THE JOURNAL OF PHYSICAL CHEMISTRY B

ARTICLE

pubs.acs.org/JPCB

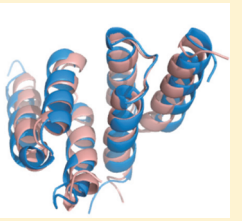
Using Chemical Shifts to Determine Structural Changes in Proteins upon Complex Formation

Andrea Cavalli,[§] Rinaldo W. Montalvao,[§] and Michele Vendruscolo*

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

ABSTRACT: Methods for determining protein structures using only chemical shift information are progressively becoming more accurate and reliable. A major problem, however, in the use of chemical shifts for the determination of the structures of protein complexes is that the changes in the chemical shifts upon binding tend to be rather limited and indeed often smaller than the standard errors made in the predictions of chemical shifts corresponding to given structures. We present a procedure that, despite this problem, enables one to use of chemical shifts to determine accurately the conformational changes that take place upon complex formation.

	Chain A		Chain B	
	NMR Free	NMR Complex	NMR Free	NMR Complex
Cheshire Free	1.2 Å	1.5 Å	1.3 Å	1.8 Å
Cheshire Complex	1.6 Å	1.0 Å	1.7 Å	1.1 Å



INTRODUCTION

It has been recently shown that chemical shifts provide sufficient information to enable the dtermination structures of small proteins in native[1−3] and transiently populated non-native states.[4] It has also been reported that chemical shifts can also be used for protein complex determination in combination with docking procedures.[5−7] These methods are particularly interesting in the light of recent advancements for NMR measurements in protein complexes using isotope labeling.[8−13]

One particularly challenging aspect of the problem of determining the structures of protein complexes using these approaches is that the changes in chemical shifts upon binding are usually small compared to the errors made by current prediction methods in calculating them.[14−17] In this work we show that, despite this problem, it is possible to exploit the information provided by chemical shifts to characterize accurately even small changes in the structures of proteins upon complex formation. This situation is similar to that encountered in the determination of protein structures using interproton distances derived from nuclear Overhauser effects,[18] where each individual distance is known with rather large uncertainty, but the simultaneous availability of many distances restricts very significantly the number of conformations compatible with the experimental information and enables one to determine a well-defined structure.

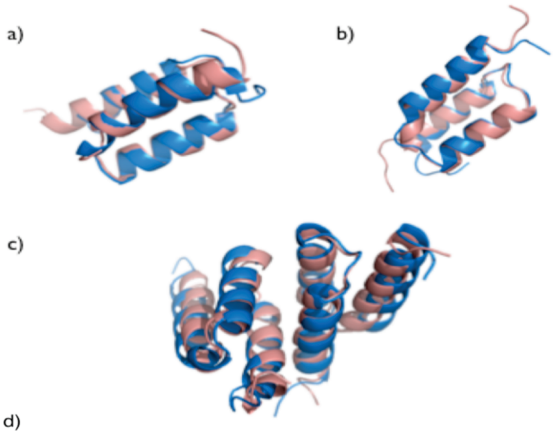
Quite generally, there are two major strategies to determine the structures of protein complexes using the information provided by chemical shifts. The first one exploits the knowledge of the structures of the free states of the proteins and uses the chemical shifts to drive a flexible docking process, allowing for conformational changes to take place when required.[5,6] The second, which is the one that we follow here, does not rely on any previous knowledge of the free structures of the component proteins, but determines directly their structures in the complex from the chemical shifts of the bound state;[7] these structures are then brought together using an essentially rigid-docking procedure. We illustrate here the second of these approaches by considering the case of an affibody−affibody complex.

METHODS

Cheshire. The structures of the individual proteins that make up a given complex are determined in this work from chemical shift information using the Cheshire method,[1] which consists of a three-phase computational procedure. In the first phase, the chemical shifts and the intrinsic secondary structure propensities of amino acid triplets are used to predict the secondary structure of the protein. In the second phase, the secondary structure predictions and the chemical shifts are used to predict backbone torsion angles for the protein. These angles are screened against a database to create a library of trial conformations of three and nine residue fragments spanning the sequence of the protein. In the third phase, a molecular fragment replacement strategy is used to bring together these fragments into low-resolution structural models; the information provided by chemical shifts is used in this phase to guide the assembly of the fragments. The resulting structures are refined with a hybrid molecular dynamics and Monte Carlo conformational search using a scoring function defined by the combination between the experimental and calculated chemical shifts, and the energy of a molecular mechanics energy.

CamDock. After the structures of the component proteins in their bound states have been determined using the Cheshire method,[1] the complex that they form is determined using the CamDock method.[6] The CamDock procedure consists of two phases, an *ab initio* generation of candidate structures for the complex by the Chord program,[6] and a structural refinement

Received: March 21, 2011
Revised: May 19, 2011
Published: June 03, 2011



d)

	Chain A		Chain B	
	NMR Free	NMR Complex	NMR Free	NMR Complex
Cheshire Free	1.2 Å	1.5 Å	1.3 Å	1.8 Å
Cheshire Complex	1.6 Å	1.0 Å	1.7 Å	1.1 Å

Figure 1. Comparison between the structures determined here form chemical shifts (blue) and those determined from standard methods (PDB 2B87, 2B88 and 2B89)[20,21] (pink): (a) ZTaq (b) Anti-ZTaq; (c) ZTaq:Anti-ZTaq complex. (d) rmsd for the structures determined here (Cheshire) and for the reference structures (NMR). ZTaq is referred to as chain A, and Anti-ZTaq as chain B.

through chemical shift restraints by the Cheshire program.[1] In the Chord method, candidate structures for protein complexes are calculated using a spherical harmonics description of the protein surface.[6] The use of spherical harmonics has several advantages over grid-based FFT docking correlation methods.[19] Most notably, rotations and translations can be carried out by operating on the initial expansion coefficients. This procedure not only results in a very quick search of the conformational space of the interacting partners, but also allows control of the degree of resolution of the shapes. The use of chemical shifts in CamDock is different from that in Haddock,[5] since the information provided by chemical shifts is not transformed into ambiguous interaction restraints. Instead, the chemical shifts corresponding to a given structure are calculated by considering a phenomenological approximation[14−17] of the secondary and tertiary interactions contributing to the chemical shifts, including distance terms, dihedral angles, ring current shifts, hydrogen bonding, and electric fields. The use of chemical shifts differs also from that in Rosetta,[7] since the chemical shifts are not only used in the preparation of the fragments but also as conformational restraints in the structural refinement procedure.

RESULTS

Affibodies constitute a class of engineered binding proteins based on the 58-residue three-helix bundle Z domain of staphylococcal protein A.[20] The structures of the two affibodies considered here (ZTaq and Anti-ZTaq) both in the free (PDB 2B88 and 2B89) and in the bound states (PDB 2B87) have been

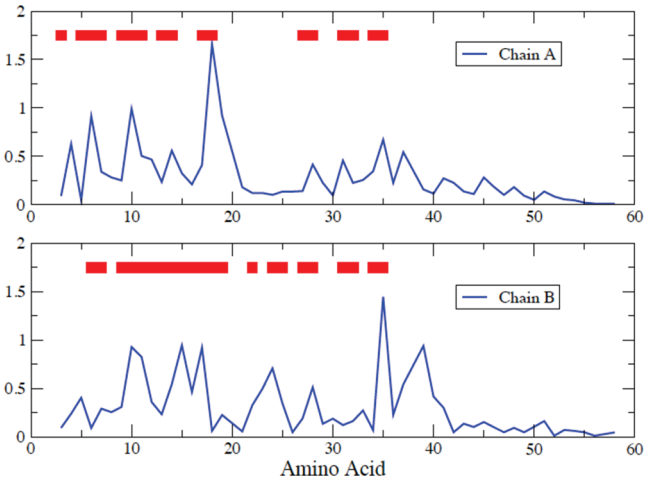


Figure 2. Backbone rmsd per residue between the free and the bound states for the structures determined here: (a) ZTaq (chain A); (b) Anti-ZTaq (chain B). The interface regions are indicated by the red bars.

previously determined by standard NMR methods,[20,21] and are used here as reference for validation purposes.

We used first the Cheshire[1] method, as described previously, to determine the structures of the free states from the chemical shifts of the free states (BMRB 6804 and 6805, Figure 1a,b). Then we used a combination of the Cheshire and CamDock[6] methods to determine the structure of the bound state from the chemical shifts of the bound states themselves (BMRB 6806, Figure 1c). This procedure involves a first step in which the Cheshire method is used to determine the structures of the individual proteins in conformations corresponding to their bound states, and then to use the CamDock method as described previously to build the complex using a rigid-body docking procedure, which is possible since the component proteins are already in their bound state conformations. Thus, in this approach, and at variance with the method that we presented originally,[6] the structures of the free states are generated for comparison, but not used as a starting point for a docking procedure. Indeed, in the procedure that we present here, the determination of the structure of the protein−protein complex is completely independent from that of the structures of the proteins in their free states. The structure of the complex determined from chemical shifts is at a 1.0 Å (backbone atoms) root-mean-square distances (rmsd) from the reference structure (Figure 1c); the corresponding RMSDs for the individual affibodies are, respectively 1.0 Å (ZTaq), and 1.2 Å (Anti-ZTaq). In the free states, the RMSDs between the structures determined from chemical shifts and the reference ones are of 1.2 Å (ZTaq) and 1.3 Å (Anti-ZTaq, Figure 1d).

We then investigated the changes in structure upon complex formation; these proteins exhibit rather small conformational changes upon complex formation. For ZTaq we found a rmsd change of 1.6 Å (NMR free ZTaq vs Cheshire complex ZTaq) or 1.5 Å (NMR complex ZTaq vs Cheshire free ZTaq). For Anti-ZTaq we found a rmsd change of 1.8 Å (NMR free Anti-ZTaq vs Cheshire complex Anti-ZTaq) or 1.7 Å (NMR complex Anti-ZTaq vs Cheshire free Anti-ZTaq). Remarkably, the differences between the free and bound structures are larger than the differences between the Cheshire and reference NMR structures (Figure 1d), thus indicating that the structure determination

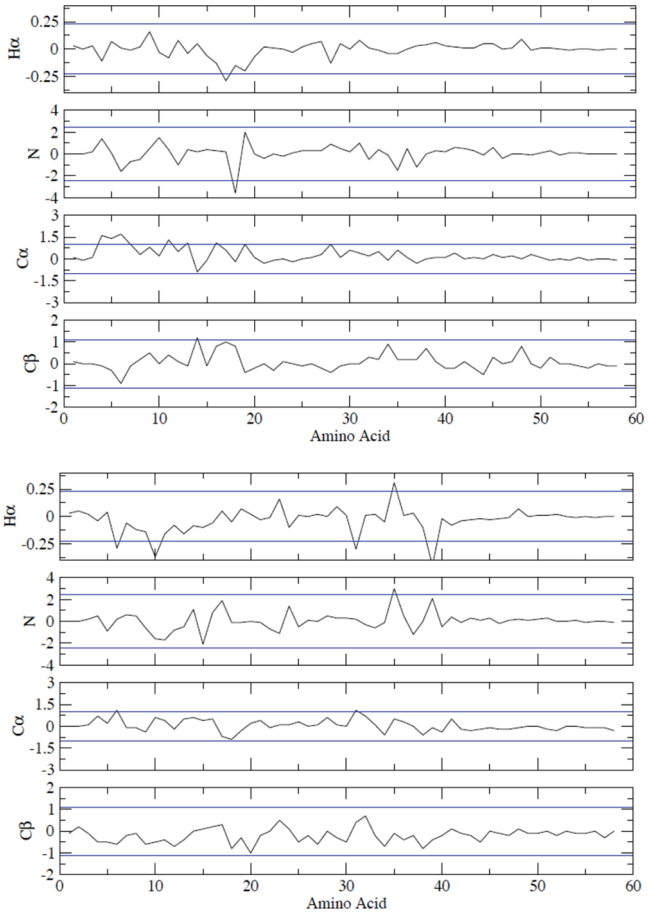


Figure 3. Comparison of the chemical shift differences for the free and the bound states (black lines): (a) ZTaq, (b) Anti-ZTaq. Most of the differences are smaller than the typical errors in the chemical shift calculations (blue horizontal lines); results are shown for method.[16] Despite this fact, the simultaneous consideration of a large number of chemical shifts enables one to characterize the conformational changes upon binding.

method that we present in this work is capable of identifying the small structural changes that take place upon binding. The specific differences between the structures are illustrated in Figure 2 as a rmsd per residue plot. Taken together these results indicate that the use of chemical shifts within the present procedure here enables one to characterize with high accuracy the conformational changes upon complex formation in the case of the ZTaq:Anti-ZTaq complex.

A question of central importance is to explain why it is possible to determine the small structural changes upon complex formation even if the differences in the experimentally measured chemical shifts for the free and the bound states are smaller than the typical error made in the predictions of the individual chemical shifts (Figure 3). Indeed, the differences between the experimental chemical shifts in the free and bound states (Figure 3, black line) are in most cases smaller that the typical errors in the calculations of the chemical shifts (Figure 3, blue lines), which are on average of 0.3 ppm for Hα atoms, 1.1 ppm for Cα atoms, 1.3 ppm for Cβ, and 2.7 ppm for N atoms.[14,16,17] Despite these differences, the results that we have presented demonstrate that the information provided by chemical shifts is sufficient to determine the changes in structure upon

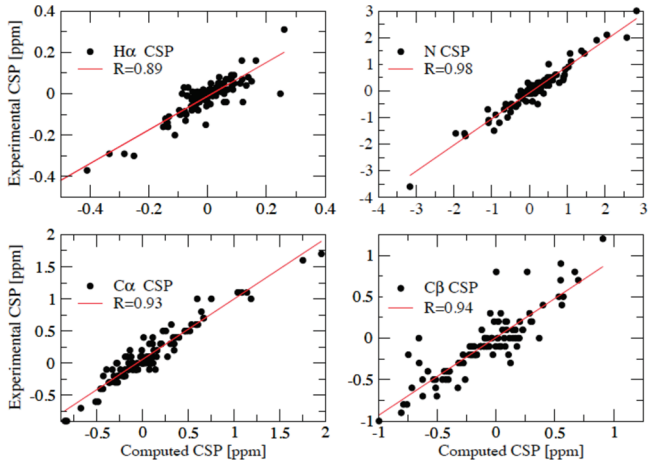


Figure 4. Comparison of the errors associated with the calculated chemical shift perturbations (CSPs) with those measured experimentally.

complex formation. We suggest that this result is possible because we have considered simultaneously a large number of chemical shifts, which provides a way to overcome the problem that the error in the prediction of individual chemical shifts is actually larger than the difference in the experimental chemical shifts of the free and bound states. As noted above, this situation is analogous to that encountered in NOE-based structure determination methods, in which the combined use of a large number of fairly loose structural reporters enables one to obtain a rather well-defined overall structure.[18] We also investigated whether, due to the cancellation of systematic errors, the errors associated with the calculated chemical shift perturbations (CSPs) are smaller than those of the calculated individual chemical shifts. By comparing the calculated CSPs based on the structure of the complex that we determined with the corresponding experimental values we found a good agreement (Figure 4), which also provides a validation for the structure itself.

CONCLUSIONS

In this study we have investigated the use of chemical to carry out the determination of the structures of protein−protein complexes in the absence of any structural information about the free states of the proteins. We used the CamDock procedure, which depending on the available chemical shift measurements can be implemented in two alternative ways. In the first, which we described previously,[6] when the structures of the free states are known, for example through X-ray crystallography or NMR spectroscopy methods, together with the chemical shifts of the complex, one can use the chemical shifts to drive a flexible docking procedure to determine the structure of the complex.[6] In the second, which we illustrated in this work, when only the chemical shifts of the complex are known, one can build the structures of the bound states of the proteins using the Cheshire[1] procedure and then use the CamDock[6] procedure to perform an essentially rigid docking. We expect these procedures to be applicable quite generally, especially considering that the structures of individual proteins can be probed within complexes using selective labeling techniques,[22,23] and to be particularly useful in cases in which there are large conformational changes upon complex formation.

dx.doi.org/10.1021/jp202647q |J. Phys. Chem. B 2011, 115, 9491–9494

We have shown that small structural changes upon complex formation can be accurately described from the corresponding changes in the chemical shifts, even when the latter are smaller than the typical errors in the predictions of their values. These results, together with recent related ones,[5−7] thus indicate that the measurement of chemical shifts in protein complexes represents a viable strategy for determining their structures.

AUTHOR INFORMATION

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

Author Contributions
§These authors made equal contributions.

ACKNOWLEDGMENT

We acknowledge support by the Leverhulme Trust (A.C., R. W.M., M.V.), BBSRC (A.C., R.W.M., M.V.), EMBO (M.V.), and the Royal Society (M.V.).

REFERENCES

(1) Cavalli, A.; Salvatella, X.; Dobson, C. M.; Vendruscolo, M. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 9615–9620.
(2) Shen, Y.; Lange, O.; Delaglio, F.; Rossi, P.; Aramini, J. M.; Liu, G. H.; Eletsky, A.; Wu, Y. B.; 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.
(3) Berjanskii, M.; Tang, P.; Liang, J.; Cruz, J. A.; Zhou, J. J.; Zhou, Y.; Bassett, E.; MacDonell, C.; Lu, P.; Lin, G. H.; Wishart, D. S. *Nucleic Acids Res.* **2009**, *37*, W670–W677.
(4) Korzhnev, D. M.; Religa, T. L.; Banachewicz, W.; Fersht, A. R.; Kay, L. E. *Science* **2010**, *329*, 1312–1316.
(5) Dominguez, C.; Boelens, R.; Bonvin, A. *J. Am. Chem. Soc.* **2003**, *125*, 1731–1737.
(6) Montalvao, R. W.; Cavalli, A.; Salvatella, X.; Blundell, T. L.; Vendruscolo, M. *J. Am. Chem. Soc.* **2008**, *130*, 15990–15996.
(7) Das, R.; Andre, I.; Shen, Y.; Wu, Y. B.; Lemak, A.; Bansal, S.; Arrowsmith, C. H.; Szyperski, T.; Baker, D. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 18978–18983.
(8) Gelis, I.; Bonvin, A.; Keramisanou, D.; Koukaki, M.; Gouridis, G.; Karamanou, S.; Economou, A.; Kalodimos, C. G. *Cell* **2007**, *131*, 756–769.
(9) Tonelli, M.; Masterson, L. R.; Cornilescu, G.; Markley, J. L.; Veglia, G. *J. Am. Chem. Soc.* **2009**, *131*, 14138–14139.
(10) Traaseth, N. J.; Verardi, R.; Veglia, G. *J. Am. Chem. Soc.* **2008**, *130*, 2400–2401.
(11) Sprangers, R.; Velyvis, A.; Kay, L. E. *Nature Methods* **2007**, *4*, 697–703.
(12) Kalodimos, C. G. *Protein Sci.* **2011**, *20*, 773–782.
(13) Sheppard, D.; Sprangers, R.; Tugarinov, V. *Prog. Nucl. Mag. Res. Spec.* **2010**, *56*, 1–45.
(14) Kohlhoff, K. J.; Robustelli, P.; Cavalli, A.; Salvatella, X.; Vendruscolo, M. *J. Am. Chem. Soc.* **2009**, *131*, 13894–13895.
(15) Lehtivarjo, J.; Hassinen, T.; Korhonen, S. P.; Perakyla, M.; Laatikainen, R. *J. Biomol. NMR* **2009**, *45*, 413–426.
(16) Neal, S.; Nip, A. M.; Zhang, H. Y.; Wishart, D. S. *J. Biomol. NMR* **2003**, *26*, 215–240.
(17) Shen, Y.; Bax, A. *J. Biomol. NMR* **2007**, *38*, 289–302.
(18) Wuthrich, K. *Science* **1989**, *243*, 45–50.
(19) Ritchie, D. W.; Kemp, G. J. L. *Proteins* **2000**, *39*, 178–194.
(20) Wahlberg, E.; Lendel, C.; Helgstrand, M.; Allard, P.; Dincbas-Renqvist, V.; Hedqvist, A.; Berglund, H.; Nygren, P. A.; Hard, T. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 3185–3190.
(21) Lendel, C.; Dogan, J.; Hard, T. *J. Mol. Biol.* **2006**, *359*, 1293–1304.
(22) Kainosho, M.; Torizawa, T.; Iwashita, Y.; Terauchi, T.; Ono, A. M.; Guntert, P. *Nature* **2006**, *440*, 52–57.
(23) Lundstrom, P.; Vallurupalli, P.; Hansen, D. F.; Kay, L. E. *Nature Prot.* **2009**, *4*, 1641–1648.