DOI: 10.1002/ange.201104905

Drug Design

An Accurate Pharmacophore Mapping Method by NMR Spectroscopy**

Yumiko Mizukoshi, Aya Abe, Takeshi Takizawa, Hiroyuki Hanzawa, Yoshifumi Fukunishi, Ichio Shimada, and Hideo Takahashi**

A wide variety of compound libraries are currently available to obtain active compounds for drug target proteins, but the affinities of initially screened compounds are usually too low and have to be improved by chemical modification of the compounds. In such cases, pharmacophore information of the compounds plays a key role for the next modification step. Structure determination of the protein–compound complex is too time-consuming to be applied for such situations. Therefore, many chemists would highly appreciate simple and accurate experimental procedures to obtain pharmacophore information.

From this viewpoint, various ligand-observed NMR spectroscopy experiments have been proposed to characterize protein–ligand interactions.[1] Among them, experiments exploiting nuclear Overhauser effects (NOEs; transferred NOE, saturation transfer difference (STD), pumped NOE, water-ligand observed by gradient spectroscopy (water-LOGSY), etc.) are widely used and are also utilized as a ligand pharmacophore (or epitope) mapping technique.[2] However, it has recently been revealed that the difference of the longitudinal relaxation of each ligand proton severely interferes with the derived pharmacophore mapping result,[3] and it is crucial to evaluate intermolecular cross-relaxation terms for accurate pharmacophore mapping. With this in mind, pharmacophore mapping by diffusion NMR spectroscopy,[4] adiabatic fast passage NOESY (AFP-NOESY),[5] and group epitope mapping considering relaxation of the ligand (GEM-CRL)[6] quantitatively exploit the obtained intermolecular cross-relaxation effect.

Herein, we propose a simple and rapid approach for pharmacophore mapping experiments, which utilizes the difference between the longitudinal relaxation rates of ligand protons with and without irradiation of the protons of the target protein.

The longitudinal relaxation of ligand proton I is represented by the modified Bloch equation [Eq. (1)]

$$\frac{d\Delta I_z}{dt} = -R_I \Delta I_z - \sum_S \sigma_{IS} \Delta S_z - \sum_X \sigma_{IX} \Delta X_z \quad (1)$$

where $\Delta I_z = I_z - I_z^0$ (I_z represents the instantaneous longitudinal magnetization, I_z^0 represents the thermal equilibrium values of I_z); ΔS_z and ΔX_z for protons S and X are defined correspondingly. R_I represents the auto-relaxation rate constant of ligand proton I, σ_{IS} represents the intramolecular cross-relaxation rate constant between ligand proton I and another ligand proton S, and σ_{IX} represents the intermolecular cross-relaxation rate constant between ligand proton I and proton X of the target protein. In the case that longitudinal relaxation rates are measured by using an inversion-recovery method, and the initial 180° inversion pulse is replaced with two consecutive 90° pulses with an appropriate phase cycling,[7] the thermal equilibrium term can be omitted from Equation (1) to give Equation (2).

$$\frac{dI_z}{dt} = -R_I I_z - \sum_S \sigma_{IS} S_z - \sum_X \sigma_{IX} X_z \quad (2)$$

When protons of the target protein are selectively irradiated and instantaneously saturated, the intermolecular cross-relaxation term [the last term of Eq. (2)] should become zero. Therefore, the longitudinal relaxation rate under the saturation of the target protons is given by Equation (3).

$$\frac{dI_z}{dt} = -R_I I_z - \sum_S \sigma_{IS} S_z \quad (3)$$

When Equation (2) is compared with (3), the difference of recovery rate of I_z with or without selective saturation of target protons generally corresponds to the intermolecular cross-relaxation term. Therefore, the relaxation-rate differences of individual ligand protons should reflect the proximities to the target protein surface, and thus the binding portions of ligand molecules could be identified by using this "difference of inversion recovery rate with and without target irradiation" (DIRECTION) methodology.

[*] Prof. Dr. H. Takahashi
Department of Supramolecular Biology
Graduate School of Nanobioscience, Yokohama City University
Suehirocho 1-7-29, Tsurumi-ku, Yokohama 230-0045 (Japan)
E-mail: hid@tsurumi.yokohama-cu.ac.jp
Dr. Y. Mizukoshi, A. Abe, Dr. Y. Fukunishi, Prof. Dr. I. Shimada, Prof. Dr. H. Takahashi
Biomedicinal Information Research Center (BIRC), National Institute of Advanced Industrial Science and Technology (AIST) (Japan)
Dr. Y. Mizukoshi, A. Abe
Japan Biological Informatics Consortium (Japan)
T. Takizawa, Dr. H. Hanzawa
Lead Discovery and Optimization Research Laboratories II
Daiichisankyo Co.Ltd. (Japan)
Prof. Dr. I. Shimada
Graduate School of Pharmaceutical Sciences
The University of Tokyo
Hongo 7-3-1, Bunkyo-ku, Tokyo 113-0033 (Japan)
E-mail: shimada@iw-nmr.f.u-tokyo.ac.jp

[**] This work was supported by grants from New Energy and Industrial Technology Development Organization.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201104905.

To perform the DIRECTION approach, we utilized the pulse sequence shown in Figure S1 in the Supporting Information. The sequence is based on that for an inversion-recovery experiment with some minor modifications. In the case of the experiment where the protons of the target protein are saturated, selective irradiation of target protons is performed at both the presaturation period (T_{presat}) and the T_1 variable recovery period (T_{var}), whereas for the experiment where the protons are not saturated, the inversion-recovery experiment is executed with irradiation at an off-resonance frequency during both periods.

First, we compared the DIRECTION approach with the original STD-based pharmacophore mapping experiment (GEM-STD). The experiments were performed with trimethoprim (TMP, 1.5 mM; Figure 1a) in the presence of bovine dihydrofolate reductase (DHFR, 60 μM dissolved in a buffer of PBS/D_2O at pH 7.2; PBS = phosphate-buffered saline), and spectra were recorded at 280 K. The TMP–DHFR interaction system was previously used to examine that the variation of T_1 of ligand protons severely interfered with the pharmacophore mapping result.[3a] Our GEM-STD result (Figure 1b) was consistent with the previous work[3a] and indicated that 7-CH_2 of TMP, which was less affected by target saturation, was not in close contact with the target protein, bovine DHFR. However, this result was inconsistent with the known binding model.[8] In the case of small compounds, the different functional groups (e.g. aromatic, methyl, methylene, methine), which have a different number of attached protons and/or different motional characters, exhibit significantly different longitudinal relaxation times, and this variation should be prone to erroneous results in the GEM-STD experiment.[3] The DIRECTION approach intends to remove this intrinsic variation of longitudinal relaxation by subtracting Equation (2) (without irradiation) from Equation (3) (with irradiation).

In the case of the DIRECTION approach, for both inversion-recovery experiments with and without irradiation of the target protons, normalized intensities of each ligand proton were plotted against T_{var} values (Figure S2 in the Supporting Information). Based on the assumption that the longitudinal relaxation rate is determined by the sum of the relaxation contributions from multiple dipolar interactions with the surrounding protons, the longitudinal relaxation of the protons can be approximated as a single-exponential process described by an apparent relaxation rate.[9] In fact, the relaxation profiles of TMP protons seem to be reasonably approximated as a single-exponential process within short T_{var} delays ($T_{var} < 1$ s; Figure S2 in the Supporting Information). The differences between apparent longitudinal relaxation rates with irradiation (R_1^{irr}) of DHFR and those without irradiation ($R_1^{non\text{-}irr}$) for each proton of TMP are shown in Figure 1c. Without being interfered by the variety of the longitudinal relaxation of each ligand proton, the DIRECTION result indicated that 7-CH_2 was the portion that was most affected by target saturation, as predicted by a known binding model.[8]

To evaluate the accuracy of the DIRECTION results, we performed a DIRECTION experiment for the interaction analysis of p38α mitogen-activated protein kinase (MAPK) and its inhibitor SB203580 (Figure 2a; $K_d = 29$ nM; $k_{on} = 2.3 \times 10^7\,M^{-1}s^{-1}$ and $k_{off} = 0.62\,s^{-1}$ at 308 K determined by surface plasmon resonance (SPR) analysis),[10] the crystal structure of which is available. In this study, the experiment was performed with 1 mM inhibitor in the presence of human p38α (25 μM dissolved in a buffer of D_2O containing 25 mM [D_{11}]tris(hydroxymethyl)aminomethane, 1 mM [D_{10}]dithiothreitol, 50 mM NaCl at pH 7.2), and spectra were recorded at 310 K.

For both inversion-recovery experiments with and without irradiation of the target protons, normalized intensities of each ligand proton were plotted against T_{var} values (Figure S3 in the Supporting Information), and the differences between the apparent longitudinal relaxation rates with irradiation of kinase p38α and those without irradiation for each proton of ligand SB203580 are shown in Figure 2b. It is notable that $R_1^{irr} - R_1^{non\text{-}irr}$ values of H1, H4/H5 (overlapping NMR signals), and H6 of SB203580 are larger than those of H2, H3, and the methyl protons. On the other hand, when a similar experi-

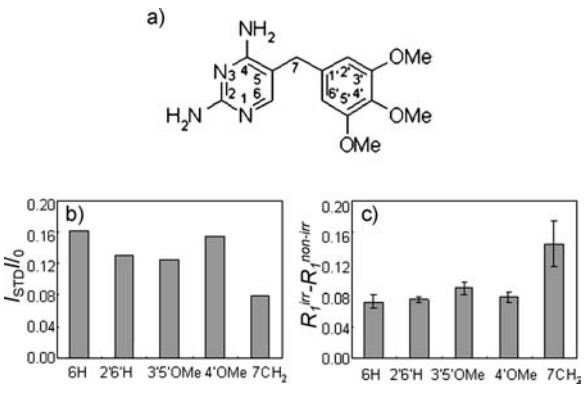

Figure 1. a) Chemical structure of TMP. b) STD-based epitope mapping of TMP binding on DHFR. The STD factor[3a] is used to evaluate the binding epitope (pharmacophore). c) DIRECTION epitope mapping of TMP binding on DHFR. The differences between the apparent longitudinal relaxation rates with irradiation of bovine DHFR (R_1^{irr}) and those without irradiation ($R_1^{non\text{-}irr}$) for each proton of TMP are shown.

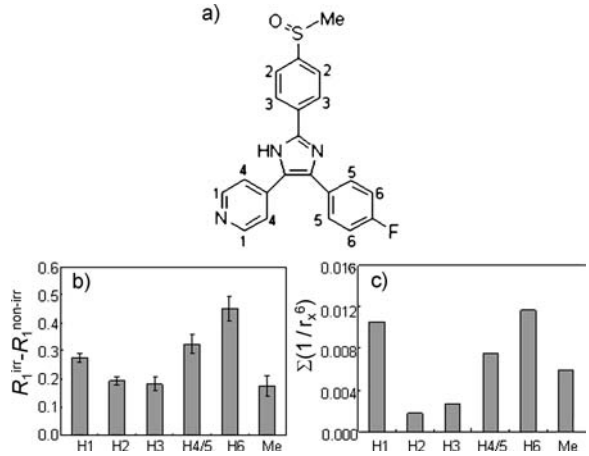

Figure 2. a) Chemical structure of p38α MAPK inhibitor SB203580. b) The results of a DIRECTION epitope mapping of SB203580 binding on p38α MAPK. c) Predicted proton density around each ligand proton (< 10 Å) from pdb 1A9U (a complex of SB203580–p38α MAPK).

ment was performed using perdeuterated p38α, $R_1^{irr}-R_1^{non\text{-}irr}$ values were too small to cause a significant difference between each ligand proton (Figure S4 in the Supporting Information). This result demonstrates that the differences of the longitudinal relaxations with and without irradiation originate directly from protons of the target molecules. This DIRECTION pharmacophore mapping experiment indicates that protons H1, H4, H5, and H6 of SB203580 are in close proximity to the protons of p38α compared with H2, H3, and methyl protons, and suggests that these two aromatic rings (a pyridine ring and fluorophenyl ring) tightly interact with the target kinase.

To validate this result, we estimated the proton density around each ligand proton. Proton–proton distances were collected based on the atom coordinates of the crystal structure of the SB203580–p38α complex (PDB ID code 1A9U) using the programs myPresto[11] and MOLMOL,[12] and the summations of $1/r^6$ (r: interatomic distance between a ligand proton and a target proton) for each ligand proton are shown in Figure 2c. The result indicates that H1, H4, H5, and H6 are located at much higher proton-density sites than H2, H3, and the methyl protons. When Figure 2b is compared with Figure 2c, the DIRECTION result is qualitatively consistent with the estimated proton density around each ligand proton.

The DIRECTION method can also be applied to the lead optimization step in drug development. We applied the method to pharmacophore mapping of a small molecule inhibitor ((1-(5-*tert*-butyl-2-methyl-2*H*-pyrazol-3-yl)-3-(4-chlorophenyl)-urea, BMU, Figure 3a) of kinase p38α, which was identified from high throughput screening (K_d = 350 nM).[13] The DIRECTION result revealed that H4 methyl protons of the pyrazole ring are less affected by the irradiation of target-proton resonances compared with H1 and H2 of the phenyl ring and H3 of the pyrazole ring (Figure 3b). Thus, this suggests that this *N*-methyl portion of the molecule would not be tightly surrounded by the protons of the target molecule (p38α); in other words, it is possible to introduce a more bulky substituted group to achieve more appropriate interaction with the target protein. In fact, the substitution of the *N*-methyl group for an *N*-phenyl group increases the affinity 40-fold (K_d = 8 nM).[13b] Thus, the DIRECTION pharmacophore mapping method could provide a guideline for lead optimization and fragment growing and/or linking in fragment-based drug design (FBDD) strategy.

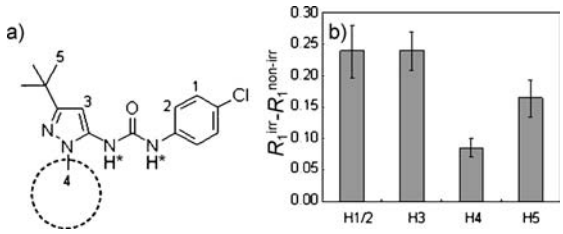

Figure 3. a) Chemical structure of BMU. Dashed circle indicates *N*-methyl portion. H* = Hydrogen atoms not observed because they exchange with the solvent D_2O. b) DIRECTION pharmacophore mapping result of BMU with p38α.

Herein, we developed a novel and robust NMR spectroscopy pharmacophore mapping approach, which is not interfered by the variation of longitudinal relaxation rates of each ligand proton. We could obtain more accurate and precise results from DIRECTION pharmacophore mapping than from the GEM-STD approach for the applications performed in this study. In theory, the initial slope of STD intensity buildup (the irradiation time dependence of the STD factor) is supposed to reflect the intermolecular cross-relaxation rate.[14] However, it is practically difficult to obtain accurate initial slopes with this approach,[6] because the sensitivity of STD peaks obtained with short irradiation times is quite low, and moreover the saturation of the protons of the target protein is not instantaneous but irradiation-time-dependent during short irradiation times (Figure S5 in the Supporting Information). A recently developed GEM-CRL experiment is another STD-based approach to obtain an accurate pharmacophore mapping result by utilizing separately measured T_1 values.[6] To apply the GEM-CRL approach, a number of simplifying approximations should be fulfilled (fast on–off equilibrium ($k_{off} > 100\ s^{-1}$), high excess of ligand (> 500-fold), long saturation time ($5 \times T_1$)); otherwise this approach would lead to an incorrect conclusion (Figure S6 in the Supporting Information). Our approach might have a substantial advantage in sensitivity and simplicity over other pharmacophore mapping approaches, including the diffusion-based approach[4] and AFP-NOESY.[5]

As is the case with other ligand-observed NMR pharmacophore mapping methods, the DIRECTION approach cannot be applied to a slowly exchanging (strong binding) ligand ($k_{off} \ll 1/T_1$). Furthermore, the strong spin-diffusion effect upon binding to a large protein would also interfere with accurate ligand pharmacophore mapping. We presume that the applicable limit of the molecular size and the exchange rate of target-protein–ligand interaction are similar to those of an STD experiment, which also utilizes the intermolecular cross-relaxation; but this hypothesis should be proven by a future theoretical simulation study for the DIRECTION approach.

The obtained pharmacophore mapping result can be directly utilized to construct a structural model of a ligand–target-protein complex, if appropriate apoprotein structures are available.[15] In cases where full-fledged structural determination of ligand–target-protein complexes are not feasible, such as in the case of difficult crystallization, this type of NMR-spectroscopy-based approach could be a powerful alternative method to solve structures of protein–ligand complexes. We have introduced DIRECTION pharmacophore mapping data into protein–ligand docking software (myPresto[11]), and the potential of our approach has been demonstrated.[16]

Received: July 14, 2011
Revised: September 21, 2011
Published online: December 30, 2011

Keywords: drug design · drug discovery · medicinal chemistry · NMR spectroscopy · protein–ligand interactions

[1] a) B. Meyer, T. Peters, *Angew. Chem.* **2003**, *115*, 890 – 918; *Angew. Chem. Int. Ed.* **2003**, *42*, 864 – 890; b) M. Pellecchia et al., *Nat. Rev. Drug Discovery* **2008**, *7*, 738 – 745; c) A. Chen, M. Shapiro, *J. Am. Chem. Soc.* **1998**, *120*, 10258 – 10259; d) M. Mayer, B. Meyer, *Angew. Chem.* **1999**, *111*, 1902 – 1906; *Angew. Chem. Int. Ed.* **1999**, *38*, 1784 – 1788; e) C. Dalvit, P. Pevarello, M. Tato, M. Veronesi, A. Vulpetti, M. Sundstroem, *J. Biomol. NMR* **2000**, *18*, 65 – 68.

[2] a) M. Mayer, B. Meyer, *J. Am. Chem. Soc.* **2001**, *123*, 6108 – 6117; b) V. M. Sánchez-Pedregal, M. Reese, J. Meiler, M. J. J. Blommers, C. Griesinger, T. Carlomagno, *Angew. Chem.* **2005**, *117*, 4244 – 4247; *Angew. Chem. Int. Ed.* **2005**, *44*, 4172 – 4175; c) C. Ludwig, P. J. Michiels, X. Wu, K. L. Kavanagh, E. Pilka, A. Jansson, U. Oppermann, U. L. Günther, *J. Med. Chem.* **2008**, *51*, 1 – 3.

[3] a) J. Yan, A. D. Kline, H. Mo, M. J. Shapiro, E. R. Zartler, *J. Magn. Reson.* **2003**, *163*, 270 – 276; b) C. A. Lepre, J. M. Moore, J. W. Peng, *Chem. Rev.* **2004**, *104*, 3641 – 3675.

[4] J. Yan, A. D. Kline, H. Mo, E. R. Zartler, M. J. Shapiro, *J. Am. Chem. Soc.* **2002**, *124*, 9984 – 9985.

[5] R. Auer, K. Kloiber, A. Vavrinska, L. Geist, N. Coudevylle, R. Konrat, *J. Am. Chem. Soc.* **2010**, *132*, 1480 – 1481.

[6] a) S. Kemper, M. K. Patel, J. C. Errey, B. G. Davis, J. A. Jones, T. D. W. Claridge, *J. Magn. Reson.* **2010**, *203*, 1 – 10; b) I. Pérez-Victoria, S. Kemper, M. K. Patel, J. M. Edwards, J. C. Errey, L. F. Primavesi, M. J. Paul, T. D. W. Claridge, B. G. Davis, *Chem. Commun.* **2009**, *39*, 5862 – 5864.

[7] a) L. E. Kay, D. A. Torchia, A. Bax, *Biochemistry* **1989**, *28*, 8972 – 8979; b) V. Sklenar, D. Torchia, A. Bax, *J. Magn. Reson.* **1987**, *73*, 375 – 379.

[8] D. A. Matthews, J. T. Bolin, J. M. Burridge, D. J. Filman, K. W. Volz, B. T. Kaufman, C. R. Beddell, J. N. Champness, D. K. Stammers, J. Kraut, *J. Biol. Chem.* **1985**, *260*, 381 – 391.

[9] a) A. Kalk, H. J. C. Berendsen, *J. Magn. Reson.* **1976**, *24*, 343 – 366; b) Z. Zheng, M. R. Gryk, M. D. Finucane, O. Jardetzky, *J. Magn. Reson. Ser. B* **1995**, *108*, 220 – 234.

[10] D. Casper, M. Bukhtiyarova, E. B. Springman, *Anal. Biochem.* **2004**, *325*, 126 – 136.

[11] Y. Fukunishi, Y. Mikami, H. Nakamura, *J. Phys. Chem. B* **2003**, *107*, 13201 – 13210.

[12] R. Koradi, M. Billeter, K. Wüthrich, *J. Mol. Graphics* **1996**, *14*, 51 – 56.

[13] a) J. Regan et al., *J. Med. Chem.* **2002**, *45*, 2994 – 3008; b) C. Pargellis et al., *Nat. Struc. Biol.* **2002**, *9*, 268 – 272.

[14] M. Mayer, T. L. James, *J. Am. Chem. Soc.* **2004**, *126*, 4453 – 4460.

[15] a) P. J. Hajduk, J. C. Mack, E. T. Olejniczak, C. Park, P. J. Dandliker, B. A. Beutel, *J. Am. Chem. Soc.* **2004**, *126*, 2390 – 2398; b) U. Schieborr, M. Vogtherr, B. Elshorst, M. Betz, S. Grimme, B. Pescatore, T. Langer, K. Saxena, H. Schwalbe, *ChemBioChem* **2005**, *6*, 1891 – 1898.

[16] Y. Fukunishi, Y. Mizukoshi, K. Takeuchi, I. Shimada, H. Takahashi, H. Nakamura, *J. Mol. Graph. Mod.* **2011**, *31*, 20 – 27.

