THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

Protein Structure Validation Using Side-Chain Chemical Shifts

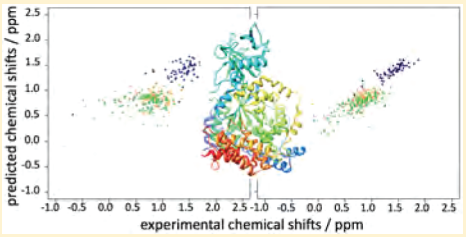
Aleksandr B. Sahakyan,[†] Andrea Cavalli,[†] Wim F. Vranken,[‡] and Michele Vendruscolo*[,†]

[†]Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, U.K.

[‡]Department of Structural Biology, VIB and Structural Biology Brussels, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussels, Belgium

Supporting Information

ABSTRACT: We present a method of assessing the quality of protein structures based on the use of side-chain NMR chemical shifts. Because these parameters are very accurate reporters of side-chain positions and are highly sensitive to tertiary structure and packing, they are particularly useful for structure validation. To analyze a given structure, we define a quality score, Q_{CS}, that compares the chemical shifts calculated from such a structure with the corresponding experimental values in a way that takes account of the errors in the calculations. The results that we report illustrate the advantages in the examination of the quality of protein structures from the perspective of side-chains.



INTRODUCTION

Owing to recent advances in genome sequencing,[1,2] the rate at which new protein-encoding genes are identified is far faster than the rate at which the structures of the corresponding proteins are determined. It is therefore important to develop methods to speed up the process of protein structure determination. Indeed, one of the major aims of structural genomics initiatives is to determine at least one representative three-dimensional structure for all known protein families.[3]

Although X-ray crystallography has a major role in these efforts, there is interest in developing methods based on nuclear magnetic resonance (NMR) spectroscopy[4−6] because they can be applied in the solution state, which closely resembles the conditions under which proteins carry out their functions, and because often proteins cannot be readily crystallized. Great advances in this direction have been made in the past 15 years, resulting in an increase in the precision and type of NMR measurements[7−12] and in the size of proteins that can be studied.[13] In this context, the introduction of the novel isotope labeling techniques[14−16] is of particular importance since the knowledge of the chemical shifts of side-chain methyl and aromatic nuclei provides access to the solution-state structure and dynamics of supramolecular complexes.[13]

To increase the role of NMR spectroscopy in structural genomics, it is very important to develop automated methods of data acquisition and processing. With such methodologies, the rate-limiting step of the structure determination procedure would become the preparation of samples.[17] Standard NMR techniques for determining protein structures consist of multiple stages, some of which can require a substantial amount of time. The stages directly linked to NMR data acquisition and analysis include data recording, assignment of the spectra, interpretation of NOE (nuclear Overhauser enhancement) signals, and structure calculation and validation. To shorten the time required by these stages, one approach is to substantially decrease the amount of data needed to resolve the structures of proteins. To this end, recent developments in methods that use chemical shifts to determine high-resolution protein structures[18−21] can become very helpful to increase the throughput of NMR strategies, since the data acquisition is minimized to the most basic experiments and interpretation procedures needed to assign the resonance signals.

Regardless of the nature of the strategies for speeding up NMR-based structure determination, the role of structure validation increases substantially upon automation of the structure determination workflow. This aspect is particularly important since NMR spectroscopy, unlike X-ray crystallography, currently still lacks consensus intrinsic measures of structural quality. Moreover, NMR structures are usually obtained with the aid of molecular mechanics force fields because NMR measurements alone are generally not sufficient by themselves to completely define the three-dimensional structure of a protein. NMR data interpretation and processing are thus potentially prone to errors. Cases are known in which the misinterpretation of even a small number of NOE cross-peaks resulted in incorrect structures.[22,23] For instance, it has been demonstrated that the lack of knowledge about the oligomeric state of a protein may result in a misinterpretation of the spectra so that homomeric protein complexes can be considered as monomeric structures.[24]

Since it is important to have structure validation tools for NMR structures that are fully based on the most basic NMR measurements, we present here an approach to validate protein structures by using only chemical shifts. By using recently developed structure-based predictors of side-chain chemical shifts,[25,26] which report on the fine details of protein side-chain positions and are very sensitive to tertiary structure and packing in proteins, we show that these NMR parameters can be used to examine the quality of protein structures from the perspective of the side-chains.

Received: December 18, 2011
Revised: March 26, 2012
Published: March 28, 2012

   dx.doi.org/10.1021/jp2122054 | J. Phys. Chem. B 2012, 116, 4754−4759

METHODS

The approach and the applications reported here are based on recently developed structure-based predictors of protein side-chain chemical shifts.[25,26] We note, however, that the approach we discuss can be used with any chemical shift prediction engine. In the predictors that we used, chemical shifts are represented as a combination of phenomenological terms that report on the influence of dihedral angle, electric field, magnetic anisotropy and ring current effects on nuclear shielding,[25,26] and nonphenomenological distance-based terms that increase the performance of the model.[25−27]

An important aspect that should be considered in establishing a general structure validation method is that any structure-based chemical shift predictor (CSP) is associated with a certain error, which is defined here as the absolute difference between the predicted and experimental chemical shifts of the given query nucleus. A key component of our method is that, having also the full profile of the CSP performance from the leave-one-out tests on a large database of proteins for each atom type, we can estimate the probability of the predictor to result in the observed error. Such probability estimates can be calculated by binning the absolute error scale and calculating the fraction of instances when the CSP results in an error within the bin range. This kind of binning would, however, decrease the number of available entries for calculating the probability estimates and, thus, the statistical significance of the resulting numbers.

To this end, we calculate the probability of the predictor to result in an error larger than the observed error, rather than the one within a certain bin. The resulting quality score for chemical shifts (Q_{CS}) thus shows the probability that the prediction error is caused by the CSP rather than inaccuracies in the protein structure under analysis. Hence, a low value of the Q_{CS} score indicates the possible presence of problems in the structure. To further increase the statistical significance of the test, we use multiple chemical shifts from a given residue to extract joint probabilities. For instance, if two methyl ^{1}H chemical shift measurements are available for a valine residue (for $H_{\gamma 1}$ and $H_{\gamma 2}$ atoms) or if two or three signals are assigned for a phenylalanine residue belonging to any of the H_δ, H_ε, and H_ζ atoms, then we calculate the joint probabilities of all the NMR resolved nuclei in the given side-chain to end up in prediction errors larger than the observed ones. For a residue in a protein with measured experimental chemical shifts for the nuclei *i*, *j*, ..., the chemical shift based structural quality factor, Q_{CS}, can thus be calculated as

$$Q_{CS} = P\left(\left|\delta_i^{exp} - \delta_i^{calc}\right| \geq \sigma_i^{calc} \cap \left|\delta_j^{exp} - \delta_j^{calc}\right| \geq \sigma_j^{calc} \cap \ldots\right) \quad (1)$$

where $|\delta_i^{exp} - \delta_i^{calc}|$ is the absolute value of the error of the chemical shift prediction for the nucleus *i* of the given residue, σ_i^{calc} is the standard error of the CSP in reproducing the experimental chemical shifts for the type of nucleus *i* in the given residue type, and the ∩ symbols indicate that all the conditions on the different nuclei should be considered simultaneously (joint probability). Values of the Q_{cs} score close to 0 indicate the likely presence of errors in the structure, while values close to 1 support the validity of the structure.

The probabilities that we used are fairly accurate because of the presence of a relatively large database on which the chemical shift predictions are benchmarked via leave-one-out tests. In the current implementation, a carefully filtered database for aromatic side-chain protons uses 1796 entries for phenylalanine and 1498 entries for tyrosine hydrogen atoms coming from 452 proteins.[26] The methyl group database used to parametrize the CH3Shift predictor uses 17873 chemical shift entries from proteins corresponding to 682 unique PDB identifiers.[25]

The method is available at http://www-vendruscolo.ch.cam.ac.uk/software.html.

RESULTS AND DISCUSSION

Chemical shifts are routinely recorded during the initial stage of NMR data processing. Moreover, chemical shifts can often be measured even from very problematic systems, such as protein aggregates[28] and intrinsically disordered proteins.[29] The advances over the past decades that have been made to better understand the nature of chemical shifts[30,31] and the developments of fast and efficient structure-based chemical shift prediction methods,[25−27,32,33] have made it possible to substantially increase the scope of chemical shifts in structural biology.[18−21] It is thus timely to further extend the use of chemical shifts for protein structure validation.

Chemical shifts are extremely sensitive to the specific structural features of protein conformations. Any change in the atomic environment of a given nucleus can significantly alter its observed chemical shift value. Therefore, any imprecision of the structure at the vicinity or in the position of the query atom will inevitably become evident from the structure-based chemical shift predictions for that nucleus. Therefore, if one can clearly differentiate between the errors that are normally expected from a specific chemical shift prediction from the apparent errors that the prediction produces, a measure of structural imprecision at the given site of the protein structure can be devised, as described in the Methods section. Previously, protein backbone chemical shifts have been used to assess structural qualities of proteins by a comparison between experimental chemical shifts and those back-calculated from the protein structures under consideration using parametrizations based on either first principles[34] or empirical[35] methods. The Q_{CS} score, which we describe in this work, takes into account the errors intrinsic to the predictor and thus exploits chemical shifts for protein structure validation in a quantitative way. In addition, side-chain ^{1}H chemical shifts are particularly suitable for protein validation purposes, since they are strongly dependent on tertiary contacts, and unlike backbone atoms, side-chains are not shielded from the surrounding by the other moieties of the amino acid residues.

A low value of the Q_{CS} score indicates a possible structural imprecision because the discrepancies between the experimental and calculated chemical shifts are larger than the intrinsic errors in the chemical shift calculations themselves (eq 1); this method, of course, assumes that there are no assignment errors in the NMR spectra.

In the following, we present a series of applications that demonstrates the usefulness of the structure validation approach that we describe in this work. This method was prompted by the initial observation that the analysis of aromatic proton chemical shifts over 452 proteins in a database of high-resolution X-ray structures identified several proteins for which the prediction quality was rather poor as assessed by the comparison with the experimental NMR measurements.[26] Because the predictions were performed through a protein-based leave-one-out tests, the predictor was not biased toward a particular protein because of that protein being involved in the

dx.doi.org/10.1021/jp2122054 | J. Phys. Chem. B 2012, 116, 4754−4759

parametrization. Examination of all the poorly performing structures revealed that all of them were different in conformation from the corresponding NMR structures obtained from the solution-state experiments from which the chemical shifts were measured.[26] The reason for the difference was either a substantial conformational rearrangement upon, for example, Ca^{2+} ion or ligand binding or the presence of missing or extra peptide segments in either the solid or solution states. These observations clearly indicated that some errors resulting from the structure-based chemical shift predictions are capable of revealing actual structural inaccuracies in the structural model or a mismatch between the experimental data and the structure that is evaluated against those data.

We first present the analysis of the Q_{CS} scores for four protein structures (Figure 1): ubiquitin (PDB id 1UBQ,[36]

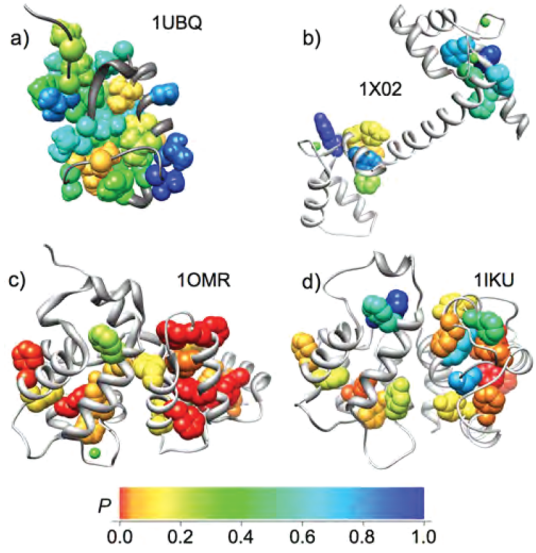


Figure 1. Examples of protein structure validation based on side-chain chemical shifts. Side-chains bearing methyl or aromatic groups are shown in space-filling representation and colored according to their Q_{SC} scores: (a) Ubiquitin (1UBQ); (b) Calmodulin (1X02); (c) X-ray structure (1OMR) and (d) NMR solution-state structure (1IKU) of Ca^{2+}-bound recoverin. Values of the Q_{cs} score close to 0 (shown in red) suggest the presence of errors in the structure, while values close to 1 (shown in blue) indicate that the structure is likely to be valid.

Figure 1a), calmodulin (1X02,[15] Figure 1b), and recoverin in its Ca^{2+}-bound state obtained from X-ray crystallography (1OMR,[37] Figure 1c) and in its Ca^{2+}-free myristoyl-bound state obtained from solution state NMR (1IKU,[38] Figure 1d). The residues that show low Q_{CS} scores are frequently clustered together, indicating the presence of a local problem in the structure. For instance, none of the NMR structural ensembles and the X-ray structure of ubiquitin that we analyzed reproduce the experimental ^1H chemical shifts for the aromatic Phe-45 ring (Figure S1 of the Supporting Information and ref 26). A map of the Q_{CS} scores obtained by analyzing the X-ray 1UBQ structure[36] (Figure 1a) shows that Leu-50, which is in the vicinity of Phe-45, is also showing low Q_{CS} scores as judged from the methyl group ^1H chemical shifts (Figure S2 of the Supporting Information). This result suggests that Phe-45 may be undergoing complex conformational fluctuations that affect its own chemical shifts as well as those of the sites close to it, and neither the X-ray structure of ubiquitin nor the NMR ensembles that we considered fully represent the dynamics of that residue. It is also interesting that the 2K39 ensemble of ubiquitin[39] shows the presence of several rotameric states for Phe-45 (Figure S1 of the Supporting Information) but does not appear to capture the correct weights of different states, since like in the other ensembles, the average predictions of the aromatic protons of Phe-45 do not agree well with the experimental chemical shifts.[26]

We found a similar patchy behavior (Figure 1b) of low Q_{CS} scores for the calmodulin 1X02 ensemble,[15] which indicates the possible presence of structural inaccuracies in this set of structures. In particular, the conformational fluctuations of the spatially neighboring Tyr-138 and Phe-89 residues might not be fully represented by the 20 structures that comprise the ensemble.[26] Another interesting case is that of recoverin,[26] for which two structural models are available, obtained from solid[37] and solution[38] states. These structures are fairly dissimilar in conformation (Figures S3 of the Supporting Information), primarily because the X-ray structure represents the Ca^{2+}-bound state. Since the NMR chemical shift measurements had been carried out for the Ca^{2+}-free solution-state of recoverin, we expect the solid-state structure to result in an overall lower Q_{CS} score spread for the aromatic residues with available ^1H chemical shift measurements (Figure 1c and d). In addition, some imprecision in Q_{CS} scores even in the Ca^{2+}-free state of recoverin can be explained by not accounting for the unconventional (myristoylated) moiety of the protein by our chemical shift predictor.

The development of our CSPs is based on a database of chemical shifts measured in solution state NMR and average structures compiled from a data set of high-resolution X-ray structures.[25,26] In many cases, side-chain chemical shifts results from complex dynamics involving different rotameric states of side-chains, where low populations of certain states with extreme values of chemical shifts can significantly alter the measured average values of the chemical shifts. However, the use of large databases of high-resolution structures enables accurate chemical shift predictions that reflect the correct relationship between structures and chemical shifts.[26] Therefore, the violations observed in the comparison between the calculated and the experimental chemical shift values from specific side-chains indicate that often an average structure does not provide a good representation of the state of a protein, either because the actual structure is different or because that particular side chain possesses a complex dynamic behavior. The latter explanation is most probably what we observe for the moiety surrounding or including the Phe-45 residue in ubiquitin. Hence, to refine the population of such states, we propose to use chemical shift predictors to generate structural restraints in molecular dynamics simulations, through which more realistic ensembles capturing the invisible states of proteins can be obtained.

To directly demonstrate that the proposed quality score is sensitive to the actual structural quality of proteins, we use an unfolding trajectory of a protein with known X-ray structure from one of our recent papers,[26] generated via a high temperature molecular dynamics simulation. Since the unfolding is done in silico and we have the snapshots of the structures along the trajectory and, hence, the structural RMSDs relative to the crystallographic structure, we can check whether the worsening of the structural quality upon unfolding is also accompanied by the worsening of the Q_{SC} score. The results are presented in Figure 2, where the negative sum of all the aromatic side-chain chemical shift-based quality

dx.doi.org/10.1021/jp2122054 | J. Phys. Chem. B 2012, 116, 4754–4759

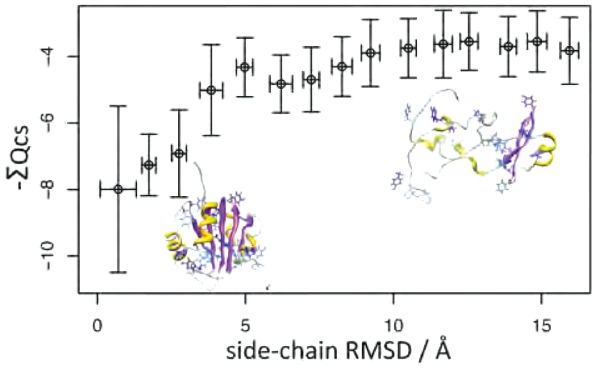


Figure 2. Relationship between the side-chain RMSDs (in Å) and the sum of all the aromatic side-chain Q_{SC} scores along the unfolding pathway of the DNA-binding domain of SV40 T-antigen. The unfolding trajectory corresponds to a 17 ns high-temperature molecular dynamics simulation.[26] Structural snapshots extracted at 7 ps intervals are analyzed. The $\sum Q_{SC}$ is changed in sign to facilitate the comparison with previous results.[26] The data are obtained by averaging all the quality scores within 1.1 Å bins of structural rmsd. The whiskers indicate the standard deviations of both the structural rmsd (x axis) and $-\sum Q_{SC}$ (y axis) within the 1.1 Å bins of structural rmsd.

scores in the examined protein, $(-\sum Q_{SC})$, are plotted against the side-chain structural root-mean-squared deviations (rmsd in angstroms), indicating that Q_{SC} indeed reports on the structural precision. It is noteworthy that the deviation of Q_{SC} at the lower structural rmsd region is relatively greater as compared with its spread in the region corresponding to the completely unfolded structures (Figure 2), which might be an indication that accounting for the structural dynamics is important for native states, and it is the averaged observable over the dynamical ensemble that results in reliable values.

We expect that structure-based methods for calculating side-chain chemical shifts as well as those that exploit such calculations for structure validation purposes will be particularly useful to examine structural models of large proteins. In these cases, side-chain chemical shifts are among the few NMR parameters that can be measured, in particular considering the recent advances in selective isotope labeling techniques for side-chains.[14−16] We have thus tested the validation technique introduced in this work against the available structural models of the largest single-chain protein studied so far by NMR spectroscopy, the 723-residue malate synthase G (MSG). Two models, 1P7T[40] and 1D8C,[41] determined by X-ray crystallography at ∼2.0 Å resolution, have been analyzed along with the 1Y8B set of 10 NMR structures[42] and the 2JQX solution structure[43] refined against NMR and small-angle X-ray scattering data. Two structures from the 1P7T PDB entry that comprise the elementary cell have been considered separately for validation (1P7T_a and 1P7T_b). The missing segments in the X-ray structures have been modeled using the Modeller program[44] with 100 different structural variants created for the missing loops of each X-ray structure. By using multiple variants for the modeled loops, we assessed the influence of structural uncertainties arising from the in silico addition of missing segments in the X-ray structures.

Figure 3 shows the correlation graphs between the experimental[45,46] and calculated chemical shifts for the structures of MSG mentioned above. Because multiple conformers are available, whiskers are included in the graphs to indicate the range of the calculated chemical shift variations,

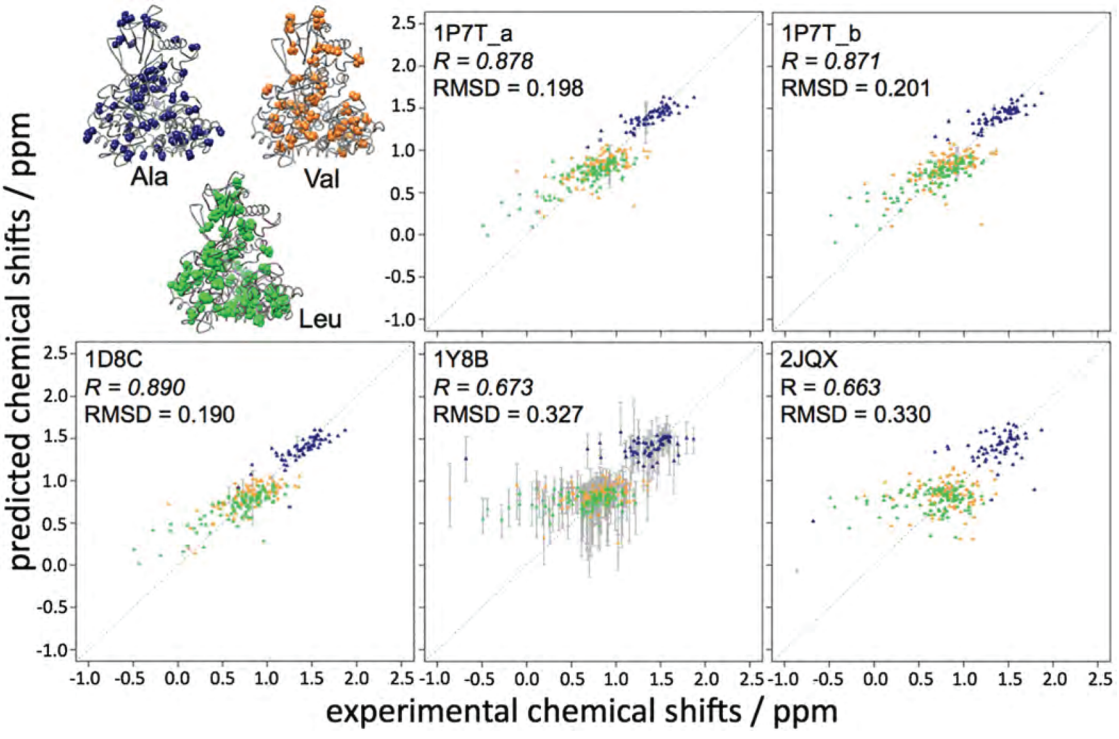


Figure 3. Comparison between predicted and experimental chemical shifts (in ppm) for side-chain methyl hydrogen atoms of alanine (dark blue), valine (orange), and leucine (green) residues of the available structures and structural ensembles of malate synthase G determined by X-ray crystallography and NMR spectroscopy. PDB codes, Pearson correlation coefficients (R) and root-mean-squared deviations (rmsd) are shown for each case. The whiskers indicate the range of the predicted chemical shifts for the models consisting of multiple structures.

dx.doi.org/10.1021/jp2122054 | J. Phys. Chem. B 2012, 116, 4754−4759

in addition to the triangles that show the average chemical shift values. The uncertainties in the modeled loop conformations of the X-ray structures of MSG affect only few chemical shifts (Figure 3a–c), for which the whiskers indicate a fairly small variance because the residues with available methyl group chemical shift measurements are distant from the modeled loops. Pearson correlation coefficients (R) and root-mean-squared deviations (rmsd) are shown on the graphs for each structural model of MSG. These results indicate that further experimental information will be required to improve the accuracy of the Ala, Val, and Leu side-chain conformations, including the fine details in their three-dimensional packing. A plot of the Q_{CS} scores along the sequence of MSG, as well as their cumulative sum, report on the overall structural quality of the different structural models (Figure 4). The heights of the bars are proportional to the Q_{CS} scores for the different residues, and the colors are from the same scale displayed in Figure 1. The transparent red bands indicate the regions where the confidence in structural inaccuracies is higher. Both the X-ray and NMR structures show inaccuracies in the representation of side-chain geometries and tertiary contacts (Figure 4). However, the NMR structures tend to give lower Q_{CS} scores in the tests of the structural quality as assessed from the side-chain perspective. Hence, the inclusion of side-chain NMR data in structure determination protocols is expected to improve the quality of the structures themselves by increasing the accuracy of NMR for side-chains that form protein interior and exterior surfaces, which are particularly important to study protein–ligand and protein–protein interactions.

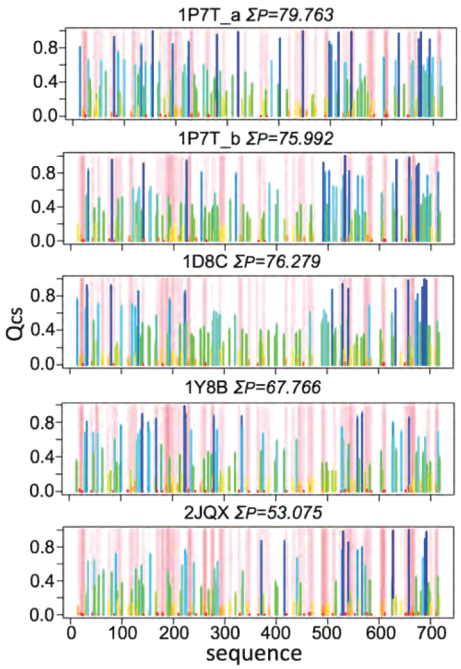


Figure 4. Q_{SC} scores plotted against the sequence index of the methyl-bearing amino acid residue in the X-ray structures and NMR ensembles of malate synthase G. The colors of the bars follow the same scale used in Figure 1. Transparent red bands identify the regions of the sequence for which the validation method predicts structural imprecision with high confidence. PDB codes and the total Q_{SC} scores are shown for each plot.

Because NMR order parameters (S^2) are also available for MSG,[47] we verified whether the errors in the chemical shift calculations and the Q_{CS} scores correlate with the S^2 values at the corresponding sites. We do not expect to find a significant correlation, since the Q_{CS} scores of methyl group ^{1}H should be primarily affected by the anisotropy of the motion and the nonlinear dynamics between different states. Indeed, we found no correlation between chemical shift prediction errors and S^2 values (Figure S4 of the Supporting Information). We then examined the correlation between the amplitude of the structural fluctuations derived from crystallographic B factors over the range of 682 proteins and the chemical shift prediction errors. Also in this case, we did not observe any significant correlation (Figure S5 of the Supporting Information).

CONCLUSIONS

We have described a method of using chemical shifts to validate protein structures and identify regions of possible structural inaccuracies. Although the method that we have presented can be readily used with any atom type for which a structure-based chemical shift predictor exists, we have focused our attention on side-chain proton chemical shifts because their values exibit a strong dependence on tertiary interactions and spatial effects. By contrast, backbone or side-chain carbon chemical shifts are prevalently determined by backbone conformation, rotameric states, and covalent interactions. Validation methods based on exclusively side-chain chemical shifts can exploit recent advances in labeling techniques, which are making it possible to measure side-chain chemical shifts for very large proteins and protein complexes by NMR.[14−16]

The availability of a chemical shift-based approach for protein structure determination may offer several opportunities to the NMR community:

(a) The method is based on NMR parameters; hence, the complications involved in the use of other experimental techniques and measurements is not required.

(b) The method uses NMR parameters that are generally measured, but not often used directly in NMR structure calculations. Indeed, NMR resonance signal assignment is the crucial first step in obtaining other parameters, such as RDCs and NOE intensities, which are used in standard methods of protein structure determination. Thus, all the measured RDC and NOE data can be used in structure determination, since it would not be necessary for some of them to be left out for further usage in structure validation.

(c) Protein structures can be analyzed from two different perspectives: those of backbone and side-chain atoms. Since backbone chemical shifts are more sensitive to the core effects and the conformation of peptide moieties, they report on the quality of the overall fold. In contrast, side-chain chemical shifts, especially the ^{1}H ones, are very sensitive to the weak interactions between spatially adjacent atoms, so they reflect accurately the fine details of the three-dimensional packing.

(d) Because chemical shifts are the most basic parameters measured in NMR spectroscopy, protein validation methods based on these parameters can become the method of choice for future high-throughput protein structure determination protocols. This strategy will decrease the number of measurements and reduce the time required for the automatic analysis of spectra and structure determination. More generally, protein structure determination and validation methods based solely

dx.doi.org/10.1021/jp2122054 | J. Phys. Chem. B 2012, 116, 4754–4759

on backbone and side-chain chemical shifts can broaden the scope of NMR spectroscopy.

■ ASSOCIATED CONTENT

Supporting Information

Five figures: a representation of the aromatic side-chains in different structural ensembles of ubiquitin, a stereo view of ubiquitin structure with methyl-bearing and aromatic amino acid side-chains colored in accordance to the Q_{cs} scores, a stereo view of two different structures of recoverin, a graph representing the absence of correlation between the generalized order parameter and CH3Shift prediction errors for methyl groups in malate synthase G, a graph representing the absence of clear dependence between the crystallographic B-factor-derived structural fluctuation amplitudes and the CH3Shift prediction errors tested on the database of 682 structures. This material is available free of charge via the Internet at http://pubs.acs.org/.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: mv245@cam.ac.uk.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

A.B.S. thanks the Herchel Smith Fund for generous support. W.F.V. acknowledges the Brussels Institute for Research and Innovation (Innoviris) Grant BB2B 2010-1-12. M.V. acknowledges funding from the Leverhulme Trust, EMBO, BBSRC, and the Royal Society.

■ REFERENCES

(1) International Human Genome Sequencing Consortium. *Nature* **2001**, *409*, 860−921.
(2) Venter, J. C.; Adams, M. D.; Myers, E. W.; Li, P. W.; Mural, R. J.; Sutton, G. G.; Smith, H. O.; Yandell, M.; Evans, C. A.; Holt, R. A.; et al. *Science* **2001**, *291*, 1304−1351.
(3) Burley, S. K.; Almo, S. C.; Bonanno, J. B.; Capel, M.; Chance, M. R.; Gaasterland, T.; Lin, D.; Sali, A.; Studier, F. W.; Swaminathan, S. *Nat. Genet.* **1999**, *23*, 151−157.
(4) Wüthrich, K. *J. Biomol. NMR* **2003**, *27*, 13−39.
(5) Bax, A. *Curr. Opin. Struct. Biol.* **1994**, *4*, 738−744.
(6) Bax, A.; Grishaev, A. *Curr. Opin. Struct. Biol.* **2005**, *15*, 563−570.
(7) Tjandra, N.; Bax, A. *J. Magn. Reson.* **1997**, *124*, 512−515.
(8) Brutscher, B. *J. Magn. Reson.* **2001**, *151*, 332−338.
(9) Korzhnev, D. M.; Skrynnikov, N. R.; Millet, O.; Torchia, D. A.; Kay, L. E. *J. Am. Chem. Soc.* **2002**, *124*, 10743−10753.
(10) Lundström, P.; Hansen, D. F.; Vallurupalli, P.; Kay, L. E. *J. Am. Chem. Soc.* **2009**, *131*, 1915−1926.
(11) Baldwin, A. J.; Kay, L. E. *Nat. Chem. Biol.* **2009**, *5*, 808−814.
(12) Korzhnev, D. M.; Religa, T. L.; Banachewicz, W.; Fersht, A. R.; Kay, L. E. *Science* **2010**, *329*, 1312−1316.
(13) Ruschak, A.; Kay, L. E. *J. Biomol. NMR* **2010**, *46*, 75−87.
(14) Goto, N. K.; Kay, L. E. *Curr. Opin. Struct. Biol.* **2000**, *10*, 585−592.
(15) Kainosho, M.; Torizawa, T.; Iwashita, Y.; Mei, T. T. M. O.; Güntert, P. *Nature* **2006**, *440*, 52−57.
(16) Tugarinov, V.; Kanelis, V.; Kay, L. E. *Nat. Protoc.* **2006**, *1*, 749−754.
(17) Heinemann, U.; Illing, G.; Oschkinat, H. *Curr. Opin. Biotechnol.* **2001**, *12*, 348−354.
(18) Cavalli, A.; Salvatella, X.; Dobson, C. M.; Vendruscolo, M. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 9615−9620.
(19) Shen, Y.; Lange, O.; Delaglio, F.; Rossi, P.; Aramini, J. M.; Liu , G.; Eletsky, A.; Wu , Y.; Singarapu, K. K.; Lemak, A.; Ignatchenko, A.; Arrowsmith, C. H.; Szyperski, T.; Montelione, G. T.; Baker, D.; Bax, A. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 4685−4690.
(20) Montalvao, R.; Cavalli, A.; Salvatella, X.; Blundell, T. M.; Vendruscolo, M. *J. Am. Chem. Soc.* **2008**, *130*, 15990−15996.
(21) Robustelli, P.; Kohlhoff, K.; Cavalli, A.; Vendruscolo, M. *Structure* **2010**, *18*, 923−933.
(22) Clore, G. M.; Omichinski, J. G.; Sakaguchi, K.; Zambrano, N.; Sakamoto, H.; Appella, E. *Science* **1995**, *267*, 1515−1516.
(23) Lambert, L. J.; Schirf, V.; Demeler, B.; Cadene, M.; Werner, M. H. *EMBO J.* **2004**, *23*, 3186.
(24) Nabuurs, S. B.; Spronk, C. A. E. M.; Vuister, G. W.; Vriend, G. *PLoS Comput. Biol.* **2006**, *2*, e9.
(25) Sahakyan, A. B.; Vranken, W. F.; Cavalli, A.; Vendruscolo, M. *J. Biomol. NMR* **2011**, *50*, 331−346.
(26) Sahakyan, A. B.; Vranken, W. F.; Cavalli, A.; Vendruscolo, M. *Angew. Chem., Int. Ed.* **2011**, *50*, 9620−9623.
(27) Kohlhoff, K. J.; Robustelli, P.; Cavalli, A.; Salvatella, X.; Vendruscolo, M. *J. Am. Chem. Soc.* **2009**, *131*, 13894−13895.
(28) Nielsen, J. T.; Bjerring, M.; Jeppesen, M. D.; Pedersen, R. O.; Pedersen, J. M.; Hein, K. L.; Vosegaard, T.; Skrydstrup, T.; Otzen, D. E.; Nielsen, N. C. *Angew. Chem., Int. Ed.* **2009**, *48*, 2118−2121.
(29) Ágoston, B. S.; Kovács, D.; Tompa, P.; Perczel, A. *Biomol. NMR Assign.* **2011**, *5*, 189−193.
(30) Oldfield, E. *J. Biomol. NMR* **1995**, *5*, 217−225.
(31) Jameson, C. J. *Annu. Rev. Phys. Chem.* **1996**, *47*, 135−169.
(32) Xu, X. P.; Case, D. A. *J. Biomol. NMR* **2001**, *21*, 321−333.
(33) Neal, S.; Nip, A. M.; Zhang, H.; Wishart, D. S. *J. Biomol. NMR* **2003**, *26*, 215−240.
(34) Vila, J. A.; Arnautova, Y. A.; Martin, O. A.; Scheraga, H. A. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 16972−16977.
(35) Berjanskii, M.; Liang, Y.; Zhou, J.; Tang, P.; Stothard, P.; Zhou, Y.; Cruz, J.; MacDonell, C.; Lin, G.; Lu, P.; Wishart, D. S. *Nucleic Acids Res.* **2010**, *38*, W633−W640.
(36) Vijay-Kumar, S.; Bugg, C. E.; Cook, W. J. *J. Mol. Biol.* **1987**, *194*, 531−544.
(37) Weiergräber, O. H.; Senin, I. I.; Philippov, P. P.; Granzin, J.; Koch, K. W. *J. Biol. Chem.* **2003**, *278*, 22972−22979.
(38) Tanaka, T.; Ames, J. B.; Harvey, T. S.; Stryer, L.; Ikura, M. *Nature* **1995**, *376*, 444−447.
(39) Lange, O. F.; Lakomek, N. A.; Farès, C.; Schröder, G. F.; Walter, K. F. A.; Becker, S.; Meiler, J.; Grubmüller, H.; Griesinger, C.; de Groot, B. L. *Science* **2008**, *320*, 1471−1475.
(40) Anstrom, D. M.; Kallio, K.; Remington, S. J. *Protein Sci.* **2003**, *12*, 1822−1832.
(41) Howard, B. R.; Endrizzi, J. A.; Remington, S. J. *Biochemistry* **2000**, *39*, 3156−3168.
(42) Tugarinov, V.; Choy, W. Y.; Orekhov, V. Y.; Kay, L. E. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 622−627.
(43) Grishaev, A.; Tugarinov, V.; Kay, L. E.; Trewhella, J.; Bax, A. *J. Biomol. NMR* **2008**, *40*, 95−106.
(44) Fiser, A.; Do, R. K.; Sali, A. *Protein Sci.* **2000**, *9*, 1753−1773.
(45) Tugarinov, V.; Kay, L. E. *J. Am. Chem. Soc.* **2003**, *125*, 13868−13878.
(46) Sheppard, D.; Guo, C.; Tugarinov, V. *J. Am. Chem. Soc.* **2009**, *131*, 1364−1365.
(47) Tugarinov, V.; Kay, L. E. *Biochemistry* **2005**, *44*, 15970−15977.