemrev.1c00730.

[190] **Sanders CR**, Prestegard JH. Magnetically Orientable Phospholipid Bilayers Containing Small Amounts of a Bile Salt Analogue, CHAPSO. Biophys J. 1990; 58(2):447–460. https://doi.org/10.1016/s0006-3495(90)82390-0.

[191] **Prosser RS**, Hwang JS, Vold RR. Magnetically Aligned Phospholipid Bilayers with Positive Ordering: A New Model Membrane System. Biophys J. 1998; 74(5):2405–2418. https://doi.org/10.1016/s0006-3495(98)77949-4.

[192] **Tycko R**, Blanco FJ, Ishii Y. Alignment of Biopolymers in Strained Gels: A New Way To Create Detectable Dipole–dipole Couplings in High-resolution Biomolecular NMR. J Am Chem Soc. 2000; 122(38):9340–9341. https://doi.org/10.1021/ja002133q.

[193] **Lakomek NA**, Walter KFA, Farès C, Lange OF, de Groot BL, Grubmüller H, Brüschweiler R, Munk A, Becker S, Meiler J, Griesinger C. Self-consistent Residual Dipolar Coupling Based Model-free Analysis for the Robust Determination of Nanosecond to Microsecond Protein Dynamics. J Biomol NMR. 2008; 41(3):139–155. https://doi.org/10.1007/s10858-008-9244-4.

[194] **Salvatella X**, Richter B, Vendruscolo M. Influence of the Fluctuations of the Alignment Tensor on the Analysis of the Structure and Dynamics of Proteins Using Residual Dipolar Couplings. J Biomol NMR. 2008; 40(1):71–81. https://doi.org/10.1007/s10858-007-9210-6.

[195] **Lindorff-Larsen K**, Best RB, DePristo MA, Dobson CM, Vendruscolo M. Simultaneous Determination of Protein Structure and Dynamics. Nature. 2005; 433(7022):128–132. https://doi.org/10.1038/nature03199.

[196] **Showalter SA**, Brüschweiler R. Quantitative Molecular Ensemble Interpretation of NMR Dipolar Couplings Without Restraints. J Am Chem Soc. 2007; 129(14):4158–4159. https://doi.org/10.1021/ja070658d.

[197] **Zweckstetter M**. NMR: Prediction of Molecular Alignment from Structure Using the PALES Software. Nat Protoc. 2008; 3(4):679–690. https://doi.org/10.1038/nprot.2008.36.

[198] **Marsh JA**, Baker JMR, Tollinger M, Forman-Kay JD. Calculation of Residual Dipolar Couplings from Disordered State Ensembles Using Local Alignment. J Am Chem Soc. 2008; 130(25):7804–7805. https://doi.org/10.1021/ja802220c.

[199] **Lange OF**, van der Spoel D, de Groot BL. Scrutinizing Molecular Mechanics Force Fields on the Submicrosecond Timescale with NMR Data. Biophys J. 2010; 99(2):647–655. https://doi.org/10.1016/j.bpj.2010.04.062.

[200] **Tjandra N**, Bax A. Direct Measurement of Distances and Angles in Biomolecules by NMR in a Dilute Liquid Crystalline Medium. Science. 1997; 278(5340):1111–1114. https://doi.org/10.1126/science.278.5340.1111.

[201] **Trigo-Mouriño P**, Navarro-Vázquez A, Ying J, Gil RR, Bax A. Structural Discrimination in Small Molecules by Accurate Measurement of Long-range Proton–carbon NMR Residual Dipolar Couplings. Angew Chem. 2011; 50(33):7576–7580. https://doi.org/10.1002/anie.201101739.

[202] **Cornilescu G**, Marquardt JL, Ottiger M, Bax A. Validation of Protein Structure from Anisotropic Carbonyl Chemical Shifts in a Dilute Liquid Crystalline Phase. J Am Chem Soc. 1998; 120(27):6836–6837. https://doi.org/10.1021/ja9812610.

[203] **Neuhaus D**, Williamson MP. The Nuclear Overhauser Effect in Structural and Conformational Analysis. Wiley; 2000.

[204] **Vögeli B**. The Nuclear Overhauser Effect from a Quantitative Perspective. Prog Nucl Magn Reson Spectrosc. 2014; 78:1–46. https://doi.org/10.1016/j.pnmrs.2013.11.001.

[205] **Smith LJ**, Sutcliffe MJ, Redfield C, Dobson CM. Structure of Hen Lysozyme in Solution. J Mol Biol. 1993; 229(4):930–944. https://doi.org/10.1006/jmbi.1993.1097.

[206] **Peter C**, Daura X, van Gunsteren WF. Calculation of NMR-relaxation Parameters for Flexible Molecules from Molecular Dynamics Simulations. J Biomol NMR. 2001; 20(4):297–310. https://doi.org/10.1023/a:1011241030461.

[207] **Wüthrich K**, Billeter M, Braun W. Pseudo-structures for the 20 Common Amino Acids for Use in Studies of Protein Conformations by Measurements of Intramolecular Proton-proton Distance Constraints with Nuclear Magnetic Resonance. J Mol Biol. 1983; 169(4):949–961. https://doi.org/10.1016/s0022-2836(83)80144-2.

[208] **Fletcher CM**, Jones DNM, Diamond R, Neuhaus D. Treatment of NOE Constraints Involving Equivalent or Non-stereoassigned Protons in Calculations of Biomacromolecular Structures. J Biomol NMR. 1996; 8(3):292–310. https://doi.org/10.1007/bf00410328.

[209] **Hoffmann F**, Mulder FAA, Schäfer LV. Accurate Methyl Group Dynamics in Protein Simulations with AMBER Force Fields. J Phys Chem B. 2018; 122(19):5038–5048. https://doi.org/10.1021/acs.jpcb.8b02769.

[210] **Redfield AG**. The Theory of Relaxation Processes. In: Waugh JS, editor. *Advances in Magnetic and Optical Resonance*, vol. 1 of Advances in Magnetic Resonance Academic Press; 1965.p. 1–32. https://doi.org/10.1016/B978-1-4832-3114-3.50007-6.

[211] **Best RB**, Vendruscolo M. Determination of Protein Structures Consistent with NMR Order Parameters. J Am Chem Soc. 2004; 126(26):8090–8091. https://doi.org/10.1021/ja0396955.

[212] **Showalter SA**, Johnson E, Rance M, Brüschweiler R. Toward Quantitative Interpretation of Methyl Side-Chain Dynamics from NMR by Molecular Dynamics Simulations. J Am Chem Soc. 2007; 129(46):14146–14147. https://doi.org/10.1021/ja075976r.

[213] **Showalter SA**, Brüschweiler R. Validation of Molecular Dynamics Simulations of Biomolecules Using NMR Spin Relaxation As Benchmarks: Application to the AMBER99SB Force Field. J Chem Theory Comput. 2007; 3(3):961–975. https://doi.org/10.1021/ct7000045.

[214] **Maragakis P**, Lindorff-Larsen K, Eastwood MP, Dror RO, Klepeis JL, Arkin IT, Jensen MO, Xu H, Trbovic N, Friesner RA, Palmer AG, Shaw DE. Microsecond Molecular Dynamics Simulation Shows Effect of Slow Loop Dynamics on Backbone Amide Order Parameters of Proteins. J Phys Chem B. 2008; 112(19):6155–6158. https://doi.org/10.1021/jp077018h.

[215] **Trbovic N**, Kim B, Friesner RA, Palmer AG. Structural Analysis of Protein Dynamics by MD Simulations and NMR Spin-relaxation. Proteins. 2008; 71(2):684–694. https://doi.org/10.1002/prot.21750.

[216] **Halle B**, Wennerström H. Interpretation of Magnetic Resonance Data from Water Nuclei in Heterogeneous Systems. J Chem Phys. 1981; 75(4):1928–1943. https://doi.org/10.1063/1.442218.

[217] **Lipari G**, Szabo A. Model-free Approach to the Interpretation of Nuclear Magnetic Resonance Relaxation in Macromolecules. 1. Theory and Range of Validity. J Am Chem Soc. 1982; 104(17):4546–4559. https://doi.org/10.1021/ja00381a009.

[218] **Nencini R**, Mantzari E, Sandelin A, Ollila S. Rapid Interpretation of Protein Backbone Rotation Dynamics Directly from Spin Relaxation Data. J Phys Chem Lett. 2024; 15(40):10204–10209. https://doi.org/10.26434/chemrxiv-2024-0w9ml.

[219] **Lipari G**, Szabo A, Levy RM. Protein Dynamics and NMR Relaxation: Comparison of Simulations with Experiment. Nature. 1982; 300(5888):197–198. https://doi.org/10.1038/300197a0.

[220] **Fenwick RB**, Dyson HJ. Classic Analysis of Biopolymer Dynamics Is Model Free. Biophys J. 2016; 110(1):3–6. https://doi.org/10.1016/j.bpj.2015.11.022.

[221] **Chen Pc**, Hologne M, Walker O, Hennig J. Ab Initio Prediction of NMR Spin Relaxation Parameters from Molecular Dynamics Simulations. J Chem Theory Comput. 2018; 14(2):1009–1019. https://doi.org/10.1021/acs.jctc.7b00750.

[222] **Ollila OHS**, Heikkinen HA, Iwaï H. Rotational Dynamics of Proteins from Spin Relaxation Times and Molecular Dynamics Simulations. J Phys Chem B. 2018; 122(25):6559–6569. https://doi.org/10.1021/acs.jpcb.8b02250.

[223] **Hoffmann F**, Mulder FAA, Schäfer LV. Predicting NMR Relaxation of Proteins from Molecular Dynamics Simulations with Accurate Methyl Rotation Barriers. J Chem Phys. 2020; 152(8):084102. https://doi.org/10.1063/1.5135379.

[224] **Nencini R**, Regnier MLG, Backlund SM, Mantzari E, Dunn CD, Ollila OHS. Probing the Dynamic Landscape of Peptides in Molecular Assemblies by Synergized NMR Experiments and MD Simulations. Commun Chem. 2024; 7(1):1–12. https://doi.org/10.1038/s42004-024-01115-4.

[225] **Bowman GR**. Accurately Modeling Nanosecond Protein Dynamics Requires At Least Microseconds of Simulation. J Comput Chem. 2016; 37(6):558–566. https://doi.org/10.1002/jcc.23973.

[226] **Jorgensen WL**, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of Simple Potential Functions for Simulating Liquid Water. J Chem Phys. 1983; 79(2):926–935. https://doi.org/10.1063/1.445869.

[227] **Wong V**, Case DA. Evaluating Rotational Diffusion from Protein MD Simulations. J Phys Chem B. 2008; 112(19):6013–6024. https://doi.org/10.1021/jp0761564.

[228] **Prompers JJ**, Brüschweiler R. General Framework for Studying the Dynamics of Folded and Nonfolded Proteins by NMR Relaxation Spectroscopy and MD Simulation. J Am Chem Soc. 2002; 124(16):4522–4534. https://doi.org/10.1021/ja012750u.

[229] **Anderson JS**, LeMaster DM. Rotational Velocity Rescaling of Molecular Dynamics Trajectories for Direct Prediction of Protein NMR Relaxation. Biophys Chem. 2012; 168-169:28–39. https://doi.org/10.1016/j.bpc.2012.05.005.

[230] **Lindorff-Larsen K**, Trbovic N, Maragakis P, Piana S, Shaw DE. Structure and Dynamics of an Unfolded Protein Examined by Molecular Dynamics Simulation. J Am Chem Soc. 2012; 134(8):3787–3791. https://doi.org/10.1021/ja209931w.

[231] **Sandelin A**, Nencini R, Yasar E, Ollila S. QEBSS: Quality Evaluation Based Simulation Selection for Analysis of Conformational Ensembles and Dynamics of Multidomain Proteins. ChemRxiv. 2024; https://doi.org/10.26434/chemrxiv-2024-h3pmt-v2.

[232] **Hoch JC**, Baskaran K, Burr H, Chin J, Eghbalnia HR, Fujiwara T, Gryk MR, Iwata T, Kojima C, Kurisu G, Maziuk D, Miyanoiri Y, Wedell JR, Wilburn C, Yao H, Yokochi M. Biological Magnetic Resonance Data Bank. Nucleic Acids Res. 2023; 51(D1):D368–D376. https://doi.org/10.1093/nar/gkac1050.

[233] **Clore GM**, Tang C, Iwahara J. Elucidating Transient Macromolecular Interactions Using Paramagnetic Relaxation Enhancement. Curr Opin Struct Biol. 2007; 17(5):603–616. https://doi.org/10.1016/j.sbi.2007.08.013.

[234] **Clore GM**, Iwahara J. Theory, Practice, and Applications of Paramagnetic Relaxation Enhancement for the Characterization of Transient Low-population States of Biological Macromolecules and Their Complexes. Chem Rev. 2009; 109(9):4108–4139. https://doi.org/10.1021/cr900033p.

[235] **Piana S**, Donchev AG, Robustelli P, Shaw DE. Water Dispersion Interactions Strongly Influence Simulated Structural Properties of Disordered Protein States. J Phys Chem B. 2015; 119(16):5113–5123. https://doi.org/10.1021/jp508971m.

[236] **Norgaard AB**, Ferkinghoff-Borg J, Lindorff-Larsen K. Experimental Parameterization of an Energy Function for the Simulation of Unfolded Proteins. Biophys J. 2008; 94(1):182–192. https://doi.org/10.1529/biophysj.107.108241.

[237] **Tesei G**, Schulze TK, Crehuet R, Lindorff-Larsen K. Accurate Model of Liquid–liquid Phase Behavior of Intrinsically Disordered Proteins from Optimization of Single-chain Properties. Proc Nat Acad Sci. 2021; 118(44):e2111696118. https://doi.org/10.1073/pnas.2111696118.

[238] **Polyhach Y**, Bordignon E, Jeschke G. Rotamer Libraries of Spin Labelled Cysteines for Protein Studies. Phys Chem Chem Phys. 2011; 13(6):2356–2366. https://doi.org/10.1039/c0cp01865a.

[239] **Michaud-Agrawal N**, Denning EJ, Woolf TB, Beckstein O. MDAnalysis: A Toolkit for the Analysis of Molecular Dynamics Simulations. J Comput Chem. 2011; 32(10):2319–2327. https://doi.org/10.1002/jcc.21787.

[240] **Gowers R**, Linke M, Barnoud J, Reddy T, Melo M, Seyler S, Domański J, Dotson D, Buchoux S, Kenney I, Beckstein O. MDAnalysis: A Python Package for the Rapid Analysis of Molecular Dynamics Simulations. In: *Proceedings of the 15th Python in Science Conference*; 2016. p. 98–105. https://doi.org/10.25080/majora-629e541a-00e.

[241] **Tesei G**, Martins JM, Kunze MBA, Wang Y, Crehuet R, Lindorff-Larsen K. DEER-PREdict: Software for Efficient Calculation of Spin-labeling EPR and NMR Data from Conformational Ensembles. PLoS Comput Biol. 2021; 17(1):e1008551. https://doi.org/10.1371/journal.pcbi.1008551.

[242] **Sasmal S**, Lincoff J, Head-Gordon T. Effect of a Paramagnetic Spin Label on the Intrinsically Disordered Peptide Ensemble of Amyloid-β. Biophys J. 2017; 113(5):1002–1011. https://doi.org/10.1016/j.bpj.2017.06.067.

[243] **Xue Y**, Skrynnikov NR. Motion of a Disordered Polypeptide Chain As Studied by Paramagnetic Relaxation Enhancements, 15N Relaxation, and Molecular Dynamics Simulations: How Fast Is Segmental Diffusion in Denatured Ubiquitin? J Am Chem Soc. 2011; 133(37):14614–14628. https://doi.org/10.1021/ja201605c.

[244] **Byrne A**, Williams DV, Barua B, Hagen SJ, Kier BL, Andersen NH. Folding Dynamics and Pathways of the Trp-cage Miniproteins. Biochemistry. 2014; 53(38):6011–6021. https://doi.org/10.1021/bi501021r.

[245] **Mark AE**, Stroet M, Nada S, Australasian Computational and Simulation Commons (ACSC); 2022. https://doi.org/10.48610/62d1f81.

[246] **Diem M**, Oostenbrink C. The Effect of Different Cutoff Schemes in Molecular Simulations of Proteins. J Comput Chem. 2020; 41(32):2740–2749. https://doi.org/10.1002/jcc.26426.

[247] **Wüthrich K**. NMR of Proteins and Nucleic Acids. Wiley; 1986.

[248] **Piana S**, Robustelli P, Tan D, Chen S, Shaw DE. Development of a Force Field for the Simulation of Single-chain Proteins and Protein–protein Complexes. J Chem Theory Comput. 2020; 16(4):2494–2507. https://doi.org/10.1021/acs.jctc.9b00251.

[249] **Borthakur K**, Sisk TR, Panei FP, Bonomi M, Robustelli P. Determining Accurate Conformational Ensembles of Intrinsically Disordered Proteins at Atomic Resolution. Nat Commun. 2025; 16(1):9036. https://doi.org/10.1038/s41467-025-64098-3, publisher: Nature Publishing Group.

[250] **Wayment-Steele HK**, El Nesr G, Hettiarachchi R, Kariyawasam H, Ovchinnikov S, Kern D. Learning Millisecond Protein Dynamics from What Is Missing in NMR Spectra. bioRxiv. 2025; https://doi.org/10.1101/2025.03.19.642801.

[251] **Hoffmann F**, Xue M, Schäfer LV, Mulder FAA. Narrowing the Gap Between Experimental and Computational Determination of Methyl Group Dynamics in Proteins. Phys Chem Chem Phys. 2018; 20(38):24577–24590. https://doi.org/10.1039/c8cp03915a, publisher: Royal Society of Chemistry.

[252] **Kümmerer F**, Orioli S, Harding-Larsen D, Hoffmann F, Gavrilov Y, Teilum K, Lindorff-Larsen K. Fitting Side-chain NMR Relaxation Data Using Molecular Simulations. J Chem Theory Comput. 2021; 17(8):5262–5275. https://doi.org/10.1021/acs.jctc.0c01338, publisher: American Chemical Society.

[253] **Ahmed MC**, Papaleo E, Lindorff-Larsen K. How Well Do Force Fields Capture the Strength of Salt Bridges in Proteins? PeerJ. 2018; 6:e4967. https://doi.org/10.7717/peerj.4967.

[254] **Debiec KT**, Gronenborn AM, Chong LT. Evaluating the Strength of Salt Bridges: A Comparison of Current Biomolecular Force Fields. J Phys Chem B. 2014; 118(24):6561–6569. https://doi.org/10.1021/jp500958r.

[255] **Debiec KT**, Cerutti DS, Baker LR, Gronenborn AM, Case DA, Chong LT. Further Along the Road Less Traveled: AMBER ff15ipq, an Original Protein Force Field Built on a Self-Consistent Physical Model. J Chem Theory Comput. 2016; 12(8):3926–3947. https://doi.org/10.1021/acs.jctc.6b00567.

[256] **Lopes PEM**, Huang J, Shim J, Luo Y, Li H, Roux B, MacKerell AD. Polarizable Force Field for Peptides and Proteins Based on the Classical Drude Oscillator. J Chem Theory Comput. 2013; 9(12):5430–5449. https://doi.org/10.1021/ct400781b.

[257] **Shi Y**, Xia Z, Zhang J, Best R, Wu C, Ponder JW, Ren P. Polarizable Atomic Multipole-Based AMOEBA Force Field for Proteins. J Chem Theory Comput. 2013; 9(9):4046–4063. https://doi.org/10.1021/ct4003702.

[258] **Muirhead H**, Perutz MF. Structure Of Hæmoglobin: A Three-dimensional Fourier Synthesis of Reduced Human Haemoglobin At 5.5 Å Resolution. Nature. 1963; 199(4894):633–638. https://doi.org/10.1038/199633a0.

[259] **Fischer M**. Macromolecular Room Temperature Crystallography. Q Rev Biophys. 2021; 54:e1. https://doi.org/10.1017/s0033583520000128.

[260] **Thorne RE**. Determining Biomolecular Structures Near Room Temperature Using X-ray Crystallography: Concepts, Methods and Future Optimization. Acta Crystallogr D. 2023; 79(1):78–94. https://doi.org/10.1107/s2059798322011652.

[261] **Fraser JS**, Clarkson MW, Degnan SC, Erion R, Kern D, Alber T. Hidden Alternative Structures of Proline Isomerase Essential for Catalysis. Nature. 2009; 462(7273):669–673. https://doi.org/10.1038/nature08615.

[262] **Keedy DA**, van den Bedem H, Sivak DA, Petsko GA, Ringe D, Wilson MA, Fraser JS. Crystal Cryocooling Distorts Conformational Heterogeneity in a Model Michaelis Complex of DHFR. Structure. 2014; 22(6):899–910. https://doi.org/10.1016/j.str.2014.04.016.

[263] **Bradford SYC**, El Khoury L, Ge Y, Osato M, Mobley DL, Fischer M. Temperature Artifacts in Protein Structures Bias Ligand-binding Predictions. Chem Sci. 2021; 12(34):11275–11293. https://doi.org/10.1039/d1sc02751d.

[264] **Altan I**, Fusco D, Afonine PV, Charbonneau P. Learning About Biomolecular Solvation from Water in Protein Crystals. J Phys Chem B. 2018; 122(9):2475–2486. https://doi.org/10.1021/acs.jpcb.7b09898.

[265] **Cerutti DS**, Case DA. Molecular Dynamics Simulations of Macromolecular Crystals. WIREs Comput Mol Sci. 2019; 9(4):e1402. https://doi.org/10.1002/wcms.1402.

[266] **Wych DC**, Aoto PC, Vu L, Wolff AM, Mobley DL, Fraser JS, Taylor SS, Wall ME. Molecular-dynamics Simulation Methods for Macromolecular Crystallography. Acta Crystallogr D. 2023; 79(1):50–65. https://doi.org/10.1107/s2059798322011871.

[267] **Meisburger SP**, Thomas WC, Watkins MB, Ando N. X-ray Scattering Studies of Protein Structural Dynamics. Chem Rev. 2017; 117(12):7615–7672. https://doi.org/10.1021/acs.chemrev.6b00790.

[268] **Bernstein FC**, Koetzle TF, Williams GJB, Meyer EF, Brice MD, Rodgers JR, Kennard O, Shimanouchi T, Tasumi M. The Protein Data Bank: A Computer-based Archival File for Macromolecular Structures. J Mol Biol. 1977; 112(3):535–542. https://doi.org/10.1016/s0022-2836(77)80200-3.

[269] **Berman HM**, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. The Protein Data Bank. Nucleic Acids Res. 2000; 28(1):235–242. https://doi.org/10.1093/nar/28.1.235.

[270] **Rose PW**, Prlić A, Bi C, Bluhm WF, Christie CH, Dutta S, Green RK, Goodsell DS, Westbrook JD, Woo J, Young J, Zardecki C, Berman HM, Bourne PE, Burley SK. The RCSB Protein Data Bank: Views of Structural Biology for Basic and Applied Research and Education. Nucleic Acids Res. 2015; 43(D1):D345–D356. https://doi.org/10.1093/nar/gku1214.

[271] **Dunbrack RL**, Karplus M. Backbone-dependent Rotamer Library for Proteins Application to Side-chain Prediction. J Mol Biol. 1993; 230(2):543–574. https://doi.org/10.1006/jmbi.1993.1170.

[272] **Lovell SC**, Word JM, Richardson JS, Richardson DC. The Penultimate Rotamer Library. Proteins. 2000; 40(3):389–408. https://doi.org/10.1002/1097-0134(20000815)40:3<389::aid-prot50>3.0.co;2-2.

[273] **Donnini S**, Groenhof G, Wierenga RK, Juffer AH. The Planar Conformation of a Strained Proline Ring: A QM/MM Study. Proteins. 2006; 64(3):700–710. https://doi.org/10.1002/prot.21006.

[274] **Grosse-Kunstleve RW**, Sauter NK, Moriarty NW, Adams PD. The Computational Crystallography Toolbox: Crystallographic Algorithms in a Reusable Software Framework. J Appl Crystallogr. 2002; 35(1):126–136. https://doi.org/10.1107/s0021889801017824.

[275] **Wall ME**. Methods and Software for Diffuse X-ray Scattering from Protein Crystals. In: Foote RS, Lee JW, editors. *Micro and Nano Technologies in Bioanalysis: Methods and Protocols* Methods in Molecular Biology, Totowa, NJ: Humana Press; 2009.p. 269–279. https://doi.org/10.1007/978-1-59745-483-4_17.

[276] **Winn MD**, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AGW, McCoy A, McNicholas SJ, Murshudov GN, Pannu NS, Potterton EA, Powell HR, Read RJ, Vagin A, Wilson KS. Overview of the CCP4 Suite and Current Developments. Acta Crystallogr D. 2011; 67(4):235–242. https://doi.org/10.1107/s0907444910045749.

[277] **Case DA**, Aktulga HM, Belfrom K, Ben-Shalom IY, Berryman JT, Brozell SR, Cerutti DS, Cheatham TE III, Cisneros GA, Cruzeiro VWD, Darden TA, Duke RE, Giambasu G, Gilson MK, Gohlke H, Goetz AW, Harris R, Izadi S, Izmailov K, Kasavajhala K, et al. Amber 2022. University of California, San Fransisco; 2022.

[278] **Pearce NM**, Krojer T, Bradley AR, Collins P, Nowak RP, Talon R, Marsden BD, Kelm S, Shi J, Deane CM, von Delft F. A Multi-crystal Method for Extracting Obscured Crystallographic States from Conventionally Uninterpretable Electron Density. Nat Commun. 2017; 8(1):15123. https://doi.org/10.1038/ncomms15123.

[279] **Liebschner D**, Afonine PV, Baker ML, Bunkóczi G, Chen VB, Croll TI, Hintze B, Hung LW, Jain S, McCoy AJ, Moriarty NW, Oeffner RD, Poon BK, Prisant MG, Read RJ, Richardson JS, Richardson DC, Sammito MD, Sobolev OV, Stockwell DH, et al. Macromolecular Structure Determination Using X-rays, Neutrons and Electrons: Recent Developments in Phenix. Acta Crystallogr D. 2019; 75(10):861–877. https://doi.org/10.1107/s2059798319011471.

[280] **Wall ME**. Internal Protein Motions in Molecular-dynamics Simulations of Bragg and Diffuse X-ray Scattering. IUCrJ. 2018; 5(2):172–181. https://doi.org/10.1107/s2052252518000519.

[281] **Krissinel E**, Henrick K. Inference of Macromolecular Assemblies from Crystalline State. J Mol Biol. 2007; 372(3):774–797. https://doi.org/10.1016/j.jmb.2007.05.022.

[282] **Trueblood KN**, Bürgi HB, Burzlaff H, Dunitz JD, Gramaccioli CM, Schulz HH, Shmueli U, Abrahams SC. Atomic Dispacement Parameter Nomenclature. Report of a Subcommittee on Atomic Displacement Parameter Nomenclature. Acta Crystallogr A. 1996; 52(5):770–781. https://doi.org/10.1107/s0108767396005697.

[283] **Williams CJ**, Headd JJ, Moriarty NW, Prisant MG, Videau LL, Deis LN, Verma V, Keedy DA, Hintze BJ, Chen VB, Jain S, Lewis SM, Arendall WB, Snoeyink J, Adams PD, Lovell SC, Richardson JS, Richardson DC. MolProbity: More and Better Reference Data for Improved All-atom Structure Validation. Prot Sci. 2018; 27(1):293–315. https://doi.org/10.1002/pro.3330.

[284] **Pearce NM**, Gros P. A Method for Intuitively Extracting Macromolecular Dynamics from Structural Disorder. Nat Commun. 2021; 12(1):5493. https://doi.org/10.1038/s41467-021-25814-x, publisher: Nature Publishing Group.

[285] **Ploscariu N**, Burnley T, Gros P, Pearce NM. Improving Sampling of Crystallographic Disorder in Ensemble Refinement. Acta Crystallogr D. 2021; 77(11):1357–1364. https://doi.org/10.1107/s2059798321010044, publisher: International Union of Crystallography.

[286] **Lang PT**, Ng HL, Fraser JS, Corn JE, Echols N, Sales M, Holton JM, Alber T. Automated Electron-density Sampling Reveals Widespread Conformational Polymorphism in Proteins. Prot Sci. 2010; 19(7):1420–1431. https://doi.org/10.1002/pro.423.

[287] **Riley BT**, Wankowicz SA, de Oliveira SHP, van Zundert GCP, Hogan DW, Fraser JS, Keedy DA, van den Bedem H. QFit 3: Protein and Ligand Multiconformer Modeling for X-ray Crystallographic and Single-particle Cryo-EM Density Maps. Prot Sci. 2021; 30(1):270–285. https://doi.org/10.1002/pro.4001.

[288] **Wankowicz SA**, Ravikumar A, Sharma S, Riley B, Raju A, Hogan DW, Flowers J, van den Bedem H, Keedy DA, Fraser JS. Automated Multiconformer Model Building for X-ray Crystallography and Cryo-EM. eLife. 2024; 12:RP90606. https://doi.org/10.7554/elife.90606, publisher: eLife Sciences Publications, Ltd.

[289] **Fraser JS**, van den Bedem H, Samelson AJ, Lang PT, Holton JM, Echols N, Alber T. Accessing Protein Conformational Ensembles Using Room-temperature X-ray Crystallography. Proc Nat Acad Sci. 2011; 108(39):16247–16252. https://doi.org/10.1073/pnas.1111325108.

[290] **Fenwick RB**, van den Bedem H, Fraser JS, Wright PE. Integrated Description of Protein Dynamics from Room-temperature X-ray Crystallography and NMR. Proc Nat Acad Sci. 2014; 111(4):E445–E454. https://doi.org/10.1073/pnas.1323440111.

[291] **Faure P**, Micu A, Pérahia D, Doucet J, Smith JC, Benoit JP. Correlated Intramolecular Motions and Diffuse X–ray Scattering in Lysozyme. Nat Struct Mol Biol. 1994; 1(2):124–128. https://doi.org/10.1038/nsb0294-124.

[292] **Clarage JB**, Romo T, Andrews BK, Pettitt BM, Phillips GN. A Sampling Problem in Molecular Dynamics Simulations of Macromolecules. Proc Nat Acad Sci. 1995; 92(8):3288–3292. https://doi.org/10.1073/pnas.92.8.3288.

[293] **Héry S**, Genest D, Smith JC. X-ray Diffuse Scattering and Rigid-body Motion in Crystalline Lysozyme Probed by Molecular Dynamics Simulation. J Mol Biol. 1998; 279(1):303–319. https://doi.org/10.1006/jmbi.1998.1754.

[294] **Meinhold L**, Smith JC. Fluctuations and Correlations in Crystalline Protein Dynamics: A Simulation Analysis of Staphylococcal Nuclease. Biophys J. 2005; 88(4):2554–2563. https://doi.org/10.1529/biophysj.104.056101.

[295] **Wall ME**, Van Benschoten AH, Sauter NK, Adams PD, Fraser JS, Terwilliger TC. Conformational Dynamics of a Crystalline Protein from Microsecond-scale Molecular Dynamics Simulations and Diffuse X-ray Scattering. Proc Nat Acad Sci. 2014; 111(50):17887–17892. https://doi.org/10.1073/pnas.1416744111.

[296] **Wych DC**, Fraser JS, Mobley DL, Wall ME. Liquid-like and Rigid-body Motions in Molecular-dynamics Simulations of a Crystalline Protein. Struct Dyn. 2019; 6(6):064704. https://doi.org/10.1063/1.5132692.

[297] **Blakeley MP**. Neutron Macromolecular Crystallography. Crystallogr Rev. 2009; 15(3):157–218. https://doi.org/10.1080/08893110902965003.

[298] **Ashkar R**, Bilheux HZ, Bordallo H, Briber R, Callaway DJE, Cheng X, Chu XQ, Curtis JE, Dadmun M, Fenimore P, Fushman D, Gabel F, Gupta K, Herberle F, Heinrich F, Hong L, Katsaras J, Kelman Z, Kharlampieva E, Kneller GR, et al. Neutron Scattering in the Biological Sciences: Progress and Prospects. Acta Crystallogr D. 2018; 74(12):1129–1168. https://doi.org/10.1107/s2059798318017503.

[299] **Word JM**, Lovell SC, Richardson JS, Richardson DC. Asparagine and Glutamine: Using Hydrogen Atom Contacts in the Choice of Side-chain Amide Orientation. J Mol Biol. 1999; 285(4):1735–1747. https://doi.org/10.1006/jmbi.1998.2401.

[300] **Word JM**, Lovell SC, LaBean TH, Taylor HC, Zalis ME, Presley BK, Richardson JS, Richardson DC. Visualizing and Quantifying Molecular Goodness-of-fit: Small-probe Contact Dots with Explicit Hydrogen Atoms. J Mol Biol. 1999; 285(4):1711–1733. https://doi.org/10.1006/jmbi.1998.2400.

[301] **Chen JCH**, Hanson BL, Fisher SZ, Langan P, Kovalevsky AY. Direct Observation of Hydrogen Atom Dynamics and Interactions by Ultrahigh Resolution Neutron Protein Crystallography. Proc Nat Acad Sci. 2012; 109(38):15301–15306. https://doi.org/10.1073/pnas.1208341109.

[302] **Blakeley MP**, Kalb-Gilboa AJ, Helliwell JR, Myles DAA. The 15k Neutron Structure of Saccharide-free Concanavalin A. Proc Nat Acad Sci. 2004; 101(47):16405–16410. https://doi.org/10.2210/pdb1xqn/pdb.

[303] **Wall ME**, Calabró G, Bayly CI, Mobley DL, Warren GL. Biomolecular Solvation Structure Revealed by Molecular Dynamics Simulations. J Am Chem Soc. 2019; 141(11):4711–4720. https://doi.org/10.1021/jacs.8b13613.

[304] **Kossiakoff AA**, Spencer SA. Neutron Diffraction Identifies His 57 As the Catalytic Base in Trypsin. Nature. 1980; 288(5789):414–416. https://doi.org/10.1038/288414a0.

[305] **Kossiakoff AA**, Spencer SA. Direct Determination of the Protonation States of Aspartic Acid-102 and Histidine-57 in the Tetrahedral Intermediate of the Serine Proteases: Neutron Structure of Trypsin. Biochemistry. 1981; 20(22):6462–6474. https://doi.org/10.1021/bi00525a027.

[306] **Mobley DL**, Gilson MK. Predicting Binding Free Energies: Frontiers and Benchmarks. Annu Rev Biophys. 2017; 46(1):531–558. https://doi.org/10.1146/annurev-biophys-070816-033654, publisher: Annual Reviews.

[307] **Mobley DL**, Heinzelmann G, Henriksen NM, Gilson MK. Predicting Binding Free Energies: Frontiers and Benchmarks (a Perpetual Review). eScholarship. 2017; https://escholarship.org/uc/item/9p37m6bq.

[308] **Cerutti DS**, Freddolino L, Duke RE, Case DA. Simulations of a Protein Crystal with a High Resolution X-ray Structure: Evaluation of Force Fields and Water Models. J Phys Chem B. 2010; 114(40):12811–12824. https://doi.org/10.1021/jp105813j.

[309] **Moreau D**. Ice Formation and Solvent Nanoconfinement in Protein Crystallography. dissertation or thesis, Cornell University; 2020.

[310] **Janowski PA**, Liu C, Deckman J, Case DA. Molecular Dynamics Simulation of Triclinic Lysozyme in a Crystal Lattice. Prot Sci. 2016; 25(1):87–102. https://doi.org/10.1002/pro.2713.

[311] **Kovalevskiy O**, Mateos-Garcia J, Tunyasuvunakool K. AlphaFold Two Years On: Validation and Impact. Proc Nat Acad Sci. 2024; 121(34):e2315002121. https://doi.org/10.1073/pnas.2315002121.

[312] **Groom CR**, Bruno IJ, Lightfoot MP, Ward SC. The Cambridge Structural Database. Acta Crystallogr B. 2016; 72(2):171–179. https://doi.org/10.1107/s2052520616003954.

[313] **Ahn HC**, Juranić N, Macura S, Markley JL. Three-dimensional Structure of the Water-insoluble Protein Crambin in Dodecylphosphocholine Micelles and Its Minimal Solvent-exposed Surface. J Am Chem Soc. 2006; 128(13):4398–4404. https://doi.org/10.1021/ja057773d.

[314] **Eisenmesser EZ**, Millet O, Labeikovsky W, Korzhnev DM, Wolf-Watz M, Bosco DA, Skalicky JJ, Kay LE, Kern D. Intrinsic Dynamics of an Enzyme Underlies Catalysis. Nature. 2005; 438(7064):117–121. https://doi.org/10.1038/nature04105.

[315] **Davis IW**, Arendall WB, Richardson DC, Richardson JS. The Backrub Motion: How Protein Backbone Shrugs When a Sidechain Dances. Structure. 2006; 14(2):265–274. https://doi.org/10.1016/j.str.2005.10.007.

[316] **Friedland GD**, Linares AJ, Smith CA, Kortemme T. A Simple Model of Backbone Flexibility Improves Modeling of Sidechain Conformational Variability. J Mol Biol. 2008; 380(4):757–774. https://doi.org/10.1016/j.jmb.2008.05.006.

[317] **Hénin J**, Lelièvre T, Shirts MR, Valsson O, Delemotte L. Enhanced Sampling Methods for Molecular Dynamics Simulations [Article V1.0]. Living J Comp Mol Sci. 2022; 4(1):1583–1583. https://doi.org/10.33011/livecoms.4.1.1583.

[318] **Wang L**, Friesner RA, Berne BJ. Replica Exchange with Solute Scaling: A More Efficient Version of Replica Exchange with Solute Tempering (REST2). J Phys Chem B. 2011; 115(30):9431–9438. https://doi.org/10.1021/jp204407d.

[319] **Zhang Y**, Zhou L, Rouge L, Phillips AH, Lam C, Liu P, Sandoval W, Helgason E, Murray JM, Wertz IE, Corn JE. Conformational Stabilization of Ubiquitin Yields Potent and Selective Inhibitors of USP7. Nat Chem Biol. 2013; 9(1):51–58. https://doi.org/10.1038/nchembio.1134.

[320] **Wiesmann C**, Barr KJ, Kung J, Zhu J, Erlanson DA, Shen W, Fahr BJ, Zhong M, Taylor L, Randal M, McDowell RS, Hansen SK. Allosteric Inhibition of Protein Tyrosine Phosphatase 1B. Nat Struct Mol Biol. 2004; 11(8):730–737. https://doi.org/10.1038/nsmb803.

[321] **Whittier SK**, Hengge AC, Loria JP. Conformational Motions Regulate Phosphoryl Transfer in Related Protein Tyrosine Phosphatases. Science. 2013; 341(6148):899–903. https://doi.org/10.1126/science.1241735.

[322] **Choy MS**, Li Y, Machado LESF, Kunze MBA, Connors CR, Wei X, Lindorff-Larsen K, Page R, Peti W. Conformational Rigidity and Protein Dynamics At Distinct Timescales Regulate PTP1B Activity and Allostery. Mol Cell. 2017; 65(4):644–658.e5. https://doi.org/10.1016/j.molcel.2017.01.014.

[323] **Yeh CY**, Izaguirre JA, Greisman JB, Willmore L, Maragakis P, Shaw DE. A Conserved Local Structural Motif Controls the Kinetics of PTP1B Catalysis. J Chem Inf Model. 2023; 63(13):4115–4124. https://doi.org/10.1021/acs.jcim.3c00286.

[324] **Crean RM**, Corbella M, Calixto AR, Hengge AC, Kamerlin SCL. Sequence – Dynamics – Function Relationships in Protein Tyrosine Phosphatases. QRB Discovery. 2024; 5:e4. https://doi.org/10.1017/qrd.2024.3.

[325] **Lexa KW**, Carlson HA. Full Protein Flexibility Is Essential for Proper Hot-spot Mapping. J Am Chem Soc. 2011; 133(2):200–202. https://doi.org/10.1021/ja1079332.

[326] **Feldman HJ**, Hogue CWV. A Fast Method to Sample Real Protein Conformational Space. Proteins. 2000; 39(2):112–131. https://doi.org/10.1002/(sici)1097-0134(20000501)39:2<112::aid-prot2>3.3.co;2-2.

[327] **Feldman HJ**, Hogue CWV. Probabilistic Sampling of Protein Conformations: New Hope for Brute Force? Proteins. 2002; 46(1):8–23. https://doi.org/10.1002/prot.1163.abs.

[328] **Ozenne V**, Bauer F, Salmon L, Huang Jr, Jensen MR, Segard S, Bernadó P, Charavay C, Blackledge M. Flexible-meccano: a Tool for the Generation of Explicit Ensemble Descriptions of Intrinsically Disordered Proteins and Their Associated Experimental Observables. Bioinformatics. 2012; 28(11):1463–1470. https://doi.org/10.1093/bioinformatics/bts172.

[329] **Estaña A**, Sibille N, Delaforge E, Vaisset M, Cortés J, Bernadó P. Realistic Ensemble Models of Intrinsically Disordered Proteins Using a Structure-encoding Coil Database. Structure. 2019; 27(2):381–391.e2. https://doi.org/10.1016/j.str.2018.10.016.

[330] **Ferrie JJ**, Petersson EJ. A Unified De Novo Approach for Predicting the Structures of Ordered and Disordered Proteins. J Phys Chem B. 2020; 124(27):5538–5548. https://doi.org/10.1021/acs.jpcb.0c02924.

[331] **Teixeira JMC**, Liu ZH, Namini A, Li J, Vernon RM, Krzeminski M, Shamandy AA, Zhang O, Haghighatlari M, Yu L, Head-Gordon T, Forman-Kay JD. IDPConformerGenerator: A Flexible Software Suite for Sampling the Conformational Space of Disordered Protein States. J Phys Chem A. 2022; 126(35):5985–6003. https://doi.org/10.1021/acs.jpca.2c03726.

[332] **Olsson MHM**, Søndergaard CR, Rostkowski M, Jensen JH. PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical pKa Predictions. J Chem Theory Comput. 2011; 7(2):525–537. https://doi.org/10.1021/ct100578z.

[333] **Jurrus E**, Engel D, Star K, Monson K, Brandi J, Felberg LE, Brookes DH, Wilson L, Chen J, Liles K, Chun M, Li P, Gohara DW, Dolinsky T, Konecny R, Koes DR, Nielsen JE, Head-Gordon T, Geng W, Krasny R, et al. Improvements to the APBS Biomolecular Solvation Software Suite. Prot Sci. 2018; 27(1):112–128. https://doi.org/10.1002/pro.3280.

[334] **Coppa C**, Bazzoli A, Barkhordari M, Contini A. Accelerated Molecular Dynamics for Peptide Folding: Benchmarking Different Combinations of Force Fields and Explicit Solvent Models. J Chem Inf Model. 2023; 63(10):3030–3042. https://doi.org/10.1021/acs.jcim.3c00138.

[335] **Kadaoluwa Pathirannahalage SP**, Meftahi N, Elbourne A, Weiss ACG, McConville CF, Padua A, Winkler DA, Costa Gomes

M, Greaves TL, Le TC, Besford QA, Christofferson AJ. Systematic Comparison of the Structural and Dynamic Properties of Commonly Used Water Models for Molecular Dynamics Simulations. J Chem Inf Model. 2021; 61(9):4521–4536. https://doi.org/10.1021/acs.jcim.1c00794.

[336] **Wang Ss**, Krumhansl JA. Superposition Assumption. II. High Density Fluid Argon. J Chem Phys. 1972; 56(9):4287–4290. https://doi.org/10.1063/1.1677858.

[337] **Mehra R**, Kepp KP. Cell Size Effects in the Molecular Dynamics of the Intrinsically Disordered Aβ Peptide. J Chem Phys. 2019; 151(8):085101. https://doi.org/10.1063/1.5115085.

[338] **Gapsys V**, de Groot BL. On the Importance of Statistics in Molecular Simulations for Thermodynamics, Kinetics and Simulation Box Size. eLife. 2020; 9:e57589. https://doi.org/10.7554/elife.57589.

[339] **Schmit JD**, Kariyawasam NL, Needham V, Smith PE. SLTCAP: A Simple Method for Calculating the Number of Ions Needed for MD Simulation. J Chem Theory Comput. 2018; 14(4):1823–1827. https://doi.org/10.1021/acs.jctc.7b01254.

[340] **Harvey SC**, Tan RKZ, Cheatham TE. The Flying Ice Cube: Velocity Rescaling in Molecular Dynamics Leads to Violation of Energy Equipartition. J Comput Chem. 1998; 19(7):726–740. https://doi.org/10.1002/(sici)1096-987x(199805)19:7<726::aid-jcc4>3.0.co;2-s.

[341] **Braun E**, Moosavi SM, Smit B. Anomalous Effects of Velocity Rescaling Algorithms: The Flying Ice Cube Effect Revisited. J Chem Theory Comput. 2018; 14(10):5262–5272. https://doi.org/10.1021/acs.jctc.8b00446.

[342] **Chodera JD**. A Simple Method for Automated Equilibration Detection in Molecular Simulations. J Chem Theory Comput. 2016; 12(4):1799–1805. https://doi.org/10.1021/acs.jctc.5b00784.

[343] **Grossfield A**, Patrone PN, Roe DR, Schultz AJ, Siderius D, Zuckerman DM. Best Practices for Quantification of Uncertainty and Sampling Quality in Molecular Simulations [Article V1.0]. Living J Comp Mol Sci. 2019; 1(1):5067–5067. https://doi.org/10.33011/livecoms.1.1.5067.

[344] **Flyvbjerg H**, Petersen HG. Error Estimates on Averages of Correlated Data. J Chem Phys. 1989; 91(1):461–466. https://doi.org/10.1063/1.457480.

[345] **Wassenaar TA**, Mark AE. The Effect of Box Shape on the Dynamic Properties of Proteins Simulated Under Periodic Boundary Conditions. J Comput Chem. 2006; 27(3):316–325. https://doi.org/10.1002/jcc.20341.

[346] **Villa A**, Fan H, Wassenaar T, Mark AE. How Sensitive Are Nanosecond Molecular Dynamics Simulations of Proteins to Changes in the Force Field? J Phys Chem B. 2007; 111(21):6015–6025. https://doi.org/10.1021/jp068580v.

[347] **Solomon I**, Bloembergen N. Nuclear Magnetic Interactions in the HF Molecule. J Chem Phys. 1956; 25(2):261–266. https://doi.org/10.1063/1.1742867.

Appendix A Effective distances probed in NOESY experiments

The NOE between two nuclei, A and B, depends primarily on their dipolar cross relaxation rate, σ_{AB}.

$$\sigma_{AB} = K \left[6J(\omega_A + \omega_B) - J(\omega_A - \omega_B) \right]$$
$$K = \left(\frac{\mu_0}{4\pi} \right)^2 \frac{\hbar^2 \gamma_A^2 \gamma_B^2}{4} \tag{17}$$

where μ_0 is the vacuum permeability, $\hbar$ is the reduced Planck constant, γ_A and γ_B are the gyromagnetic ratios of the nuclei, and ω_A and ω_B are the Larmor frequencies of the nuclei. $J(\omega)$ is a spectral density function associated with the relative geometry of the two spins, whose correlation function is determined by the dynamics of the lattice. It essentially describes the spectrum of the power available from the lattice to activate the dipolar cross relaxation pathway. Thus, the spectral density function is given by the Fourier transformation of the correlation function for the internuclear vector in the lab frame, $C(t)$:

$$J(\omega) = 2 \int_0^\infty dt C(t) \cos(\omega t) \tag{18}$$

It is often assumed [217] that internal motions are separable from overall tumbling so that the correlation function can be written as

$$C(t) = C_O(t) C_I(t) \tag{19}$$

where the correlation function for overall tumbling is

$$C_O(t) = \frac{1}{5} \exp \left(-\frac{t}{\tau_C} \right) \tag{20}$$

and the correlation function for internal motions is

$$C_I(t) = \left\langle \frac{3 \cos^2 \theta_{AB}(t) - 1}{2 R_{AB}(0)^3 R_{AB}(t)^3} \right\rangle \tag{21}$$

Here, τ_C is the rotational diffusion time constant for isotropic tumbling, $R_{AB}(t)$ is the distance between the nuclei at time t, and $\theta_{AB}(t)$ is the angle between the internuclear vectors at time 0 and time t in the molecular frame. Note that some authors use a different convention in which $C_O(t)$ is defined without the numerical factor of $\frac{1}{5}$, which is absorbed into the constant factor for the dipolar cross relaxation rate (K in Eq. 17). Therefore, the equations in this appendix may differ from the equations in certain publications by numerical factors.

As discussed in Section 3.1.4, NOESY experiments provide effective mean distances between nuclei. We next discuss two common approaches to determining effective distances and the conditions under which they apply. For a more detailed treatment, see Neuhaus & Williamson [203] and Vögeli [204]. For simplicity, we consider the case of a homonuclear NOESY experiment, where—to good approximation—$\omega_A = \omega_B = \omega$, so the dipolar cross relaxation rate reduces to

$$\sigma_{AB} = K\left[6J(2\omega) - J(0)\right] \tag{22}$$

For a rigid molecule that tumbles isotropically with a rotational diffusion time constant τ_C, the spectral density function is given by

$$J(\omega) = \frac{2}{5}\left(\frac{1}{R_{AB}^6}\right)\frac{\tau_C}{1 + \omega^2\tau_C^2} \tag{23}$$

where R_{AB} is the constant distance between the nuclei. In this case, the dipolar cross relaxation rate becomes

$$\sigma_{AB} = \frac{2}{5}\frac{K}{R_{AB}^6}\left(\frac{6\tau_C}{1 + 4\omega^2\tau_C^2} - \tau_C\right) \tag{24}$$

For a non-rigid molecule, it is often assumed that internal motions (i.e., intramolecular motions) are uncorrelated with molecular tumbling and can be characterized by a single exponential decay with a time constant τ_{int}. Under these assumptions, called the Lipari-Szabo or model-free approximation [217], the spectral density is given by

$$J(\omega) = \frac{2}{5}\left\langle\frac{1}{R_{AB}^6}\right\rangle\left(S^2\frac{\tau_C}{1 + \omega^2\tau_C^2} + \left(1 - S^2\right)\frac{\tau_{\text{tot}}}{1 + \omega^2\tau_{\text{tot}}^2}\right) \tag{25}$$

where

$$\frac{1}{\tau_{\text{tot}}} = \frac{1}{\tau_C} + \frac{1}{\tau_{\text{int}}} \tag{26}$$

and

$$S^2 = \left\langle\frac{1}{R_{AB}^6}\right\rangle^{-1}\frac{4\pi}{5}\sum_{m=-2}^{2}\left\langle\frac{Y_{2m}(\theta_{AB}, \phi_{AB})}{R_{AB}^3}\right\rangle^2 \tag{27}$$

Here, $Y_{2m}(\theta_{AB}, \phi_{AB})$ are the second-order spherical harmonics defined using the spherical coordinates (θ_{AB}, ϕ_{AB}) of the internuclear displacement in the molecular frame. S^2 is an order parameter that is determined by the plateau value of the correlation function $C(t)$ at long times and characterizes the orientational freedom of the internuclear displacement. The order parameter varies between one—when the internuclear displacement is rigid in the molecular frame—and zero—when the internuclear displacement tumbles isotropically in the molecular frame. In the Lipari-Szabo model [217], the dipolar cross relaxation rate becomes

$$\begin{aligned}\sigma_{AB} = \frac{2K}{5}\left\langle\frac{1}{R_{AB}^6}\right\rangle\Bigg[&S^2\left(\frac{6\tau_C}{1 + 4\omega^2\tau_C^2} - \tau_C\right)\\ &+(1 - S^2)\left(\frac{6\tau_{\text{tot}}}{1 + 4\omega^2\tau_{\text{tot}}^2} - \tau_{\text{tot}}\right)\Bigg]\end{aligned} \tag{28}$$

This expression is difficult to use in practice because the time constant for internal motions—and thus τ_{tot}—cannot be easily measured. Fortunately, we can simplify this expression by considering limiting cases when internal motions are

either much slower or much faster than molecular tumbling. In the limit where internal motions are much slower than molecular tumbling, i.e. $\tau_{\text{int}} \gg \tau_C$, $\tau_{\text{tot}} \approx \tau_C$ such that the terms containing the order parameter S^2 cancel and the dipolar cross relaxation rate loses its dependence on the angular degrees of freedom.

$$\sigma_{AB} = \frac{2K}{5}\left\langle\frac{1}{R_{AB}^6}\right\rangle\left(\frac{6\tau_C}{1 + 4\omega^2\tau_C^2} - \tau_C\right) \tag{29}$$

This expression looks like the expression for a rigid molecule (Eq. 24) with the internuclear distance probed by the NOE enhancement replaced by the following effective distance.

$$R_{AB}^{\text{slow}} = \left\langle\frac{1}{R_{AB}^6}\right\rangle^{-1/6} \tag{30}$$

Thus, in the limit when internal motions are much slower than molecular tumbling, the NOE averages the variation in internuclear distance due to internal motions raised to the reciprocal sixth power and is independent of the orientation of the internuclear displacement.

In the opposite limit, i.e., when internal motions are much faster than molecular tumbling ($\tau_{\text{int}} \ll \tau_C$), the term in Equation 25 containing $(1 - S^2)$ becomes negligible, so

$$\sigma_{AB} = \frac{2K}{5}\left\langle\frac{1}{R_{AB}^6}\right\rangle S^2\left(\frac{6\tau_C}{1 + 4\omega^2\tau_C^2} - \tau_C\right) \tag{31}$$

In this case, the effective distance probed by the NOE enhancement is

$$R_{AB}^{\text{fast,ang}} = \left(\frac{4\pi}{5}\sum_{m=-2}^{2}\left\langle\frac{Y_{2m}(\theta_{AB}, \phi_{AB})}{R_{AB}^3}\right\rangle^2\right)^{-1/6} \tag{32}$$

The effective distance can be simplified further by assuming that the distances R_{AB} exhibit greater variations than the angles in the spherical harmonic functions. Under this assumption, the angular dependence is ignored, and the effective distance is chosen to be

$$R_{AB}^{\text{fast}} = \left(\left\langle\frac{1}{R_{AB}^3}\right\rangle^2\right)^{-1/6} = \left\langle\frac{1}{R_{AB}^3}\right\rangle^{-1/3} \leq R_{AB}^{\text{fast,ang}} \tag{33}$$

where the inequality comes from the observation that the spherical harmonic functions are normalized such that the angular dependence in Eq. 32 must reduce the magnitude of the quantity averaged over and thus increase the effective distance.

Appendix B The Redfield equations for NMR spin relaxation

Spin relaxation rates in proteins are often measured for spins coupled through a covalent bond, and a heteronuclear NOE can also be measured for the same bond. The Redfield equations [210] describe how the spin relaxation rates and NOE depend on the dynamics of the bond within a molecule. For the backbone ^{15}N amide bond in proteins, the Redfield equations are

$$R_1 = \frac{1}{T_1} = \frac{K}{R_{NH}^6} \left(6J(\omega_H + \omega_N) + 3J(\omega_N) + J(\omega_H - \omega_N)\right)$$
$$+ \frac{(\omega_N \Delta\sigma)^2}{15} J(\omega_N)$$

$$R_2 = \frac{1}{T_2} = \frac{1}{2}\frac{K}{R_{NH}^6} \left(6J(\omega_H + \omega_N) + 3J(\omega_N) + J(\omega_H - \omega_N)\right) \tag{34}$$
$$+ 6J(\omega_H) + 4J(0) + \frac{(\omega_N \Delta\sigma)^2}{90}\left(3J(\omega_N) + 4J(0)\right)$$

$$\text{NOE} = \frac{K}{R_{NH}^6} \left(6J(\omega_H + \omega_N) + J(\omega_H - \omega_N)\right) \times \frac{\gamma_H}{\gamma_N} T_1$$

where, similar to Appendix A,

$$K = \left(\frac{\mu_0}{4\pi}\right)^2 \frac{\hbar^2 \gamma_N^2 \gamma_H^2}{4} \tag{35}$$

and μ_0 is the vacuum permeability, $\hbar$ is the reduced Planck constant, γ_N and γ_H are the gyromagnetic ratios of the nuclei, ω_N and ω_H are the Larmor frequencies of the nuclei, R_{NH} is the amide bond length (assumed constant), and $J(\omega)$ is the spectral density function (Eq. 18). $\Delta\sigma$ is the chemical shift anisotropy for the ^{15}N spin, usually taken to be 160 ppm.

The length of the N-H bond is typically assumed to be constant, so that the correlation function for internal motions (Eq. 21) is given by the rotational correlation function

$$C_{\text{rot}}(t) = \left\langle \frac{3}{2}\cos\theta_{AB}(t) - \frac{1}{2} \right\rangle \tag{36}$$

Under the Lipari-Szabo or model-free approximation [217], in which overall molecular tumbling and intramolecular motions are assumed to be uncorrelated and each characterized by a single exponential decay, the spectral density function becomes

$$J(\omega) = \frac{2}{5}\left(S^2 \frac{\tau_C}{1 + \omega^2\tau_C^2} + \left(1 - S^2\right)\frac{\tau_{\text{tot}}}{1 + \omega^2\tau_{\text{tot}}^2}\right) \tag{37}$$

where S^2 (Eq. 27) is the order parameter that characterizes the orientational freedom of the internuclear vector, τ_C is the time constant for overall molecular tumbling, τ_{int} is the time constant for intramolecular motion, and τ_{tot} is given by Eq. 26.

Appendix C The Solomon-Bloembergen equation for PRE

The relationship between structure, dynamics, and the transverse PRE Γ_2 is rigorously described by the Solomon-Bloembergen equation [347].

$$\Gamma_2 = \frac{K}{2}\left(3J(\omega_H) + 4J(0)\right) \tag{38}$$

where

$$K = \left(\frac{\mu_0}{4\pi}\right)^2 \frac{\gamma_H^2 g^2 \mu_B^2}{4} \tag{39}$$

and μ_0 is the vacuum permeability, g is the electron g-factor, μ_B is the magnetic moment of the free electron, γ_H is the proton gyromagnetic ratio, ω_H is the Larmor frequency of the proton nucleus, and $J(\omega)$ is the spectral density function. As introduced in Appendices A and B, it is common to assume the Lipari-Szabo approximation [217] in order to write the spectral density using an order parameter S^2.

$$J(\omega) = \frac{2}{5}\left\langle \frac{1}{R_{eH}^6} \right\rangle \left(S^2 \frac{\tau_C}{1 + \omega^2\tau_C^2} + \left(1 - S^2\right)\frac{\tau_{\text{tot}}}{1 + \omega^2\tau_{\text{tot}}^2}\right) \tag{40}$$

where R_{eH} is the distance between the amide proton and the paramagnetic electron.