TABLE II
Dissociation constants

All data are the averages and estimated deviation of at least two independent measurements.

	K_d (μM)	Fold decrease
S5a UIM binding to wild type and mutant HR23B UbL		
Wild type	3.4 ± 0.3	
L8A	>1000	>300
I47A	16.6 ± 3.1	4.9
V71H	>1000	>300
M73V	1.8 ± 0.3	0.5
T75A	5.1 ± 0.7	1.5
HR23B UbL binding to wild type and mutant S5a UIM		
Wild type	3.4 ± 0.3	
L278S	15.8 ± 0.5	4.6
E283A	12.1 ± 1.0	3.6
A290S	109 ± 14	32
S294A	213 ± 24	63

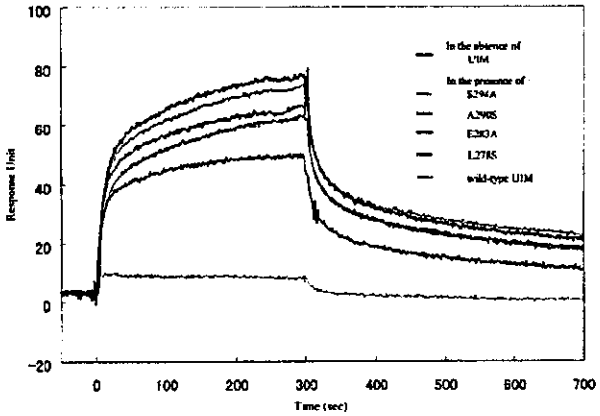

FIG. 4. **The effect of mutations on the affinity of S5a UIM for tetraubiquitin examined by a surface plasmon resonance-based competition assay.** Resonance curves were measured for 0.1 μM tetraubiquitin binding to immobilized GST-UIM in the presence or absence of 40 μM His-tagged wild type UIM or its point mutants are shown. The graph indicates that the mutants of UIM, S294A, A290S, E283A, and L278S, have weaker binding affinities for tetraubiquitin than the wild type UIM does.

terminal UIM, it still binds (10). This observation suggests that the electrostatic interaction mediated by Glu^{283} of UIM makes a nonessential contribution to the UIM-ubiquitin interaction.

The contact area of the UIM-UbL complex is 474 Å^2, as defined by the calculated change in the solvent-accessible surface area of UbL upon UIM binding. The area is comparable with the average values of 517 ± 83 and 617 ± 66 Å^2 calculated for peptide complexes of ten SH3 domains and nine SH2 domains, respectively. In the previous chemical shift perturbation experiments of the UbL domains from PLIC-2 and HR23B, NMR signals of residues close to but outside (as well as at) the interface of our UIM-HR23B complex structure also exhibited substantial changes upon binding to S5a (18 and 30). For example, the signals of Lys^{79}, Lys^{82}, Ile^{102}, and Lys^{103} of PLIC-2 exhibited substantial chemical shift perturbations upon S5a binding (18). However, the corresponding residues of the UbL domain of HR23B, Lys^{51}, Asn^{54}, Val^{74}, and Thr^{75}, make no direct contribution to the UIM interface of HR23B UbL. The noninvolvement of Thr^{75} in UIM binding was confirmed by mutagenesis; its substitution by alanine causes no apparent change in the affinity of UbL for UIM (Table II). We

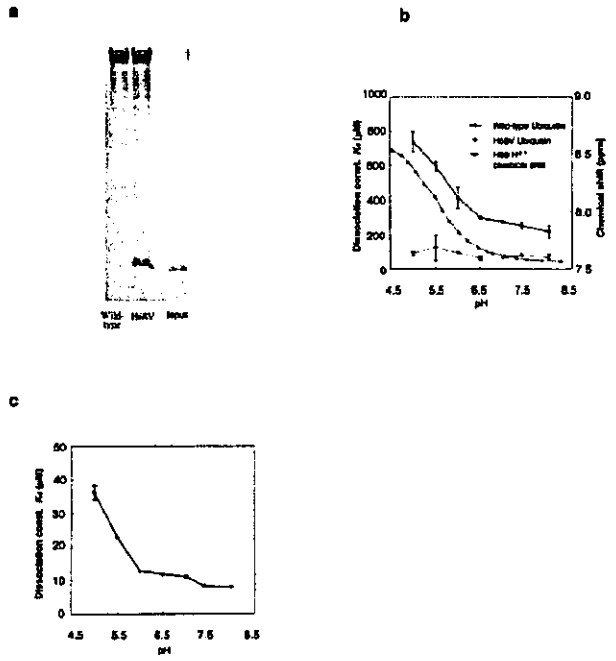

FIG. 5. **The UIM-ubiquitin interaction.** *a*, mutation of His^{68} to valine (H68V) enhances the binding affinity of ubiquitin to the UIM of S5a, shown by GST pull-down assay. The bound proteins were analyzed by GSH-mediated pull-down assay coupled with SDS-polyacrylamide gel electrophoresis and silver staining. The input lane represents 10% of the wild type ubiquitin used in the experiment. *b*, correlation between the affinity of ubiquitin for UIM and the protonation of ubiquitin His^{68}. Shown is the pH dependence of the interaction between UIM and wild type or H68V ubiquitin and the $^1H^{\varepsilon 1}$ chemical shift of His^{68} (*green*). The pK_a value of His^{68}, obtained by fitting the $H^{\varepsilon 1}$ and $H^{\delta 2}$ chemical shift data at pH range from 4.1 to 8.2, is 5.5. *c*, the pH dependence of interaction between UBA from yeast Dsk2p and ubiquitin.

assume that the NMR signals of PLIC-2 residues in the proximity of the interface were probably affected by local conformational changes or long range electrostatic effects on S5a binding. Such effects are often seen in chemical shift perturbation experiments (45) and were observed in the interaction between the UbL of HR23B and UIM (30). The UIM-UbL interaction mode proposed by based on chemical shift perturbation experiments Ryu *et al.* (30) is different from that observed in our complex structure. They assumed that residues from Glu^{283} to Gln^{296} or Gly^{297} adopts a helix, based on main chain chemical shift values of the UbL-bound UIM; that Glu^{284} and Glu^{285} of this helix locate near Lys^{48} and Lys^{51} of UbL; and that Leu 295 is positioned near the hydrophobic patch including Leu^8 of UbL. These interactions would place the helix of UIM on the UbL surface in the opposite orientation to that observed in our structure.

Role of His^{68} of Ubiquitin in UIM Binding—Our data suggest that the protonation of His^{68} of ubiquitin elicits a pH-dependent interference of the access of ubiquitin to UIM and UBA. In yeast, the substitution of this residue with alanine has been shown to lead to, albeit weak, sensitivity to cold in yeast growth (42).

The *in vivo* function of the observed pH dependence of UIM and UBA binding of ubiquitin mediated by His^{68} is unclear. Ubiquitin tags serve as a sorting signal for vesicular transport via endosomes, the *trans*-Golgi network, and multivesicular bodies (4, 46). The pH-sensitive change in their binding to downstream effectors or modulators is reminiscent of cargo adaptors that are involved in vesicular transport, such as the KDEL and mannose 6-phosphate receptors, which bind and

dissociate cargo molecules in a pH-dependent manner and thereby deliver them unidirectionally between cellular compartments at various luminal pH values (47). However, there is no evidence that ubiquitin tags move into the acidic lumenal space of vesicles. Alternatively, it may be possible that acidic membrane components affect the protonation state of this histidine in monoubiquitin tags and thus mediate their binding to UIMs of endocytic factors. Many endocytic factors with a UIM also contain ENTH, VHS, or FYVE domains that are located near the UIMs in terms of primary structure (4, 46). These domains bind to acidic lipids, such as phosphoinositides and phosphatidic acid, and thereby bring the adjacent UIMs into proximity with negatively charged membrane components, which could create a low pH environment near the membrane. Proteins, such as bacterial colicin A and pheromone-binding protein from the silk moth, are suggested to undergo conformational transitions at or near negatively charged lipid interfaces (48, 49, 50). These observations raise the intriguing possibility that His^{68} of ubiquitin may serve as a pH sensor to regulate the access of ubiquitin to UIM and UBA of downstream effectors.

His^{68} likely also regulates the conformation of Lys^{48}-linked polyubiquitin chains. The conformation of diubiquitin switches from an open to closed state as the pH increases from 4.5 to 6.8 (51). The crystal structure of diubiquitin indicates that the hydrophobic patch at the UIM interface also serves as the intersubunit interface, which includes the side chain of His^{68} (52). The same surface has been suggested to function as the intersubunit interface in the closed conformation in solution (51). Thus, the protein surface including the hydrophobic patch and His^{68} functions as the intersubunit interface of diubiquitin at higher pH but not at low pH, as occurs in ubiquitin interactions with both UIM and UBA (51). Our pH titration experiment showed that the pH dependences of the $H^{\epsilon 1}$ and $H^{\delta 2}$ chemical shifts of both subunits of diubiquitin are similar to that of monoubiquitin, shown in Fig. 5*b*, and thus the histidine side chains of diubiquitin are mostly protonated at pH 4.5 and mostly deprotonated at pH 6.8.[2] These observations raise the possibility that His^{68} is involved in regulating the intersubunit association, as well as in regulating ubiquitin-UIM and ubiquitin-UBA interactions. Notably, the same surfaces of the distal two ubiquitin subunits of tetraubiquitin function as intersubunit interfaces (51).

This histidine is conserved in the UbL domains of PLIC-2 and Parkin. A triple point mutation in the UbL domain of PLIC-2 (I75A, A77S, and H99A) abolished its binding to the proteasome (18). The mutated residues correspond to Ile^{44}, Ala^{46}, and His^{68} of ubiquitin, all of which are thought to be important for UIM binding. Therefore, the detrimental effect of the triple mutation of PLIC-2 cannot be attributed solely to substitution of the histidine.

UIM Consensus Sequence—The structure of the UIM-UbL complex, along with the binding data from the S5a UIM, UbL, and ubiquitin mutants presented here, enables us to revise the consensus sequence for ubiquitin binding by UIM. On the basis of the sequence alignment of S5a and endocytic factors (7), it was previously proposed to be ϕ-*X*-*X*-Ala-*X*-*X*-*X*-Ser-*X*-*X*-Ac, where ϕ and Ac denote a large hydrophobic residue and an acidic residue, respectively. Our data now show that, whereas Glu^{283}, Ile^{287}, Ala^{290}, and Ser^{294} of S5a contribute to UbL and ubiquitin recognition, Gly^{297}, which is located at the position of the C-terminal Ac in the proposed consensus, makes no contact with UbL and is disordered in the structure of the complex.

Residues shown to be important for UbL binding by our data are highly conserved in the UIMs of endocytic factors that have been either shown or implied to bind a ubiquitin tag (Fig. 1*b*). The most important residues, alanine and serine at positions 290 and 294 in S5a, respectively, are conserved among these factors, and a hydrophobic residue at position 287 is also well conserved. In contrast, the conservation of the acidic residue at position 283 is lower than those in UIM sequences (7). This residue seems to make a smaller contribution than the others, as revealed by our mutagenesis data (Table II and Fig. 4; see "UIM Interface of UbL"). These observations allow us to revise the consensus sequence of UIM to (Ac)-*X*-*X*-*X*-ϕ-*X*-*X*-Ala-*X*-*X*-*X*-Ser-*X*-*X*-(Ac). Fisher *et al.* (40) have also shown that glutamic acid cluster at residues 259–262 and residues Ala^{266} and Ser^{270} of Hrs UIM are important for its affinity for ubiquitin (see Fig. 1*b* for the sequence of Hrs and Ref. 37). They also showed that a mutation of Glu^{273} of Hrs UIM caused a decreased affinity, although the importance of the corresponding residue in S5a UIM, Gly^{297}, is not observed in our structure.

Conclusion—We have reported the structure of a UIM bound to a UbL. Our structural and mutational data indicate that the contact sites in UbL are highly conserved in ubiquitin, but ubiquitin also presents a histidine residue at the interface. Thus, this study provides a structural basis for the interaction between ubiquitin and UIM-containing downstream effectors. Future experiments need to be directed at establishing the functional significance of the observed pH dependence of ubiquitin binding to UIM and UBA.

Acknowledgments—We thank A. Ohno for providing the UBA domain of Dsk2p, H. Kobayashi for providing the expression plasmid for the UBA domain of Dsk2p, and K. Iwai for providing the ubiquitin-activating enzyme (E1). We thank K. Tanaka for discussion.

Addendum—After submission of this manuscript, the solution structures of the UbL of HR23A in complexed with S5a UIM (53) and the Vps27 UIM-ubiquitin complex (54) were published. The structure of HR23A UbL-UIM is similar to that of HR23B-UIM shown in this paper. The possibility of the interference of ubiquitin His^{68} for UIM binding was also discussed based on the structure. In contrast, Swanson *et al.* (54) showed that alanine substitution of His^{68} of ubiquitin impairs the binding of ubiquitin to the N-terminal UIM of Vps27.

[2] T. Tenno, unpublished data.

REFERENCES

1. Conaway, R. C., Brower, C. S., and Conaway, J. W. (2002) *Science* **296**, 1254–1258
2. Sun, Z. W., and Allis, C. D. (2002) *Nature* **418**, 104–108
3. Hoege, C., Pfander, B., Moldovan, G. L., Pyrowolakis, G., and Jentsch, S. (2002) *Nature* **419**, 135–141
4. Hicke, L. (2001) *Cell* **106**, 527–530
5. Hicke, L. (2001) *Nat. Rev. Mol. Cell. Biol.* **2**, 195–201
6. Buchberger, A. (2002) *Trends Cell Biol.* **12**, 216–221
7. Hofmann, K., and Falquet, L. (2001) *Trends Biochem. Sci.* **26**, 347–350
8. Polo, S., Sigismund, S., Faretta, M., Guidi, M., Capua, M. R., Bossi, G., Chen, H., De Camilli, P., and Di Fiore, P. P. (2002) *Nature* **416**, 451–455
9. Katzmann, D. J., Odorizzi, G., and Emr, S. D. (2002) *Nat. Rev. Mol. Cell. Biol.* **3**, 893–905
10. Young, P., Deveraux, Q., Beal, R. E., Pickart, C. M., and Rechsteiner, M. (1998) *J. Biol. Chem.* **273**, 5461–5467
11. Elsasser, S., Gali, R. R., Schwickart, M., Larsen, C. N., Leggett, D. S., Muller, B., Feng, M. T., Tubing, F., Dittmar, G. A., and Finley, D. (2002) *Nat. Cell Biol.* **4**, 725–730
12. van Nocker, S., Sadis, S., Rubin, D. M., Glickman, M., Fu, H., Coux, O., Wefes, I., Finley, D., and Vierstra, R. D. (1996) *Mol. Cell. Biol.* **16**, 6020–6028
13. Lam, Y. A., Lawson, T. G., Velayutham, M., Zweier, J. L., and Pickart, C. M. (2002) *Nature* **416**, 763–767
14. Bilodeau, P. S., Urbanowski, J. L., Winistorfer, S. C., and Piper, R. C. (2002) *Nat. Cell Biol.* **4**, 534–539
15. Shih, S. C., Katzmann, D. J., Schnell, J. D., Sutanto, M., Emr, S. D., and Hicke, L. (2002) *Nat. Cell Biol.* **4**, 389–393
16. Shekhtman, A., and Cowburn, D. (2002) *Biochem. Biophys. Res. Commun.* **296**, 1222–1227
17. Raiborg, C., Bache, K. G., Gillooly, D. J., Madshus, I. H., Stang, E., and Stenmark, H. (2002) *Nat. Cell Biol.* **4**, 394–398
18. Walters, K. J., Kleijnen, M. F., Goh, A. M., Wagner, G., and Howley, P. M. (2002) *Biochemistry* **41**, 1767–1777
19. Sakata, E., Yamaguchi, Y., Kurimoto, E., Kikuchi, J., Yokoyama, S., Yamada, S., Kawahara, H., Yokosawa, H., Hattori, N., Mizuno, Y., Tanaka, K., and Kato, K. (2003) *EMBO Rep.* **4**, 301–306
20. Tashiro, M., Okubo, S., Shimotakahara, S., Hatanaka, H., Yasuda, H., Kainosho, M., Yokoyama, S., and Shindo, H. (2003) *J. Biomol. NMR* **25**, 153–156
21. Schauber, C., Chen, L., Tongaonkar, P., Vega, I., Lambertson, D., Potts, W.,

and Madura, K. (1998) *Nature* **391**, 715–718
22. Hiyama, H., Yokoi, M., Masutani, C., Sugasawa, K., Maekawa, T., Tanaka, K., Hoeijmakers, J. H., and Hanaoka, F. (1999) *J. Biol. Chem.* **274**, 28019–28025
23. Guzder, S. N., Sung, P., Prakash, L., and Prakash, S. (1998) *J. Biol. Chem.* **273**, 31541–31546
24. Gillette, T. G., Huang, W., Russell, S. J., Reed, S. H., Johnston, S. A., and Friedberg, E. C. (2001) *Genes Dev.* **15**, 1528–1539
25. Suzuki, T., Park, H., Kwofie, M. A., and Lennarz, W. J. (2001) *J. Biol. Chem.* **276**, 21601–21607
26. van Laar, T., van der Eb, A. J., and Terleth, C. (2002) *Mutat. Res.* **499**, 53–61
27. Chen, L., and Madura, K. (2002) *Mol. Cell. Biol.* **22**, 4902–4913
28. Beal, R., Deveraux, Q., Xia, G., Rechsteiner, M., and Pickart, C. (1996) *Proc. Natl. Acad. Sci. U. S. A.* **93**, 861–866
29. Withers-Ward, E. S., Mueller, T. D., Chen, I. S., and Feigon, J. (2000) *Biochemistry* **39**, 14103–14112
30. Ryu, K. S., Lee, K. J., Bae, S. H., Kim, B. K., Kim, K. A., and Choi, B. S. (2003) *J. Biol. Chem.* **278**, 36621–36627
31. Cavanagh, J., Fairbrother, W. J., Palmer, A. G., III, and Skelton, N. J. (1996) *Protein NMR Spectroscopy*, Academic Press, San Diego, CA
32. Cornilescu, G., Delaglio, F., and Bax, A. (1999) *J. Biomol. NMR* **13**, 289–302
33. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. D Biol. Crystallogr.* **54**, 905–921
34. Pearlman, D. A., Case, D. A., Caldwell, J. W., Ross, W. S., Cheatham, T. E., III, DeBolt, S., Ferguson, D., Seibel, G., and Kollman, P. (1995) *Comp. Phys. Commun.* **91**, 1–41
35. Xia, B., Tsui, V., Case, D. A., Dyson, H. J., and Wright, P. E. (2002) *J. Biomol. NMR* **22**, 317–331
36. Koradi, R., Billeter, M., and Wuethrich, K. (1996) *J. Mol. Graph.* **14**, 51–55
37. Laskowski, R. A., Rullmann, J. A. C., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) *J. Biomol. NMR* **8**, 477–486
38. Funakoshi, M., Sasaki, T., Nishimoto, T., and Kobayashi, H. (2002) *Proc. Natl. Acad. Sci. U. S. A.* **99**, 745–750
39. Cornilescu, G., Marquardt, J. L., Ottiger, M., and Bax, A. (1998) *J. Am. Chem. Soc.* **120**, 6836–6837
40. Fisher, R. D., Wnag, B., Alam, S. L., Higginson, D. S., Robinson, H., Sundquist, W. I., and Hill, C. P. (2003) *J. Biol. Chem.* **278**, 28976–28984
41. Beal, R. E., Toscano-Cantaffa, D., Young, P., Rechsteiner, M., and Pickart, C. M. (1998) *Biochemistry* **37**, 2925–2934
42. Sloper-Mould, K. E., Jemc, J. C., Pickart, C. M., and Hicke, L. (2001) *J. Biol. Chem.* **276**, 30483–30489
43. Mueller, T. D., and Feigon, J. (2002) *J. Mol. Biol.* **319**, 1243–1255
44. Kang, R. S., Daniels, C. M., Francis, S. A., Shih, S. C., Salerno, W. J., Hicke, L., and Radhakrishnan, I. (2003) *Cell* **113**, 621–630
45. Takahashi, H., Nakanishi, T., Kami, K., Arata, Y., and Shimada, I. (2000) *Nat. Struct. Biol.* **7**, 220–223
46. Johnson, E. S. (2002) *Nat. Cell Biol.* **4**, E295–298
47. Olson, L. J., Zhang, J., Dahms, N. M., and Kim, J. J. (2002) *J. Biol. Chem.* **277**, 10156–10161
48. Merrill, A. R., Steer, B. A., Prentice, G. A., Weller, M. J., and Szabo, A. G. (1997) *Biochemistry* **36**, 6874–6884
49. van der Goot, F. G., Gonzalez-Manas, J. M., Lakey, J. H., and Pattus, F. (1991) *Nature* **354**, 408–410
50. Horst, R., Damberger, F., Luginbuhl, P., Guntert, P., Peng, G., Nikonova, L., Leal, W. S., and Wuthrich, K. (2001) *Proc. Natl. Acad. Sci. U. S. A.* **98**, 14374–14379
51. Varadan, R., Walker, O., Pickart, C., and Fushman, D. (2002) *J. Mol. Biol.* **324**, 637–647
52. Cook, W. J., Jeffrey, L. C., Carson, M., Zhijian, C., and Pickart, C. M. (1992) *J. Biol. Chem.* **267**, 16467–16471
53. Mueller, T., and Feigon, J. (2003) *EMBO J.* **22**, 4634–4645
54. Swanson, K., Kang, R. S., Stamenova, S.D., Hicke, L., and Radhakrishnan, I. (2003) *EMBO J.* **22**, 4579–4606

NUCLEOSIDES, NUCLEOTIDES & NUCLEIC ACIDS
Vol. 23, Nos. 1 & 2, pp. 255–262, 2004

New Sequential-Assignment Routes of Nucleic Acid NMR Signals Using a [5′-^{13}C]-Labeled DNA Dodecamer†

Etsuko Kawashima,[1,*] Takeshi Sekine,[1] Kaoru Umabe,[1] Kazuo Kamaike,[1] Toshimi Mizukoshi,[2] Nobuhisa Shimba,[2] Ei-ichiro Suzuki,[2] and Chojiro Kojima[3]

[1]School of Pharmacy, Tokyo University of Pharmacy and Life Science, Horinouchi, Hachioji, Tokyo, Japan
[2]Central Research Laboratories of Ajinomoto Co. Inc., Suzuki-cho, Kawasaki-ku, Kawasaki-shi, Japan
[3]Graduate School of Biological Science, Nara Institute of Science and Technology, Ikoma, Nara, Japan

ABSTRACT

NMR signal assignments for DNA oligomers have been performed by the well-established sequential assignment procedures based on NOESY and COSY. The H4′/H5′/H5″ resonance region is congested and difficult to analyze without the use of isotope-labeled DNA oligomers. Here a DNA dodecamer constructed with 2′-deoxy[5′-^{13}C]ribonucleotides, 5′-d(*C*G*C*G*A*A*T*T*C*G*CG)-3′ (*N = [5′-^{13}C]Nucleotide), was prepared in an effort to analyze the H4′/H5′/H5″ resonance region by 2D ^{1}H-^{13}C HMQC-NOESY. In the C5′ and H1′ resonance region, weak and strong cross peaks for C5′(i)-H1′(i) and C5′(i)-H1′(i-1), respectively, were found, thus enabling the sequential assignment within this region. A similar sequential assignment route was found between C5′ and H2″. Proton pair distances evaluated from the canonical B-DNA as well as A-DNA indicated that these

†In honor and celebration of the 70th birthday of Professor Leroy B. Townsend.

*Correspondence: Etsuko Kawashima, School of Pharmacy, Tokyo University of Pharmacy and Life Science, Horinouchi, Hachioji, Tokyo, 192-0392 Japan; Fax: +81-426-76-3973; E-mail: kawasima@ps.toyaku.ac.jp.

DOI: 10.1081/NCN-120027832
1525-7770 (Print); 1532-2335 (Online)

sequential-assignment routes on a 2D ^{1}H-^{13}C HMQC-NOESY spectrum work for most nucleic acid stem regions.

Key Words: NMR; [5′-^{13}C]DNA; C5′ resonance assignment; C5′-H1′ sequential assignment; C5′-H2″ sequential assignment; 2D ^{1}H-^{13}C HMQC-NOESY.

INTRODUCTION

An investigation of the conformational diversity of the sugar-phosphate backbone and/or the sugar moieties in nucleic acids is important in the elucidation of nucleic acid-protein or -drug recognition processes. The development of an efficient synthetic method for 2′-deoxy[5′-^{2}H]ribonucleosides (5′*S*:5′*R* = ca. 2:1)[1,2] has fueled expectations that they would facilitate the unambiguous assignment of both the H5′ and H5″ signals of an oligodeoxyribonucleotide and the analysis of sugar-phosphodiester backbone conformations by NMR spectroscopy. Through the use of 2D ^{1}H-^{31}P HSQC spectroscopy on deuterium-labeled nucleotides, Ono and co-workers[3] achieved the assignment of the methylene protons at C5′ of a DNA dodecamer. On the other hand, Kojima and co-workers[4] obtained the 15 ^{3}J coupling constants between H4′ and H5′/H5″ due to the simplified spin systems through the use of NOESY and DQF-COSY NMR analyses. We assumed that more precise analyses of the DNA backbone structure including distance information should be possible by using nucleotides labeled with carbon-13 at the 5′ position, in terms of heteronuclear multidimensional NMR spectroscopy. Conformational analysis by NMR using a [5′-^{13}C]DNA-oligonucleotide, however, has never been reported. Thus, following the development of an efficient method for the synthesis of 2′-deoxy[5′-^{13}C]ribonucleosides,[5,6] we now report on the construction of a [5′-^{13}C]DNA-oligonucleotide and its conformational analysis by HMQC-NOESY NMR. In relation to the C5′ resonance assignment, we found several new sequential-assignment routes on a 2D ^{1}H-^{13}C HMQC-NOESY spectrum. Details of the results obtained are described herein.

RESULTS AND DISCUSSION

The 2′-deoxy[5′-^{13}C]nucleosides[5,6] were converted into the corresponding 3′-phosphoramidite derivatives according to the method of Ono and co-workers.[7–9] These were used for the construction of 5′-d(*^{1}C*^{2}G*^{3}C*^{4}G*^{5}A*^{6}A*^{7}T*^{8}T*^{9}C*^{10}G*^{11}C^{12}G)-3′ by the solid-phase phosphoramidite method[10] using G-CPG at the 3′ terminus. The residue number used here is shown on the left shoulder. Purification of the oligonucleotide was performed according to the method of Kyogoku and co-workers.[11]

All ^{1}H and ^{31}P assignments for this oligomer at 30°C have been reported by Hare and co-workers,[12] Kellogg and Schweitzer,[13] and Ono and co-workers.[3] For the assignment of each C5′ resonance, the reported ^{1}H chemical shifts[3,12,13] provided useful information. Through the use of the ^{1}H-^{13}C HSQC spectrum combined with the ^{1}H assignments, the 5′-^{13}C signals were easily assigned for 7 of the 11 residues: C1, T7, C9, C3, A6, G4 and A5 (Figure 1).

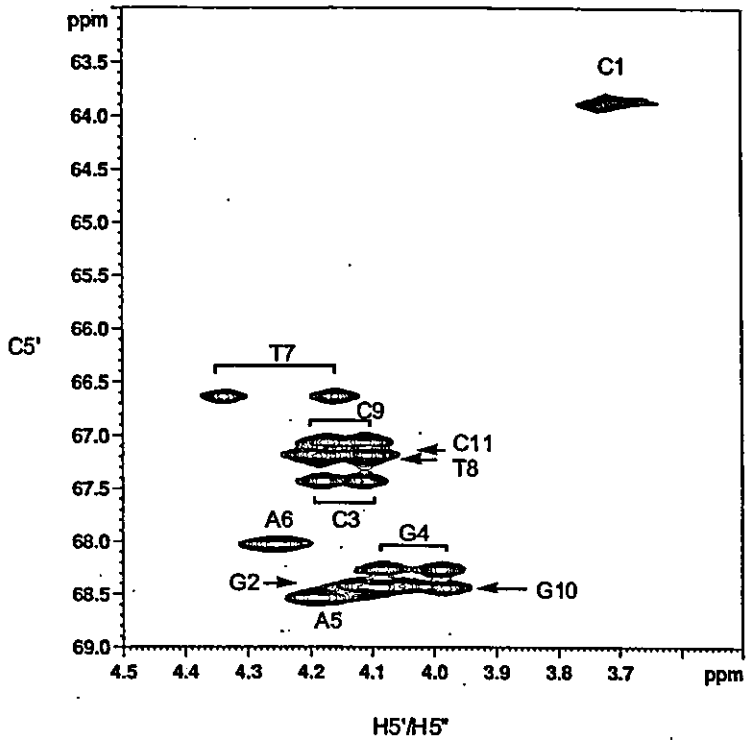

Figure 1. C5′–H5′ spectral region of the ^{1}H-^{13}C HSQC spectrum of the [5′-^{13}C]-labeled DNA dodecamer [5′-d(*^{1}C*^{2}G*^{3}C*^{4}G*^{5}A*^{6}A*^{7}T*^{8}T*^{9}C*^{10}G*^{11}C^{12}G)-3′]. Resonance assignments are given by residue names.

Another spectrum was required to complete the assignments due to the overlapping of the ^{1}H and ^{13}C signals. In Figure 2, the C5′ (w1) and H1′ (w2) region of the ^{1}H-^{13}C HMQC-NOESY spectrum is shown. In this region, the C5′-H1′ sequential connectivity was found, as indicated by the sequential-walk line between the inter-residue H1′(i-1)-C5′(i) (strong) and intra-residue H1′(i)-C5′(i) cross-peaks (weak). Application of this sequential-walk enabled all of the 5′-^{13}C signals to be unambiguously assigned: 64.0 (C1), 66.7 (T7), 67.2 (C9), 67.3 (T8 and C11, overlap), 67.5 (C3), 68.1 (A6), 68.4 (G4), 68.5 (G2 and G10, overlap), and 68.6 (A5), ppm (residue name), respectively. The assignment of each H1′ was also achieved using the sequential-walk line between the inter-residue H1′(i-1)-C5′(i) (strong) and intra-residue H1′(i)-C5′(i) cross-peaks (weak) (Figure 2).

A similar sequential connectivity was found between the C5′ and H2″ signals, although the intensity of the cross peak was much weaker than that of C5′-H1′. Based on the 5′-^{13}C assignments, several intra-residue NOE cross-peaks were assigned: H6/8(i)-C5′(i), H2′(i)-C5′(i) and H3′(i)-C5′(i). In particular, the H6/8(i)-C5′(i) region of the ^{1}H-^{13}C HMQC-NOESY spectrum was sparse, and thus H6 and H8 were easily assigned. The inter- and/or intra-residue H4′-C5′ NOE cross-peaks were also observed, but were not completely assigned due to signal overlapping. These intra-residue NOEs

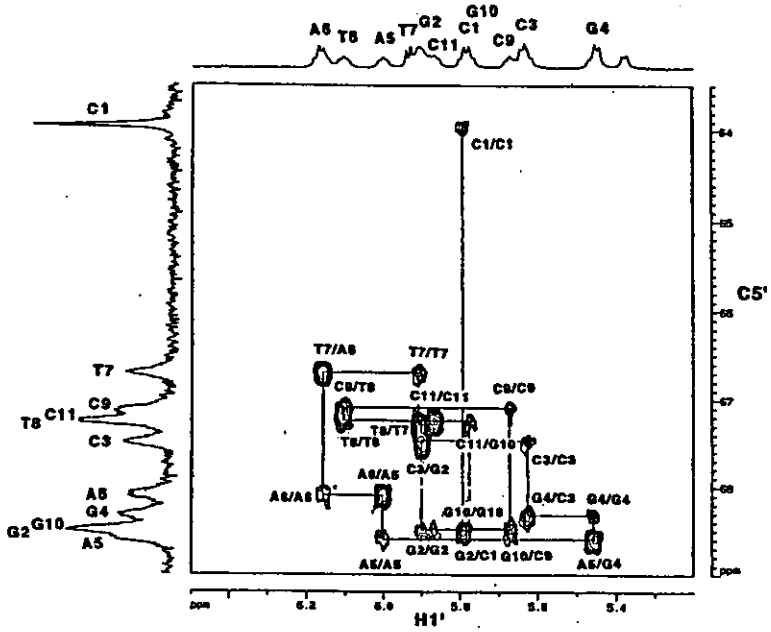

Figure 2. C5′–H1′ spectral region of the 2D ^{1}H-^{13}C HMQC-NOESY spectrum of the [5′-^{13}C]-labeled DNA dodecamer. Resonance assignments are given by the residue names. Lines in the spectrum indicate the sequential assignment routes.

were used to confirm the sequential assignments made by the H1′-C5′ and H2″-C5′ cross-peaks.

The H1′-C5′ and H2″-C5′ sequential-walks found here were analyzed by examining proton-pair distances including H4′/H5′/H5″ as shown in Tables 1 and 2 for canonical B- and A-DNA, respectively. The intra- and inter-residue H1′–H5′, H2″–H5′, H2″–H5″, H3′–H5′, H3′–H5″, H4′–H5′ and H4′–H5″ distances were always shorter than 5 Å. That is, NOEs can be observed for these pairs in both the B- and A-DNA conformations (see Tables 1 and 2). Thus 7 sequential assignment routes were potentially available in the NOESY spectrum. Since the H5′ and H5″ signals could not be distinguished in our 2D HMQC-NOESY, 3 out of 7 routes were degenerate, while 4 routes were available. In fact, two of these, C5′–H1′ and C5′–H2″, were found in our HMQC-NOESY spectrum, though the other two, C5′–H3′ and C5′–H4′, could not identified in our spectrum due to signal overlapping. New sequential assignment routes seem to be powerful for nucleic acids with either the A or B conformation. NOESY-based sequential assignments have been employed even for ^{13}C-enriched DNA and RNA oligomers. This is because sequential assignment through small ^{13}C–^{31}P *J*-couplings does not work well in some cases. Therefore, the NOESY-based sequential assignment routes found here are quite important.

The H4′/H5′/H5″ region (3.9 ~ 4.4 ppm, ^{1}H) in the NOESY spectrum contains ample structural information, as previously pointed out by Wijmenga and co-workers[14] and Kojima and co-workers.[4] Selective ^{13}C enrichment at the 5′ methylene

Table 1. Distances of the proton pairs including H4′/H5′/H5″ below 5 Å in the canonical B-DNA conformation.[a]

	H8/H6	H5	M[#]	H1′	H2′	H2″	H3′	H4′	H5′	H5″
Intra-residue										
H4′ (i)	4.9/4.6	–	–	3.6	3.9	4.1	2.7	†	2.6	2.3
H5′ (i)	3.5/3.3	5.0	~5	4.4	3.7	4.9	3.7	2.6	†	1.8
H5″ (i)	4.4/4.2	–	–	–	3.9	5.0	2.9	2.3	1.8	†
Inter-residue										
H4′ (i + 1)[b]	–/–	–	–	4.2	–	–	–	–	–	–
H4′ (i-1)[b]	–/–	–	–	–	–	–	–	–	3.8	4.1
H5′ (i + 1)[b]	–/–	–	–	1.7	4.3	3.2	4.5	3.8	–	–
H5′ (i-1)[b]	–/–	–	–	–	–	–	–	–	–	–
H5″ (i + 1)[b]	–/–	–	–	3.3	–	4.0	4.7	4.1	–	–
H5″ (i-1)[b]	–/–	–	–	–	–	–	–	–	–	–

[a]Distances in Å.
[b]Inter-residue proton pair has two combinations for each residue, that is, (i-1)th residue to (i)th residue and (i)th to (t + 1)th. (i) refers to the focusing residue. (i-1) and (i + 1) refer to the former and later rows, respectively.
[#]: Methyl proton.
–: Distance longer than 5 Å.
†: Identical proton.

Table 2. Distances of the proton pairs including H4′/H5′/H5″ below 5 Å in the canonical A-DNA conformation.[a]

	H8/H6	H5	M[#]	H1′	H2′	H2″	H3′	H4′	H5′	H5″
Intra-residue										
H4′ (i)	4.3/4.0	–	–	3.3	3.8	2.8	3.0	†	2.5	2.4
H5′ (i)	3.6/3.3	–	–	4.6	–	4.9	3.7	2.5	†	1.8
H5″ (i)	4.2/3.9	–	–	–	–	4.7	3.1	2.4	1.8	†
Inter-residue										
H4′ (i + 1)[b]	–/–	–	–	–	4.3	4.0	–	–	–	–
H4′ (i-1)[b]	–/–	–	–	–	–	–	–	–	4.0	5.0
H5′ (i + 1)[b]	–/–	–	–	3.6	2.7	1.6	4.3	4.0	–	–
H5′ (i-1)[b]	–/–	–	–	–	–	–	–	–	–	–
H5″ (i + 1)[b]	–/–	–	–	–	3.7	3.1	4.8	5.0	–	–
H5″ (i-1)[b]	–/–	–	–	–	–	–	–	–	–	–

[a]Distances in Å.
[b]Inter-residue proton pair has two combinations for each residue, that is, (i-1)th residue to (i)th residue and (i)th to (t + 1)th. (i) refers to the focusing residue. (i-1) and (i + 1) refer to the former and later rows, respectively.
[#]: Methyl proton.
–: Distance longer than 5 Å.
†: Identical proton.

position makes it possible to distinguish the H5′/5″ signals from others, especially from H4′. The unambiguous assignment of the heavily overlapped H4′/H5′/H5″ signals has been the main impediment in the complete assignment of the NOESY cross-peaks for DNA. Therefore, it is important that the 5′-^{13}C enriched DNA sample completely overcomes this overlapping problem. Additionally, the use of two ^{13}C-edited NOESY spectra, where the H5′/5″ signals are selectively picked up or eliminated, can yield the complete NOESY data sets without any overlapping between the H4′ and H5′/5″ resonances.

Using a uniformly ^{13}C-enriched sample, it would not be difficult to distinguish the H5′/5″ signals from those of H4′, since the ^{13}C chemical shifts are different for C4′ (80 ~ 85 ppm) and C5′ (60 ~ 65 ppm). However, the transverse relaxation time of the 5′-^{13}C signals is much shorter than the other carbons, thus the constant-time approach, that can refocus the ^{13}C–^{13}C *J*-couplings, does not work well. As a result, the apparent linewidths of the 5′-^{13}C signals for the uniformly ^{13}C-enriched sample will be much broader, and the new sequential assignment routes reported here might not work well.

CONCLUSION

New sequential assignment routes using 5′-^{13}C signals, C5′–H1′ and C5′–H2″, were found in the 2D ^{1}H-^{13}C HMQC-NOESY spectrum. Using these new routes, all of the 5′-^{13}C and the H1′/H2″ signals were completely and easily assigned. The salient feature of these assignment routes was the use of the 5′-^{13}C signals, rather than H5′ and H5″, to reduce the complexity of the NOESY spectrum. These sequential assignment routes will become a powerful tool in the second or third sequential-walk routes for most nucleic acids, and in the assignment of the H4′/H5′/H5″ signals.

EXPERIMENTAL SECTION

Sample preparation. The 2′-deoxy[5′-^{13}C]nucleoside derivatives (N^6-benzoyl-2′-deoxyadenosine, N^2-acetyl-O^6-diphenylcarbamoyl-2′-deoxyguanosine, N^4-benzoyl-2′-deoxycytidine, and thymidine), reported by Kawashima and co-workers,[5,6] were converted into the corresponding 3′-phosphoramidite derivatives according to the method of Ono and co-workers.[7–9] 5′-d(*^{1}C*^{2}G*^{3}C*^{4}G*^{5}A*^{6}A*^{7}T*^{8}T*^{9}C*^{10}G*^{11}C^{12}G)-3′ was then synthesized on a DNA synthesizer (Applied Biosystems Inc., ABI 391) by the solid-phase phosphoramidite method[10] using G-CPG (purchased from Applied Biosystems, Inc.) at the 3′ terminus. The [5′-^{13}C]DNA-oligonucleotide was subsequently purified after the removal of the protecting group. The purification procedure employed was that previously reported by Kyogoku and co-workers.[11] The residue number used here is shown on the left shoulder. The NMR sample was dissolved in D_2O and kept in a 5 mm tube. The double-strand concentration was estimated to be 1.5 mM from the UV absorbance.

NMR experiments. The two dimensional ^{1}H-^{13}C proton-detected heteronuclear single-quantum correlation spectroscopy (HSQC) spectrum and the ^{1}H-^{13}C HMQC-NOESY[15] spectrum were recorded on a Bruker DMX 600 spectrometer operating at a

600 MHz ^{1}H frequency at 30°C, where the spectral widths were 2000 and 6000 Hz with 256 and 2048 recording points in the ^{13}C (t1) and ^{1}H (t2) dimensions, respectively. The $\pi/4$ shifted sine bell window function was applied and zero-filled to 1024 (^{13}C) and 4096 (^{1}H) points. The ^{13}C-edited $^{1}H-^{1}H$ NOESY by Otting and co-workers[16] was recorded with identical parameters except for the t1 dimension, recorded at 6000 Hz with 512 recording points at the ^{1}H frequency. The pulse repetition delay time was 2 sec, and 80 ~ 112 scans were employed for each t1 increment. The phase sensitive detection in t1 was performed by the TPPI procedure reported by Marion and Wüthrich.[17] The ^{1}H and ^{13}C chemical shifts were calculated relative to the resonance of solvent H_2O (4.68 ppm) for HMQC[18] and from TSP (3-(trimethylsilyl)propionic acid) for the HMQC and NOESY spectra.

ACKNOWLEDGMENTS

The authors gratefully appreciate discussions with Professor Emeritus Yoshiharu Ishido, a Grant-in-aid for Scientific Research (C) (No. 14572012) from the Ministry of Education, Science, Culture, and Sports, and the Hayashi Memorial Foundation for Female Natural Science.

REFERENCES

1. Kawashima, E.; Toyama, K.; Ohshima, K.; Kainosho, M.; Kyogoku, Y.; Ishido, Y. Novel synthesis of 2′-deoxy[5′-^{2}H]ribonucleoside derivatives from 5′-*O*-Ac-2′-deoxy-5′-PhSe-ribonucleoside derivatives. Tetrahedron Lett. **1995**, *36*, 6699–6700.
2. Kawashima, E.; Toyama, K.; Ohshima, K.; Kainosho, M.; Kyogoku, Y.; Ishido, Y. A novel approach to diastereoselective synthesis of 2′-deoxy[5′-$^{2}H_1$]ribonucleoside derivatives by reduction of the corresponding 5′-*O*-Acetyl-2′-deoxy-5′-phenylselenoribonucleoside derivatives with a $Bu_3Sn^{2}H$—Et_3B system. Chirality **1997**, *9*, 435–442.
3. Ono, A.; Makita, T.; Tate, S.; Kawashima, E.; Ishido, Y.; Kainosho, M. C5′ methylene proton signal assignment of DNA/RNA oligomers labeled with C5′-monodeuterated nucleosides by $^{1}H-^{31}P$ HSQC spectroscopy. Magn. Reson. Chem. **1996**, *34*, S40–S46.
4. Kojima, C.; Kawashima, E.; Toyama, K.; Ohshima, K.; Ishido, Y.; Kainosho, M.; Kyogoku, Y. Stereospecific assignment of H5′ and H5″ in a (5′*R*)-/(5′*S*)-deuterium-labeled DNA decamer for $^{3}J_{HH}$ determination and unambiguous NOE assignments. J. Biomol. NMR **1998**, *11*, 103–109.
5. Kawashima, E.; Umabe, K.; Sekine, T. Synthesis of [5′-^{13}C]ribonucleosides and 2′-deoxy[5′-^{13}C]ribonucleosides. J. Org. Chem. **2002**, *67*, 5142–5151.
6. Sekine, T.; Kawashima, E.; Ishido, Y. Efficient synthesis of D-[5-^{13}C]ribose from D-ribose and its conversion into [5′-^{13}C]nucleosides. Tetrahedron Lett. **1996**, *37*, 7757–7760.
7. Ono, A.; Ts'O, P.O.P.; Kan, I.S. Triplex formation of oligonucleotides containing 2′-*O*-methylpseudoisocytidine in substitution for 2′-deoxycytidine. J. Am. Chem. Soc. **1991**, *113*, 4032–4033.

8. Ono, A.; Ts'O, P.O.P.; Kan, I.S. Triplex formation of an oligonucleotide containing 2'-*O*-methylpseudoisocytidine with a DNA duplex at neutral pH. J. Org. Chem. **1992**, *57*, 3225–3230.
9. Adams, S.P.; Kavka, K.S.; Wykes, E.J.; Holder, S.B.; Galluppi, G.R. Hindered dialkylamino nucleoside phosphite reagents in the synthesis of two DNA 51-mers. J. Am. Chem. Soc. **1983**, *105*, 661–663.
10. Beaucage, S.L.; Caruthers, M.H. Deoxynucleoside phosphoramidites. A new class of key intermediates for deoxypolynucleotide synthesis. Tetrahedron Lett. **1981**, *22*, 1859–1862.
11. Kyogoku, Y.; Kojima, C.; Lee, S.J.; Tochio, H.; Suzuki, N.; Matsuo, H.; Shirakawa, M. Induced structural changes in protein-DNA complexes. Methods Enzymol. **1995**, *261*, 524–541.
12. Hare, D.R.; Wemmer, D.E.; Chou, S.H.; Drobny, G.; Reid, B.R. Assignment of the non-exchangeable proton resonances of d(C-G-C-G-A-A-T-T-C-G-C-G) using two-dimensional nuclear magnetic resonance methods. J. Mol. Biol. **1983**, *171*, 319–336.
13. Kellogg, G.W.; Schweitzer, B.I. Two-dimensional ^{31}P-driven NMR procedures for complete assignment of backbone resonances in oligodeoxyribonucleotides. J. Biomol. NMR **1993**, *3*, 577–593.
14. Wijmenga, S.S.; Mooren, M.M.W.; Hilbers, C.W. NMR of nucleic acids, from spectrum to structure. In *NMR in Macromolecules*; Roberts, G.C., Ed.; IRL Press: Oxford, 1995; 217–288.
15. Fesik, S.W.; Zuiderweg, E.R.P. Heteronuclear three-dimensional NMR spectroscopy. A strategy for the simplification of homonuclear two-dimensional NMR spectra. J. Magn. Reson. **1988**, *78*, 588–593.
16. Otting, G.; Senn, H.; Wagner, G.; Wüthrich, K. Editing of 2D ^{1}H NMR spectra using X half-filters. Combined use with residue-selective ^{15}N labeling of proteins. J. Magn. Reson. **1986**, *70*, 500–505.
17. Marion, D.; Wüthrich, K. Application of phase sensitive two-dimensional correlated spectroscopy (COSY) for measurements of ^{1}H–^{1}H spin–spin coupling constants in proteins. Biochem. Biophys. Res. Commun. **1983**, *113*, 967–974.
18. Wishart, D.S.; Bigam, C.G.; Yao, J.; Abildgaard, F.; Dyson, H.J.; Oldfield, E.; Markley, J.L.; Sykes, B.D. ^{1}H, ^{13}C and ^{15}N chemical shift referencing in biomolecular NMR. J. Biomol. NMR **1995**, *6*, 135–140.

Received August 26, 2003
Accepted October 14, 2003

Paramagnetic NMR study of Cu^{2+}–IDA complex localization on a protein surface and its application to elucidate long distance information

Makoto Nomura[a], Toshitatsu Kobayashi[a], Toshiyuki Kohno[b], Kenichiro Fujiwara[c], Takeshi Tenno[d], Masahiro Shirakawa[c], Itsuko Ishizaki[a], Kazuo Yamamoto[e], Toshifumi Matsuyama[e], Masaki Mishima[a], Chojiro Kojima[a,*]

[a] *Laboratory of Biophysics, Graduate School of Biological Science, Nara Institute of Science and Technology, 8916-5 Takayama, Ikoma, Nara 630-0192, Japan*
[b] *Mitsubishi Kagaku Institute of Life Sciences (MITILS), Machida, Tokyo 194-8511, Japan*
[c] *Graduate School of Integrated Science, Yokohama City University, Tsurumi, Yokohama 230-0054, Japan*
[d] *Graduate School of Science and Engineering, Ehime University, Matsuyama, Ehime 790-8577, Japan*
[e] *Division of Cytokine Signaling, Department of Molecular Microbiology and Immunology, Nagasaki University Graduate School of Biomedical Sciences, 1-12-4 Sakamoto, Nagasaki 852-8523, Japan*

Received 10 March 2004; revised 12 April 2004; accepted 13 April 2004

Available online 27 April 2004

Edited by Thomas L. James

Abstract **The paramagnetic metal chelate complex Cu^{2+}-iminodiacetic acid (Cu^{2+}–IDA) was mixed with ubiquitin, a small globular protein. Quantitative analyses of ^{1}H and ^{15}N chemical shift changes and line broadenings induced by the paramagnetic effects indicated that Cu^{2+}–IDA was localized to a histidine residue (His68) on the ubiquitin surface. The distances between the backbone amide proton and the Cu^{2+} relaxation center were evaluated from the proton transverse relaxation rates enhanced by the paramagnetic effect. These correlated well with the distances calculated from the crystal structure up to 20 Å. Here, we show that a Cu^{2+}–IDA is the first paramagnetic reagent that specifically localizes to a histidine residue on the protein surface and gives the long-range distance information.**

Keywords: NMR; Surface histidine; Cu^{2+}–IDA; Paramagnetic relaxation; Pseudo-contact shift; Non-metal protein

1. Introduction

In the field of biological NMR, long-range distance constraints are required for large-size multi-domain proteins and high-throughput structure determinations. Paramagnetic metals have useful properties, such as paramagnetic relaxation enhancement and pseudo-contact shift, both of which affect the NMR signals over long-range distances. For metal proteins and metal–drug–DNA complexes, these paramagnetic effects have been successfully utilized as long-range constraints in structure determinations [1–7].

Kay and co-workers [8,9] developed a method to evaluate the distance between the backbone amide proton and Cu^{2+} using the proton transverse relaxation rate enhanced by the paramagnetic effect for non-metal proteins. The method relies on localizing the paramagnetic ion to a modified target protein such as possessing a specially designed ATCUN motif tag [8,9]. Dvoretsky et al. [10] used the amide proton longitudinal relaxation rates for Mn^{2+} that was localized by mutating an amino acid of the target protein to cysteine followed by covalent modification of this residue using thiol-reactive EDTA. Both methods rely on mutation of the target protein to localize the paramagnetic metal, a procedure that may result in structural changes of the target protein.

Cu^{2+}-iminodiacetic acid (Cu^{2+}–IDA), a metal chelate complex, is an established reagent used in the area of immobilized metal ion affinity chromatography (IMAC) [11]. Immobilized Cu^{2+}–IDA on a solid support specifically binds to histidine residues ($K_d = 10^{-5}$–10^{-4} M) [11–14], and has been used to purify various protein and DNA molecules [11,15]. Other applications involving the use of Cu^{2+}–IDA IMAC have dealt with investigations concerning the structure–function relationship of proteins, which have functional histidine residues [16,17]. Although there are many applications, the precise localization of Cu^{2+}–IDA on the protein surface has not been well characterized [11,12].

Here, the localization of Cu^{2+}–IDA on three different protein surfaces was examined using chemical shift changes, line broadenings and ^{1}H transverse relaxation rate enhancements induced by the paramagnetic effects of Cu^{2+}. The proteins examined were ubiquitin, which has one histidine residue (His68) on the surface, and two other proteins that contain a C-terminal poly histidine tag, the ubiquitin-like domain of human HR23B (HR23B UbL) and the DNA binding domain of human interferon regulatory factor 4 (IRF4 DBD). The potential acquisition of long-range distance constraints was also examined.

2. Materials and methods

2.1. Sample preparation

Recombinant human ubiquitin was expressed in *Escherichia coli* BL21(DE3) as a GST-fusion protein. Following the addition of isopropyl β-D-thiogalactoside, protein expression was induced for 8 h at 30 °C in minimal medium containing ^{15}N-labeled ammonium chloride. Cells were suspended in 50 mM Tris–HCl (pH 8.0), 100 mM KCl and 1 mM Pefabloc SC (Roche). The suspension was lysed by sonication,

* Corresponding author. Fax: +81-743-72-5579.
E-mail address: kojima@bs.aist-nara.ac.jp (C. Kojima).

0014-5793/$22.00
doi:10.1016/j.febslet.2004.04.023

ultra-centrifuged, and the supernatant was loaded onto a glutathione-Sepharose 4B column (Amersham). Elution was performed with 50 mM Tris–HCl (pH 8.0), 100 mM KCl and 30 mM glutathione. The GST–ubiquitin fraction was loaded onto a Superdex 26/60 75 pg (Amersham) gel filtration column equilibrated with 50 mM phosphate (pH 6.5), 50 mM KCl and 1 mM Pefabloc SC. The eluted GST–ubiquitin fraction was then cleaved with PreScission protease (Amersham) and loaded onto a gel filtration column. The final NMR buffer consisted of 50 mM phosphate (pH 6.7) and 200 mM KCl. The expression and purification of HR23B UbL and IRF4 DBD was performed as previously described [18].

2.2. NMR spectroscopy

The stock solution of Cu^{2+}–IDA consisted of 12.5 mM $CuCl_2$, 13.8 mM IDA, 50 mM phosphate and 200 mM KCl (pH 2.2). Following the addition of this stock solution to 0.1 or 1 mM ubiquitin, 0.1 mM HR23B UbL, and 0.5 mM IRF4 DBD, the pH of the sample solution was adjusted to pH 6.7. The ratio of Cu^{2+}–IDA/protein was 0.4 for each unless otherwise noted.

All NMR data were obtained on a Bruker DRX800 spectrometer operating at 800 MHz with a triple-resonance probe head equipped with a XYZ-gradient unit. The ^1H–^{15}N HSQC spectra were collected at 293 K, with 2048 (t2) times 150 (t1) complex points and spectral widths of 14 kHz (^1H) and 1.6 kHz (^{15}N) with carrier positions at 4.7 and 121.8 ppm, respectively. A $\pi/2$ shifted sine-bell window function was applied before zero-filling and Fourier transformation. The processing was performed using the NMRPipe package [19], while chemical shifts, linewidths and peak volumes were measured using Sparky software [20]. The transverse relaxation rates were obtained using a ^1H–^{15}N HSQC based spin echo experiment with a shaped refocusing pulse which was successfully used to obtain long-range distance information with the ATCUN Cu^{2+} system [8]. The relaxation delay times were 8.5, 11, 13.5, 16, 18.5, 21, 23.5, 28.5, 31, 33.5, 38.5, and 48.5 ms for Cu^{2+}–IDA–ubiquitin, and 8.5, 11, 13.5, 16, 18.5, 21, 23.5, 28.5, 33.5, 38.5, 48.5, and 68.5 ms for free ubiquitin.

2.3. Data analyses

Linewidth and chemical shift change. Linewidths were measured for 1 mM ubiquitin samples with and without 0.4 mM Cu^{2+}–IDA. The differences in linewidth were then calculated. When the Cu^{2+}–IDA–ubiquitin peak was not found compared to free ubiquitin ($S/N < 5$), the difference in linewidth was set to the maximum value of 50 Hz. Chemical shift differences of samples with and without Cu^{2+}–IDA were calculated and averaged for three independent experiments. The differences in linewidth and chemical shift values were mapped to the PDB (1UBI).

Paramagnetic ^1H relaxation enhancements. Three independent relaxation rate measurements of the 1 mM ubiquitin sample complexed with 0.4 mM Cu^{2+}–IDA were performed. The differences between the transverse relaxation rates with and without Cu^{2+}–IDA were calculated and averaged. The mean values of three experiments were fitted to the relation $R_{2M} = a/r^6 + b$ by non-linear fitting using Mathematica 4.2 (Wolfram) with error bar dependent weights, where R_{2M} is the transverse relaxation rate enhanced by the paramagnetic effects, r is the distance between the target proton and the relaxation center, and a and b are constants. The amide proton position was obtained from the PDB (1UBI) for the residue with the generalized order parameter $S^2 > 0.7$ [21] and the relaxation center position (assumed as Cu^{2+} atom position) was optimized.

3. Results

3.1. Localization and interaction of Cu^{2+}–IDA with ubiquitin

^1H–^{15}N HSQC experiments were performed in an effort to evaluate the interaction of Cu^{2+}–IDA with ubiquitin. Titration experiments using Cu^{2+}–IDA with 0.1 and 1 mM ubiquitin solutions resulted in dose-dependent and site-specific line broadenings and shifts. The localization of Cu^{2+}–IDA on the protein was not due to the presence of ionic interactions, but due to some type of coordination to the protein [16,22], since the addition of 200 mM KCl had no effect. The linewidth differences between Cu^{2+}–IDA–ubiquitin and free ubiquitin were mapped to the 3D structure of ubiquitin as shown in Fig. 1 (left), and strongly suggest that Cu^{2+}–IDA bound to ubiquitin around His68. The chemical shift differences up to 30 ppb (25 Hz) were also mapped (Fig. 1, right), supporting the presence of a specific binding site close to His68. At pH 5.0, these line broadenings and shifts were not observed even though the Cu^{2+}–IDA is stable at this pH [16,22]. This result is consistent with the notion that the imidazole nitrogen of histidine ($pK_a = 6.04$) acts as the potential electron donor for Cu^{2+}–IDA [16].

Cu^{2+} has four coordination sites. IDA can coordinate to three of these, while the protein coordinates to the remaining site [23]. Other metals such as Fe^{3+}, Co^{2+}, and Ni^{2+} can have four coordination sites, which could be chelated by IDA. ^1H–^{15}N HSQC spectra for these paramagnetic metal–IDA–ubiquitin complexes showed no significant line broadenings and chemical shift differences for ^1H, however small chemical shift differences (less than one-tenth of those for Cu^{2+}) were observed for ^{15}N around His68. This result is consistent with the notion that the localization of the paramagnetic metal–IDA complex is around His68.

3.2. Paramagnetic relaxation enhancement analysis

Transverse relaxation rates were measured using spin echo experiments based on ^1H–^{15}N HSQC spectra in order to obtain long-range distance information. Distances from Cu^{2+} to each amide proton were calculated using the crystal structure (PDB ID, 1UBI) and transverse relaxation rates, and were designated as the calculated and experimental distances, respectively. Using 49 out of 71 amide proton relaxation rates, the Cu^{2+} position was optimized using the Mathematica software to minimize the difference between those distances. The distances from the optimized Cu^{2+} position to N^δ and N^ϵ of the His68 imidazole were 5.6 ± 0.6 and 4.5 ± 1.0 Å, respectively (Fig. 2). The B-factors for N^δ and N^ϵ were 27 and 28

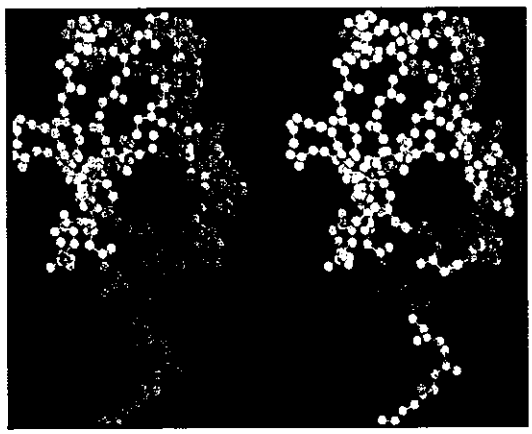

Fig. 1. Line broadenings (left) and peak shifts (right) induced by the Cu^{2+}–IDA complex mapped on the ubiquitin structure. With increasing broadenings and shifts, the color changes from blue to red. Maximum, medium and minimum values (50, 25, and 0 Hz for broadenings, and 30, 15, and 0 ppb for shifts) are colored in red, green and blue, respectively. Intermediate colors such as yellow (between green and red) are given gradually. In the peak shifts (right), white represents disappeared peaks. These pictures were drawn by RASMOL 2.6 [29].

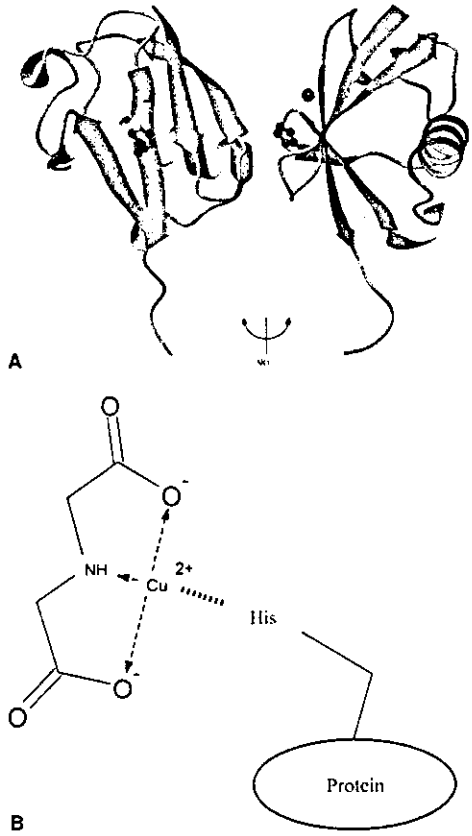

Fig. 2. The ribbon structure of ubiquitin with Cu^{2+} (A), and the chemical structure of IDA and its coordination to Cu^{2+} (B). Side chain of His68 is drawn as a ball-stick model with the two nitrogen atoms colored in green. Cu^{2+} is shown as the gold sphere. The atomic coordinate of Cu^{2+} is optimized, minimizing the calculated and observed proton transverse relaxation rates. The ribbon pictures were prepared by MOLMOL 2k.1 [30].

Å^2, respectively, and much larger than those for the backbone atoms of His68 (2–6 Å^2). This indicates that the imidazole ring is flexible and that those coordinates are less reliable. Additionally, the delocalization of the paramagnetic electron might modulate the relaxation center position [2]. Thus, it is not clear from the distance measurements if His68 does indeed bind the Cu^{2+}–IDA.

The correlation plot between the calculated and experimental distances is shown in Fig. 3. The correlation coefficient R was 0.75. The observed transverse relaxation enhancement was up to 20 Å and the mean error from the calculated values was approximately 3.6 Å. This result highlights the potential in acquiring long-range distance information using the Cu^{2+}–IDA complex, although the precision was not very high.

Relaxation experiments were repeated on samples containing a change in Cu^{2+}–IDA concentration from 0.4 to 0.2 or 0.8 mM. The relaxation data determined were fitted to the theoretical equation and were well converged. The converged coordinates of Cu^{2+} were quite similar to each other for the three data sets (RMSD = 1.3 Å). The experimental distances were correlated with the calculated ones, where $R = 0.67$ and 0.80 for the 0.2 and 0.8 mM samples, respectively.

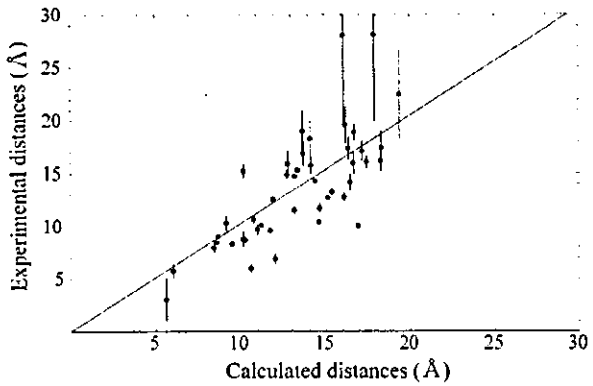

Fig. 3. Correlation plot between calculated and experimental distances from Cu^{2+} to each backbone amide proton of ubiquitin. Calculated distances were taken from the crystal structure. The experimental distances were obtained from the paramagnetic enhancement of the proton transverse relaxation rates.

4. Discussion

The binding between ubiquitin and Cu^{2+}–IDA was evaluated further by EPR. EPR spectra were recorded by a JEOL JES-TE300 X-band spectrometer at 77 K using 2.5 mM Cu^{2+}–IDA. The spectra significantly changed in the presence and absence of ubiquitin ($g_{\|} = 2.27$ and 2.32, $A_{\|} = 15.8$ and 15.9, and $g_{\perp} = 2.06$ and 2.07, respectively). In the ubiquitin titration series, Cu^{2+}–IDA spectra were explained by the weighted sum of only two spectra, i.e., free and 1:1 complex. EPR splitting patterns of Cu^{2+}–IDA complexed with a histidine were quite similar to ubiquitin complex. These data strongly suggest that Cu^{2+}–IDA coordinates to ubiquitin at one specific site, His68.

The localization of Cu^{2+}–IDA on two other proteins was investigated to reveal whether Cu^{2+}–IDA specifically binds to the surface histidine residue. The ^1H–^{15}N HSQC spectra of HR23B UbL and IRF4 DBD with and without Cu^{2+}–IDA clearly demonstrated the site-specific line broadenings and shifts around the histidine residues. Broadening ratios for all proteins examined here are given in Fig. 4. Distance information was not elucidated for HR23B UbL and IRF4 DBD, since the coordinates were not available for the free proteins. A cysteine residue is the potential binding site of Cu^{2+}–IDA, but not examined in this study because of two difficulties. First, Cu^{2+}–IDA complex cannot be used with dithiothreitol (DTT) since a Cu^{1+}–DTT complex is formed [24]. Second, a free cysteine residue decreases protein stability, and the mixture of the Cu^{2+}–IDA and a cysteine forms aggregates. Thus, *N*-ethylmaleimide was used to chemically modify the surface cysteine residue of IRF4 DBD in our experiments. Other paramagnetic metal chelate complexes, such as Mn^{2+}–NTA and Gd^{3+}–EDTA, did not bind specifically to the histidine residues for three proteins possessing a poly hisitidine tag: HR23B UbL, IRF4 DBD and the S5a UbL binding region [18]. These results are consistent with hard-soft-acid-base (HSAB) theory [25–27].

Characteristic properties of the Cu^{2+}–IDA complex, that includes highly specific localization to the histidine residue and the capability of obtaining distance information, offer clear advantages in the analysis of macromolecular intermolecular interactions such as protein–protein, protein–DNA and

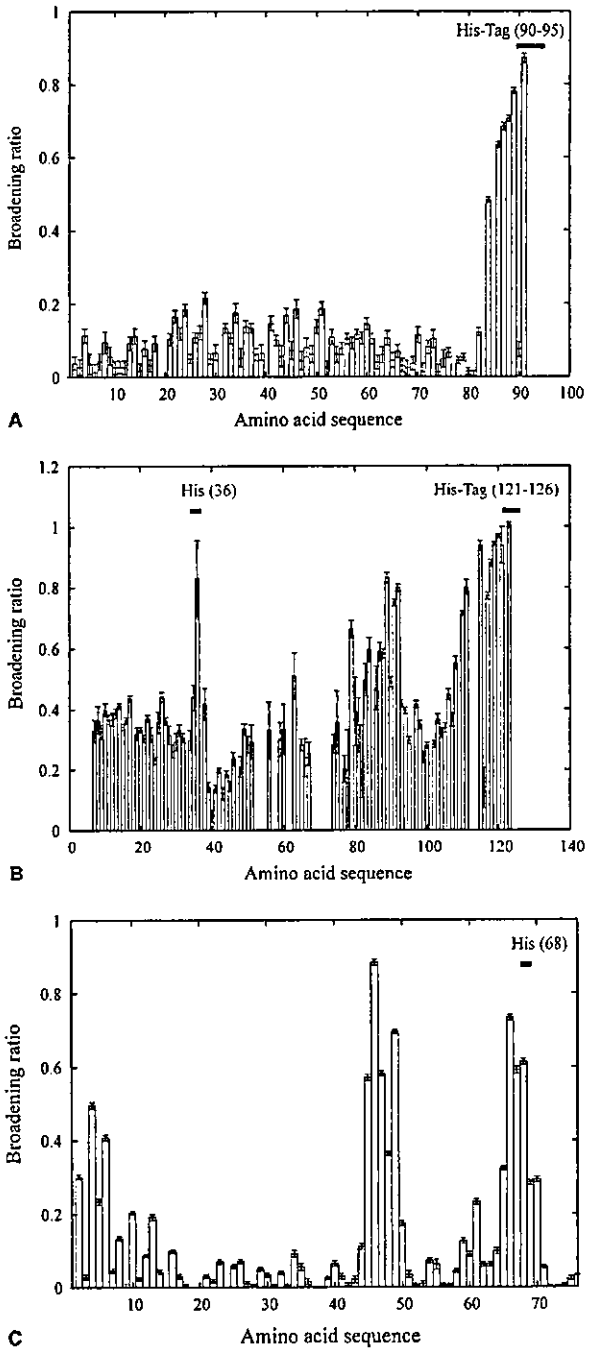

Fig. 4. The broadening ratios of HR23B UbL (A), IRF4 DBD (B), and ubiquitin (C) were plotted against the amino acid sequence. The broadening ratio is defined as $\sqrt{(I-I_0)^2}/\sqrt{I_0^2}$, where I and I_0 represent the peak height with and without Cu^{2+}–IDA (1 equiv.), respectively. The 79th to 92th residues of IRF4 DBD were spatially close to His36 in the modeled structure.

protein–ligand interactions. Hansen et al. [28] employed the paramagnetic longitudinal proton relaxation enhancement of a metal protein to detect transient intermolecular protein–protein interactions. Mal et al. [9] utilized the ATCUN Cu^{2+} system to detect protein–protein interactions. Therefore, transient interactions showing few intermolecular NOEs can be the subject of investigation using Cu^{2+}–IDA provided that a surface histidine residue is located near the interface.

Another potential application of the Cu^{2+}–IDA complex can be in the investigation of membrane proteins. Membrane proteins are usually solubilized within a detergent micelle and it is difficult to construct an expression system. Distance information in general is insufficient,and any specific modification appears daunting,to say the least. We applied the Cu^{2+}–IDA system to a membrane protein solubilized within a detergent,and the ^{1}H–^{15}N HSQC spectrum showed site-specific line broadenings and shifts (Okuda et al., unpublished results). These results indicate the potential application of the Cu^{2+}–IDA system in the study of membrane protein structures.

In conclusion, all examined paramagnetic metal chelate complexes were not localized to a specific site on the protein surface with exception of Cu^{2+}–IDA,that was specifically localized to His68 on the ubiquitin surface. This feature of Cu^{2+}–IDA was successfully utilized to elucidate long-range distance information from proton transverse relaxation enhancements. Cu^{2+}–IDA could also be used in conjunction with poly histidine tagged proteins,but it may be difficult to obtain precise distance constraints for proteins possessing many histidine residues. Application of the Cu^{2+}–IDA method to the investigation of surface histidine residues is a technique worthy of consideration, since surface histidine residues are frequently found in many proteins and can play important roles in many biochemical reactions and interactions.

Acknowledgements: We thank H. Okuda for his assistance, Prof. J.L. Markley and Prof. C. Griesinger for their critical comments, and Prof. S. Suzuki for his helpful comments on EPR analyses. This work was supported in part by Grants-in-Aid for 21st Century COE Research and Scientific Research, a Grant for the National Project on Protein Structural and Functional Analyses from MEXT (the Japanese Ministry of Education, Culture, Sports, Science and Technology), a Kaneko Narita Research Grant of Protein Research Foundation and a Grant from the Yamada Science Foundation.

References

[1] Bertini, I., Donaire, A., Jimenez, B., Luchinat, C., Parigi, G., Piccioli, M. and Poggi, L. (2001) J. Biomol. NMR 21, 85–98.
[2] Bertini, I., Luchinat, C. and Parigi, G. (2001) Solution NMR of Paramagnetic Molecules. Elsevier, Amsterdam.
[3] Bertini, I., Luchinat, C. and Piccioli, M. (2001) Methods Enzymol. 339, 314–340.
[4] Gochin, M. (1997) J. Am. Chem. Soc. 119, 3377–3378.
[5] Gochin, M. (1998) J. Biomol. NMR 12, 243–257.
[6] Kikuchi, J., Iwahara, J., Kigawa, T., Murakami, Y., Okazaki, T. and Yokoyama, S. (2002) J. Biomol. NMR 22, 333–347.
[7] Arnesano, F., Banci, L., Bertini, I., Felli, I.C., Luchinat, C. and Thompsett, A.R. (2003) J. Am. Chem. Soc. 125, 7200–7208.
[8] Donaldson, L.W., Skrynnikov, N.R., Choy, W.Y., Muhandiram, D.R., Sarkar, B., Forman-Kay, J.D. and Kay, L.E. (2001) J. Am. Chem. Soc. 123, 9843–9847.
[9] Mal, T.K., Ikura, M. and Kay, L.E. (2002) J. Am. Chem. Soc. 124, 14002–14003.
[10] Dvoretsky, A., Gaponenko, V. and Rosevear, P.R. (2002) FEBS Lett. 528, 189–192.
[11] Gaberc-Porekar, V. and Menart, V. (2001) J. Biochem. Biophys. Methods 49, 335–360.
[12] Yip, T.T., Nakagawa, Y. and Porath, J. (1989) Anal. Biochem. 183, 159–171.
[13] Hutchens, T.W. and Yip, T.T. (1990) Anal. Biochem. 191, 160–168.
[14] Mrabet, N.T. (1992) Biochemistry 31, 2690–2702.

[15] Murphy, J.C., Jewell, D.L., White, K.I., Fox, G.E. and Willson, R.C. (2003) Biotechnol. Progr. 19, 982–986.
[16] Berna, P.P., Mrabet, N.T., Van Beeumen, J., Devreese, B., Porath, J. and Vijayalakshmi, M.A. (1997) Biochemistry 36, 6896–6905.
[17] Jiang, K.Y., Pitiot, O., Anissimova, M., Adenier, H. and Vijayalakshmi, M.A. (1999) Biochim. Biophys. Acta 1433, 198–209.
[18] Fujiwara, K., Tenno, T., Sugasawa, K., Jee, J.G., Ohki, I., Kojima, C., Tochio, H., Hiroaki, H., Hanaoka, F. and Shirakawa, M. (2004) J. Biol. Chem. 279, 4760–4767.
[19] Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J. and Bax, A. (1995) J. Biomol. NMR 6, 277–293.
[20] Goddard, T.D. and Kneller, D.G. (1999) Sparky 3. University of California, San Francisco.
[21] Schneider, D.M., Dellwo, M.J. and Wand, A.J. (1992) Biochemistry 31, 3645–3652.
[22] Sulkowski, E. (1987) in: Protein Purification: Micro to Macro (Burgess, R.R., Ed.), pp. 149–162, Alan R. Liss Inc, New York.
[23] Hochuli, E., Dobeli, H. and Schacher, A. (1987) J. Chromatogr. 411, 177–184.
[24] Kr zel, A., Lesniak, W., Jezowska-Bojczuk, M., Mlynarz, P., Brasun, J., Kozlowski, H. and Bal, W. (2001) J. Inorg. Biochem. 84, 77–88.
[25] Aime, S., D'Amelio, N., Fragai, M., Lee, Y.M., Luchinat, C., Terreno, E. and Valensin, G. (2002) J. Biol. Inorg. Chem. 7, 617–622.
[26] Kemple, M.D., Ray, B.D., Lipkowitz, K.B., Prendergast, F.G. and Nageswara Rao, B.D. (1988) J. Am. Chem. Soc. 110, 8275–8787.
[27] Petros, A.M., Mueller, L. and Kopple, K.D. (1990) Biochemistry 29, 10041–10048.
[28] Hansen, D.F., Hass, M.A., Christensen, H.M., Ulstrup, J. and Led, J.J. (2003) J. Am. Chem. Soc. 125, 6858–6859.
[29] Sayle, R.A. and Milner-White, E.J. (1995) Trends Biochem. Sci. 20, 374.
[30] Koradi, R., Billeter, M. and Wuthrich, K. (1996) J. Mol. Graph. 14, 51–55, 29–32.

Nucleic Acids Symposium Series No. 48 105-106

Solution NMR study of DNA recognition mechanism of IRF4 protein

Itsuko Ishizaki[1], Makoto Nomura[1], Kazuo Yamamoto[2], Toshifumi Matsuyama[2], Masaki Mishima[1] and Chojiro Kojima[1]

[1]Laboratory of Biophysics, Graduate School of Biological Science, Nara Institute of Science and Technology, 8916-5 Takayama, Ikoma, Nara 630-0192, Japan and [2]Division of Cytokine Signaling, Department of Molecular Microbiology and Immunology, Nagasaki University Graduate School of Biomedical Science, 1-12-4 Sakamoto, Nagasaki 852-8523, Japan

ABSTRACT

Transcription factor IRF-4 prefers the DNA sequence including CCGAAA. The consensus sequence of the IRF family proteins is NNGAAA, and all crystal structures indicate the NN region does not interact with IRF proteins directly. Here the sequence preference of IRF-4 was investigated by NMR and fluorescence anisotropy as an example of the indirect sequence recognition. The ^{1}H-^{15}N HSQC spectra of the IRF-4/DNA complex containing the CCGAAA sequence indicated that the 1:1 complex was formed. The dissociation constants (K_d) for two DNA oligomers containing CCGAAA and GGGAAA were determined by fluorescence anisotropy, but their difference was very small.

INTRODUCTION

IRF-4 is a lymphoid and myeloid restricted member of the IRF transcriptional family.[1,2] PU.1/IRF-4/DNA ternary complex has been cocrystallized and analyzed biochemically,[3] however, PU.1 is not related to many physiological functions of IRF-4.[2,4-7] Recognition sequences of IRF-4 have been determined (Yoshida *et al.*, personal communication) using the selected and amplified binding sequences (SAAB) method.[8] Most of them contain the CCGAAA sequences, though the consensus sequence of IRF family proteins is known as NNGAAA. In the crystal structures of IRF/DNA complexes including the PU.1/IRF-4/DNA ternary complex (GTGAAA), NN sequence does not have direct contacts with IRF family proteins. In this study, we have investigated the sequence preference of IRF-4 physicochemically using heteronuclear two-dimensional NMR and fluorescence anisotropy.

MATERIALS AND METHODS

Recombinant human IRF-4 DNA binding domain (20-137) was expressed in *Escherichia coli* BL21(DE3) strain with minimal medium containing ^{15}N-labeled ammonium chloride. All DNA oligomers were purchased from Fasmac Corp. (Atsugi, Japan). All NMR spectra were measured using a Bruker DRX800 NMR spectrometer operating at 800 MHz. The K_d values were determined by the fluorescence anisotropy measurement using Beacon 2000 (Pan Vera Corp., Madison) with oligonucleotides labeled with hexachlorofluorescein. The K_d values were calculated from the fluorescence anisotropy data using Prism 3.0 software (GraphPad, San Diego).

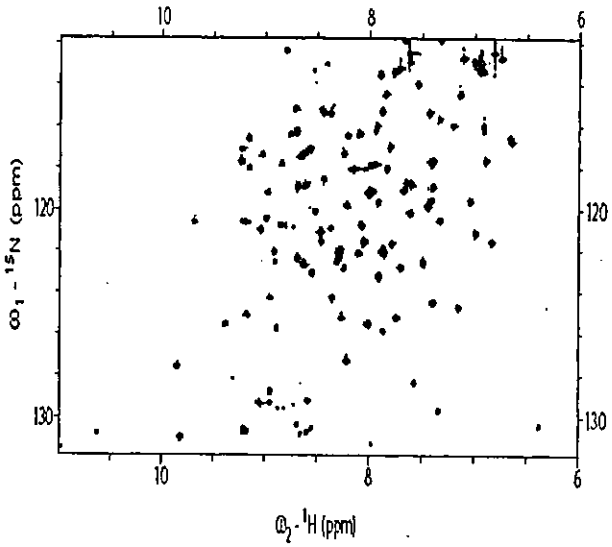

Figure 1. ^{1}H-^{15}N HSQC spectrum of the IRF-4/DNA complex recorded by 800 MHz NMR at 30 °C. DNA oligomer containing CCGAAA sequence was used. The 0.2 mM NMR sample contains 50 mM sodium phosphate (pH 6.7), 100 mM sodium chloride, 0.1 mM EDTA and 1 mM DTT.

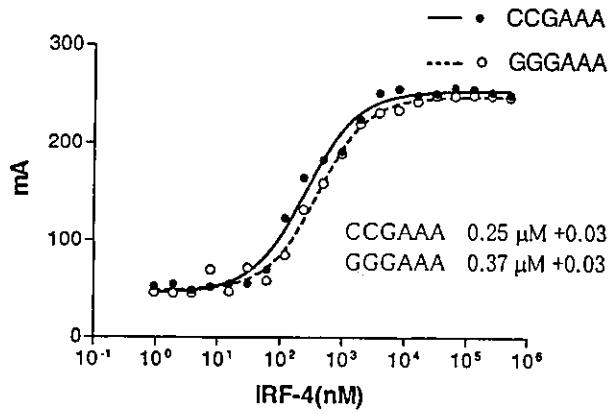

Figure 2. Florescence anisotropy of the fluorescein-labeled DNA with different concentration of IRF-4 recorded by Beacon 2000 at 22 °C. Filled and open circles show the data for the DNA oligomers containing CCGAAA and GGGAAA sequences, and the solid and dotted lines are best fitted to calculate K_d values, respectively. The sample contains 10 nM DNA with 25 mM sodium phosphate buffer (pH 6.5) and 1 mM DTT. Protein concentration is changed from 1 nM to 0.5 mM.

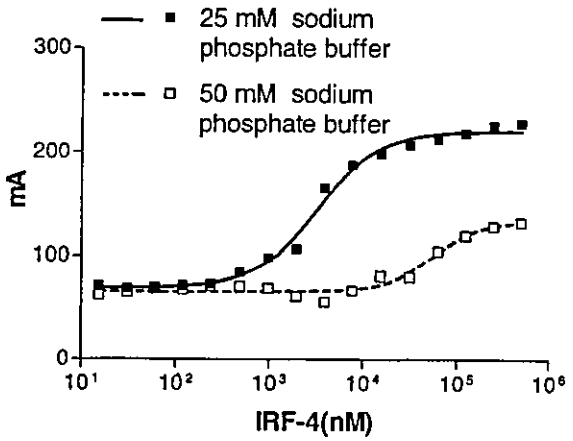

Figure 3. Salt concentration dependence of the IRF-4/DNA complex formation. The measurement condition was same to figure 2. Filled and open squares show the data for 25 mM and 50 mM sodium phosphate buffer, respectively. The solid line is identical to that shown in Figure 2.

RESULT AND DISCUSSION

The DNA oligomer containing the CCGAAA sequence was titrated to the ^{15}N enriched IRF-4 DNA binding domain, and both the ^{1}H-^{15}N HSQC spectra and the imino proton spectra were recorded. Changing the ratio of IRF-4 to DNA from 1:0 to 1:1, the spectra changed drastically. In Figure 1, the ^{1}H-^{15}N HSQC spectrum of IRF-4/DNA complex is shown. The spectra recorded at the ratio 1:1 and 1:2 were similar each other, and no significant shift was detected. Thus the DNA oligomer containing the CCGAAA sequence forms the 1:1 complex with IRF-4.

To evaluate the affinity of IRF-4 to two DNA sequences containing CCGAAA and GGGAAA, Kd values were determined by the fluorescent anisotropy measurement (Figure 2). The determined K_d values were 0.25 μM and 0.37 μM for the CCGAAA and GGGAAA sequences, respectively. This difference was very small, but larger than the experimental error. The increase of the salt concentration attenuates the affinity significantly (Figure 3). NOESY spectra of two DNA sequences containing CCGAAA and GGGAAA showed that they were the canonical B form conformation and no special feature was found. These results indicate the difference of two DNA oligomers including the CCGAAA and GGGAAA sequence is very small at both free and complex states. Thus the small difference of the indirect recognition site, such as DNA dynamics, may reflect the sequence preference of IRF-4 drastically. The detailed calorimetric study may solve this question.

REFERENCES

1. Mamane, Y., Sharma, S., Grandvaux, N., Hernandez, E. and Hiscott, J. (2002) *Annu. Rev. J. Interferon Cytokine Res.* **22**, 135-143.

2. Maracki, S. and Fenton, M. J. (2002) *Annu. Rev. J. Interferon Cytokine Res.* **22**, 121-133.

3. Escalante, C. R., Brass, A. L., Pongbala, J. M. R., Shatova, E., Shen, L., Singh, H. and Aggarwal, A. K. (2002) *Mol. Cell* **10**, 1097-1105.

4. Gupta, S. Jiang, M., Anthony, A. and Pernis, A.B. (1999) *J. Exp. Med.* **190**, 1837-1848.

5. Yamagata, T., Nishida, J., Tanaka, T., Sakai, R., Mitani, K., Yoshida, M., Taniguti, T., Yazaki, Y. and Hirai, H. (1996) *Mol. Cell. Biol.* **16**, 1283-1294.

6. Marecki, S. Atchison, M. L. and Fenton, M. J. (1999) *J. Immunol.* **163**, 2713-2722.

7. Nelson, N., Marks, M. S., Driggers, P. H. and Ozato, K. (1993) *Mol. Cell. Biol.* **13**, 588-599.

8. Blackwell, T. K., Kretzner, L., Blackwood, E. M., Eisenman, R. N. and Weintraub, H. (1990) *Science* **250**, 1149-1151.