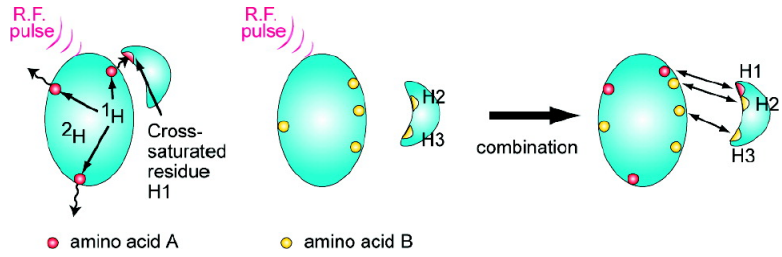
Article

Amino Acid Selective Cross-Saturation Method for Identification of Proximal Residue Pairs in a Protein#Protein Complex

Shunsuke Igarashi, Masanori Osawa, Koh Takeuchi, Shin-ichiro Ozawa, and Ichio Shimada

J. Am. Chem. Soc., **2008**, 130 (36), 12168-12176 • DOI: 10.1021/ja804062t • Publication Date (Web): 16 August 2008



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

J|A|C|S
ARTICLES

Published on Web 08/16/2008

Amino Acid Selective Cross-Saturation Method for Identification of Proximal Residue Pairs in a Protein−Protein Complex

Shunsuke Igarashi,[†] Masanori Osawa,[†] Koh Takeuchi,[†] Shin-ichiro Ozawa,[†] and Ichio Shimada*[,†,‡]

Graduate School of Pharmaceutical Sciences, The University of Tokyo, Hongo, Bunkyo-ku, Tokyo 113-0033, Japan and Biomedicinal Information Research Center (BIRC), National Institute of Advanced Industrial Science and Technology (AIST), Aomi Koto-ku, Tokyo 135-0064, Japan

Received May 30, 2008; E-mail: shimada@iw-nmr.f.u-tokyo.ac.jp

Abstract: We describe an NMR-based approach, the amino acid selective cross-saturation (ASCS) method, to identify the pairs of the interface residues of protein−protein complexes. ASCS uses a "cross-saturation (CS)-donor" protein, in which only one amino acid is selectively ^{1}H-labeled in a ^{2}H-background, and a "CS-acceptor" protein with uniform ^{2}H,^{15}N labeling. Irradiation of the ^{1}H-labeled amino acid, which exists only in the donor, decreases the intensity of the ^{1}H−^{15}N HSQC signals of the acceptor residues proximal to the ^{1}H-labeled CS-source residue(s) through the CS phenomenon. Given the three-dimensional structure of each protein in the complex, but not the complex structure, the combinatorial analysis of multiple ASCS results specify the CS-source residue(s), based on the spatial complementarity between the CS-source residues on the CS donor and the cross-saturated amide protons on the acceptor. NMR investigations of the labeling selectivity and efficiency in an *E. coli* host, which are critical for ASCS, revealed that Ala, Arg, His, Ile, Leu, Lys, Met, Phe, Pro, Trp, and Tyr are selectively labeled with a high ^{1}H/^{2}H ratio. The observation of the ASCS was then confirmed using the known structure of the yeast ubiquitin (Ub) and yeast ubiquitin hydrolase 1 (YUH1). Conversely, reasonable candidates for the CS-source residues were suggested by the analysis of the ASCS results, with reference to the individual structures of YUH1 and Ub. The pairwise distance information between the CS-source residues and the cross-saturated amide groups obtained by ASCS will be useful for modeling protein−protein complexes.

Introduction

Advances in structural biology techniques have greatly accelerated the structural determinations of proteins and their complexes, and thus increasing numbers of three-dimensional structures are now available in the Protein Data Bank.[1] However, an individual structure can provide only limited information about its functions, since the structure itself yields few clues about the specific target recognition. Therefore, structural analyses of protein−protein interactions are required to reveal the mechanism of their functions. We previously developed an NMR technique, the cross-saturation (CS) method, which identifies the interface residues of a protein−protein complex based on the internuclear dipolar−dipolar interactions.[2,3] The CS method uses differential isotope labeling of the two interacting proteins. One protein, called the CS donor, is unlabeled, and the other, the CS acceptor, is uniformly labeled with ^{2}H and ^{15}N. Irradiation of the aliphatic protons with RF pulses from approximately −2 to 4 ppm directly saturates the aliphatic resonances of the donor molecule, and the saturation immediately diffuses to the entire donor molecule. The saturation also transfers to the amide protons of the acceptor at the interface of the complex. This intermolecular saturation transfer, i.e., the CS, is detected as a signal intensity reduction in the ^{1}H−^{15}N HSQC spectrum of the acceptor residues at the interface of the complex.

In principle, however, the CS method provides no information about the binding residues of the donor, and thus on the relative orientation of the two proteins in the CS method, because the donor protons are subjected to saturation by the irradiation and subsequent spin diffusion to the whole donor molecule. In order to obtain pairwise information about the interacting residues between the donor and acceptor molecules, we have developed an amino acid selective cross-saturation (ASCS) method, in which the CS-donor protein is amino acid selectively ^{1}H-labeled in a ^{2}H-background.

This amino acid selective ^{1}H-labeling in a ^{2}H-background applied to ASCS requires highly selective and efficient ^{1}H-labeling. Although the metabolism of the carbon and nitrogen atoms of amino acids in *E. coli* has been extensively studied,[4,5] hydrogen atom metabolism has not been clarified comprehen-

[†] The University of Tokyo.
[‡] National Institute of Advanced Industrial Science and Technology (AIST).

(1) Berman, H.; Henrick, K.; Nakamura, H. *Nat. Struct. Biol.* **2003**, *10*, 980.
(2) Shimada, I. *Methods Enzymol.* **2005**, *394*, 483–506.
(3) Takahashi, H.; Nakanishi, T.; Kami, K.; Arata, Y.; Shimada, I. *Nat. Struct. Biol.* **2000**, *7*, 220–3.
(4) Waugh, D. S. *J. Biomol. NMR* **1996**, *8*, 184–92.
(5) Lian, L. Y.; Middleton, D. A. *Prog. Nucl. Magn. Reson. Spectrosc.* **2001**, *39*, 171–190.

10.1021/ja804062t CCC: $40.75

sively.[6−12] Therefore, we first carried out the NMR analysis of the ^{1}H-labeling selectivity and efficiency for a number of amino acids, using thioredoxin (Trx) as a test case.

To validate the ASCS method, we chose the complex of yeast ubiquitin (Ub) and the C90S mutant of yeast ubiquitin hydrolase 1 (YUH1).[13] ASCS experiments for several amino acids exhibited the selective CS from the labeled amino acid on YUH1 to the proximal Ub amide protons.

Conversely, we have tried to identify the CS-source "residues" to satisfy all of the ASCS results, using the distribution of the labeled amino acid on the YUH1 surface and the structural topology of the cross-saturated amide protons on Ub. Manual or systematic search for the spatial complementarity yielded the reasonable combinations of the CS-source residues for each cross-saturated amide proton. The pairwise distance information between the CS sources and the corresponding cross-saturated amide protons obtained by the ASCS method would provide a good initial model for protein−protein docking.

Materials and Methods

Protein Expression and Purification. The DNA fragment encoding Trx was inserted into the pACYC184 plasmid. Trx was expressed in the *E. coli* C41 strain and was purified by chromatography on DEAE Sephadex A-25 (GE Healthcare) and MonoQ 5/50 GL (GE Healthcare) columns.

The plasmids encoding Ub and wild type YUH1 were provided by Dr. Toshiyuki Kohno. The DNA encoding Ub was inserted into the pET26b plasmid, which was used to transform to the *E. coli* BL21 strain. The expression of the uniformly ^{2}H,^{15}N-labeled Ub was induced by isopropyl-β-D-(−)-thiogalactopyranoside (IPTG) in M9 minimal medium, containing 1.0 g of ^{15}N-labeled ammonium chloride, 2.0 g of D-glucose-d_6, and 1.0 g of ^{2}H,^{15}N-Celtone Base powder (Spectra Stable Isotopes) in 1.0 L of ^2H_2O (ISOTEC). Protein purification was performed, according to the previous report.[14]

The C90S mutation of YUH1 was introduced by the QuikChange (Stratagene) method. The protein was expressed in the *E. coli* C41 strain and was purified as previously described, with several modifications.[14] Briefly, the protein was purified to homogeneity by chromatography on DEAE Sepharose Fast Flow (GE Healthcare) and MonoQ 5/50 GL anion exchange columns, followed by fractionation on a HiLoad 26/60 Superdex 75 pg size exclusion column (GE Healthcare).

Amino Acid Selective ^{1}H-Labeling. Amino acid selective ^{1}H-labeling in a ^{2}H-background was carried out for Trx and the C90S mutant of YUH1. Cells were cultured at 37 °C in M9 minimal medium, containing 1.0 g of ammonium chloride, 2.0 g of D-glucose-d_6, and 0.5 g of ^{2}H-Celtone Base powder in 1.0 L of ^2H_2O. For Trx expression, uniform ^{15}N-labeling was also introduced, using ^{15}N-ammonium chloride and ^{2}H,^{15}N-Celtone Base powder. IPTG (1.0 mM) was added for the induction, when the optical density at 600 nm reached 0.8. At the same time, a 5-fold concentration of the target ^{1}H-amino acid, as compared to that of the ^{2}H-amino acid included in ^{2}H-Celtone Base Powder, was added (http://www.spectrastableisotopes.com/Catalog/ Celtone_Powder_Information.aspx). Cells were harvested after 6 h of induction. Eighteen amino acids (Ala, Arg, Asp, Cys, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr, and Val) were investigated to determine which of these are suitable for the labeling.

NMR Investigation of the Labeling Selectivity and Efficiency. The selectivity and efficiency of the amino acid selective labeling were examined by obtaining the NMR spectra of approximately 1.0 mM Trx in 20 mM sodium phosphate buffer (pH 5.7), containing 100 mM NaCl, 0.10 mM 3-(trimethylsilyl)-propionic acid-d_4 sodium salt (TSP), and 99.8%(v/v) ^2H_2O or 10% (v/v) ^2H_2O. NMR spectra were observed on a Bruker Avance 600 MHz spectrometer equipped with a cryogenic probe or a Bruker DRX 400 MHz spectrometer at 35 °C. The assignments of the Trx NMR signals were previously reported in the literature.[15−17]

The labeling selectivity was examined by the ^{1}H−^{13}C HSQC spectra for the natural abundance of ^{13}C atoms. ^{1}H/^{2}H ratios of the side chain hydrogen atoms were evaluated by the intensity of the ^{1}H signals, as compared to those of 0.10 mM TSP in the ^{1}H 1D spectrum. ^{1}H/^{2}H ratios for H_α signals overlapping with the H_2O signal were estimated by the isotope shift of the signals in the ^{1}H−^{15}N HSQC spectra.[18]

CS Experiments. The samples for the CS and ASCS experiments were a mixture of 1.2 mM differently labeled C90S mutant of YUH1 as the CS-donor molecule and 1.0 mM uniformly ^{2}H,^{15}N-labeled Ub as the CS acceptor. For the conventional CS and its negative control experiment, unlabeled and uniformly ^{2}H-labeled C90S mutant of YUH1 were used, respectively. The ASCS experiments were carried out using the Ala, Arg, Leu, Met, Phe, Pro, or Tyr selectively ^{1}H-labeled C90S mutant of YUH1. The sample buffer contained 50 mM sodium phosphate (pH 6.5) and 50 mM NaCl in 20% H_2O/80% ^2H_2O. CS and ASCS experiments were performed at 30 or 10 °C. Irradiation was carried out for 1.5 or 4.5 s by a WURST-2 decoupling scheme. The bandwidth of the WURST-2 saturation was ∼4.5 kHz, in which the maximum RF amplitude was set to 0.17 kHz. Since the saturation frequency was set at 0.83 ppm, direct saturation by the WURST-2 ranges from −3.0 to 4.6 ppm. This saturation range includes almost 70% for H_α and more than 99% of the aliphatic protons except for H_α in the BMRB database. We confirmed that the water signal was not affected by the WURST-2. The recycling delay was set at 3.5 s. Recording time was 3−12 h per sample.

Structural Model of the Ub-C90S Mutant of YUH1 Complex. The crystal structure of the complex of human Ub and wild type YUH1 (1cmx)[13] was used for the validation of the ASCS results. The coordinates of the hydrogen atoms were generated using MOLMOL 2k.2.[19] It should be noted that the global structure of the C90S mutant of YUH1 is identical to that of the wild type, as verified by the small difference in their ^{1}H−^{15}N HSQC spectra (data not shown). Although yeast and human Ub have three residue differences, where Ser19, Asp24, and Ser28 in yeast Ub are replaced with Pro19, Glu24, and Ala28 in human Ub, the structural differences are reportedly limited to these three residues, as suggested by their difference Fourier map.[14] For the search of the CS-source residues, each of the individual structures in the complex was used.

(6) Matthews, H. R.; Matthews, K. S.; Opella, S. J. *Biochim. Biophys. Acta* **1977**, *497*, 1–13.
(7) LeMaster, D. M. *FEBS Lett.* **1987**, *223*, 191–6.
(8) Arrowsmith, C. H.; Pachter, R.; Altman, R. B.; Iyer, S. B.; Jardetzky, O. *Biochemistry* **1990**, *29*, 6332–41.
(9) Wishart, D. S.; Sykes, B. D.; Richards, F. M. *Biochim. Biophys. Acta* **1993**, *1164*, 36–46.
(10) Curley, R. W., Jr.; Panigot, M. J.; Hansen, A. P.; Fesik, S. W. *J. Biomol. NMR* **1994**, *4*, 335–40.
(11) Tugarinov, V.; Kanelis, V.; Kay, L. E. *Nat. Protoc.* **2006**, *1*, 749–54.
(12) Rajesh, S.; Nietlispach, D.; Nakayama, H.; Takio, K.; Laue, E. D.; Shibata, T.; Ito, Y. *J. Biomol. NMR* **2003**, *27*, 81–6.
(13) Johnston, S. C.; Riddle, S. M.; Cohen, R. E.; Hill, C. P. *EMBO J.* **1999**, *18*, 3877–87.
(14) Sakamoto, T.; Tanaka, T.; Ito, Y.; Rajesh, S.; Iwamoto-Sugai, M.; Kodera, Y.; Tsuchida, N.; Shibata, T.; Kohno, T. *Biochemistry* **1999**, *38*, 11634–42.
(15) Dyson, H. J.; Holmgren, A.; Wright, P. E. *Biochemistry* **1989**, *28*, 7074–87.
(16) Chandrasekhar, K.; Krause, G.; Holmgren, A.; Dyson, H. J. *FEBS Lett.* **1991**, *284*, 178–83.
(17) Chandrasekhar, K.; Campbell, A. P.; Jeng, M. F.; Holmgren, A.; Dyson, H. J. *J. Biomol. NMR* **1994**, *4*, 411–32.
(18) Takeuchi, K.; Ng, E.; Malia, T. J.; Wagner, G. *J. Biomol. NMR* **2007**, *38*, 89–98.
(19) Koradi, R.; Billeter, M.; Wuthrich, K. *J. Mol. Graphics* **1996**, *14*, 29–32.

Calculation of the Maximum Deviation to Evaluate the Spatial Complementarity. Given a set of the CS-source candidate residues and the corresponding cross-saturated amide protons, their spatial complementarity was evaluated as follows. First, the center of gravity of the side chain protons of each CS-source candidate residues are defined as a representative point. Supposing there are three pairs of the CS-source residues and the cross-saturated protons, a triangle of the representative points of the CS-source candidates and a triangle of the cross-saturated amide protons are formed. Then, the centers of gravity of the two triangles are overlaid on the coordinate origin, and one triangle was rotated along the Euler axes at 10° steps. At each step, the distances between the corresponding vertices of the triangles are calculated, among which the longest distance is defined as the maximum deviation. The orientation with the best spatial complementarity is obtained when the maximum deviation gives the smallest value during the rotational search.

In general cases, in which N pairs of the CS-source candidates and the cross-saturated amide protons are supposed, the rotational search of the two N-gonal shapes can be carried out similarly to calculate the best relative orientation with the smallest maximum deviation.

Results

Amino Acid Selective ^{1}H Labeling. Since the ASCS method identifies the interacting residues without direct observation of the CS-donor protein NMR resonances, the amino acid selective ^{1}H-labeling must be strictly controlled in order to ensure that the CS is exclusively from the labeled amino acids. Therefore, the labeling selectivity and efficiency for a number of amino acids were quantitatively examined by NMR spectroscopy.

^{1}H−^{13}C HSQC spectra using ^{13}C nuclei at the natural abundance were observed to evaluate the labeling selectivity. As control experiments, we confirmed that the unlabeled Trx showed all signals (Figure 1A). Meanwhile, the uniformly ^{2}H-labeled Trx signals suggested that the ^{1}H ratios are no more than 3% (Figure 1B). Amino acid selectively labeled Trx samples were then prepared for 18 amino acids, Ala, Arg, Asp, Cys, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr, and Val. ^{1}H−^{13}C HSQC spectra for Arg-selective ^{1}H-labeled Trx in a ^{2}H-background ([^{1}H-Arg] Trx) and [^{1}H-Asp] Trx are shown in Figure 1C and 1D, respectively, as examples. [^{1}H-Arg] Trx exhibited a set of signals from an Arg residue, Arg73, which is the only Arg residue in Trx, indicating that the ^{1}H-labeling was selective for Arg (Figure 1C). On the other hand, [^{1}H-Asp] Trx showed many signals from Asp and other amino acids (Figure 1D), showing that the ^{1}H-labeling was not selective for Asp. Likewise, we found that the selective labeling was feasible for Ala, Arg, Cys, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Trp, and Tyr but was difficult for Asp, Gln, and Ser. Limited scrambling was observed from Thr to Ile H_δ and from Val to Leu H_γ and H_δ, respectively (Table 1).

The amount of ^{1}H at each hydrogen position of the labeled amino acid was then quantified (Table 1). The side-chain hydrogen atoms were well labeled with ^{1}H for Ala, Arg, His, Ile, Leu, Lys, Met, Phe, Pro, Trp, and Tyr, although the ^{1}H ratio for H_α was relatively low. Cys and Gly exhibited 30−40% ^{1}H-labeling efficiency.

In summary, the amino acid selective ^{1}H-labeling was applicable to Ala, Arg, His, Ile, Leu, Lys, Met, Phe, Pro, Tyr, Trp, Cys, and Gly but not to Asp, Gln, Ser, Thr, and Val. Especially, the former 11 amino acids were labeled with high efficiency.

Outline of the ASCS Method. The conventional CS method, in which the CS donor and the acceptor are unlabeled and uniformly ^{2}H,^{15}N-labeled, respectively, is shown in Figure 2A.

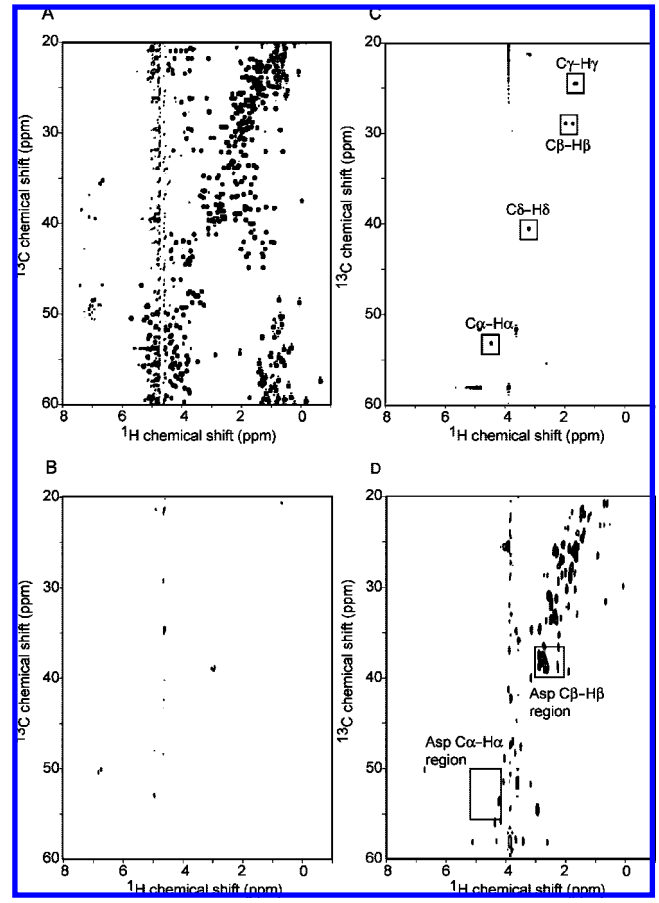


Figure 1. ^{1}H−^{13}C HSQC spectra at natural abundance of ^{13}C using a cryogenic probe. (A) Unlabeled Trx. (B) Uniformly ^{2}H-labeled Trx. (C) [^{1}H-Arg] Trx. The signals in the boxes are from the Arg residues of Trx. (D) [^{1}H-Asp] Trx. Regions for Asp C_α−H_α and C_β−H_β in Trx are boxed. The C_α−H_α signals overlapping with the water signal were eliminated by the baseline correction.

The CS is observed as the signal intensity reduction of the amide protons of the acceptor at the interface of the complex. However, no structural information is obtained for the CS donor, due to the saturation of the whole donor molecule.

In the ASCS method, on the other hand, the CS donor is amino acid selectively ^{1}H-labeled in a ^{2}H-background, while the CS acceptor is uniformly ^{2}H,^{15}N-labeled. With this labeling method, CS originates from the ^{1}H-labeled amino acid residues on the donor. Figure 2B schematically depicts the amino acid X, Y, and Z selective CS, where the acceptor amide protons, H1, H2, and H3 and H4, are cross-saturated, respectively. As shown in Figure 2C, the distribution of the CS-source residues of amino acid type X, Y, and Z well overlap that of the cross-saturated amide protons on the CS-acceptor molecule at the binding interface. Therefore, given each three-dimensional structure of the CS donor and acceptor and the combinatorial analysis of multiple ASCS results using several differently labeled donors, identification of the CS-source residue(s) on the donor is possible.

Observation of the CS from the Amino Acid Selectively ^{1}H-Labeled Protein. The observation of the ASCS was validated using a known structure of the complex of Ub and C90S mutant of YUH1.[13] The C90S mutant of YUH1 was found to bind Ub with a dissociation constant of 43 nM (data not shown). Hereafter, we refer to the C90S mutant of YUH1 as YUH1.

Table 1. Selectivity and Efficiency of Amino Acid Selective ^{1}H-Labeling

amino acid	selectivity	percentage of ^{1}H[a]				
		H_α	H_β	H_γ	H_δ	others
Ala	selective	44	78			
Arg	selective	96	84	86	74	
His	selective	51	67			77 (H_δ1), 63 (H_ϵ1)
Ile	selective	<3	90		100	86 (H_γ12,H_γ13), 100 (H_γ2)
Leu	selective	10	100	100	97	
Lys	selective	72	93	93	93	75 (H_ϵ)
Met	selective	40	71	94		80 (H_ϵ)
Phe	selective	7	81		97	97 (H_ϵ)
Pro	selective	54	79	79	56	
Trp	selective	10	45			75 (H_δ1), 77 (H_ϵ3), 77 (H_ζ2), 82 (H_ζ3), 79 (H_η2)
Tyr	selective	<3	56	74	78	
Cys	selective	<3	30			
Gly	selective	37				
Thr	metabolized	83	100	86		81(Ile H_δ)[c]
Val	metabolized	11	70	68		40(Leu H_γ)[c], 51(Leu H_δ)[c]
Asp	metabolized	54	N/A[b]	N/A[b]		
Gln	metabolized	<3	N/A[b]			
Ser	metabolized	<3	N/A[b]			

[a] The errors for these values are estimated to be ~10%. [b] Signal was not assigned due to the degeneracy of the signals. [c] Percentage of ^{1}H in other amino acid.

First, the CS experiments were performed using unlabeled YUH1 (the conventional CS method, Figure 3A) and uniformly ^{2}H-labeled YUH1 as a negative control (Figure 3B). While the ^{2}H-labeled YUH1 showed no signal intensity reduction, the conventional CS experiment provided signal intensity reductions for residues 7−11, 36, 40−42, 47, and 69−75 of Ub (Figure 3A). These residues were mapped on the Ub−YUH1 complex structure (Figure 3C), indicating that the interface with YUH1 was properly identified.

ASCS experiments were then carried out for Ala, Arg, Leu, Pro, Tyr, Met, or Phe selectively ^{1}H-labeled YUH1 in complex with uniformly ^{2}H,^{15}N-labeled Ub, respectively. Some of these amino acids exist in the Ub interface: Ala227, Leu58, Leu168, Pro10, and Tyr167 of YUH1 are located within 4 Å, and Met215 and Phe47 are located within 4 to 5 Å of any amide proton of Ub (Figure 4). [^{1}H-Arg] YUH1 was also investigated as a negative control, since no Arg residues exist in the Ub-binding site on YUH1. Figure 5 shows each of the ASCS results. The CS was observed from [^{1}H-Ala] YUH1 to Leu71 and Leu73 of Ub (Figure 5A); from [^{1}H-Leu] YUH1 to Leu73, Arg74, and Gly75 (Figure 5B); from [^{1}H-Pro] YUH1 to Gly75 (Figure 5C); and from [^{1}H-Tyr] YUH1 to Gly75 (Figure 5D). Although no signal intensity reduction was observed for Met and Phe under the same conditions as those for the other amino acids (data not shown), small intensity reductions were observed when changing the saturation duration from 1.5 to 4.5 s and the temperature from 30 to 10 °C (Figure 5E, F). No CS was detected from Arg to Ub (Figure 5G).

Search for the CS-Source Residues. Supposing that the complex structure is not available but both protein structures are available, how can we investigate the interaction model? To address this question, we tried to identify the Ub binding site on YUH1 by analyzing the Ala, Leu, Pro, and Tyr selective ASCS results with reference to the individual structures of Ub and YUH1. The saturation of Pro residues on [^{1}H-Pro] YUH1 and Tyr residues on [^{1}H-Tyr] YUH1 caused the signal reduction at Gly75 of Ub (Figure 5C, D), indicating that the CS-source Pro and Tyr residues are located in close proximity to the YUH1 surface. The combinations of proximal Pro and Tyr residues (<7 Å) found on the YUH1 surface were Pro10/Tyr167, Pro15/Tyr37, Pro49/Tyr37, and Pro188/Tyr190. The Ub molecule was then placed on the YUH1 surface, to allow Gly75 of Ub to approach each of the four Pro/Tyr sites. While keeping the amide proton of Gly75 of Ub in proximity of each Pro/Tyr site, various orientations of Ub were examined to satisfy other ASCS results, i.e., the CS-source Ala residue(s) close to the cross-saturated Leu71 and Leu73 (Figure 5A) and the CS-source Leu residue(s) close to Leu73, Arg74, and Gly75 (Figure 5B) of Ub. As a result, this manual analysis identified Pro10 and Tyr167 as the CS source for Gly75 of Ub; Ala218 and/or Ala227 for Leu71 and Leu73 of Ub; and Leu58, Leu165, and/or Leu229 for Leu73, Arg74, and Gly75 of Ub.

Systematically, all possible combinations of the CS-source residues were evaluated, based on the maximum deviation of the CS-source residues and the cross-saturated amide protons (see Materials and Methods). Based on the ASCS results, shown in Figure 5A to 5D, we extracted 2 Ala residues, 3 Leu residues, 1 Tyr residue, and 1 Pro residue from the YUH1 surface as the CS-source residues. Since YUH1 has 10 Ala, 17 Leu, 11 Pro, and 4 Tyr residues, the number of combinations of CS-source candidates is $10^2 \times 17^3 \times 4^1 \times 11^1 = 21\,617\,200$. For each combination of CS-source candidates, the best relative orientation of Ub and YUH1 was then searched so that the geometry of the selected CS-source candidate on the YUH1 structure could overlap the geometry of the cross-saturated amide protons on Ub. The CS-source candidates were sorted by the maximum deviation in ascending order, and the first part of the listing is shown in Table 2. The top eight among 20 million combinations were the same candidate sets as those obtained by the aforementioned manual analysis, which exist at the interface with Ub in the crystal structure (Figure 6).

Discussion

Selectivity and Efficiency of the Amino Acid Selective ^{1}H-Labeling. We have shown experimentally that amino acid selective ^{1}H-labeling is feasible, using an *E. coli* protein expression system for Ala, Arg, Cys, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Trp, and Tyr. The scrambling of the side-chain protons of these amino acids was almost negligible when the protein expression was induced for 6 h. Indeed, an induction for 18 h showed unchanged selectivity, as compared to the 6 h induction for Arg-selective ^{1}H-labeling (data not shown). This indicates that a longer induction may be possible for the ^{1}H-selective labeling of some amino acids.

Val and Thr are not applicable for ASCS by the present labeling method, since strong NMR signals for the H_γ and H_δ atoms of Leu and the H_δ atoms of Ile were observed for the samples intended to introduce the Val and Thr selective ^{1}H-labeling, respectively. According to the metabolism map based on the ^{15}N analysis,[4] Val and Leu are converted from α-ketoisovalerate, and the reverse reaction from Val to α-ketoisovalerate is likely to occur. Hence, our results of the one-way scrambling from Val to Leu might be explained by the pathway via α-ketoisovalerate. Also, the conversion from Thr to Ile appears to be one way on the metabolism map, which is consistent with our results.

Asp, Gln, and Ser were significantly metabolized to other amino acids. The labeling of Asn and Glu would not be selective, because interconversions occur between Asp and Asn and between Glu and Gln.[4,5]

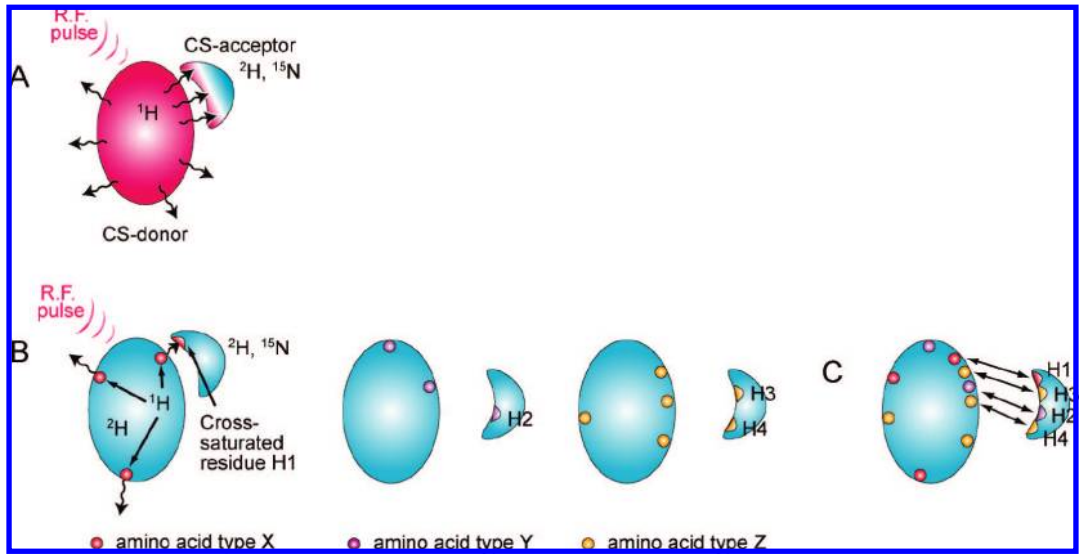


Figure 2. (A) Schematic representation of conventional CS methods. The CS donor is unlabeled, and the CS acceptor is uniformly ^{2}H,^{15}N labeled.[3] (B) Schematic representation of the ASCS method. The CS-donor molecules are amino acids, X, Y, and Z selectively ^{1}H-labeled in a ^{2}H-background on each panel, respectively, while the CS-acceptor molecule is labeled in the same way as the conventional CS shown in (A). The cross-saturated amide protons on the acceptor are depicted as H1, H2, H3, and H4, on the acceptor. (C) Combinatorial approach to determine the proximal residue pairs. The CS-source residues can be identified by combining several ASCS results, based on the spatial complementarity of the CS-source candidates and the cross-saturated amide protons.

For the amino acids that could not be selectively labeled by this method, auxotrophic mutants preventing the metabolism of a specific amino acid could be applicable.[4,5] For example, the DL39 strain, which requires Asp, Ile, Leu, Phe, Tyr, and Val, may suppress Leu and Ile synthesis from Val and Thr, respectively, and thus, it might be effective for Val and Thr selective labeling.[20] Alternatively, a cell-free expression system that excludes amino acid metabolizing enzymes would be suitable for the selective labeling, as mentioned in the literature.[21−27]

The labeling efficiency at each hydrogen position was evaluated for the selectively labeled amino acids (Table 1). It is known that H_α is easily replaced by ^{2}H when cells are grown in ^2H_2O-based medium.[18,28] This may cause the low incorporation efficiency of ^{1}H for H_α. Although most of the amino acids exhibited more than 50% incorporation of ^{1}H into their side chains, Cys showed only 30% labeling. This is probably due to the metabolism of other ^{2}H-amino acids included in the ^{2}H Celtone Base Powder. Indeed, Cys is known to be efficiently biosynthesized from Ser.[4] It should be noted that the low efficiency of the ^{1}H-labeling might weaken or diminish the CS effect, even for the interacting residues. Therefore, the experimental conditions, such as the duration and power of the saturation and the experimental temperature, should be optimized carefully when no CS is observed.

Amino Acid Selective ^{1}H-Labeling To Obtain the Structural Information. This amino acid selective ^{1}H-labeling in a ^{2}H-background was originally utilized in SOS-NMR,[29] which is based on the saturation transfer from the ^{1}H-labeled amino acid of the protein to its small molecular ligand. With reference to the three-dimensional structure of the protein, the ligand binding site on the labeled protein is identified by integrating several amino acid selective saturation transfer results.[30] This approach possesses the advantage that the time-consuming NMR assignment of the donor is not necessary, and thus, SOS-NMR is an efficient method to obtain structural information about the recognition of a small molecular ligand.[29]

However, substantial difficulties are encountered when applying the SOS-NMR method to the identification of a protein−protein interface, especially in the selective excitation of the saturation donor protein, since SOS-NMR uses an unlabeled, i.e., uniformly ^{1}H-labeled, molecule as the saturation acceptor.

In the ASCS method developed here, the CS donor is amino acid selectively ^{1}H-labeled in a ^{2}H-background, while the CS acceptor is uniformly ^{2}H,^{15}N-labeled, enabling the selective excitation of the CS donor by irradiating the aliphatic signals of the selectively ^{1}H-labeled amino acid.[2,3]

Given the three-dimensional structure of each protein of the complex, but not the complex structure, the multiple ASCS results will facilitate the specification of the CS-source residue(s) for each cross-saturated amide proton. Thus, the binding site on the CS donor is identified as the cluster of the CS-source residues. In addition, the pairwise information on the proximal residue pairs can be used to construct the structural model of the protein−protein complex.

(20) LeMaster, D. M.; Richards, F. M. *Biochemistry* **1988**, *27*, 142–50.
(21) Ozawa, K.; Headlam, M. J.; Schaeffer, P. M.; Henderson, B. R.; Dixon, N. E.; Otting, G. *Eur. J. Biochem.* **2004**, *271*, 4084–93.
(22) Kigawa, T.; Muto, Y.; Yokoyama, S. *J. Biomol. NMR* **1995**, *6*, 129–34.
(23) Yabuki, T.; Kigawa, T.; Dohmae, N.; Takio, K.; Terada, T.; Ito, Y.; Laue, E. D.; Cooper, J. A.; Kainosho, M.; Yokoyama, S. *J. Biomol. NMR* **1998**, *11*, 295–306.
(24) Klammt, C.; Lohr, F.; Schafer, B.; Haase, W.; Dotsch, V.; Ruterjans, H.; Glaubitz, C.; Bernhard, F. *Eur. J. Biochem.* **2004**, *271*, 568–80.
(25) Morita, E. H.; Shimizu, M.; Ogasawara, T.; Endo, Y.; Tanaka, R.; Kohno, T. *J. Biomol. NMR* **2004**, *30*, 37–45.
(26) Shi, J.; Pelton, J. G.; Cho, H. S.; Wemmer, D. E. *J. Biomol. NMR* **2004**, *28*, 235–47.
(27) Staunton, D.; Schlinkert, R.; Zanetti, G.; Colebrook, S. A.; Campbell, I. D. *Magn. Reson. Chem.* **2006**, *44*, S2–9.
(28) Lohr, F.; Katsemi, V.; Hartleib, J.; Gunther, U.; Ruterjans, H. *J. Biomol. NMR* **2003**, *25*, 291–311.
(29) Hajduk, P. J.; Mack, J. C.; Olejniczak, E. T.; Park, C.; Dandliker, P. J.; Beutel, B. A. *J. Am. Chem. Soc.* **2004**, *126*, 2390–8.
(30) Mayer, M.; Meyer, B. *Angew. Chem., Int. Ed.* **1999**, *38*, 1784–8.

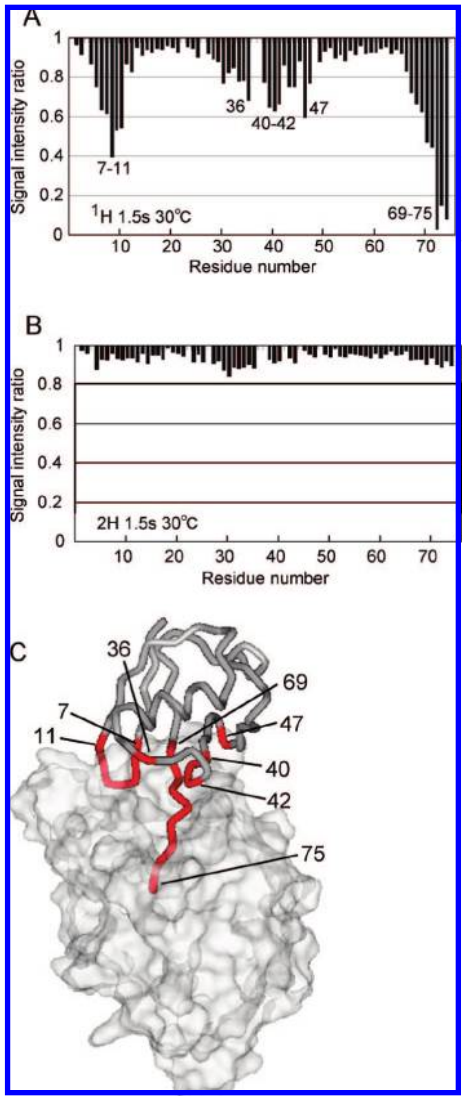


Figure 3. CS from YUH1 to Ub. (A) Conventional CS experiment using unlabeled YUH1. Residues 7−11, 36, 40−42, 47, and 69−76 of Ub were significantly cross-saturated. (B) CS experiment using uniformly ^{2}H-labeled YUH1. No residue was cross-saturated. (C) Mapping of the residues identified in (A) on the crystal structure of the Ub−YUH1 complex. Ub and YUH1 are shown as a backbone tube model and a molecular surface model, respectively. Cross-saturated residues are shown in red.

Validation of the ASCS Experimental Results. As expected from the crystal structure of the Ub−YUH1 complex, where Ala227, Leu58, Leu165, Pro10, and Tyr167 exist within 4 Å from the Ub amide(s) (Figure 4), the [^{1}H-Ala], [^{1}H-Leu], [^{1}H-Pro], and [^{1}H-Tyr] YUH1 samples showed the CS to the Ub amide proton(s) (Figure 5). These results suggest that the CS from the selectively ^{1}H-labeled residues on YUH1 was successfully observed.

In the cases of [^{1}H-Phe] and [^{1}H-Met] YUH1, no amide proton on Ub was cross-saturated under the same experimental conditions as those for the amino acids mentioned above (the duration of the saturation for 1.5 s and experimental temperature at 30 °C), although one of the protons of Phe47, Phe226, and Met215 are located within 4.0, 4.3, and 5.1 Å from the amide protons of Gly47, Thr9, and Thr9, respectively (Figures 4E, F). This might be due to the rapid local motion of these side chains, which attenuates the dipolar-dipolar interactions. Thus, the upper limit of ASCS under the current conditions seems to exist between 4 and 5 Å.

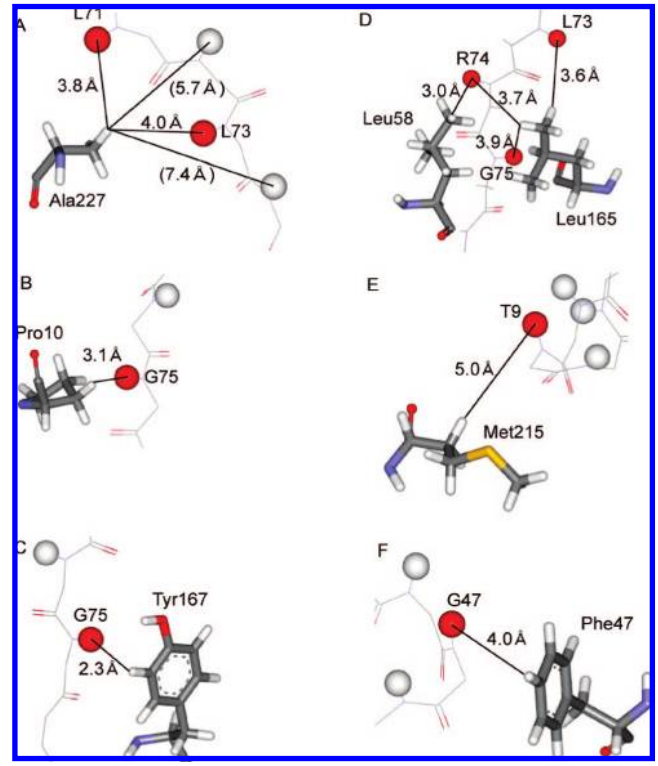


Figure 4. Distances between the proximal CS-source residues and the amide protons of Ub. (A) Ala, (B) Pro, (C) Tyr, (D) Leu, (E) Met, and (F) Phe residues on YUH1 within 5 Å of any amide proton of Ub, are shown as stick models. Ub amide hydrogen atoms are shown in spheres. Cross-saturated amide hydrogen atoms are shown in red. Main chain atoms of Ub are shown as line models. The most proximal residue on YUH1 is connected by a line, and the distance is indicated. Values in parentheses in (A) indicate distances between an amide hydrogen atom without CS and the most proximal hydrogen atom of Ala.

Longer saturation times and/or the lower temperatures might be chosen to facilitate dipolar−dipolar interactions. Indeed, the saturation time of 4.5 s at 10 °C enabled the observation of the CS from Met (5.0 Å) and Phe (4.0 Å) (Figure 5E, F), although other amide protons also should decrease signal intensities, presumably due to spin diffusion. Therefore, the experimental conditions such as the saturation time and the experimental temperature should be optimized, and the resultant optimal condition should be applied to all the samples with different amino acid selective ^{1}H-labeling. On the contrary, the lack of a cross-saturated signal under one set of experimental conditions does not necessarily mean that the labeled amino acid does not participate in the binding interface. SOS-NMR, which uses the NOE from the selectively labeled protein to the unlabeled small molecular ligand in order to define the ligand binding site, applies the exclusive distance constraint that the selectively labeled amino acid is considered as being outside the ligand binding site when no NOE is observed. Such exclusive information was not utilized here to identify the binding interface, considering the possibility of improving the experimental conditions and/or the insufficient labeling efficiency.

Identification of the CS Donor Residues. The ASCS results provide the information on the "amino acid type" of the CS source residue(s) on the donor as well as the cross-saturated "amide proton" of the CS acceptor, where the residue pair exists within 4−5 Å. However, there are a number of the same amino acid residues on the donor surface. Thus, if the CS-source "residues" are specified, then the pairwise information of the

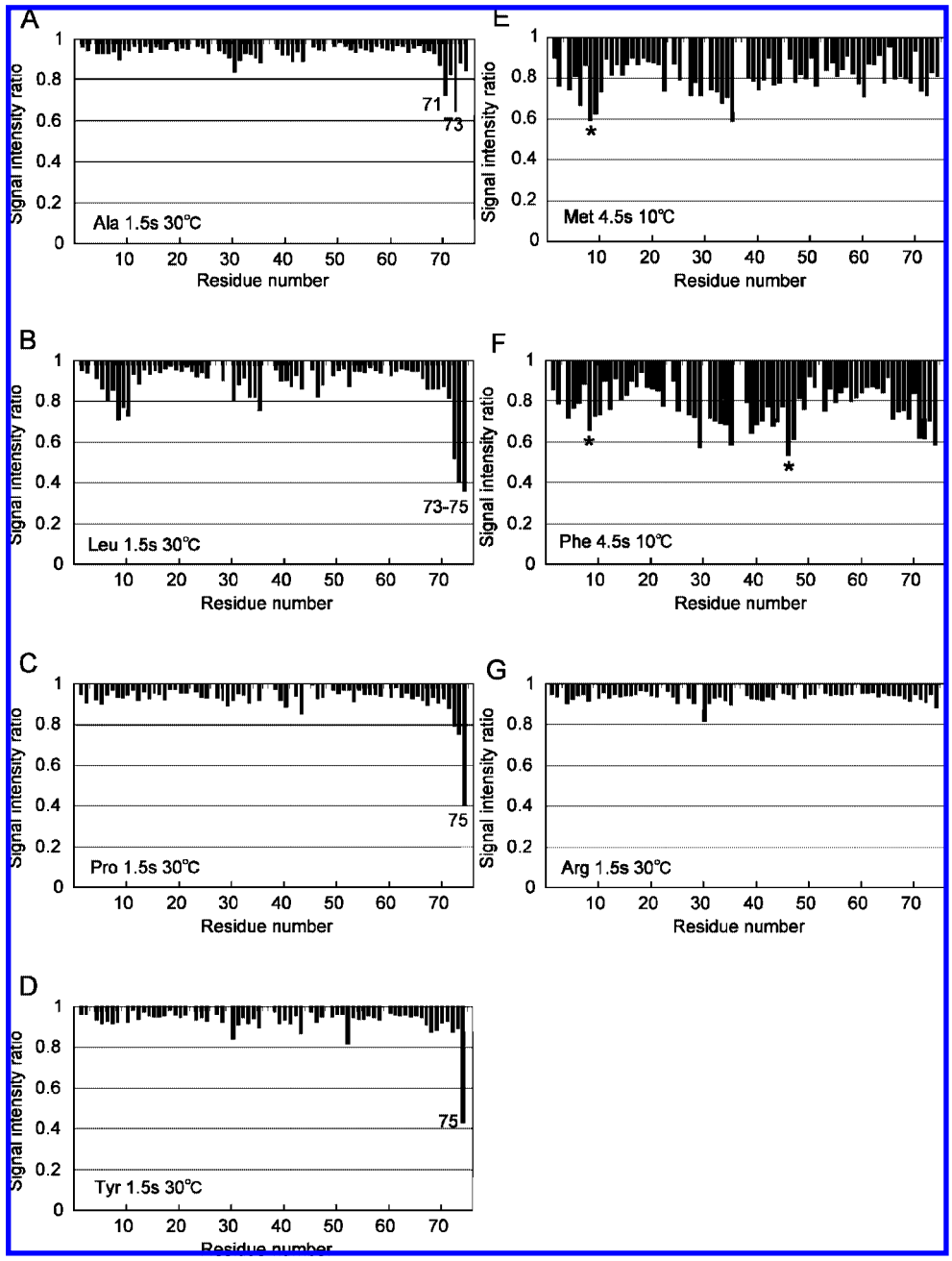


Figure 5. (A) Ala, (B) Leu, (C) Pro, (D) Tyr, (E) Met, (F) Phe, and (G) Arg-selective CS experiments. (A−D and G) ASCS experiments were performed at 30 °C by irradiating for 1.5 s. Numbers show cross-saturated Ub residues. (E and F) ASCS experiments were performed at 10 °C by irradiating for 4.5 s. Asterisks indicate the Ub residues within 4−5 Å from the ^{1}H-labeled amino acid residues of YUH.

CS-source residues and the cross-saturated amide proton can be used as the intermolecular structural constraints for building a complex model.

Here, the manual analysis of the ASCS results successfully identified the binding site on YUH1. The systematic search, based on the three-dimensional spatial complementarity evaluated by maximum deviation, yielded the possible combinations of the CS-source residues, where the eight combinations out of 21 617 200 with the smallest maximum deviation are from the residues in the Ub interface.

The search was based on the maximum deviations, rather than the rmsd values, although a small rmsd value is one of the good criteria for the spatial complementarity. Since rmsd calculation averages all deviations, it could result in the small rmsd value, even when one of the deviation is too large to satisfy the experimental results of ASCS. Therefore, the maximum deviation is chosen to evaluate the spatial complementarity and the satisfaction of the ASCS results simultaneously.

It should be noted that the maximum deviations in Table 2 are not the real distances between hydrogen atoms, where the coordinates of the CS-donor residues are replaced by the center of gravity of the hydrogen atoms. In order to identify the CS-source residues, these candidates have to be verified by the "structural complementarity" of Ub and YUH1, using molecular dynamics simulations. Docking protocols using the ASCS results as structural constraints are currently being developed. As shown in Figure 6, however, the candidates proposed by the ASCS results mostly describe the Ub binding site on YUH1, which should provide good initial models for the protein−protein docking simulations.

Table 2. Listing of the CS-Source Candidates on YUH1 for the Cross-Saturated Amide Protons of Ub

cross-saturated amide proton of Ub	L71	L73	L73	R74	G75	G75	G75	maximum deviation
CS-source residues on YUH1	A227	A227	L58	L165	L58	Y167	P10	4.809
	A227	A227	L165	L165	L58	Y167	P10	4.827
	A227	A227	L165	L58	L58	Y167	P10	4.892
	A227	A227	L58	L58	L58	Y167	P10	5.031
	A227	A227	L58	L58	L165	Y167	P10	5.065
	A227	A227	L58	L165	L165	Y167	P10	5.121
	A227	A227	L165	L58	L165	Y167	P10	5.121
	A227	A227	L165	L165	L165	Y167	P10	5.151
	A218	A227	L165	L165	L165	Y167	P10	5.365
	A218	A227	L58	L165	L165	Y167	P10	5.372
	A218	A227	L165	L58	L165	Y167	P10	5.372
	A218	A227	L165	L165	L58	Y167	P10	5.372
	A46	A46	L45	L48	L48	Y37	P49	5.393
	A46	A46	L48	L45	L48	Y37	P49	5.393
	A227	A227	L229	L165	L58	Y167	P10	5.405

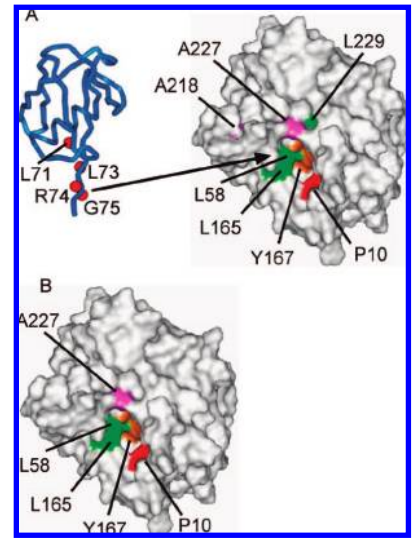


Figure 6. (A) Left: Cross-saturated amide protons (red spheres) in four amino acid selective CS experiments. The Ub backbone is shown as a blue tube model. Right: Candidates of proximal residues on YUH1 to cross-saturated Ub residues. Ala, Pro, Tyr, and Leu are colored pink, red, orange, and green, respectively. (B) Residue proximal to cross-saturated residues on the crystal structure of Ub−YUH1.

ASCS Method for the Structural Analysis of the Protein−Protein Interaction. To date, docking simulation software still cannot readily build the complex structures of macromolecules without information about the interacting sites, even when each structure component of the complex is available.[31,32] Thus, the experimental information about the binding sites and the interacting residue pairs obtained by the conventional CS and present ASCS methods will facilitate the construction of the macromolecular complex model.

Other NMR techniques such as chemical shift perturbation (CSP) and intermolecular NOE have been used to determine the intermolecular interaction sites.[33] The CSP method provides the interface on both proteins using the backbone assignments for both proteins. However, the determination of the relative orientation of the two proteins is difficult when using interface information. In contrast, the NOE analysis reveals proximal atom pairs, which contribute to the precise calculation of the complex structure, including the relative rotational orientation. A recently reported approach utilized sparse, but unambiguous, intermolecular NOEs to dock protein−protein complexes.[34] However, the NOE analysis additionally requires the side-chain assignments, which are time-consuming and often impossible to obtain for larger systems.

More efficient approaches have been developed, where the assignments of the NMR resonances are not required. For example, the fast mapping method,[35] which is based on CSP, uses the amino acid selectively ^{15}N-labeled protein. Without the backbone assignment, this method can rapidly reveal the binding interfaces of the two interacting protein but not their binding orientation.

In contrast, the ASCS method not only identifies the binding interface of a protein−protein complex but also determines the relative orientation of the two interacting proteins, without NMR assignment of the CS donor. Therefore, in a case where the assignment of the CS acceptor in complex with the CS donor, the *E. coli* protein expression system, and the free form structures of the CS acceptor and the CS donor are applicable, the ASCS method presented here is a general and practical approach for the structural analysis of protein−protein interactions, which yields pairwise information to construct a complex model.

In the ASCS experiments, it is required to quantitatively measure the intensity of each peak in ^{1}H−^{15}N HSQC or TROSY spectra under 20% H_2O/80% D_2O solvent conditions. Therefore, the sample concentration of more than 100 μM would be needed, when a 600 MHz NMR spectrometer with a cryogenic probe is used for the measurements. The ASCS method has similar limitations as the conventional CS method in terms of sensitivity. Indeed, the measurement time for ASCS (Figure 5) was the same as that for conventional CS (Figure 3).

Application to the TCS Method. In the case of a system with weak interactions, the transferred cross-saturation (TCS) method is applicable, since it utilizes the chemical exchange processes between the free and bound states.[36] Since TCS can analyze large proteins with molecular weights of more than 100 K for the CS donor,[36−41] the application of the amino acid selective TCS (ASTCS) could reveal the binding interfaces on such large proteins.

We have carried out ASTCS experiments using the wild type YUH1 and Ub, with a dissociation constant of 18 μM. The results were similar to the ASCS results presented here (data not shown), indicating the feasibility of the ASTCS for larger complexes (>100 kDa) that is not feasible to determine the structure with conventional NMR techniques.

(31) Gray, J. J. *Curr. Opin. Struct. Biol.* **2006**, *16*, 183–93.
(32) Bonvin, A. M. *Curr. Opin. Struct. Biol.* **2006**, *16*, 194–200.
(33) Zuiderweg, E. R. *Biochemistry* **2002**, *41*, 1–7.
(34) Tang, C.; Clore, G. M. *J. Biomol. NMR* **2006**, *36*, 37–44.
(35) Reese, M. L.; Dotsch, V. *J. Am. Chem. Soc.* **2003**, *125*, 14250–1.
(36) Nakanishi, T.; Miyazawa, M.; Sakakura, M.; Terasawa, H.; Takahashi, H.; Shimada, I. *J. Mol. Biol.* **2002**, *318*, 245–9.
(37) Takeuchi, K.; Yokogawa, M.; Matsuda, T.; Sugai, M.; Kawano, S.; Kohno, T.; Nakamura, H.; Takahashi, H.; Shimada, I. *Structure* **2003**, *11*, 1381–92.
(38) Nishida, N.; Sumikawa, H.; Sakakura, M.; Shimba, N.; Takahashi, H.; Terasawa, H.; Suzuki, E. I.; Shimada, I. *Nat. Struct. Biol.* **2003**, *10*, 53–58.
(39) Takeuchi, K.; Takahashi, H.; Sugai, M.; Iwai, H.; Kohno, T.; Sekimizu, K.; Natori, S.; Shimada, I. *J. Biol. Chem.* **2004**, *279*, 4981–7.
(40) Yokogawa, M.; Takeuchi, K.; Shimada, I. *J. Am. Chem. Soc.* **2005**, *127*, 12021–7.
(41) Ichikawa, O.; Osawa, M.; Nishida, N.; Goshima, N.; Nomura, N.; Shimada, I. *EMBO J.* **2007**, *26*, 4168–76.

Choice of Amino Acids for the Selective Labeling. Which amino acids should be primarily considered for an ASCS experiment? Arg, Cys, His, Ile, Leu, Met, Phe, Trp, Tyr, and Val reportedly show a propensity to exist at protein−protein interfaces.[42] Among these amino acids, Arg, Leu, and Val frequently occur in the primary sequence. Therefore, ASCS may be more likely to be observed for Arg- and Leu-selective labeled proteins (Val is not suitable for the selective labeling), and thus these amino acids are useful to optimize the conditions for the ASCS experiments. The rest of the amino acids (Cys, His, Ile, Met, Phe, Trp, and Tyr) could be major determinants to identify the CS source residues, due to their low occurrence. Altogether, we recommend starting with a few amino acids, among the latter seven amino acids, as well as Arg and Leu.

Conclusion

In summary, we have described an NMR approach, the ASCS method, to determine protein binding sites using only the main chain amide proton assignment of the acceptor protein, by the selective labeling of the donor molecule. The selectivity and efficiency of the amino acid selective ^{1}H-labeling, which are critical for ASCS, were quantitatively investigated. ASCS experiments on the known structure of the complex of Ub−YUH1 provided CS profiles that were consistent with the structure. The combination of ASCS results from several different amino acids could provide pairwise information for the proximal residues between the CS sources and the CS acceptors as well as the binding site on the selectively labeled protein. This information would be useful to build a complex model using existing protein−protein docking protocols.

Acknowledgment. We thank Dr. Toshiyuki Kohno for the plasmids used to produce the Ub and YUH1 proteins. This work was supported in part by grants from the Japan New Energy and Industrial Technology Development Organization (NEDO) and the Ministry of Economy, Trade, and Industry (METI) (to I.S.), a Grant-in-Aid for Scientific Research on Priority Areas from the Japanese Ministry of Education, Culture, Sports, Science, and Technology (to M.O. and I.S.), and the Takeda Science Foundation (to M.O.).

JA804062T

(42) Chen, H.; Zhou, H. X. *Proteins* **2005**, *61*, 21–35.