DOI: 10.1002/ange.200501209

Mapping Protein–Protein Interfaces on the Basis of Proton Density Difference**

*Xiaogang Sui, Yingqi Xu, Janel L. Giovannelli, Nancy T. Ho, Chien Ho, and Daiwen Yang**

A protein normally functions through interacting with its partners. In order to understand the molecular basis for biochemical and biological processes, it is important to identify the contact surfaces of interacting molecules or even to characterize the structures of protein complexes. Since it is much easier and faster to define interacting surfaces by NMR spectroscopy than to determine the full structures, a number of methods have been developed to map protein interactions.[1] The traditional approach based on chemical-shift perturbation is easy to do,[2] but its precision is low. In cases where the entire protein changes conformation, this approach fails as a mapping device but can indicate the presence of allosteric processes.[3] It cannot be applied to map the intersubunit interface of oligomeric proteins in which the monomeric form is unstable. There exist methods based on the paramagnetic line-broadening effect with the use of site-specific spin labeling[4] or paramagnetic ions,[5] but they are applicable only to systems in which spin labeling is available or conformation change upon forming a complex does not increase relaxation protection from paramagnetic ions.

Amide-proton exchange has also been used to map protein interfaces,[6] but it is not a reliable tool because the change of exchange rates depends on local and global structural changes upon the formation of the complexes. Intermolecular NOESY analysis can be used to map the interactions, but it is difficult to completely suppress intra-moleuclar NOEs.[7] A more sophisticated method based on saturation transfer from an acceptor to a reporting protein, called cross-saturation,[8] has been proposed. It provides more reliable information about the binding interface, but highly

[*] X. G. Sui, Dr. Y. Q. Xu, Prof. D. W. Yang
Department of Biological Sciences
Department of Chemistry
National University of Singapore
14 Science Drive 4, Singapore 117543 (Singapore)
Fax: (+65) 6779-2486
E-mail: dbsydw@nus.edu.sg
J. L. Giovannelli, N. T. Ho, Prof. C. Ho
Department of Biological Sciences
Carnegie Mellon University
Pittsburgh, PA 15213 (USA)

[**] This research was supported by a Young Investigator Award from the Biomedical Research Council (BMRC) and the Agency for Science, Technology and Research A*Star of Singapore to D.W.Y. and by a grant from the National Institutes of Health (R01HL-24525) to C.H.

Supporting information for this article is available on the WWW under http://www.angewandte.org or from the author.

deuterated ^{15}N-labeled protein is required for the suppression of spin diffusion in the reporting protein. Incomplete deuteration of the reporting protein may result in the identification of false interaction sites. In addition, to reduce spin diffusion among amide protons, especially for α-helical proteins, it is necessary to do experiments in a solvent mixture of $H_2O/^2H_2O$ with $<30\%$ H_2O. This is equivalent to lowering the effective sample concentration by a factor of about three or more. Thus, this technique may not work for proteins with moderate solubility or that are difficult to obtain fully deuterated.

Herein, we propose a novel strategy to map amino acid residues involved in protein–protein interactions at interfaces on the basis of the dependence of NMR relaxation on proton density, by using two moderately deuterated samples in which the reporting protein in a protein–protein complex is ^{2}H,^{15}N-labeled while the acceptor protein is either unlabeled or ^{2}H-labeled. Figure 1 shows the pulse scheme used to measure

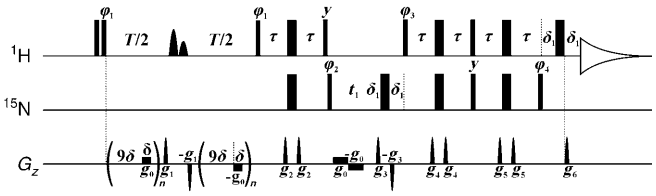

Figure 1. Pulse scheme used; all narrow (wide) bars represent 90° (180°) rectangular pulses. The first shaped 180° pulse (1 ms, REBURP, centered at 9.1 ppm) selectively inverts the signals of amide protons, the second shaped 180° pulse (2 ms, rectangle) that of H_2O (G_z represents pulsed field z gradients). The delays used are: $\tau = 2.3$ ms, $\delta = 0.5$ ms, and $\delta_1 = 1.1$ ms. Radiation damping during the relaxation delay T is suppressed by gradient g_0. The phase cycling employed is: $\varphi_1 = x,x,-x,-x$; $\varphi_2 = x,-x$; $\varphi_3 = y$; $\varphi_4 = x$; rec $= x,-x$. The durations and strengths of the sine-shaped gradients are: $g_1 = 2$ ms, 20 G cm^{-1}; $g_2 = 1$ ms, 17.5 G cm^{-1}; $g_3 = 1$ ms, 40 G cm^{-1}; $g_4 = 1$ ms, 22.5 G cm^{-1}; $g_5 = 1$ ms, 35 G cm^{-1}; and $g_6 = 1$ ms, 8 G cm^{-1}. The strength of g_0 is 1.5 G cm^{-1}. Quadrature detection in F_1 uses the enhanced sensitivity pulse gradient method, where for each t_1 separate data sets are recorded for (g_6, φ_3, φ_4) and ($-g_6$, φ_3+180°, φ_4+180.

selective longitudinal relaxation rates of amide protons in a deuterated reporting protein A, which binds to its acceptor protein B. The effect of phase alternation of φ_1 is to alternately store magnetization along the $-Z$ axis and the $+Z$ axis at the very beginning of the relaxation period so that the longitudinal magnetization relaxes as $\exp(-RT)$.[9] In this way, the measurement of relaxation rates will be independent of the interscan delay. Suppression of cross-relaxation between amide and aliphatic protons is achieved by selective inversion of only amide and aromatic protons, while reduction of amide magnetization loss due to exchange with H_2O is achieved by maintaining the proton magnetization of the amide and H_2O along the same direction during the entire relaxation period. A single scan TROSY scheme[10] is used to obtain TROSY-HSQC data.

The initial relaxation rate of an amide proton ($R^1(H_Z^N)$) under selective inversion is dominated by dipolar interactions with its surrounding aliphatic protons and it can be approximated as shown in Equation (1), where τ_c is the molecular

$$R^1(H_Z^N) \approx 0.1\,\tau_c \left(\frac{\mu_0 \hbar \gamma_H^2}{4\pi}\right)^2 \left\{ \sum_{j=1}^{m} \frac{S_j^2}{r_{HHAj}^6} + \sum_{k=1}^{n} \frac{S_k^2}{r_{HHBk}^6} \right\} + R_{ex} \qquad (1)$$

tumbling time; μ_0 is the permeability in vacuum; $\hbar = h/2\pi$ and h is Planck's constant; γ_H is the proton gyromagnetic ratio; r_{HHAj} is the distance between the amide proton and the jth residual aliphatic proton in protein A; r_{HHBk} is the distance between the amide proton and the kth residual aliphatic proton in protein B; S_j^2 and S_k^2 are the order parameters describing the motional amplitude of an H–H$_j$ and an H–H$_k$ vector, respectively; m and n are the total numbers of the respective aliphatic protons in proteins A and B interacting with the amide proton; and R_{ex} is the relaxation contribution from the unsuppressed NH–H_2O exchange effect, which is minimized by preventing radiation damping and keeping H_Z^N and H_2O magnetization parallel.

If protein A is perdeuterated (that is, $m = 0$), amide protons located at the binding interface have larger R^1 values than other amide protons since they are much closer to the protons in protein B. In principle, amino acid residues in the interacting region can be identified based on R^1 values by using a single sample consisting of a perdeuterated reporting protein and an unlabeled acceptor. However, incomplete deuteration in aliphatic protons and the unsuppressed NH–H_2O exchange effect (R_{ex}) can lead to the incorrect identification of interacting sites. In order to overcome this drawback, we propose herein the use of two samples in $H_2O/^2H_2O$ (95:5) solution: 1) ^{15}N,^{2}H-labeled protein A complexed with unlabeled protein B and 2) ^{15}N,^{2}H-labeled protein A complexed with ^{2}H-labeled protein B. The difference of R^1 values for a given amide proton in the two samples [Eq. (2), in which n_1 and n_2 are the numbers of aliphatic

$$\Delta R^1 \propto \sum_{k=1}^{n_1} S_k^2 r_{HHBk}^{-6} - \sum_{j=1}^{n_2} S_j^2 r_{HHBj}^{-6} \qquad (2)$$

protons in protein B in samples 1 and 2, respectively] is mainly determined by the location of the amide (that is, r_{HHBk}) but is independent of the deuteration percentage of protein A (m) and the R_{ex} value. Although incomplete deuteration of protein B in the second sample affects the amplitude of ΔR^1, it cannot result in false binding sites. With the assumption that $S^2 = 1$ and of perdeuteration of protein B in sample 2 (that is, $n_2 = 0$), the ΔR^1 value can be used to estimate an upper limit of the effective distance between an amide proton in protein A and its proximal protons in protein B (r_{eff}) [Eq. (3)]. Even in the case of 85% deuteration of protein B

$$r_{eff} = [\sum r_{HHBk}^{-6}]^{-1/6} \qquad (3)$$

and $S^2 = 0.3$, the actual r_{eff} is $\approx 25\%$ shorter than its upper limit, a fact implying that effects of the local dynamics and deuteration rate of protein B can be neglected (see Figure S1 in the Supporting Information). On the basis of r_{eff} values, one can model the structure of a protein complex, provided that individual structures are known.

We have applied the above-mentioned strategy to map the intersubunit interface of recombinant human adult hemoglobin in the carbonmonoxide form (rHbCO A, 1 mM, pH 7.0)

on an 800 MHz NMR spectrometer. Here, proteins A and B correspond to the α and β chains, respectively, and the α chain is $^{15}N,^{2}H$-labeled. The average deuteration percentage was ≈85% for both chains. As expected, the decay of H_Z^N often follows a multiexponential form due to strong cross-relaxation among amide protons.[11] Nevertheless, it is possible to obtain initial relaxation rates for both samples and then to calculate ΔR^1 values. Since we are interested in only the ΔR^1 value, an alternative way to obtain it is from the dependence of the intensity ratio of a given proton in two samples ($I_1(t)$/$I_2(t)$) on the relaxation delay (which is shown by numerical simulations; see Figure S1 in the Supporting Information), for which $I_1(t)$ ($I_2(t)$) is the relative intensity at time t with respect to that at the first delay for sample 1 with unlabeled β chains (sample 2 with ^{2}H-labeled β chains).

Figure 2 shows the decays of a number of representative amide protons. Amides not involved in subunit contacts have shown no dependence of I_1/I_2 on the relaxation delay. Amides involved in weak contacts display nearly monoexponential profiles during a period of 500 ms, while those located in close-contact regions show nearly monoexponential profiles within the first 150 ms. Amide protons that involve weak or no direct subunit contact but that are close to the amides with strong subunit interactions display slower decay or no decay in the initial period and then faster decay in the later stage (F36 in Figure 2). This is due to the two-step "spin-diffusion" process. Fortunately, the spin-diffusion effect decreases dramatically with the steps and it has been found to be negligible for multistep processes for rHbCO A within a relaxation delay of up to 400 ms.

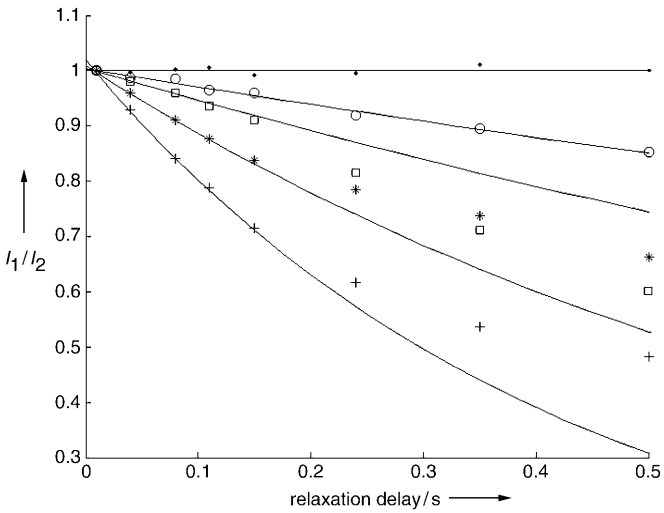

Figure 2. The dependence of the intensity ratio of normalized peak intensities in two samples (I_1/I_2) on the relaxation delay for a number of representative residues of rHbCO A (K60, ◆; V96, ○; F36, □; R31, *; A123, +). The normalized intensity for a given amide in one sample ($I(t)$) was obtained by taking the ratio of the peak intensities at delay t and the first delay (10 ms). Experimental data and initial decay profiles are indicated by symbols and solid lines, respectively. The total experimental time was ≈22 h.

Amino acid residues involved in subunit contacts can be simply identified from the ratios of normalized peak intensities (I_1/I_2) at one desired delay (for example, 250 ms; Figure 3a). Alternatively, they can be identified more confidently from ΔR^1 values (Figure 3b). The difference in the initial rate (ΔR^1) of an amide proton was estimated by fitting I_1/I_2 to a monoexponential equation in three steps: 1) calculation of the relaxation rate with all data points; 2) removal of the point with the longest delay in the existing data; and 3) repetition of step 2 until the absolute difference of the relaxation rates obtained with m+1 and m data points is smaller than the sum of the fitting errors of these two fittings, where m is the number of data points and $m > 3$. The fitting error was determined by using the Jackknife method.[12] The relaxation rate and fitting error obtained with m+1 points were reported. Amide protons in the α chain with a ΔR^1 value smaller than 0.7 times the average ΔR^1 value (≈0.1 s^{-1}, the noise level) are considered to have no interaction with the β chain. The results are consistent with the contact sites derived from the crystal structure of the R2 state of human normal adult hemoglobin (PDB entry 1BBB). With the use of a cross-saturation experiment on sample 1 in $H_2O/^2H_2O$ (30:70) solution, many residues in the contact region were also identified.

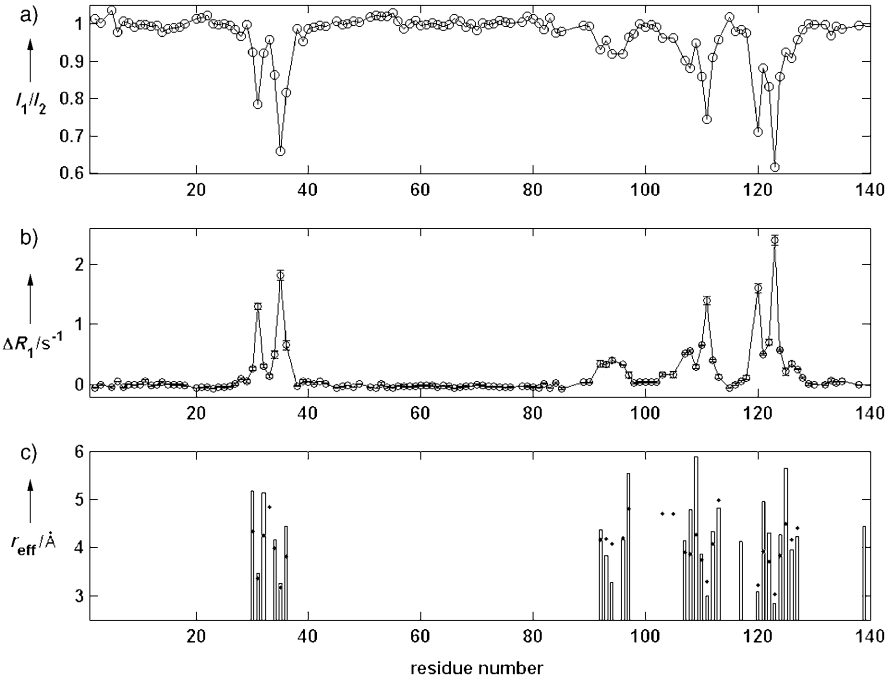

Figure 3. a) Peak intensity ratios of amides in two samples (I_1/I_2) against the residue number of the α chain of human normal adult hemoglobin; desired delay: 250 ms; first delay: 10 ms. b) Difference of initial relaxation rates (ΔR^1) between samples 1 and 2. Errors are indicated with vertical bars. c) Comparison of effective distances derived from the ΔR^1 values (◆) and from the crystal structure of the R2 state of rHbCO A (bar). The r_{eff} value is set to zero when it is larger than 5 and 6 Å for the NMR-derived data and the crystal-structure-based data, respectively.

However, the $\alpha_1\beta_2$ interface (residues 92–97) cannot be identified because of the strong noise level caused by spin diffusion among many residual protons in the α chain (see Figure S2 in the Supporting Information).

Figure 3c shows the upper limits of effective distances (r_{eff}) estimated from the ΔR^1 values and a τ_c value of 32 ns (τ_c was established from ^{15}N relaxation times). The distances agree well with those calculated from the R2 structure except for a few residues (see Table S1 in the Supporting Information). The r_{eff} value for α94D estimated from the ΔR^1 value (4.07 Å) is significantly larger than that from the R2 crystal structure (3.28 Å) because the $\alpha_1\beta_2$ interface is more dynamic in solution than in the crystal state. The dynamics of this region has been shown previously by a relaxation study of β37W which is in contact with α94D.[13] The discrepancy for α117F (in a loop) and α139K (in the C terminus) is also due to different mobility in the solution and crystal states. The effective distances for α33F, α103H, and α105L are significantly shorter from the present method than from the crystal structure. This can be explained by a slight difference in subunit arrangement and side-chain packing in the solution and crystal states. The slight difference of the $\alpha_1\beta_1$ interface in solution and in the crystal state has previously been evidenced by the different solvent exchange rates of side-chain protons of α103H and α122H in the T and R states.[14]

Although the experiment proposed here requires deuteration of both the reporting protein and its binding partner, only moderate deuteration that is easily achievable is required. Residues located in the protein–protein interface can be mapped out simply from the ratios of normalized amide peak intensities at one desired delay in two samples. The effect of protein-concentration difference in the two samples is removed by normalizing the peak intensities of each sample at the desired delay with respect to those at a very short delay (for example, 10 ms). In order to obtain upper limits of effective distances for the amides at contact sites, one can measure the initial relaxation rate differences (ΔR^1) in the two samples. The distance information together with other experimental data such as residual dipolar couplings can be used to model protein complex structures with docking techniques.[2,15]

Received: April 6, 2005
Revised: June 2, 2005
Published online: July 20, 2005

Keywords: heme proteins · interfaces · NMR spectroscopy · protein–protein interactions · structure elucidation

[1] E. R. P. Zuiderweg, *Biochemistry* **2002**, *41*, 1.
[2] G. M. Clore, C. D. Schwieters, *J. Am. Chem. Soc.* **2003**, *125*, 2902.
[3] Y. Chen, J. Reizer, M. H. Saier, W. J. Fairbrother, P. E. Wright, *Biochemistry* **1993**, *32*, 32.
[4] T. Scherf, R. Hiller, J. Anglister, *FASEB J.* **1995**, *9*, 120.
[5] S. Arumugam, C. L. Hemme, N. Yoshida, K. Suzuki, H. Nagase, M. Berjanskii, B. Wu, S. R. Van Doren, *Biochemistry* **1998**, *37*, 9650.
[6] G. A. Mueller, A. M. Smith, M. D. Chapman, G. S. Rule, D. C. Benjamin, *J. Biol. Chem.* **2001**, *276*, 9359.
[7] a) C. Zwahlen, P. Legault, S. J. F. Vincent, J. Greenblatt, R. Konrat, L. E. Kay, *J. Am. Chem. Soc.* **1997**, *119*, 6711; b) C. Eichmüller, W. Schüler, R. Konrat, B. Kräutler, *J. Biomol. NMR* **2001**, *21*, 107.
[8] H. Takahashi, T. Nakanishi, K. Kami, Y. Arata, I. Shimada, *Nat. Struct. Biol.* **2000**, *7*, 220.
[9] L. K. Nicholson, L. E. Kay, D. M. Baldisseri, J. Arango, P. E. Young, A. Bax, D. A. Torchia, *Biochemistry* **1992**, *31*, 5253.
[10] J. Weigelt, *J. Am. Chem. Soc.* **1998**, *120*, 10778.
[11] R. L. Vold, R. R. Vold, *Prog. Nucl. Magn. Reson. Spectrosc.* **1978**, *12*, 79.
[12] A. F. Bissell, R. A. Ferguson, J. Roy. , *Stat. Soc. D-Sta.* **1975**, *24*, 79.
[13] Y. Yuan, V. Simplaceanu, J. A. Lukin, C. Ho, *J. Mol. Biol.* **2002**, *321*, 863.
[14] M. R. Mihailescu, I. M. Russu, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 3773.
[15] C. Dominguez, R. Boelens, A. M. J. J. Bonvin, *J. Am. Chem. Soc.* **2003**, *125*, 1731.

