

# Comparing Atomistic Simulation Data With the NMR Experiment: How Much Can NOEs Actually Tell Us?

Bojan Zagrovic and Wilfred F. van Gunsteren\*

Laboratory of Physical Chemistry, Department of Chemistry and Applied Biosciences, ETH Hönggerberg, Zürich, Switzerland

**ABSTRACT** Simulated molecular dynamics trajectories of proteins and nucleic acids are often compared with nuclear magnetic resonance (NMR) data for the purposes of assessing the quality of the force field used or, equally important, trying to interpret ambiguous experimental data. In particular, nuclear Overhauser enhancement (NOE) intensities or atom–atom distances derived from them are frequently calculated from the simulated ensembles because the distance restraints derived from NOEs are the key ingredient in NMR-based protein structure determination. In this study, we ask how diverse and nonnative-like an ensemble of structures can be and still match the experimental NOE distance upper bounds well. We present two examples in which simulated ensembles of highly nonnative polypeptide structures (an unfolded state ensemble of the villin headpiece and a high-temperature denatured ensemble of lysozyme) are shown to match fairly well the experimental NOE distance upper bounds from which the corresponding native structures were derived. For example, the unfolded ensemble of villin headpiece, which is on average  $0.90 \pm 0.13$  nm root-mean-square deviation away from the native NMR structure, deviates from the experimental restraints by only 0.027 nm on average. However, this artificially good agreement is largely a consequence of 1) the highly nonlinear effects of  $r^{-6}$  (or  $r^{-3}$ ) averaging and 2) focusing only on the experimentally observed set of NOE bounds. Namely, in addition to the experimentally observed NOEs, both simulated ensembles (especially the villin ensemble) also predict a large number of NOEs, which are not seen in the experiment. If these are taken into account, the agreement between simulation and experiment gets markedly worse, as it should, given the nonnative nature of the underlying simulated ensembles. In light of the examples given, we conclude that comparing experimental NOE distance restraints with large simulated ensembles provides just by itself only limited information about the quality of simulation. *Proteins* 2006; 63:210–218. © 2006 Wiley-Liss, Inc.

**Key words:** NOE upper bounds; molecular dynamics simulations; predicting NMR observables

## INTRODUCTION

Molecular dynamics simulations have been developed with the aim to provide a microscopic picture in terms of space and time behind the ensemble- and time-averaged phenomena that are typically seen in the experiment. The relationship between simulation and experiment is one of mutual interdependence. On the one hand, the molecular force fields and the sampling methods used for simulations are calibrated and tested against experimental data. On the other hand, simulations are often used for interpreting experimental results in terms of atomistic models. Nuclear magnetic resonance (NMR) is one of the premier experimental methods for studying the structure and dynamics of biomolecules.<sup>1–3</sup> Naturally, the NMR observables such as chemical shifts, <sup>3</sup>J-coupling frequencies, nuclear Overhauser enhancement (NOE) intensities, relaxation parameters, and residual dipolar couplings or quantities derived from these such as NOE distance upper bounds and order parameters are often calculated from simulated trajectories and compared with their experimental counterparts.<sup>4–21</sup>

The NOE intensities are particularly important in this respect, because they directly depend on the interatomic distances in the molecules studied. The prescription for calculating NOEs and distance bounds derived from them based on a given ensemble of simulated structures is straightforward:

1. locate the pairs of nuclei that show an NOE in the experiment;
2. follow their dynamics;
3. use  $r^{-6}$  (or  $r^{-3}$ ) averaging of the internuclear distance  $r$  to obtain predicted average distances which can be compared with the corresponding NMR-derived distances.<sup>8,13,18,21,22</sup>

Alternatively, a computationally much more complex relaxation matrix approach<sup>23,24</sup> can be used to obtain NOE

---

Grant sponsor: EMBO postdoctoral fellowship; Grant sponsor: National Center of Competence in Research (NCCR) in Structural Biology of the Swiss National Science Foundation; Grant sponsor: National Institutes of Health; Grant number: R01GM62868.

\*Correspondence to: Wilfred F. van Gunsteren, Laboratory of Physical Chemistry, Department of Chemistry and Applied Biosciences, ETH Hönggerberg, Zürich, 8093, Switzerland. E-mail: wfvgn@igc.phys.chem.ethz.ch

Received 22 June 2005; Revised 3 October 2005; Accepted 28 October 2005

Published online 19 January 2006 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.20872

intensities that can be directly compared with experimentally observed intensities.

The strong dependence of the NOE intensity on the internuclear distance introduces a known complication in that the ensemble-averaged NOE values are strongly dominated by the short distances in the ensemble. In other words,  $r^{-6}$  averaging of the internuclear distances (or  $r^{-3}$ , depending on the relative intramolecular mobility compared with overall tumbling motion<sup>25</sup>) heavily biases the average toward the low-end tail of the internuclear distance distribution. As a consequence, markedly different distributions of distances can yield very similar averages. Because the experiment measures only the average, it might be difficult to decide whether a given simulated ensemble samples from the correct distribution, even if the average values do match. Different aspects of this problem have been discussed before.<sup>1,8,26</sup>

There is another, somewhat more subtle artifact that can arise when comparing simulated and experimental NOE intensities. Namely, one often focuses only on the pairs of nuclei that show an NOE in the experiment. However, it is also important to ask whether a given simulation perhaps predicts certain NOEs that are not seen or not identified in the experiment. A successful simulation should account for the presence of the experimentally seen intensities, but should also not predict intensities that are not there in the experiment and for which no reason can be found for their absence. Structural information derived from the absence of NOEs has already been used in protein structure determination and validation, but only fairly infrequently.<sup>27-31</sup> When it comes to simulations, if one focuses exclusively on positive information coming from the experiment, it is possible to get agreement with simulation even if the simulated ensemble is clearly unrealistic.

In this study, we analyzed the effects of nonlinear distance averaging inherent in NMR as well as the problem of NOE overprediction using two different simulated ensembles: an unfolded state ensemble of the 36-residue villin headpiece<sup>32-34</sup> and a high-temperature denatured ensemble of 129-residue hen egg-white lysozyme.<sup>35</sup> We asked how diverse and nonnative-like an ensemble of structures can be and still match the experimental NOE distance upper bounds well, and analyzed the underlying reasons.

## METHODS

### Simulations

#### *Villin simulations*

Using a worldwide distributed computer cluster,<sup>36,37</sup> we have simulated thousands of tens of nanoseconds long, independent folding trajectories for the villin headpiece molecule. All simulations were started from the extended conformation ( $\phi = -135^\circ$ ,  $\psi = 135^\circ$ ) with N-acetyl and C-amino caps, each simulation initiated with a different random-number seed. The simulations, run using the Tinker biomolecular simulation package, involved Langevin dynamics in implicit GB/SA solvent<sup>38</sup> (velocity damping parameter of  $\gamma = 91 \text{ ps}^{-1}$ , to match that of water) with

a 2-fs integration step, at 300 K. Bond lengths were constrained using RATTLE.<sup>39</sup> No cutoffs were used for electrostatics. The protein was modeled using the OPLS-UA force field.<sup>40</sup> The structures were saved for analysis every 1 ns of simulated time. The analysis given here was conducted on a composite ensemble consisting of 3,999 structures at the 33-ns time-point. The simulations were conducted on approximately 10,000 processors as a part of the ongoing Folding@Home distributed computing project (folding.stanford.edu), and involved a total of about quarter of a trillion ( $2.5 * 10^{11}$ ) integration steps. This corresponds to approximately 1,000 single CPU years (500 MHz). Further details about the simulation are given elsewhere.<sup>34</sup>

#### *Lysozyme simulations*

Using an in-house Pentium cluster, we have simulated 362 independent high-temperature unfolding simulations of lysozyme using the GROMOS 53A6 force field<sup>21,41</sup> and the GROMOS MD package.<sup>42</sup> The simulations were run in vacuo at a constant temperature of 700 K for 50 ps each. All simulations were started from the experimental NMR structure of lysozyme (1E8L, first model<sup>35</sup>) after energy minimization and removal of all cysteine bridges, each using a different random-number seed. The temperature was kept constant by coupling the system to an external temperature bath with a relaxation time of 0.1 ps.<sup>43</sup> Nonbonded interactions were treated using a triple range scheme with 0.08- and 1.4-nm cutoffs. The electrostatic interactions were treated using the reaction-field method with a cutoff of 1.4 nm. The translational motion and rotation around the center of mass were removed at the beginning of each simulation. Nonbonded pair lists were updated every five steps, and the integration step size was 2 fs in all simulations. The bonds including hydrogen atoms were constrained using the SHAKE algorithm<sup>44</sup> with a relative geometric accuracy of  $10^{-4}$ . The analysis was performed on the ensemble consisting of all 362 structures at 50 ps after energy minimization.

For both villin and lysozyme simulations, the secondary structure analysis was performed using DSSP.<sup>45</sup> Percentage native secondary structure is given as the fraction of residues that are in the same DSSP secondary structure category in a given structure as in the corresponding experimentally derived native structure. The root-mean-square deviation (RMSD) analysis was performed using all nonhydrogen atoms for both translational/rotational fit and final calculation of deviation.

#### **Comparison With the NMR NOE Distance Upper Bounds**

The experimental XPLOR-formatted NOE distance restraint sets for villin (1VII) and lysozyme (1E8L) were taken from the Protein Data Bank.<sup>46</sup> Because OPLS-UA and GROMOS force fields use the united-atom formalism, when calculating interproton distances and averages based on simulation, the following approach was taken.<sup>42</sup> In the case that a hydrogen atom of interest was not represented explicitly in the simulation, then either: a) a virtual atom

was constructed (CH, CH<sub>2</sub>) based on standard geometries, or b) a pseudo atom was constructed and a standard pseudo atom correction was applied to the NOE upper bound (cases in which NOE upper bounds were assigned to more than one proton). For a nonstereospecifically assigned CH<sub>2</sub> group, a pseudo atom correction of 0.09 nm was added to the upper bound. For methyl groups, the correction was 0.1 nm. For the methyl protons in an isopropyl group, a correction of 0.22 nm was applied, and for unassigned H $\delta$  and H $\epsilon$  atoms in a flipping benzene ring, 0.21 nm. These corrections are only marginally different from the pseudo atom corrections used in the lysozyme structure determination<sup>35</sup> and the ones proposed by Wuethrich.<sup>2</sup> The corrections that we used were derived based on the standard bond lengths and angles in GRO-MOS.<sup>41,42</sup> We did not apply additional multiplicity corrections to the NOE upper bounds.<sup>47,48</sup>

The averaging of individual distances in simulation (say, between hydrogen atoms  $i$  and  $j$ ) was done according to  $r_{i,j}^{avg} = \langle r_{i,j}^{-6} \rangle^{-1/6}$  for villin and  $r_{i,j}^{avg} = \langle r_{i,j}^{-3} \rangle^{-1/3}$  for lysozyme, where the average extends over the entire simulated ensemble. Tropp<sup>25</sup> has shown that in the case of relatively small molecules in which the timescale of internal fluctuations is longer than the overall tumbling time, the NOE intensities between hydrogen atoms  $i$  and  $j$  average according to  $\langle r_{i,j}^{-6} \rangle^{-1/6}$ , whereas if the internal fluctuations are significantly faster compared with the tumbling time of the molecule, they average according to  $\langle r_{i,j}^{-3} \rangle^{-1/3}$ .<sup>25</sup> Given the size of the villin headpiece (36 residues), it likely belongs in the former category and was treated as such, whereas lysozyme (129 residues) was treated using the latter formalism. In the most important cases, values using both types of averaging are given in the text for comparison. The  $r^{-6}$  assumption is typically used when transforming measured intensities to distance bounds in the context of experimental structure determination.<sup>49</sup>

A violation of the NOE upper bound for a particular pair of hydrogen atoms  $i$  and  $j$  is calculated as  $v_{i,j} = \langle r_{i,j}^{-p} \rangle^{-1/p} - nmr_{i,j}$ , where the average goes over the entire simulated ensemble ( $p = 3$  or  $6$ ). The violation  $v_{i,j}$  is considered zero if  $\langle r_{i,j}^{-p} \rangle^{-1/p} \leq nmr_{i,j}$  because negative deviations cannot be considered violations. Average violations,  $v(nmr, \langle sim \rangle)$  =  $\langle v_{i,j} \rangle = \frac{1}{N} \sum_{i,j} v_{i,j}$  were also calculated by averaging the individual violations over all  $N$  experimentally determined NOE upper bounds. In general, in the text an average violation of the type  $v(A,B)$  denotes the average violation of distance (or upper bound) set B from distance (or upper bound) set A, averaged over all distances/upper bounds in the set A. For instance, average violation of the distances based on the NMR-derived model structure from the theoretically predicted NOEs, averaged over all theoretically predicted NOEs, is denoted as  $v(\langle sim \rangle, nmr)$ . Furthermore, the average over all of the average violations of each individual member of the simulated ensembles with respect to the experimental upper-bounds is denoted as  $\langle v(nmr, sim) \rangle$ .

When calculating theoretically predicted NOEs based on our simulated structures, we selected all pairs of relax-

ation centers separated by  $<0.55$  nm in the  $r^{-3}$ - or  $r^{-6}$ -averaged distance sets and assumed these pairs would result in an NOE (the caveats of this approach are discussed in the Discussion section). The value of 0.55 nm is typically used in the NMR refinement as the upper distance cutoff corresponding to the weak NOEs (and was used as such in the case of lysozyme refinement quoted here<sup>35</sup>). In the villin structure, there are 193 hydrogen atoms (or relaxation centers such as methyl groups or isopropyl groups) that could potentially contribute to an NOE. This number is 622 in the case of lysozyme. The centers that in the experimental distance bounds were treated as nonstereospecific (e.g., CH<sub>2</sub> groups) were also treated as such in the course of our NOE prediction, whereas all stereospecifically assigned hydrogen atoms were treated as such in our simulations as well.

## RESULTS

As the first example, we look at the unfolded state ensemble of the villin headpiece. Villin headpiece is a 36-residue, three-helix bundle protein and is one of the fastest folding polypeptides known.<sup>50,51</sup> The unfolded state ensemble was generated by simulating a large number of independent trajectories started from an extended conformation for approximately 30 ns each. This time is significantly shorter than the folding time of villin (4.3  $\mu$ s<sup>52</sup>), assuring that what we are looking at is the kinetically defined unfolded state of the molecule. The simulated ensemble at 33 ns is compact (average radius of gyration  $\langle Rg \rangle = 1.0 \pm 0.1$  nm which is comparable to the native value of 0.96 nm) with little secondary structure (27% native-like secondary structure on average; 15% of the native-like helical structure) or permanent tertiary contacts [ $\langle \text{RMSD} \rangle = 0.90 \pm 0.13$  nm from the NMR villin structure, Fig. 1(A)].<sup>34,53</sup> However, if one compares the experimental NOE upper distance bounds for the native villin structure with the  $r^{-6}$ -averaged distances from the simulated unfolded state, one gets agreement for many NOEs (Fig. 2). Despite the simulated ensemble being very diverse and nonnative-like, 77% of the experimental restraints are matched with 0-nm violation and 90% with a violation of  $<0.1$  nm (427 of 474 restraints). Only 2% of the restraints are violated by  $>0.3$  nm [Fig. 2(A)]. When averaged over all experimentally determined NOE upper bounds, the simulated ensemble matches the experiment with an average deviation of only 0.027 nm [Fig. 2(B), leftmost bar] (this number climbs to 0.063 nm if one uses  $r^{-3}$ -averaging of distances). This is true even though the individual structures comprising the ensemble deviate significantly more from the experimental upper bounds [ $0.13 \pm 0.02$  nm on average, Fig. 2(B), central bar].

If one looks at the relationship between the all-heavy-atom RMSD from the native villin structure and the average violations of the NMR upper bounds for all the structures in the unfolded ensemble, one sees an approximately linear dependence [Fig. 2(C)]. Fitting a line constrained to pass through the origin to this data set allows one to estimate what all-heavy-atom RMSD would be predicted for a structure that violated the NMR upper

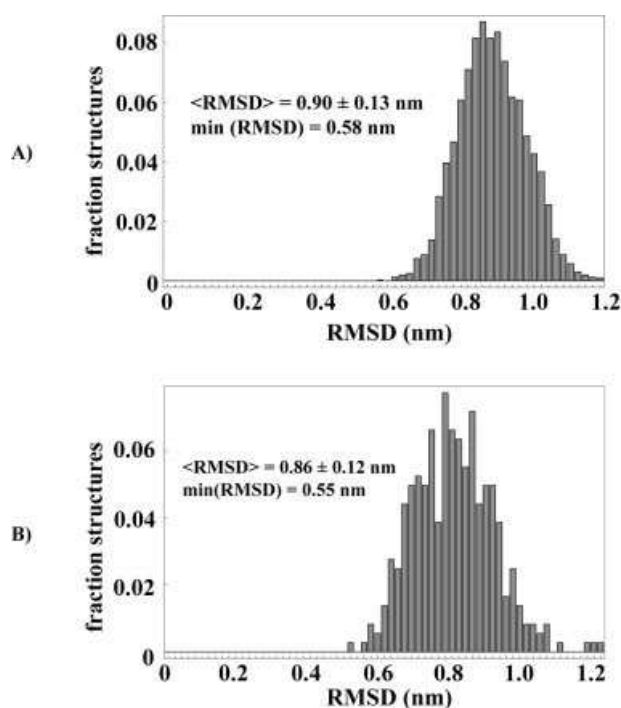


Fig. 1. Distribution of the all-heavy-atom RMSD from the NMR-derived model structures for the simulated unfolded-state ensemble of (A) villin, and (B) lysozyme. The mean and the standard deviation of the distributions are given on the left together with the lowest RMSD value in the distribution. The histograms were binned using 0.02-nm bins.

bounds to the same degree as the set of average distances from our simulation (0.027 nm). The inexactness and simplicity of this analysis notwithstanding, it is striking to see that the equivalent RMSD would be only 0.19 nm.

The explanation for such good match between the simulated unfolded ensemble of villin and the experimental distance upper bounds for native villin is twofold. The first answer lies in the highly nonlinear effect of  $r^{-6}$  averaging. Namely, for each particular pair of relaxation centers, it is enough that in just a small fraction of the entire ensemble the distance between the two centers is low, and the entire ensemble average of that particular distance will be low and will match the experimental distance restraint well. Burgi et al.<sup>26</sup> have discussed this problem on the level of individual distance distributions, but here is an example of how this effect manifests itself on the level of a large ensemble of polypeptide structures. It is relatively easy to match the experimental NOE restraints when dealing with a diverse, relatively compact set of structures.

Because the average distances are dominated by the low extremes in the distribution, many ensemble-averaged distances in the entire  $r^{-6}$ -averaged distance map will be low, and hence indicative of a potential NOE. In other words, the average distance map will contain significantly more contracted distances compared with distance maps based on individual structures. Herein lies the second explanation for the artificially good match between theory and experiment. Namely, when comparing simulation and experiment, we focused only on the list of experimentally derived restraints and looked for deviations from those in

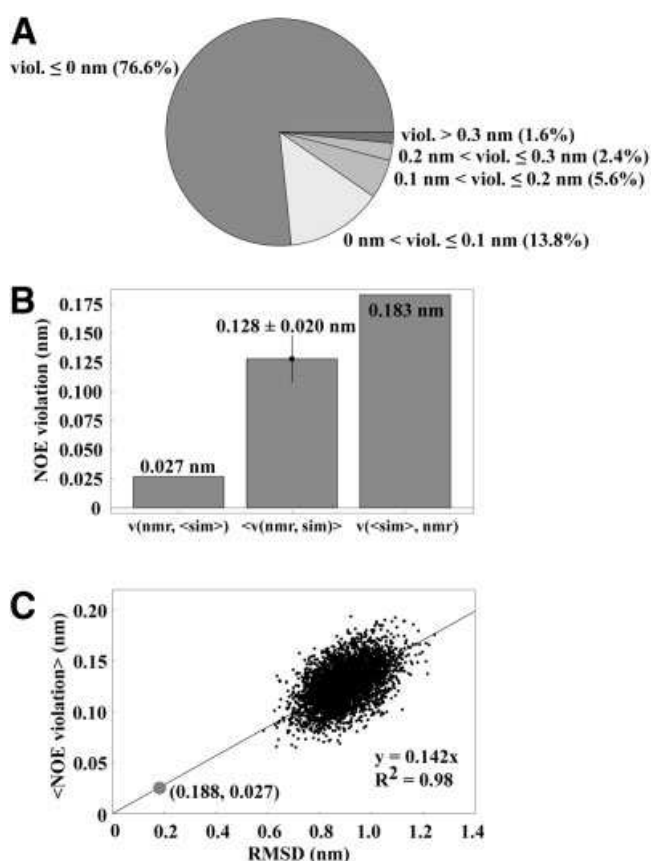


Fig. 2. **A:** Pie chart of the violations of the experimentally derived distance upper bounds for the  $r^{-6}$ -averaged distance map of the villin ensemble. **B:** Violation of the experimental restraints for the  $r^{-6}$ -averaged simulated distances averaged over all experimentally determined NOE pairs,  $v(nmr, \langle sim \rangle)$ ; the average violation of the experimental upper bounds for the individual members of the simulated ensemble (and standard deviation),  $\langle v(nmr, sim) \rangle$ ; violation of the distances calculated from the NMR-derived model structure from the theoretically predicted NOEs, averaged over all theoretically predicted NOEs,  $v(\langle sim \rangle, nmr)$ . **C:** The relationship between the average violation from the experimentally derived distance upper bounds and the positional RMSD from the NMR-derived model structure of villin (1VII) for all 3,999 simulated unfolded structures. The line is a linear fit constrained to pass through the origin (note that this leads to a redefinition of the  $R^2$  metric). The dot indicates the equivalent RMSD (0.188 nm) for an imaginary structure whose average deviation from the experimentally derived distance upper bounds would be the same as the corresponding deviation of the  $r^{-6}$ -averaged distance map in our simulations (0.027 nm).

the average simulated distance map. What happens if we look at all pairs of protons and ask for which of them the simulation would predict an NOE? Compared with the 474 experimentally seen NOEs, our simulation predicts a total of >4,700 NOEs (Fig. 3). In other words, >25% of all relaxation center pairs (4,756 of 18,528) in the molecule are predicted to be <0.55 nm apart after  $r^{-6}$  ensemble averaging.

In fact, given how many NOEs it predicts, the unfolded state ensemble of villin acts almost as a “shape-shifter” known from science-fiction literature. Namely, no matter which structure (i.e., a subset of experimental restraints) one confronts it with, the unfolded ensemble matches it well. In Figure 4 we show one particular structure taken

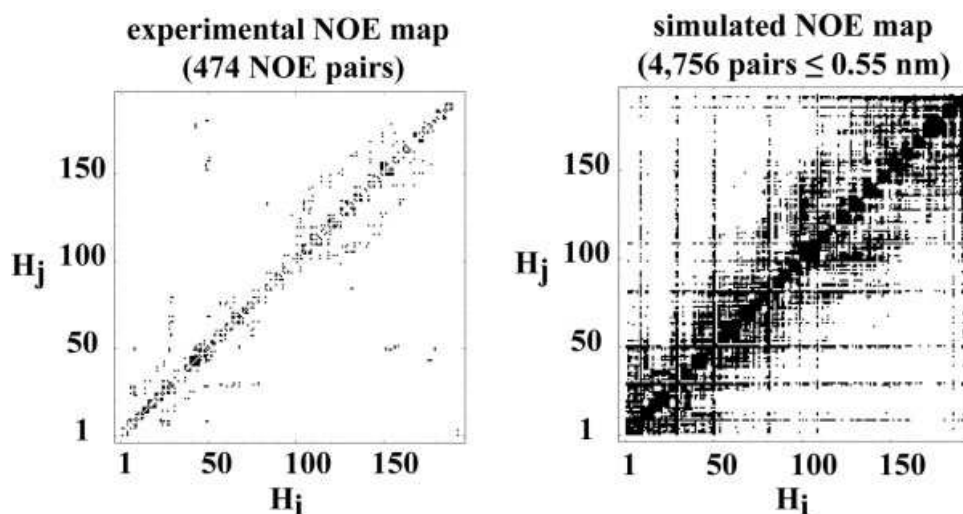


Fig. 3. Comparison of the experimentally measured NOE map (left) and the one derived based on the  $r^{-6}$ -averaged distance map of the simulated villin ensemble (right; all proton pairs separated by  $<0.55$  nm are depicted by black dots).

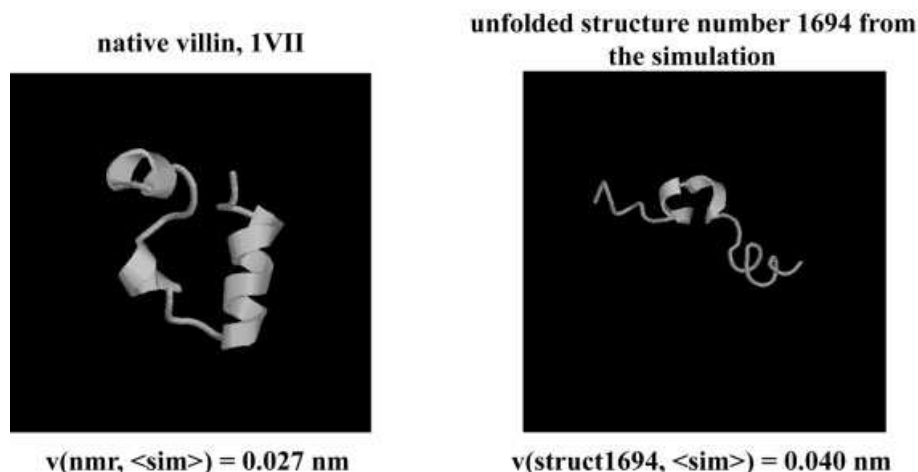


Fig. 4. NMR-derived model structure of villin (left). The  $r^{-6}$ -averaged distance map based on the simulated unfolded-state ensemble of villin violates the 474 experimental distance restraints by 0.027 nm (left). Single structure of villin from the simulated unfolded-state ensemble (right). The violation of the predicted NMR distance restraints derived from this structure for the  $r^{-6}$ -averaged distance map based on the simulated unfolded-state ensemble is only 0.040 nm. The all-heavy-atom positional RMSD between the native villin structure and the structure on the right is 1.25 nm.

from the simulated ensemble of unfolded structures. Its RMSD from the native villin is 1.25 nm and it is obvious that their dissimilarity could hardly be greater. However, the deviation between the NOE upper bounds predicted based on this structure and the ones based on the  $r^{-6}$ -averaged unfolded state distance map is only 0.040 nm. Of course, this is true only if one restricts one's attention solely to the pairs of protons that exhibit NOEs in the test structure.

The second example that we discuss deals with the high-temperature denatured state ensemble of lysozyme. It should be noted that the nonphysical nature of the simulation procedure is not relevant at all, because our aim was just to generate a diverse, highly nonnative set of structures. Indeed, the resulting ensemble at 50 ps is on

average  $0.86 \pm 0.12$  nm RMSD away from the native lysozyme structure [Fig. 1(B)] with only a small fraction of native secondary structure (20% of native helical content). Nevertheless, the experimental NOE-derived distance restraints are matched quite well (Fig. 5). Out of 1,632 experimental restraints, 1,191 (73%) are matched with 0-nm deviation by the  $r^{-3}$ -averaged distance map from the simulation at 50 ps. Furthermore, 90% of the experimental restraints are matched with  $<0.1$  nm violation, whereas only 3% are violated by  $>0.3$  nm [Fig. 5(A)]. The average violation of the experimentally derived NOE upper bounds for the  $r^{-3}$ -averaged distance map at 50 ps is only 0.036 nm, when averaged over all experimental restraints [Fig. 5(B), leftmost bars]. This number decreases to 0.014 nm if one uses  $r^{-6}$  averaging of distances. This is true even

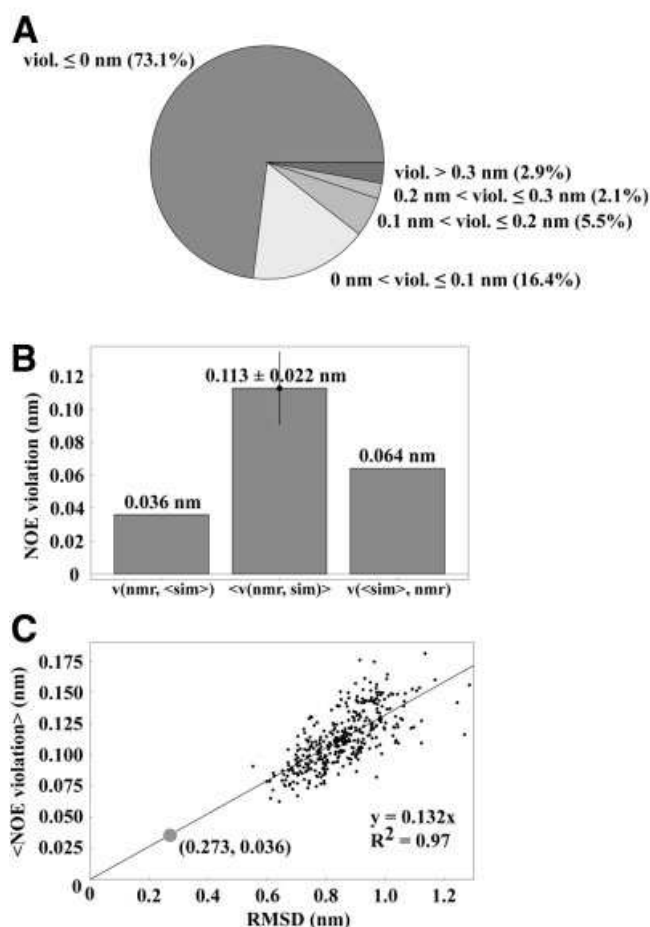


Fig. 5. **A:** Pie chart of the deviations from the experimentally derived distance upper bounds for the  $r^{-3}$ -averaged distance map of the simulated lysozyme ensemble; **B:** deviation of the  $r^{-3}$ -averaged simulated distances from the experimental upper bounds averaged over all experimentally determined NOE pairs,  $v(nmr, \langle sim \rangle)$ ; the average deviation of the individual members of the simulated ensembles (and standard deviation) from the experimental upper bounds,  $\langle v(nmr, sim) \rangle$ ; deviation of the distances calculated from the NMR-derived model structure from the theoretically predicted NOEs, averaged over all theoretically predicted NOEs,  $v(\langle sim \rangle, nmr)$ . **C:** The relationship between the average deviation from the experimentally derived distance upper bounds and the all-heavy-atom positional RMSD from the NMR-derived model structure of lysozyme (1LYM) for all 362 simulated structures. The line is a linear fit constrained to pass through the origin. The dot indicates the equivalent RMSD (0.273 nm) for an imaginary structure whose average deviation from the experimentally derived distance upper bounds would be the same as the corresponding deviation of the  $r^{-3}$ -averaged distance map in our simulations (0.036 nm).

though the deviation of the individual members of the simulated ensemble from the experimental restraints is significantly greater (e.g.,  $0.113 \pm 0.022$  nm on average for the ensemble at 50 ps) [Fig. 5(B), central bar]. If one again uses the simple assumption of linear dependence of the RMSD from the native lysozyme structure and the average violation of NOE upper bounds and calculates the equivalent all-heavy-atom RMSD for the average distance map from the simulations, one gets a low value of 0.27 nm [Fig. 5(C)].

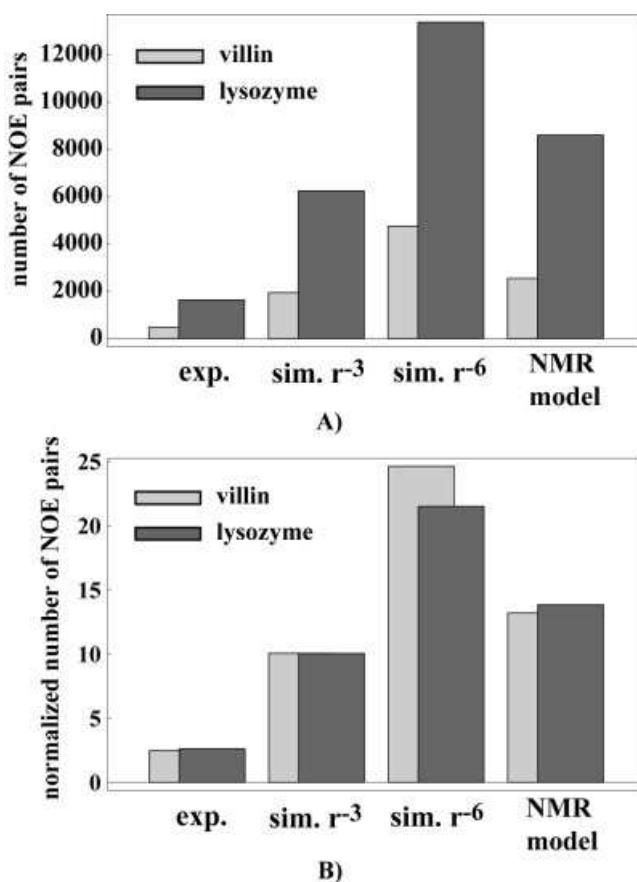


Fig. 6. **A:** The number of NOE upper bounds for villin and lysozyme based on: the NMR experiment (exp.), the  $r^{-3}$ - and the  $r^{-6}$ -averaged distance maps calculated for the simulated denatured ensembles (sim.  $r^{-3}$  and  $r^{-6}$ ), the reported NMR-derived model structures (NMR model). **B:** The same as in A with each value being normalized by dividing by the total number of relaxation centers in the molecule (villin: 193; lysozyme: 622).

As suggested by the previous example with villin, this conclusion can be explained by the properties of  $r^{-3}$  averaging of distances. Furthermore, the picture changes significantly if one looks at *all* of the NOEs that are predicted by the simulation. Figure 6(A) shows the number of experimental upper bound restraints compared with the number of NOE pairs predicted by simulation for both villin and lysozyme. Whereas in the experiment 474 NOE intensities were reported for villin (1,632 for lysozyme), our simulations predict a total of 1,941 NOEs using  $r^{-3}$  averaging or 4,756 NOEs using  $r^{-6}$  averaging (6,252 and 13,385 for lysozyme, respectively). Not surprisingly, the great majority of the experimentally observed NOEs are included in the simulated sets, but because these predict many more other NOEs, one would be hard pressed to ascribe them significant predictive capability.

## DISCUSSION

Oostenbrink et al.<sup>21</sup> have recently looked at the native state simulations of lysozyme and noted that, with the 53A6 GROMOS force field, the experimental NOE upper bounds are violated by only 0.018 nm on average in an

ensemble that deviates from the experimental structure by a backbone positional RMSD of 0.25 nm on average. In the example given here, even though the average backbone positional RMSD climbs to 0.71 nm [or 0.86 nm for the all-heavy-atom positional RMSD, Fig. 1(B)] with low native secondary structure content (20% of the native  $\alpha$ -helical content), the deviation from experimental restraints changes to 0.036 nm only. This deviation is, for example, comparable in absolute terms to the average violations for the simulated native DNA dodecamer from the corresponding experimental restraints (0.032 nm), reported in the same study.<sup>21</sup> The simulations of the native states of lysozyme and DNA dodecamer represent state-of-the-art level in molecular dynamics, and based on other evidence can be thought to be accurate.<sup>21</sup> This illustrates how the average deviation of upper bounds gives by itself only limited information about the actual quality of the simulated ensemble. However, this is also true when it comes to individual NOE deviations as well. For example, whereas the fraction of the violations  $>0.1$  nm in the case of native lysozyme simulations is 5%,<sup>21</sup> this number changes to only 10% in the case of the highly denatured ensemble from our simulations. Both of these percentages are relatively high compared with what is typically achieved in NMR structure refinement, but it is noteworthy that they should be in approximately the same range, given how different the two simulated ensembles are.

What can be done when comparing averages from large simulated ensembles with experimental NOE upper bounds to avoid some of the pitfalls exemplified here? First, enumerating all predicted NOE pairs based on average internuclear distances from simulation (as in Fig. 3) should always be done. The presence of a pair of hydrogen atoms at a short distance in the simulated ensemble need not necessarily mean that an NOE between the two will be detected in experiment (see discussion below). However, a simulation that predicts an exorbitant number of such pairs needs to raise one's suspicion.

Second, assuming that a "correct" NMR-derived model structure is available, the degree of overprediction of the number of NOEs in the simulation can be detected and quantified in the following manner. Instead of asking how much a set of distances from simulation deviates from the experimental restraints averaged over all of the experimental restraints, one may calculate the opposite measure. Namely, one can calculate how much the distances from the experimentally derived model structure deviate from what one predicts to be the distance restraints from simulation, averaged over all of the latter restraints.<sup>30</sup> For example, the distance map predicted based on the NMR structure of villin deviates on average  $>0.18$  nm from the  $r^{-6}$ -averaged distance map from the simulations, compared with only 0.027 nm using the reverse measure [Fig. 2(B)]. If the number of NOEs predicted from simulation was significantly less than what was actually seen, this figure would be much closer to the initial value of 0.027 nm. Indeed, in the case of lysozyme, because of less severe  $r^{-3}$  averaging of distances, the number of predicted NOEs based on simulation relative to the number of NOEs seen

in the experiment is significantly less than in the case of villin [Fig. 6(A)]. Consequently, whereas the  $r^{-3}$ -averaged simulated distance map deviates from the experimental restraints by 0.036 nm on average, the reverse measure of deviation gives a relatively close value of 0.064 nm [Fig. 5(B)].

In a typical NMR structural experiment, the set of NOEs that is measured and/or reported is significantly smaller than the set of all pairs of nuclei which are within 0.55 nm or so from each other in the reported structure.<sup>29</sup> For instance, the number of NOE intensities that would be predicted based on the native NMR-derived model structures of villin and lysozyme using our approach would be 2,547 and 8,609, respectively, which is significantly fewer than the 474 and 1,632 distance restraints, respectively, reported in the experiment [Fig. 6(A)]. The reasons for this discrepancy are multifaceted.<sup>1,2,29,49</sup> One important reason is spectral overlap and incomplete assignment. Second, spin diffusion has a significant role in relaxation in macromolecules and can lead to "washing out" of certain intensities. Third, NMR relaxation rates, and consequently NOE cross peak intensities, are affected by the intramolecular motions and their relationship with the overall tumbling. Finally, nonspecific leakage contributions from other relaxation mechanisms such as for example those arising from the interactions with the solvent can also affect the appearance of an NOE. Because of all of these reasons, negative information (i.e., the absence of an NOE between two relaxation centers) is only rarely used as a source of information for distance-based refinement.<sup>27,28,54</sup>

In the case of villin, the number of pairs of relaxation centers that are close to each other ( $<0.55$  nm) in the reported 1VII structure is significantly less than the number of predicted NOEs seen in the average distance maps from the simulations using  $r^{-6}$  averaging [2,547 versus 4,756, Fig. 6(A)]. Therefore, in this case there is a significant number of "new" NOEs that appear solely because of  $r^{-6}$  averaging and our argument about the simulations overpredicting the number of NOEs still stands. However, if one looks at the corresponding numbers in the lysozyme simulations, one sees that the  $r^{-3}$ -averaged distance map predicts fewer NOE pairs than does the distance map based on the native 1E8L structure (6,252 versus 8,609). This is also true in the case of villin if  $r^{-3}$  averaging is used (1,941 NOE pairs in the simulation, compared with 2,547 based on the NMR model). Therefore, in this case one cannot speak of a significant overprediction of NOEs because the same (and more) happens in the case of the native NMR-derived model structures. However, even here the  $r^{-3}$  averaging of the distances from simulations brings about low deviation from experimental NOE bounds on average.

For larger molecules, the relative fraction of all pairs of relaxation centers that are in close proximity to each other (and can hence exhibit an NOE) is less than for a smaller molecule. This is a direct consequence of the fact that the number of potential pairs scales with  $N^2$  whereas the number of those in close proximity scales only with  $N$ ,

where  $N$  is the number of relaxation centers. Indeed, in the case of villin, NOEs are predicted for 25% of all possible pairs of relaxation centers, whereas the equivalent number in the case of lysozyme is only 7%. Based on this, one might be tempted to conclude that the degree of overprediction of NOEs is less severe in the case of lysozyme. However, in Figure 6(B) we compare the total number of NOE pairs for the two simulations normalized by the total number of relaxation centers in the two molecules. When represented in this way, it is obvious that the two simulations are actually quite similar to each other in terms of the deviation from experiment when it comes to the number of NOE pairs they predict. We suggest that such normalization should always be performed when comparing two sets of simulations to avoid the size effects discussed here.

The calculation of average distances from simulation was performed herein according to simple  $\langle r_{i,j}^{-p} \rangle^{-1/p}$  averaging, where  $p$  is 3 or 6. The reverse of this approach is routinely used in NMR structure refinement when the measured intensities are related to distance bounds.<sup>2,49</sup> However, when it comes to calculating NOE intensities, this approach is a significant simplification compared with the full relaxation matrix approach.<sup>23</sup> Nevertheless, Feenstra et al.<sup>55</sup> have shown that using simple  $\langle r_{i,j}^{-6} \rangle^{-1/6}$  averaging results in only marginal errors compared with the more exact relaxation matrix approach, because the extent of sampling of phase space makes a bigger difference than the nature of the method for calculating NOE intensities. Also, because of its simplicity  $r_{i,j}^{avg} = \langle r_{i,j}^{-p} \rangle^{-1/p}$  is probably more frequently used, which was another reason why in this study we have used it. We expect that all of the arguments and criticisms given here should hold as well in the case of the average distances obtained using more rigorous methods. One advantage of the full relaxation matrix approach is that it allows one to predict the absence of specific cross peaks in the spectra, as long as it is caused by the intricacies of the relationship between the intramolecular motions and the overall tumbling motion.<sup>56</sup>

To conclude, we have shown how two very diverse and nonnative ensembles of simulated structures match experimental NOE upper bounds fairly well. These two examples should serve as a warning for simulators comparing their results with experiments, but also for experimentalists interested in understanding the limitations of the experimental observables they use to determine structure.

#### ACKNOWLEDGMENTS

The authors thank the thousands of Folding@Home contributors, without whom villin simulations would not be possible. A complete list of contributors can be found at <http://folding.stanford.edu>. B.Z. acknowledges support from EMBO postdoctoral fellowship. This work was financially supported by grants from the National Center of Competence in Research (NCCR) in Structural Biology of the Swiss National Science Foundation (W.v.G.; lysozyme simulations) as well as by an NIH grant (villin simulations), which is gratefully acknowledged.

#### REFERENCES

- Jardetzky O, Roberts GC. NMR in molecular biology. New York: Academic Press; 1981.
- Wuethrich K. NMR of proteins and nucleic acids: New York: John Wiley & Sons; 1986.
- Cavanagh J, Fairbrother WJ, Palmer AG, Skelton NJ. Protein NMR spectroscopy. New York: Academic Press; 1996.
- Levy RM, Dobson CM, Karplus M. Dipolar NMR relaxation of nonprotonated aromatic carbons in proteins. Structural and dynamical effects. *Biophys J* 1982;39(1):107–113.
- van Gunsteren WF, Berendsen HJ, Geurtsen RG, Zwinderman HR. A molecular dynamics computer simulation of an eight-base-pair DNA fragment in aqueous solution: comparison with experimental two-dimensional NMR data. *Ann NY Acad Sci* 1986;482:287–303.
- 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.
- Smith PE, van Schaik RC, Szyperski T, Wuthrich K, van Gunsteren WF. Internal mobility of the basic pancreatic trypsin inhibitor in solution: a comparison of NMR spin relaxation measurements and molecular dynamics simulations. *J Mol Biol* 1995;246(2):356–365.
- Daura X, Antes I, van Gunsteren WF, Thiel W, Mark AE. The effect of motional averaging on the calculation of NMR-derived structural properties. *Proteins* 1999;36(4):542–555.
- Schneider TR, Brunger AT, Nilges M. Influence of internal dynamics on accuracy of protein NMR structures: derivation of realistic model distance data from a long molecular dynamics trajectory. *J Mol Biol* 1999;285(2):727–740.
- Stocker U, Spiegel K, van Gunsteren WF. On the similarity of properties in solution or in the crystalline state: a molecular dynamics study of hen lysozyme. *J Biomol NMR* 2000;18(1):1–12.
- 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.
- Wong KB, Clarke J, Bond CJ, et al. Towards a complete description of the structural and dynamic properties of the denatured state of barnase and the role of residual structure in folding. *J Mol Biol* 2000;296(5):1257–1282.
- Pfeiffer S, Fushman D, Cowburn D. Simulated and NMR-derived backbone dynamics of a protein with significant flexibility: a comparison of spectral densities for the betaARK1 PH domain. *J Am Chem Soc* 2001;123(13):3021–3036.
- Case DA. Molecular dynamics and NMR spin relaxation in proteins. *Acc Chem Res* 2002;35(6):325–331.
- Wu X, Wang S, Brooks BR. Direct observation of the folding and unfolding of a beta-hairpin in explicit water through computer simulation. *J Am Chem Soc* 2002;124(19):5282–5283.
- Fersht AR, Daggett V. Protein folding and unfolding at atomic resolution. *Cell* 2002;108(4):573–582.
- Bruschweiler R. New approaches to the dynamic interpretation and prediction of NMR relaxation data from proteins. *Curr Opin Struct Biol* 2003;13(2):175–183.
- Zhou R. Free energy landscape of protein folding in water: explicit vs. implicit solvent. *Proteins* 2003;53(2):148–161.
- Soares TA, Daura X, Oostenbrink C, Smith LJ, van Gunsteren WF. Validation of the GROMOS force-field parameter set 45A3 against nuclear magnetic resonance data of hen egg lysozyme. *J Biomol NMR* 2004;30(4):407–422.
- Lindorff-Larsen K, Best RB, Depristo MA, Dobson CM, Vendruscolo M. Simultaneous determination of protein structure and dynamics. *Nature* 2005;433(7022):128–132.
- Oostenbrink C, Soares TA, van der Vegt NF, van Gunsteren WF. Validation of the 53A6 GROMOS force field. *Eur Biophys J* 2005;34(4):273–284.
- Withka JM, Swaminathan S, Srinivasan J, Beveridge DL, Bolton PH. Toward a dynamical structure of DNA: comparison of theoretical and experimental NOE intensities. *Science* 1992;255(5044):597–599.
- Keepers JW, James TL. A theoretical study of distance determinations from NMR. Two-dimensional nuclear Overhauser effect spectra. *J Magn Reson* 1984;57:404–426.
- Boelens R, Koning TMG, Kaptein R. Determination of biomolecu-

- lar structures from proton-proton NOEs using a relaxation matrix approach. *J Mol Struct* 1988;173:299–311.
25. Tropp J. Dipolar relaxation and nuclear Overhauser effects in nonrigid molecules: the effect of fluctuating internuclear distances. *J Chem Phys* 1980;72:6035–6043.
  26. Burgi R, Pitera J, van Gunsteren WF. Assessing the effect of conformational averaging on the measured values of observables. *J Biomol NMR* 2001;19(4):305–320.
  27. de Vlieg J, Boelens R, Scheek RM, Kaptein R, van Gunsteren WF. Restrained molecular dynamics procedure for protein tertiary structure determination from NMR data: a Lac repressor headpiece structure based on information on J-coupling and from presence and absence of NOE's. *Isr J Chem* 1986;27:181–188.
  28. Wilcox GR, Fogh RH, Norton RS. Refined structure in solution of the sea anemone neurotoxin ShI. *J Biol Chem* 1993;268(33):24707–24719.
  29. Dorelejers JF, Ravés ML, Rullmann T, Kaptein R. Completeness of NOEs in protein structure: a statistical analysis of NMR. *J Biomol NMR* 1999;14(2):123–132.
  30. Gronwald W, Kirchhofer R, Gorler A, et al. RFAC, a program for automated NMR R-factor estimation. *J Biomol NMR* 2000;17(2):137–151.
  31. Grishaev A, Llinas M. CLOUDS, a protocol for deriving a molecular proton density via NMR. *Proc Natl Acad Sci USA* 2002;99(10):6707–6712.
  32. McKnight CJ, Matsudaira PT, Kim PS. NMR structure of the 35-residue villin headpiece subdomain. *Nat Struct Biol* 1997;4(3):180–184.
  33. Duan Y, Kollman PA. Pathways to a protein folding intermediate observed in a 1-microsecond simulation in aqueous solution. *Science* 1998;282(5389):740–744.
  34. Zagrovic B, Snow CD, Shirts MR, Pande VS. Simulation of folding of a small alpha-helical protein in atomistic detail using worldwide-distributed computing. *J Mol Biol* 2002;323(5):927–937.
  35. Schwalbe H, Grimshaw SB, Spencer A, et al. A refined solution structure of hen lysozyme determined using residual dipolar coupling data. *Protein Sci* 2001;10(4):677–688.
  36. Zagrovic B, Sorin EJ, Pande V. Beta-hairpin folding simulations in atomistic detail using an implicit solvent model. *J Mol Biol* 2001;313(1):151–169.
  37. Snow CD, Nguyen H, Pande VS, Gruebele M. Absolute comparison of simulated and experimental protein-folding dynamics. *Nature* 2002;420(6911):102–106.
  38. Qiu D, Shenkin PS, Hollinger FP, Still WC. The GB/SA continuum model for solvation. A fast analytical method for the calculation of approximate Born radii. *J Phys Chem* 1997(101):3005–3014.
  39. Andersen HC. Rattle: a “velocity” version of the Shake algorithm for molecular dynamics calculations. *J Comp Phys* 1983;52:24–34.
  40. Jorgensen WL, Tirado-Rives J. The OPLS potential functions for proteins: energy minimizations for crystals of cyclic peptides and crambin. *J Am Chem Soc* 1988;110:1666–1671.
  41. Oostenbrink C, Villa A, Mark AE, van Gunsteren WF. A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6. *J Comp Chem* 2004;25(13):1656–1676.
  42. van Gunsteren WF, Billeter SR, Eising AA, et al. Biomolecular simulation: the GROMOS96 manual and user guide. Zurich: Biomos; 1996.
  43. Berendsen HJC, Postma JP, van Gunsteren WF, DiNola A, Haak JR. Molecular dynamics with coupling to an external bath. *J Chem Phys* 1984;81:3684–3690.
  44. Ryckaert JP, Ciccotti G, Berendsen HJC. Numerical-integration of Cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J Comp Phys* 1977;23:327–341.
  45. Kabsch W, Sander C. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* 1983;22(12):2577–2637.
  46. Berman HM, Westbrook J, Feng Z, et al. The Protein Data Bank. *Nucleic Acids Res* 2000;28(1):235–242.
  47. Constantine KL, Madrid M, Banyai L, Trexler M, Patthy L, Llinas M. Refined solution structure and ligand-binding properties of PDC-109 domain b. A collagen-binding type II domain. *J Mol Biol* 1992;223(1):281–298.
  48. Fletcher CM, Jones DNM, Diamond R, Neuhaus D. Treatment of NOE constraints involving equivalent or nonstereoassigned protons in calculations of biomacromolecular structures. *J Biomol NMR* 1996;8:292–310.
  49. Roberts GC. NMR of macromolecules: a practical approach. Oxford: IRL Press; 1993.
  50. Eaton WA, Munoz V, Hagen SJ, et al. Fast kinetics and mechanisms in protein folding. *Annu Rev Biophys Biomol Struct* 2000;29:327–359.
  51. Kubelka J, Hofrichter J, Eaton WA. The protein folding ‘speed limit.’ *Curr Opin Struct Biol* 2004;14(1):76–88.
  52. Kubelka J, Eaton WA, Hofrichter J. Experimental tests of villin subdomain folding simulations. *J Mol Biol* 2003;329(4):625–630.
  53. Zagrovic B, Snow C, Khaliq S, Shirts M, Pande V. Native-like mean structure in the unfolded ensemble of small proteins. *J Mol Biol* 2002;323(1):153–164.
  54. Summers MF, South TL, Kim B, Hare DR. High-resolution structure of an HIV zinc fingerlike domain via a new NMR-based distance geometry approach. *Biochemistry* 1990;29(2):329–340.
  55. Feenstra KA, Peter C, Scheek RM, van Gunsteren WF, Mark AE. A comparison of methods for calculating NMR cross-relaxation rates (NOESY and ROESY intensities) in small peptides. *J Biomol NMR* 2002;23(3):181–194.
  56. 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.