4780–4827; (*d*) H. Ren, F. Xiao, K. Zhan, Y.-P. Kim, H. Xie, Z. Xia and J. Rao, *Angew. Chem., Int. Ed.*, 2009, **48**, 9658–9662; (*e*) S. Mizukami, S. Watanabe, Y. Hori and K. Kikuchi, *J. Am. Chem. Soc.*, 2009, **131**, 5016–5017; (*f*) D. Maurel, S. Banala, T. Laroche and K. Johnsson, *ACS Chem. Biol.*, 2010, **5**, 507–516; (*g*) C. Uttamapinant, K. A. White, H. Baruah, S. Thompson, M. Fernández-Suárez, S. Puthenveetil and A. Y. Ting, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**, 10914–10919; (*h*) D. Srikun, A. E. Albers, C. I. Nam, A. T. Iavarone and C. J. Chang, *J. Am. Chem. Soc.*, 2010, **132**, 4455–4465.

2 (*a*) D. Summerer, S. Chen, N. Wu, A. Deiters, J. W. Chin and P. G. Schultz, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**, 9785–9789; (*b*) C. Smith, *Nature*, 2007, **4**, 755–759; (*c*) A. Olichon and T. Surrey, *J. Biol. Chem.*, 2007, **282**, 36314–36320; (*d*) W. Sandtner, F. Bezanilla and A. M. Correa, *Biophys. J.*, 2007, **93**, L45–L47; (*e*) J. Artus and A.-K. Hadjantonakis, in *Modern Research and Educational Topics in Microscopy*, ed. A. Mendez-Vilas and J. Diaz, Formatex, 2007, pp. 190–202; (*f*) M. Z. Lin and L. wang, *Physiology*, 2008, **23**, 131–141; (*g*) R. W. Watkins, L. D. Lavis, V. M. Kung, G. V. Los and R. T. Raines, *Org. Biomol. Chem.*, 2009, **7**, 3969–3975.

3 (*a*) B. A. Griffin, S. R. Adams and R. Y. Tsien, *Science*, 1998, **281**, 269–272; (*b*) B. A. Griffin, S. R. Adams, J. Jones and R. Y. Tsien, *Methods Enzymol.*, 2000, **327**, 565–578; (*c*) S. R. Adams, R. E. Campbell, L. A. Gross, B. R. Martin, G. K. Walkup, Y. Yao, J. Llopis and R. Y. Tsien, *J. Am. Chem. Soc.*, 2002, **124**, 6063–6076; (*d*) C. Hoffman, G. Gaietta, M. Bunemann, S. R. Adams, S. Oberdolff-Maass, B. Behr, J.-P. Vilardaga, R. Y. Tsien and M. H. Ellisman, *Nat. Methods*, 2005, **12**, 453–462; (*e*) O. Tour, S. R. Adams, R. A. Kerr, R. M. Meijer, T. J. Sejnowski, R. W. Tsien and R. Y. Tsien, *Nat. Chem. Biol.*, 2007, **3**, 423–431; (*f*) S. R. Adams and R. Y. Tsien, *Nat. Protoc.*, 2008, **3**, 1527–1534.

4 (*a*) G. V. Los, A. Darzins, N. Karassina, C. Zimprich, R. Learish, M. G. McDougall, L. P. Encell, R. Friedman-Ohana, M. Wood, G. Vidugiris, K. Zimmerman, P. Otto, D. H. Klaubert and K. V. Wood, *Cell Notes*, 2005, **11**, 2–6; (*b*) G. V. Los, L. P. Encell, M. G. McDougall, D. D. Hartzell, N. Karassina, C. Zimprich, M. G. Wood, R. Learish, R. F. Ohana, M. Urh, D. Simpson, J. Mendez, K. Zimmerman, P. Otto, G. Vidugiris, J. Zhu, A. Darzins, D. H. Klaubert, R. F. Bulleit and K. V. Wood, *ACS Chem. Biol.*, 2008, **3**, 373–382; (*c*) J. Schroder, H. Benink, M. Dyba and G. V. Los, *Biophys. J.*, 2009, **96**, L01–L03; (*d*) R. F. Ohana, L. P. Encell, K. Zhao, D. Simpson, M. R. Slater, M. Urh and K. V. Wood, *Protein Expression Purif.*, 2009, **68**, 110–120.

5 (*a*) A. Keppler, S. Gendreizig, T. Gronemeyer, H. Pick, H. Vogel and K. Johnsson, *Nat. Biotechnol.*, 2003, **21**, 86–89; (*b*) A. Keppler, H. Pick, C. Arrivoli, H. Vogel and K. Johnsson, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 9955–9959; (*c*) M. Kindermann, I. Sielaff and K. Johnsson, *Bioorg. Med. Chem. Lett.*, 2004, **14**, 2725–2728; (*d*) N. Johnsson and K. Johnsson, *ACS Chem. Biol.*, 2007, **2**, 31–38; (*e*) A. Gautier, A. Juilleratk, C. Heinis, I. R. Corrêa Jr., M. Kindermann, F. Beaufils and K. Johnsson, *Chem. Biol.*, 2008, **15**, 128–136; (*f*) M. Bannwarth, I. R. Corrêa Jr., M. Sztretye, S. Pouvreau, C. Fellay, A. Aebischer, L. Royer, E. Ríos and K. Johnsson, *ACS Chem. Biol.*, 2009, **4**, 179–190; (*g*) M. A. Brun, K.-T. Tan, E. Nakata, M. J. Hinner and K. Johnsson, *J. Am. Chem. Soc.*, 2009, **131**, 5873–5884.

6 (*a*) A. Ojida, K. Honda, D. Shinmi, S. Kiyonaka, Y. Mori and I. Hamachi, *J. Am. Chem. Soc.*, 2006, **128**, 10452–10459; (*b*) H. Nonaka, S. Tsukiji, A. Ojida and I. Hamachi, *J. Am. Chem. Soc.*, 2007, **129**, 15777–15779.

7 Y. Sekine-Aizawa and R. L. Huganir, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 17114–17119.

8 T. L. Halo, J. Appelbaum, E. M. Hobert, D. M. Balkin and A. Schepartz, *J. Am. Chem. Soc.*, 2009, **131**, 438–439.

9 (*a*) L. W. Miller, J. Sable, P. Goelet, M. P. Sheetz and V. W. Cornish, *Angew. Chem., Int. Ed.*, 2004, **43**, 1672–1675; (*b*) L. W. Miller, Y. Cai, M. P. Sheetz and V. W. Cornish, *Nat. Methods*, 2005, **2**, 255–257; (*c*) N. T. Calloway, M. Choob, A. Sanz, M. P. Sheetz, L. W. Miller and V. W. Cornish, *ChemBioChem*, 2007, **8**, 767–774; (*d*) S. S. Gallagher, J. E. Sable, M. P. Sheetz and V. W. Cornish, *ACS Chem. Biol.*, 2009, **4**, 547–556; (*e*) H. E. Rajasekhar, D. R. Reddy, S. Mohandessi, N. G. Butlin and L. W. Miller, *Angew. Chem., Int. Ed.*, 2009, **48**, 4990–4992.

10 C. Szent-Gyorgyi, B. F. Schmidt, Y. Creeger, G. W. Fisher, K. L. Zakel, S. Adler, J. A. Fitzpatrick, C. A. Woolford, Q. Yan, K. V. Vasilev, P. B. Berget, M. P. Bruchez, J. W. Jarvik and A. Waggoner, *Nat. Biotechnol.*, 2008, **26**, 235–240.

11 S. Watanabe, S. Mizukami, Y. Hori and K. Kikuchi, *Bioconjugate Chem.*, 2010, **21**, 2320–2326.

12 (*a*) D. Maurel, L. Comps-Agrar, C. Brock, M.-L. Rives, E. Bourrier, M. A. Ayoub, H. Bazin, N. Tinel, T. Durroux, L. Prézeau, E. Trinquet and J.-P. Pin, *Nat. Methods*, 2008, **5**, 561–567; (*b*) F. Kampmeier, M. Ribbert, T. Nachreiner, S. Dembski, F. Beaufils, A. Brecht and S. Barth, *Bioconjugate Chem.*, 2009, **20**, 1010–1015.

13 K. K. Sadhu, S. Mizukami, S. Watanabe and K. Kikuchi, *Chem. Commun.*, 2010, **46**, 7403–7405.

14 (*a*) G. Guillaume, M. Vanhove, J. Lamotte-Brasseur, P. Ledent, M. Jamin, B. Joris and J.-M. Frère, *J. Biol. Chem.*, 1997, **272**, 5438–5444; (*b*) H. Adachi, T. Ohta and H. Matsuzawa, *J. Biol. Chem.*, 1991, **266**, 3186–3191.

15 G. Punte and B. E. Rivero, *Acta Crystallogr., Sect. C: Cryst. Struct. Commun.*, 1989, **45**, 1952–1957.

16 K. Kumar and P. R. Carey, *J. Chem. Phys.*, 1975, **63**, 3697–3707.

17 (*a*) E. L. Kapinus and I. I. Dilung, *Chem. Phys. Lett.*, 1990, **174**, 75–79; (*b*) U. Wenzel and H.-G. Lohmannsroben, *J. Photochem. Photobiol., A*, 1996, **96**, 13–18; (*c*) S. Santabarbara, R. Barbato, G. Zucchelli, F. M. Garlaschi and R. S. Jennings, *FEBS Lett.*, 2001, **505**, 159–162; (*d*) K.-S. Focsaneanu and J. C. Scaiano, *Photochem. Photobiol. Sci.*, 2005, **4**, 817–821; (*e*) H. Li, J. Kang, L. Ding, F. Lu and Y. Fang, *J. Photochem. Photobiol., A*, 2008, **197**, 226–231.

18 Y. Oseki, M. Fujitsuka, D. W. Cho, A. Sugimoto, S. Tojo and T. Majima, *J. Phys. Chem. B*, 2005, **109**, 19257–19262.

19 (*a*) R. Livingston, L. Thompson and M. V. Ramarao, *J. Am. Chem. Soc.*, 1952, **74**, 1073–1075; (*b*) J. C. Nnadi, A. W. Peters and S. Y. Wang, *J. Am. Chem. Soc.*, 1972, **94**, 712–716; (*c*) M. Sirish and B. G. Maiya, *J. Photochem. Photobiol., A*, 1994, **77**, 189–200; (*d*) E. Tyystjärvi, N. King, M. Hakala and E.-M. Aro, *J. Photochem. Photobiol., B*, 1999, **48**, 142–147; (*e*) C. Zhang, *J. Phys. Chem. A*, 2006, **110**, 14029–14035.

20 S. Vijayakumar, G. Ravishanker, R. F. Pratt and D. L. Beveridge, *J. Am. Chem. Soc.*, 1995, **117**, 1722–1730.

Chemical Science

View Online
Dynamic Article Links

Cite this: DOI: 10.1039/c1sc00071c

www.rsc.org/chemicalscience

EDGE ARTICLE

^{19}F MRI detection of β-galactosidase activity for imaging of gene expression†

Shin Mizukami,[ab] Hisashi Matsushita,[a] Rika Takikawa,[a] Fuminori Sugihara,[c] Masahiro Shirakawa[d] and Kazuya Kikuchi*[ab]

Received 3rd February 2011, Accepted 23rd March 2011
DOI: 10.1039/c1sc00071c

Imaging of gene expression by magnetic resonance imaging (MRI) yields direct information regarding living systems that cannot be obtained *via* other methods. In this study, we report the rational design and synthesis of a novel ^{19}F MRI probe that detects β-galactosidase (β-gal) activity, enabling the imaging of gene expression in cells. The ^{19}F MRI signal of the probe was quenched by the intramolecular paramagnetic resonance enhancement from a Gd^{3+} ion. A contrivance was made in the probe structure to recover the ^{19}F MRI signal after hydrolysis by β-gal with a following self-immolative reaction. This ^{19}F MRI signal change was observed in the physiological aqueous condition. The probe could also detect β-gal activity in fixed HEK293T cells. In conclusion, this new probe enables the ^{19}F MRI detection of cellular gene expression. The probe design strategy is also expected to lead to the development of MRI probes for a wide variety of hydrolase activities.

Introduction

Imaging of gene expression gives us various information such as the expression timing of target proteins, gene transfer efficiency, and detection of a disease-related gene expression. To monitor gene expression by various methods, reporter proteins[1–3] are useful. Fluorescence detection of gene expression by using fluorescent proteins is particularly important because fluorescence measurement has several advantages including sensitivity, convenience, spatiotemporal resolution, *etc.* However, the poor transmission of fluorescence is one of the limitations for the *in vivo* application. Use of magnetic resonance imaging (MRI)[4] is one way to solve the problem, because MRI yields high-resolution images of deep regions of living animal bodies. Therefore, MRI is currently considered to be one of the most promising techniques for *in vivo* investigation of physiological events.[5]

Recently, several smart ^{1}H MRI probes for visualizing gene expression *via* β-galactosidase activity have been reported.[6] In principle, however, such ^{1}H MRI signal enhancement needs to be discriminated from the background ^{1}H MRI signals of water, fatty acids, and other biomolecules. To avoid this limitation, we have focused on the use of ^{19}F MRI. ^{19}F, as well as ^{1}H, is one of the most highly sensitive nuclei for NMR spectroscopy and MRI,[7] and almost no intrinsic ^{19}F MRI signals are observed in animal bodies. Thus, ^{19}F MRI probes that can visualize biological events have been increasingly reported.[8] We have also developed off-on switching ^{19}F MRI probes to detect protease activity[9] on the basis of paramagnetic relaxation enhancement (PRE),[10] a phenomenon in which the relaxation of nuclei is enhanced near paramagnetic molecules.

By expanding this probe principle, we here report a novel ^{19}F MRI probe that detects cellular gene expression. β-galactosidase (β-gal) was chosen as the reporter protein for gene expression, because it has several advantages where reporter proteins are concerned.[3,11] The advantages are as follows: (1) induction of β-gal synthesis occurs over a large dynamic range, (2) β-gal is tolerated and functional in many organisms including mammals, (3) various substrates of β-gal are available or easily synthesized, (4) many assay methods that use β-D-galactopyranoside-coupled aglycones are available, and (5) there is almost no intrinsic β-gal activity in mammalian cells. Therefore, β-gal is one of the most widely used reporter proteins for imaging of gene expression. Through the detection of β-gal activity, we tried ^{19}F MRI detection of cellular gene expression.

Results

Probe design concept, synthesis and physical properties

Several probes have been developed that can detect β-gal activity.[3,12] X-gal is one of the most widely used probes among them. Such β-D-galactopyranoside-coupled aromatic compounds are known to be the substrates of β-gal, and several fluorescent probes for β-gal have been developed. Taking this substrate

[a]*Division of Advanced Science and Biotechnology, Graduate School of Engineering, Osaka University, 2-1 Yamadaoka, Suita, Osaka, 565-0871, Japan. E-mail: kkikuchi@mls.eng.osaka-u.ac.jp; Fax: (+81) 6-6879-7875*
[b]*Immunology Frontier Research Center (IFReC), Osaka University, Osaka, 565-0871, Japan*
[c]*International Graduate School of Arts and Sciences, Yokohama City University, Kanagawa, 230-0045, Japan*
[d]*Graduate School of Engineering, Kyoto University, Kyoto, 615-8510, Japan*

† Electronic supplementary information (ESI) available: Synthesis of compounds, representative HPLC chromatograms and ^{19}F NMR. See DOI: 10.1039/c1sc00071c

recognition property of β-gal into consideration, we designed a ^{19}F MRI probe Gd-DFP-Gal for detecting β-gal activity by combining the paramagnetic relaxation enhancement (PRE) based probe design principle that we previously developed[9] with the structures of conventional β-gal probes (Fig. 1). The transverse relaxation time T_2 of the ^{19}F nucleus near Gd^{3+} is expected to be reduced by the PRE from Gd^{3+}, which has seven unpaired electrons in its 4f orbital. Thus, the T_2 of the trifluoromethyl (CF_3) group of Gd-DFP-Gal was expected to be strongly reduced.

Another designed function of Gd-DFP-Gal is the self-immolative property that can be induced by enzymatic cleavage. When Gd-DFP-Gal is hydrolyzed by β-gal, the probe is expected to be automatically converted to the corresponding quinone methide by the successive elimination of the substituent at the benzyl position.[13] Thus, the T_2 of the trifluoromethyl group extends after the β-galactoside bond is cleaved because of the cancellation of the intramolecular PRE. MRI signal intensity (*i.e.*, the peak height of the NMR signal) is proportional to $\exp(-t/T_2)$, where t is the echo time in the spin-echo method. Thus, the T_2 extension leads to an increase in the MRI signal. On the basis of this principle, we expected that the originally quenched ^{19}F MRI signal of Gd-DFP-Gal would emerge upon the enzyme reaction.

Gd-DFP-Gal was synthesized in five steps (Scheme S1, ESI†). Details of the synthetic procedure are described in the Supporting Information.† As we expected, the NMR peak of Gd-DFP-Gal was not observed, although that of the Gd-free probe DFP-Gal was a sharp single peak (Fig. S2, ESI†). Disappearance of the ^{19}F NMR peak of Gd-DFP-Gal indicates that the T_2 was markedly reduced because of the strong intramolecular PRE.

In vitro detection of β-gal activity by ^{19}F NMR and ^{19}F MRI

Because the relaxation times of Gd-DFP-Gal were dramatically reduced, it was expected that the enzymatic degradation of Gd-DFP-Gal would induce the recovery of the disappeared ^{19}F NMR peak. Gd-DFP-Gal was incubated with β-gal at 37 °C in the reaction buffer (pH 7.3) containing 5% D_2O, and the time course of the ^{19}F NMR peak was monitored (Fig. 2a). A single peak appeared at around 16 ppm (internal standard: sodium trifluoroacetate) and increased in a time-dependent manner. As the progress of the enzyme reaction was confirmed by RP-HPLC, the peak of Gd-DFP-Gal disappeared and a new peak appeared (Fig. 2b). The new peak was identified to be 4-hydroxymethyl-3-trifluoromethylphenol (compound **2**) from the retention time. In addition, the ESI-MS of the HPLC peak fraction gave a molecular weight identical to **2** (m/z = 192). No other noticeable peaks in the reaction solution HPLC diagram suggest that Gd-DFP-Gal was converted to **2** by β-galactosidase activity with nearly complete efficiency.

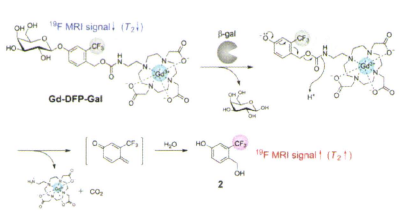


Fig. 1 Structure of Gd-DFP-Gal and the principle for the ^{19}F MRI detection of β-gal activity.

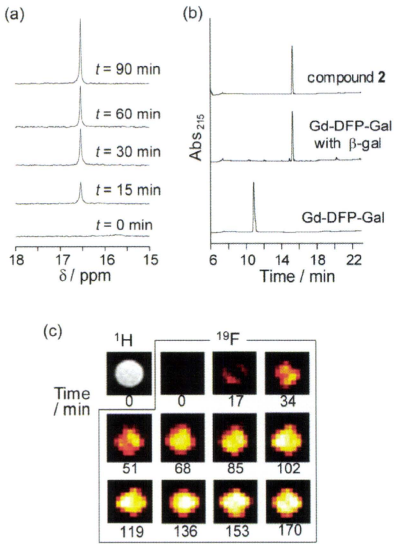


Fig. 2 Detection of β-gal activity by Gd-DFP-Gal. (a) Time-dependent ^{19}F NMR spectral change of Gd-DFP-Gal (1 mM) under incubation with β-gal. Sodium trifluoroacetate was used as the internal standard (0 ppm). (b) Confirmation of the enzymatic cleavage by RP-HPLC (eluent: H_2O–acetonitrile containing 0.1% TFA). (c) Time course of the density-weighted ^{19}F MR phantom images of Gd-DFP-Gal (1 mM) at 37 °C after β-gal was added.

The relaxation times T_1 and T_2 of the reaction sample became 0.306 s and 0.086 s, respectively, after the enzyme reaction. Both of them showed considerable extension compared to those of Gd-DFP-Gal, probably due to the cancellation of the intramolecular PRE from Gd^{3+}. These values are still less than those of the Gd^{3+}-free probe DFP-Gal: 1.293 s for T_1 and 0.271 s for T_2. When the relaxation times of Gd-DFP-Gal were measured at various probe concentrations after the enzymatic cleavage, both T_1 and T_2 extended as the concentration decreased (Fig. S3 and Table S1, ESI†). This concentration dependency of the relaxation times indicates that the intermolecular PRE is effective under the experimental condition even after the enzyme reaction is complete. To confirm the probe specificity, Gd-DFP-gal was incubated with other similar enzymes, α-galactosidase and β-glucuronidase. However, ^{19}F NMR signals of Gd-DFP-gal were not recovered by incubation with such enzymes (Fig. S4, ESI†).

To demonstrate the possibility of further application, ^{19}F MRI detection of β-gal activity was performed using Gd-DFP-Gal. ^{19}F MRI phantom images were measured using an 11.7 T MRI instrument. Gd-DFP-Gal was mixed with *Escherichia coli* β-gal before being poured into a 1-mm-inner radius capillary. The density-weighted MR images were then captured by the fast spin-echo method. As expected from the ^{19}F NMR results, Gd-DFP-Gal showed no ^{19}F MRI signals in the absence of β-gal. After the probe was mixed with β-gal, however, the ^{19}F MRI signals gradually increased in a time-dependent manner (Fig. 2c). Without addition of the enzyme, the MRI image did not show any signals for several hours (data not shown). These results demonstrate that this novel mechanism-based probe Gd-DFP-Gal enables the specific ^{19}F MRI detection of β-gal activity.

^{19}F NMR and ^{19}F MRI detection of β-gal expression in fixed HEK293T cells

Next, the applicability of Gd-DFP-Gal to the detection of intracellular gene expression was confirmed. β-gal was expressed in HEK293T cells, and the cells were fixed with formaldehyde and detergent. Then, Gd-DFP-Gal was incubated with the cells, and β-gal activity in the medium supernatant was analyzed by ^{19}F NMR (Fig. 3a). As a result, incubation of Gd-DFP-Gal only with the cells expressing β-gal induced the clear increase of a single ^{19}F NMR peak (Fig. 3b).

Then, the ^{19}F MRI detection of β-gal gene expression was attempted. HEK293T cells expressing or not expressing β-gal were cultured on 7-mm-diameter glass vessels. After the fixation of the cells, Gd-DFP-Gal (final conc.: 1 mM) was added into the glass vessels, and the cells were incubated at 37 °C for 2 h. The vessels were stacked in an 8 mm NMR tube, as shown in Fig. 3b, and the ^{1}H and ^{19}F MR images were captured. Although both vessels showed indistinguishable signal intensity in ^{1}H MRI (Fig.3c left), only the vessel that included HEK293T cells expressing β-gal showed remarkable ^{19}F MRI signals (Fig. 3c right). These results indicate that Gd-DFP-Gal can specifically detect gene expression in fixed HEK293T cells by means of reporter β-gal activity.

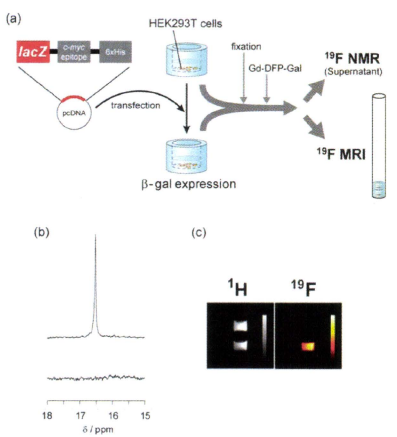


Fig. 3 ^{19}F NMR and ^{19}F MRI detection of gene expression in HEK293T cells. (a) Illustration of the experimental procedures for the ^{19}F NMR and ^{19}F MRI measurements. (b) ^{19}F NMR spectra of the culture medium containing 1 mM Gd-DFP-Gal incubated with fixed cells expressing (top) or not expressing (bottom) β-gal. (c) ^{1}H (left) and ^{19}F (right) MR images of culture vessels containing 1 mM Gd-DFP-Gal fixed cells. Color scale bars were inserted in the images.

Discussion

From the point of probe design strategy, development of Gd-DFP-Gal is an important step forward in molecular imaging studies. Our previous ^{19}F MRI probe that detects protease activity also utilized the cancellation of the intramolecular PRE for signal switching.[9] In these cases, the ^{19}F atoms and the paramagnetic ions were conjugated to each other at the opposite end of the probes (Fig. 4a). The MRI signals were enhanced by enzymatic cleavage of the substrate linker. Although this strategy works for a wide variety of hydrolases such as other proteases, endonucleases and phosphodiesterases, it could not be applied to several hydrolases such as phosphatases that have a substrate-binding pocket covering one end of the substrate. In order to detect these enzyme activities by MRI, we expanded the probe design concept by exploiting a self-immolative reaction, as presented in this study (Fig. 1). We now have the ability to design ^{19}F MRI probes for a broader range of hydrolases by choosing either of the ways illustrated in Fig. 4.

Although imaging of gene expression in fixed cells by using Gd-DFP-Gal, there are two obstacles for future perspective to *in vivo* imaging. One is the membrane permeability of the probes. Since the new probe did not permeate cell membrane, the cells needed to be fixed with formaldehyde for imaging. However, use of cell-penetrating peptides,[14] which enabled the incorporation of proteins into live cells, may dissolve the problem.

The other is the sensitivity. Generally, the sensitivity of ^{19}F MRI probes is worse than ^{1}H MRI probes. This is because ^{1}H MRI visualizes many water molecules around the probe

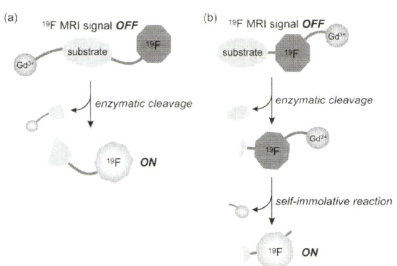


Fig. 4 Two ^{19}F MRI probe design strategies using PRE cancellation by (a) enzymatic cleavage of the substrate linker, and (b) enzyme activity-induced self-immolative reaction.

molecules, although ^{19}F MRI probes give only the probe signals. Concerning the problem, improvement of both probes and instruments will contribute to solve it. About the probe sensitization, we are under the development and would report elsewhere in future.

Conclusions

In summary, we succeeded in the imaging of gene expression in mammalian cells using a novel ^{19}F MRI probe that detects β-gal activity. The probe design concept is based on the MRI signal quenching by PRE. Also, by exploiting a self-immolative organic reaction in the probe design, we could overcome the limitation of our previous probe design, in which the ^{19}F atoms and the paramagnetic ions should be located at the opposite end of the probes. New design strategy will lead to the development of the probes for a wider variety of hydrolases such as phosphatases. Although *in vivo* imaging of gene expression by ^{19}F MRI is still challenging, the further progress of the probe properties will contribute to the solution of the difficult and significant subject.

Experimental section

^{19}F NMR relaxation time measurements

Samples were prepared at 500 μM concentration in 10 mM Tris buffer (pH 7.3) containing 10 mM magnesium chloride and 5% D_2O. The longitudinal relaxation time T_1 was measured by an inversion recovery method and the transverse relaxation time T_2 was measured by the spin-echo method.

Enzyme reaction

Gd-DFP-Gal was dissolved at 500 μM in 10 mM Tris buffer (pH 7.3) containing 10 mM magnesium chloride and 5% D_2O. Samples (500 μL) were incubated with β-gal (5.03 U) at 37 °C for 2 h. The reaction progress was monitored by ^{19}F NMR and RP-HPLC using an octadecyl silane (ODS) column. For the ^{19}F MRI experiment, 1 mM Gd-DFP-Gal was dissolved in 10 mM Tris buffer solution (pH 7.3) containing 10 mM magnesium chloride and 5% D_2O. Samples with or without β-gal (1.2 mU) were filled into glass capillaries (inner diameter: approximately 1 mm; Hischmann Laborgerate). The capillaries were then inserted into an 8 mm NMR tube and the ^{19}F MRI were measured.

Cellular experiments

HEK293T cells were grown at 37 °C in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum (FBS), 100 U mL^{-1} penicillin G, and 100 mg mL^{-1} streptomycin in a humidified atmosphere with 5% CO_2. The cells were plated at 1.2×10^6 cells in 60 mm dishes or 1.2×10^5 cells cm^{-2} on 24-well plates. Next, the cells were transfected with pcDNA™4/TO/*myc*-His/*lacZ*° plasmid using Lipofectamine 2000, and the cells were incubated for 24 h at 37 °C in a CO_2 incubator. After the cells were washed three times with phosphate-buffered solution (PBS), they were incubated with trypsin-EDTA at 37 °C for 5 min under 5% CO_2.

For ^{19}F NMR analysis, the cells were cultured with 1 mM Gd-DFP-Gal for 2 h at 37 °C in the reaction buffer (10 mM Tris-sodium buffer (pH 7.3) and 10 mM magnesium chloride) on 24-well plates. Then, the supernatants were moved into NMR tubes and the ^{19}F NMR spectra were measured.

For ^{19}F MRI analysis, the cells were moved onto 7 mm (outer diameter) glass vessels (Hilgenberg GmbH), and were incubated for 7 h at 37 °C in DMEM with 10% FBS. After the cells were washed three times with PBS, they were incubated with 3.7% formaldehyde for 10 min at room temperature. Then, cells were washed three times with PBS and incubated with 1 mM Gd-DFP-Gal for 2 h at 37 °C in the reaction buffer (Tris-sodium buffer (pH 7.3) and 10 mM magnesium chloride). The vessels were put into an 8 mm NMR tube, and the ^1H and ^{19}F MRI were measured.

Acknowledgements

We thank Dr Tetsuro Kokubo at Yokohama City University for the use of the MRI instrument, and Dr Haruhiko Bito (University of Tokyo), Dr Hiroyuki Okuno (University of Tokyo) and Dr Shin-ichi Muramatsu (Jichi Medical University) for the helpful discussion. This research is supported by Ministry of Education, Culture, Sports, Science and Technology–Japan (Grant No. 21685019 and 20675004), by the Japan Society for the Promotion of Science (JSPS) through its Funding Program for World-Leading Innovative R&D on Science and Technology (FIRST Program), by Ministry of Health, Labour and Welfare–Japan, and by the New Energy and Industrial Technology Development Organization (NEDO) of Japan. S.M. acknowledges the Inamori Foundation.

Notes and references

1 R. Y. Tsien, *Annu. Rev. Biochem.*, 1998, **67**, 509.
2 C. H. Contag and M. H. Bachman, *Annu. Rev. Biomed. Eng.*, 2002, **4**, 235.
3 (*a*) G. P. Nolan, S. Fiering, J.-F. Nicholas and L. A. Herzenberg, *Proc. Natl. Acad. Sci. U. S. A.*, 1988, **85**, 2603; (*b*) I. G. Serebriiskii and E. A. Golemis, *Anal. Biochem.*, 2000, **285**, 1.
4 (*a*) A. Jasanoff, *Trends Neurosci.*, 2005, **28**, 120; (*b*) D. E. Sosnovik and R. Weissleder, *Curr. Opin. Biotechnol.*, 2007, **18**, 4.
5 (*a*) R. Weissleder and M. J. Pittet, *Nature*, 2008, **452**, 580; (*b*) J. L. Major and T. J. Meade, *Acc. Chem. Res.*, 2009, **42**, 893.
6 (*a*) A. Y. Louie, M. M. Hüber, E. T. Ahrens, U. Rothbächer, R. Moats, R. E. Jacobs, S. E. Fraser and T. J. Meade, *Nat. Biotechnol.*, 2000, **18**, 321; (*b*) Y. T. Chang, C. M. Cheng, Y. Z. Su, W. T. Lee, J. S. Hsu, G. C. Liu, T. L. Cheng and Y. M. Wang, *Bioconjugate Chem.*, 2007, **18**, 1716; (*c*) E. L. Que, D. W. Domaille and C. J. Chang, *Chem. Rev.*, 2008, **108**, 1517; (*d*) W. Cui, L. Liu, V. D. Kodibagkar and R. P. Mason, *Magnetic Resonance in Medicine*, 2010, **64**, 65.
7 (*a*) J. Yu, V. D. Kodibagkar, W. Cui and R. P. Mason, *Curr. Med. Chem.*, 2005, **12**, 819; (*b*) C. Belle, C. Beguin, S. Hamman and J.-L. Pierre, *Coord. Chem. Rev.*, 2009, **253**, 963.
8 (*a*) W. Cui, P. Otten, Y. Li, K. S. Koeneman, J. Yu and R. P. Mason, *Magn. Reson. Med.*, 2004, **51**, 616; (*b*) M. Higuchi, N. Iwata, Y. Matsuba, K. Sato, K. Sasamoto and T. C. Saido, *Nat. Neurosci.*, 2005, **8**, 527; (*c*) V. D. Kodibagkar, J. Yu, L. Liu, H. P. Hetherington and R. P. Mason, *Magn. Reson. Imaging*, 2006, **24**, 959; (*d*) K. Tanaka, N. Kitamura, K. Naka and Y. Chujo, *Chem. Commun.*, 2008, 6176; (*e*) P. Porcari, S. Capuani, E. D'Amore, M. Lecce, A. La Bella, F. Fasano, R. Campanella, L. M. Migneco, F. S. Pastore and B. Maraviglia, *Phys. Med. Biol.*, 2008, **53**, 6979; (*f*) K. Tanaka, N. Kitamura, Y. Takahashi and Y. Chujo, *Bioorg. Med. Chem.*, 2009, **17**, 3818; (*g*) Y. Takaoka, T. Sakamoto, S. Tsukiji, M. Narazaki, T. Matsuda, H. Tochio, M. Shirakawa and I. Hamachi, *Nat. Chem.*, 2009, **1**, 557; (*h*) K. Tanabe, H. Harada, M. Narazaki, K. Tanaka, K. Inafuku,

H. Komatsu, T. Ito, H. Yamada, Y. Chujo, T. Matsuda, M. Hiraoka and S. Nishimoto, *J. Am. Chem. Soc.*, 2009, **131**, 15982.
9 (*a*) S. Mizukami, R. Takikawa, F. Sugihara, Y. Hori, H. Tochio, M. Wälchli, M. Shirakawa and K. Kikuchi, *J. Am. Chem. Soc.*, 2008, **130**, 794; (*b*) S. Mizukami, R. Takikawa, F. Sugihara, M. Shirakawa and K. Kikuchi, *Angew. Chem., Int. Ed.*, 2009, **48**, 3641.
10 L. Helm, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2006, **49**, 45.
11 C. V. Hall, P. E. Jacob, G. M. Ringold and F. Lee, *J. Mol. Appl. Genet.*, 1983, **2**, 101.
12 (*a*) J. P. Horwitz, J. Chua, R. J. Curby, A. J. Tomson, M. A. DaRooge, B. E. Fisher, J. Mauricio and I. Klundt, *J. Med. Chem.*, 1964, **7**, 574; (*b*) A. B. Pardee, F. Jacob and J. Monod, *J. Mol. Biol.*, 1959, **1**, 165.
13 (*a*) D. Shabat, R. J. Amir, A. Gopin, N. Pessah, M. Shamis and W.-M. Dai, *Chem.–Eur. J.*, 2004, **10**, 2626; (*b*) J. A. Duimstra, F. J. Femia and T. J. Meade, *J. Am. Chem. Soc.*, 2005, **127**, 12847; (*c*) T. Komatsu, K. Kikuchi, H. Takakusa, K. Hanaoka, T. Ueno, M. Kamiya, Y. Urano and T. Nagano, *J. Am. Chem. Soc.*, 2006, **128**, 15946; (*d*) N.-H. Ho, R. Weissleder and C.-H. Tung, *ChemBioChem*, 2007, **8**, 560.
14 E. Vives, J. Schmidt and A. Pelegrin, *Biochim. Biophys. Acta*, 2008, **1786**, 126.

Biomol NMR Assign
DOI 10.1007/s12104-011-9301-x

ARTICLE

^{1}H, ^{13}C, and ^{15}N resonance assignment of the first PDZ domain of mouse ZO-1

Yoshitaka Umetsu · Natsuko Goda · Ryo Taniguchi · Kaori Satomura · Takahisa Ikegami · Mikio Furuse · Hidekazu Hiroaki

Received: 7 January 2011 / Accepted: 12 March 2011

Abstract Zonula occludens-1 (ZO-1) is a scaffolding molecule critical to the formation of intercellular adhesion structures, such as tight junctions (TJs) and adherens junctions (AJs). ZO-1 contains three PDZ domains followed by a GUK domain and a ZU5 domain. The first PDZ of ZO-1 (ZO-1(PDZ1)) serves as a protein–protein interaction module and interacts with the C-termini of almost all claudins to initiate the formation of a belt-like structure on the lateral membranes, thereby promoting TJ formation. It has been recently reported that approximately 15% of all PDZ domains bind phosphoinositides, and ZO-1(PDZ1) is the one of these. Here we report the ^{15}N, ^{13}C, and ^{1}H chemical shift assignments of the first PDZ domain of mouse ZO-1. The resonance assignments obtained in this work may contribute in clarifying the interplay between the two binary interactions, ZO-1(PDZ1)–claudins and ZO-1(PDZ1)–phospholipids, and suggesting a novel regulation mechanism underlying the formation and maintenance of cell–cell adhesion machinery downstream of the phospholipid signaling pathways.

Keywords Cell–cell adhesion · Tight junction · ZO-1; PDZ domain · Phosphoinositide

Y. Umetsu · N. Goda · R. Taniguchi · K. Satomura · H. Hiroaki
Division of Structural Biology Graduate School of Medicine, Kobe University, Kobe, Hyogo, Japan

K. Satomura · M. Furuse · H. Hiroaki
Targeted Protein Research Program (JST-TPRP), Kobe University, Kobe, Hyogo, Japan

T. Ikegami
Institute for Protein Research, Osaka University, Suita, Osaka, Japan

M. Furuse · H. Hiroaki (✉)
Global-COE (Center of Excellence) Program for Integrative Membrane Biology, Kobe University, Kobe, Hyogo, Japan
e-mail: hiroakih@med.kobe-u.ac.jp

M. Furuse
Division of Cell Biology, Graduate School of Medicine, Kobe University, Kobe, Hyogo, Japan

Biological context

The tight junction (TJ), or zonula occludens, is the most apical intercellular junctional complex found in epithelial and endothelial cells. It is believed that TJs maintain the major epithelial and endothelial barrier functions, thus maintaining the unique composition of chemical and biological substances at the apical and basolateral spaces of the cell layer (Furuse et al. 1993; Ikenouchi et al. 2007). TJs are responsible for the formation of barriers regulating the passage of solutes and cells through the paracellular space. In addition to their semipermeable barrier and/or gate function, TJs are also involved in some signal transduction pathways, i.e., they regulate their own assembly and barrier function, as well as transmit signals from the paracellular space to the cell interior by coordinating a variety of signaling and trafficking molecules that regulate cell differentiation, proliferation, and polarity (Tsukita et al., 2001; Ikenouchi et al. 2005; Martin-Padura et al. 1998).

The extracellular part of TJs is composed of four families of integral membrane proteins, claudins (Furuse et al. 1998; Morita et al. 2003), occludin (Furuse et al. 1993), junctional adhesion molecules (JAMs) (Martin-Padura et al. 1998), and tricellulin (Ikenouchi et al. 2005). Most of these membrane proteins are linked to each other by PDZ (PSD95/Discs large/ZO-1) domain-containing scaffold proteins. The PDZ

Published online: 24 March 2011

domain is compact and globular. It is the most abundant peptide-binding module in the human genome, in which there may be as many as 440 PDZ domains in 259 different proteins (Letunic et al. 2004). In TJs, membrane-associated guanylate kinase (MAGUK) proteins, ZO-1, and ZO-2 are localized exclusively in polarized epithelia (Gonzalez-Mariscal et al. 2000; Gonzalez-Mariscal et al. 2003).

Here we describe the NMR assignments of the first PDZ domain of ZO-1, ZO-1(PDZ1). As well as most other PDZ domain-containing proteins, ZO-1 and ZO-2 are multiple tandem PDZ domains that harbor other protein-binding modules, such as SH3 and ZU5 domains, thereby functioning as a scaffold. One of the pivotal interactions in junctional complexes is that between the first PDZ domains from ZO-1 or ZO-2 and the cytosolic C-terminal PDZ-binding motif of claudins (Itoh et al. 1999). More recently, approximately 15% of PDZ domains were shown to interact with phospholipids directly (Wu et al. 2007; Zimmermann 2006), and ZO-1(PDZ1) and ZO-1(PDZ2) are examples in the case (Meerschaert et al. 2009). Thus, the resonance assignments of ZO-1(PDZ1) contribute to the study of the interactions between ZO-1(PDZ1) and claudins or other ligands, such as phosphoinositides.

Methods and experiments

Sample preparation

The expression vector for the recombinant GST-tagged form of mouse ZO-1(PDZ1) (residues 18–110) was constructed using PRESAT-vector methodology (Tenno et al. 2004). Isotopically labeled protein for NMR spectroscopy was generated in *Escherichia coli* BL21(*DE3*) from 1 L M9 minimal medium culture grown at 20°C in the presence of [^{15}N]-NH_4Cl and [^{13}C]-glucose as the sole nitrogen and carbon sources, respectively. The harvested cells were resuspended in lysis buffer (50 mM Tris–HCl, pH 7.5, 150 mM NaCl) and disrupted by sonication. The supernatant was applied to a DEAE-Sepharose (GE Healthcare) column and then affinity purified by Glutathione Sepharose 4 Fast Flow (GE Healthcare) chromatography. The GST tag was removed by PreScission protease on beads. Next, the ZO-1(PDZ1) was purified by gel filtration using a HiLoad 26/60 Superdex 75 pg (GE Healthcare). Using this purification protocol, ^{15}N-labeled or ^{13}C, ^{15}N-doubly-labeled ZO-1(PDZ1) was prepared. The purified protein was concentrated to 0.73 mM and dialyzed with 22 mM MES (pH5.9). The protein solution (313.5 μL) and 16.5 μL 99% D_2O were mixed, resulting in 0.7 mM ZO-1(PDZ1) in 20 mM MES (pH 5.9) in 95% H_2O–5% D_2O. ^{14}N-amino-acid-selective-inversely-labeled ^{15}N-labeled-ZO-1(PDZ1) was also prepared. (Arg^-)- and (Lys^-)-^{15}N-labeled-ZO-1(PDZ1) were prepared from 0.5 L ^{15}N-enriched M9 media supplemented with ^{14}N-Lys or ^{14}N-Arg, respectively. (The details of the protocol are in preparation.)

NMR spectroscopy

NMR experiments were performed on a Bruker AvanceDRX (600 MHz) or an Avance III (600 MHz) NMR spectrometer (Bruker), the latter was equipped with a cryogenic triple-resonance probe. For the assignment of backbone ^{1}H, ^{13}C, and ^{15}N resonances, HNCA, HN(CO)CA, HNCACB, CBCA(CO)NH, HNCO, HN(CA)CO, and 3D ^{15}N-edited-NOESY-HSQC spectra were recorded. For side chain resonance assignment, 2D constant-time ^{1}H-^{13}C HSQC, 3D ^{13}C-edited-NOESY-HSQC, HCCH-TOCSY, CC(CO)NNH, and HCC(CO)NNH spectra were recorded. All NMR spectra were recorded at 298 K. All spectra were processed using NMRPipe (Delaglio et al. 1995) and analyzed using the program SPARKY (Goddard and Kneller 2004).

Assignments and data deposition

The sequential assignment of backbone signals was initiated by the automatic program MARS (Jung and Zweckstetter 2004), resulting in approximately 60% completion. However, subsequent iterative MARS runs could not increase further assignment gains beyond 65%. We then carefully traced the sequential connectivities based on triple-resonance data sets. Further help with assignment was provided by the ^{15}N-edited NOESY spectra. These backbone assignments were confirmed by a visual inspection of ^{15}N-HSQC spectra of (Lys^- -/(Arg^-)- inversely labeled ZO-1(PDZ1).

Following a sequential assignment procedure, 94.7% of the $^1H^N$, ^{15}N resonances of the backbone amide groups (90 out of the 95 non-Pro residues) were assigned (Fig. 1). In addition, 93.0% of H^α (93 out of 100 residues), 94.0% of $^{13}C^\alpha$ (94 out of 100 residues), and 94.3% of $^{13}C^\beta$ (82 out of 87 residues) resonances were assigned. The secondary structure of ZO-1(PDZ-1) was predicted by comparing the chemical shifts with random coil values using TALOS program. This result and our preliminary structural analysis of ZO-1(PDZ-1) indicate six β-strands (Glu10-Leu16, Ile26-Ser29, Val45-Val49, Arg64-Val68, Val71-Ser72, and Ala92-Arg98) and an α-helix (Glu77-Lys87), in good agreement with the structural elements of crystal structure of ZO-1(PDZ-1) (PDB: 2H3 M) (Appleton et al. 2006). The sequential correlations of the backbone resonances of His36, Phe37, Gln38, and Ser39 were missing. On the ^{1}H-^{15}N HSQC spectra, there was only one unassigned main-chain amide proton resonance, which may correspond to one of these residues. The residues are involved in the long loop

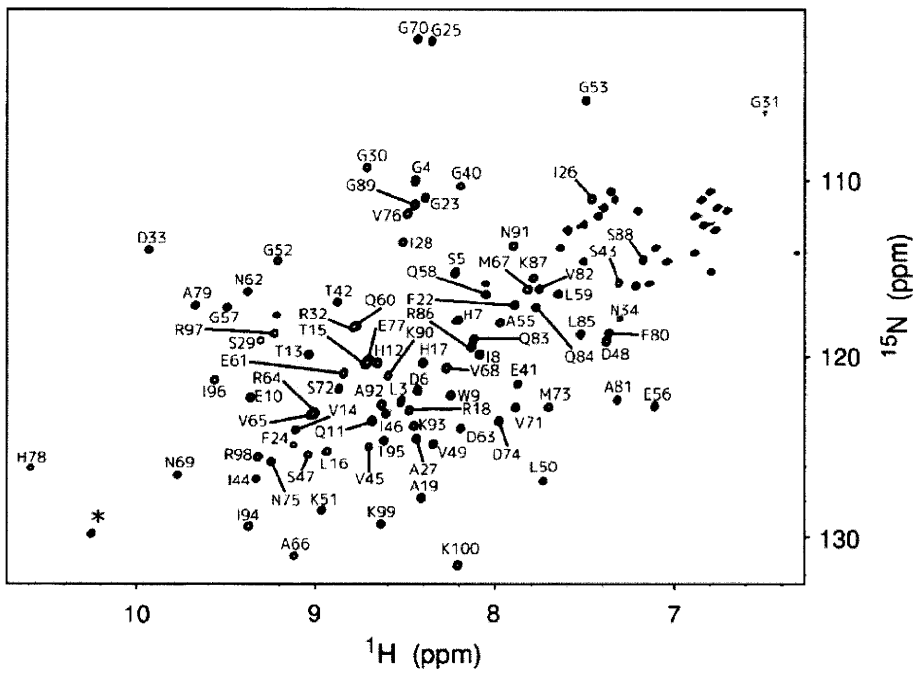


Fig. 1 A portion of the ^{15}N-^{1}H-HSQC spectrum of the first PDZ domain of mouse ZO-1 protein illustrating a number of the assigned backbone ^{15}N resonances. Asterisk indicates the Trp9 side chain

between $\beta 2$ and $\beta 3$. In the crystal structure of ZO-1(PDZ1), the loop adopted a compact L-like structure and was closed toward the $\beta 2$ surface. However, part of the loop had some contact with the neighboring molecule in the crystal lattice, suggesting that this folded structure may not be natural in solution. In our preliminary analysis of the solution structure of ZO-1(PDZ1), this folded loop structure was not observed (manuscript in preparation). It is likely that the loop, including the four "missing residues," adopts a loose, extended conformation in certain structural equilibrium.

Because of the sample preparation from the GST-fusion protein construct, the NMR sample must contain the extra seven residues (GPLGSDH) preceding the N-terminus in mouse ZO-1 (residues 18–110) in the sample. Thus, in the chemical shift data and Fig. 1, the residues 1–7 correspond to GPLGSDH from the expression vector, whereas the residues 8–100 are equivalent to the residues 18–110 of the full length mouse ZO-1. The assigned chemical shifts of ZO-1(PDZ1) have been deposited in the BioMagResBank under accession number 11424.

Acknowledgments This work was supported by a Grant-in-Aid from Bioinformatics Research and Development (BIRD) and Target Protein Research Program (TPRP) of the Japan Scientific and Technology Cooperation (JST).

References

Appleton BA, Zhang Y, Wu P, Yin JP, Hunziker W, Skelton NJ, Sidhu SS, Wiesmann C (2006) Comparative structural analysis of the Erbin PDZ domain and the first PDZ domain of ZO-1. Insights into determinants of PDZ domain specificity. J Biol Chem 281:22312–22320

Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR 6:277–293

Furuse M, Hirase T, Itoh M, Nagafuchi A, Yonemura S, Tsukita S, Tsukita S (1993) Occludin: a novel integral membrane protein localizing at tight junctions. J Cell Biol 123:1777–1788

Furuse M, Fujita K, Hiiragi T, Fujimoto K, Tsukita S (1998) Claudin-1 and -2: novel integral membrane proteins localizing at tight junctions with no sequence similarity to occludin. J Cell Biol 141:1539–1550

Goddard TD, Kneller DG (2004) Sparky 3. University of California, San Francisco

Gonzalez-Mariscal L, Betanzos A, Avila-Flores A (2000) MAGUK proteins: structure and role in the tight junction. Semin Cell Dev Biol 11:315–324

Gonzalez-Mariscal L, Betanzos A, Nava P, Jaramillo BE (2003) Tight junction proteins. Prog Biophys Mol Biol 81:1–44

Ikenouchi J, Furuse M, Furuse K, Sasaki H, Tsukita S, Tsukita S (2005) Tricellulin constitutes a novel barrier at tricellular contacts of epithelial cells. J Cell Biol 171:939–945

Ikenouchi J, Umeda K, Tsukita S, Furuse M, Tsukita S (2007) Requirement of ZO-1 for the formation of belt-like adherens junctions during epithelial cell polarization. J Cell Biol 176:779–786

Itoh M, Furuse M, Morita K, Kubota K, Saitou M, Tsukita S (1999) Direct binding of three tight junction-associated MAGUKs, ZO-1, ZO-2, and ZO-3, with the COOH termini of claudins. J Cell Biol 147:1351–1363

Jung YS, Zweckstetter M (2004) Mars–robust automatic backbone assignment of proteins. J Biomol NMR 30:11–23

Letunic I, Copley RR, Schmidt S, Ciccarelli FD, Doerks T, Schultz J, Ponting CP, Bork P (2004) SMART 4.0: towards genomic data integration. Nucleic Acids Res 32(Database issue):142–144

Martin-Padura I, Lostaglio S, Schneemann M, Williams L, Romano M, Fruscella P, Panzeri C, Stoppacciaro A, Ruco L, Villa A,

Simmons D, Dejana E (1998) Junctional adhesion molecule, a novel member of the immunoglobulin superfamily that distributes at intercellular junctions and modulates monocyte transmigration. J Cell Biol 142:117–127

Meerschaert K, Tun MP, Remue E, De Ganck A, Boucherie C, Vanloo B, Degeest G, Vandekerckhove J, Zimmermann P, Bhardwaj N, Lu H, Cho W, Gettemans J (2009) The PDZ2 domain of zonula occludens-1 and -2 is a phosphoinositide binding domain. Cell Mol Life Sci 66:3951–3966

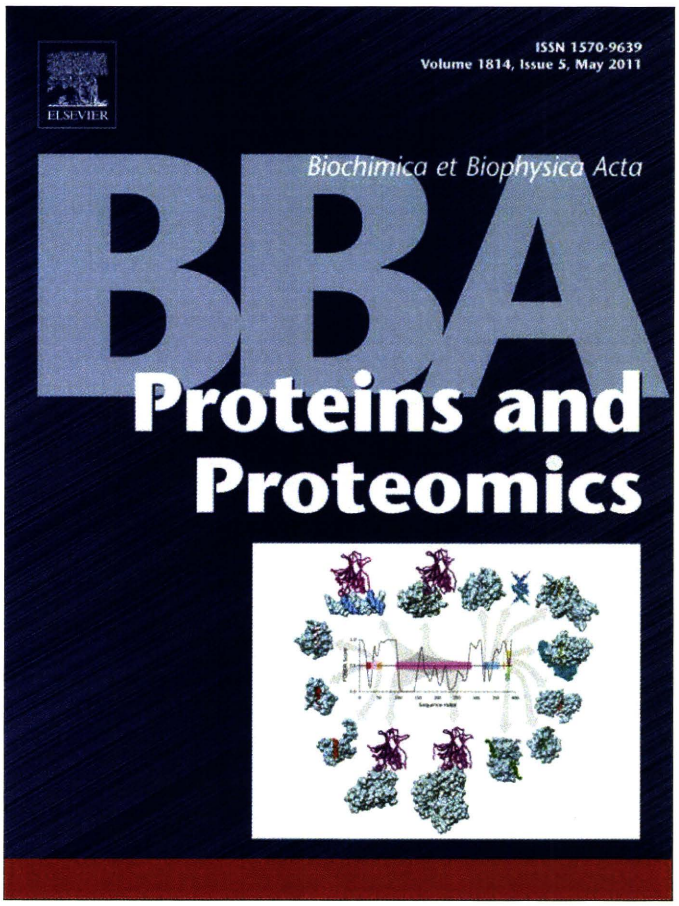
Morita K, Sasaki H, Furuse K, Furuse M, Tsukita S, Miyachi Y (2003) Expression of claudin-5 in dermal vascular endothelia. Exp Dermatol 12:289–295

Tenno T, Goda N, Tateishi Y, Tochio H, Mishima M, Hayashi H, Shirakawa M, Hiroaki H (2004) High-throughput construction method of expression vector of peptides for NMR study suited for isotopic labeling. Protein Eng. Des. Sel. 17:305–314

Tsukita S, Furuse M, Itoh M (2001) Multifunctional strands in tight junctions. Nat Rev Mol Cell Biol 2:285–293

Wu H, Feng W, Chen J, Chan LN, Huang S, Zhang M (2007) PDZ domains of Par-3 as potential phosphoinositide signaling integrators. Mol Cell 28:886–898

Zimmermann P (2006) The prevalence and significance of PDZ domain-phosphoinositide interactions. Biochim Biophys Acta 1761:947–956



Biochimica et Biophysica Acta 1814 (2011) 724–730

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbapap

Structural difference of vasoactive intestinal peptide in two distinct membrane-mimicking environments

Yoshitaka Umetsu [a], Takeshi Tenno [a,b], Natsuko Goda [a], Masahiro Shirakawa [b], Takahisa Ikegami [c], Hidekazu Hiroaki [a,d,*]

[a] Division of Structural Biology, Graduate School of Medicine, Kobe University, 7-5-1 Kusunoki-cho, Chuo-ku, Kobe, Hyogo 650-0017, Japan
[b] Department of Molecular Engineering, Graduate School of Engineering, Kyoto University, Katsura, Nishikyo-ku, Kyoto 615-8530, Japan
[c] Institute of Protein Research, Osaka University, Suita, Osaka 565-0871, Japan
[d] Global Center of Excellence Program for Integrative Membrane Biology, Kobe University, Kobe, Hyogo, Japan

ARTICLE INFO

Article history:
Received 14 January 2011
Received in revised form 9 March 2011
Accepted 16 March 2011
Available online 23 March 2011

Keywords:
Vasoactive intestinal peptide
Pituitary adenylate cyclase-activating peptide
G protein-coupled receptor
NMR structure
Dodecylphosphocholine micelle

ABSTRACT

Vasoactive intestinal peptide (VIP) is a 28-amino acid neuropeptide which belongs to a glucagon/secretin superfamily, the ligand of class II G protein-coupled receptors. Knowledge for the conformation of VIP bound to membrane is important because the receptor activation is initiated by membrane binding of VIP. We have previously observed that VIP-G (glycine-extended VIP) is unstructured in solution, as evidenced by the limited NMR chemical shift dispersion. In this study, we determined the three-dimensional structures of VIP-G in two distinct membrane-mimicking environments. Although these are basically similar structures composed of a disordered N-terminal region and a long α-helix, micelle-bound VIP-G has a curved α-helix. The side chains of residues Phe[6], Tyr[10], Leu[13], and Met[17] found at the concave face form a hydrophobic patch in the micelle-bound state. The structural differences in two distinct membrane-mimicking environments show that the micelle-bound VIP-G localized at the water–micelle boundary with these side chains toward micelle interior. In micelle-bound PACAP-38 (one of the glucagon/secretin superfamily peptide) structure, the identical hydrophobic residues form the micelle-binding interface. This result suggests that these residues play an important role for the membrane binding of VIP and PACAP.

1. Introduction

Vasoactive intestinal peptide (VIP), a member of the glucagon/secretin superfamily, is a 28-amino acid neuropeptide that is evolutionarily well conserved from fish (cod) and frogs to humans. In mammals, except guinea pigs, the sequence identity is at least 85% [1]. VIP widely presents in the central and peripheral nervous systems. It acts in a wide range of physiological and pathological processes related to development, growth, and the control of neuronal and endocrine cells; it also functions in the digestive, respiratory, reproductive, and cardiovascular systems. Furthermore, VIP plays key roles in cancers, immune responses, and circadian rhythm [2]. It is a class II G protein-coupled receptor (GPCR) ligand that acts through interaction with two receptor subtypes (VPAC1 and VPAC2) [3]. VIP shares 68% homology with its closely related homolog, pituitary adenylate cyclase-activating polypeptide (PACAP), another secretin family member peptide [4,5], and both VIP and PACAP bind to VPAC1 and VPAC2 with equivalent affinities [6]. On binding of VIP to VPAC1 or VPAC2, the cyclic adenosine 5′-phosphate level increases significantly [7], while adenylate cyclase [8] and phospholipase C [9] are also activated, followed by divergent downstream effects through various transcription factors.

Recently, key biological functions of VPAC1 and VPAC2 in innate immune responses have emerged [10], and thus, the therapeutic potential of VIP agonists as well as VIP itself is promising [11]. The increased number of studies supporting the therapeutic potential of VIP includes the murine models of pancreatitis [12], human Crohn's disease (TNBS-induced colitis) [13,14], bacterial sepsis (survival model) [15], and human rheumatoid arthritis (RA; collagen-induced model) [16]. In the first example of the pancreatitis model, a selective VPAC1 agonist rather than VIP itself provides therapeutic benefits because of the opposing action of VPAC2 against VPAC1. In addition, in the other three cases (Crohn's disease, sepsis survival, and RA), the potential anti-inflammatory effect of VIP itself works well, most probably through suppression of inflammatory cytokines such as TNF-alpha and IL-6.

In addition to the therapeutic use of VIP, specific molecular interactions between VIP–VIP receptors are also useful in diagnosis

Abbreviations: NMR, nuclear magnetic resonance; TRX, thioredoxin; PACAP, pituitary adenylate cyclase-activating polypeptide; VIP, vasoactive intestinal peptide; HSQC, heteronuclear single quantum coherence; DPC, dodecylphosphocholine; GPCR, G protein-coupled receptor; RA, rheumatoid arthritis; SAR, structure–activity relationship; TFE, trifluoroethanol

* Corresponding author at: Division of Structural Biology, Graduate School of Medicine, Kobe University, 7-5-1 Kusunoki-cho, Chuo-ku, Kobe, Hyogo 650-0017, Japan. Tel.: +81 78 382 5813; fax: +81 78 382 5816.
E-mail address: hiroakih@med.kobe-u.ac.jp (H. Hiroaki).

1570-9639/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.bbapap.2011.03.009

of several cancers [17,18]. In cancers such as breast and prostate, an elevated expression of VPAC1 is observed [19]. Thus, accumulation of radio-labeled VIP analogs, such as ^{99m}Tc-labeled VIP and ^{64}Cu-labeled TP3939, enables visualization of cancers by positron emission tomography. VPAC receptors temporarily downregulate by endocytosis when VIP binds to them [20]. This mechanism has encouraged researchers to develop VIP analogs as imaging agents.

The 3D structure of VIP and its stability highlight the possible development of drug delivery systems with an improved therapeutic potency of VIP and its analogs. Several VIP formulations with increased biochemical and physicochemical stability have been studied extensively [21–23]. A typical variation for therapeutic VIPs is their incorporation into phospholipids or liposomes. Such variations generally double the activity of VIP. The primary benefit of the lipid/liposome formulation is to protect VIP from enzymatic degradation in vivo. In addition, a structural transition from a random coil to an α-helix on lipid and/or liposome binding is the rate-limiting step at the ligation of VIP to the receptors. A number of structural studies have shown that PACAP [24,25] as well as other class II GPCR ligands [26–29] adopt α-helical conformations when they bind to the receptor. A pre-existing α-helical conformation is also important in increasing the biological activities of the peptides [30], which is partly explained by a two-step model demonstrating the PACAP–PAC1R interaction [24]. Thus, the solution structure of VIP in the absence of its receptors and presence of phospholipids is another important issue.

In this study, we focused on the structural analysis of VIP in order to enable us to develop potent VIP analogs using a structure–activity relationship (SAR) study. Previously, the solution structure of VIP in 30% trifluoroethanol (TFE) solution was determined by the ^{1}H nuclear magnetic resonance (NMR) technique but not using isotopically labeled peptide [31], thereby showing limited structural convergence. In addition, the structural coordinates were not commonly available, thus limiting SAR studies by other groups. Here we determined the solution structures of VIP-G (28-residue VIP with extra C-terminal glycine, an equivalent peptide of the biosynthetic precursor of VIP before C-terminal amidation) in two different conditions, 50% MeOH, and aqueous buffer containing dodecylphosphocholine (DPC) micelles. Use of isotopically labeled VIP enabled analysis of the NMR spectra in the presence of a large number of DPC micelles in solution. Here we discuss the differences between the two structures in detail.

2. Materials and methods

2.1. Preparation of ^{15}N- and ^{13}C/^{15}N-labeled VIP-G

The bacterial expression system with VIP-G as the fusion protein of thioredoxin, followed by a 6-poly histidine tag, factor Xa cleavage site, and VIP-G [thioredoxin (TRX)_VIP-G], was as described previously [32]. For ^{15}N and ^{13}C/^{15}N labeling of the fusion protein TRX_VIP-G, M9 medium containing 0.5 g/L ^{15}N-NH_4Cl as the sole nitrogen source and 4 g/L of ^{12}C -glucose or 2 g/L of $^{13}C_6$-glucose, respectively, were used. *Escherichia coli* BL21(*DE3*) harboring pET-TRX_VIP-G was grown in 100 mL of M9 medium containing 50 μg/mL ampicillin for 16 h. The cells were transferred into 0.9 L cultures of the same medium in 5-L baffled flasks. Isopropyl-D-thiogalactopyranoside induction was performed as described earlier, except that the cells were incubated for 6 h at 30 °C prior to harvesting.

BL21(*DE3*) cells from 1 L of M9 medium were pelleted, resuspended in 30 mL of 50 mM Tris–HCl (pH 7.5), 0.15 M NaCl, and 10 mM mercaptoethanol, and disrupted by sonication. Cell debris were removed by centrifugation, and the extracts were then passed through a 4-mL column of DEAE-Sepharose. The extracts, cleared by either osmotic disruption or sonication, were loaded onto a 4-mL column of fast-flow chelating Sepharose, previously charged with 50 mM $NiSO_4$, and equilibrated in 20 mM imidazole and 50 mM Tris–HCl (pH 7.5). The column was then washed with 50 mM imidazole and 50 mM Tris–HCl (pH 7.5), followed by fusion protein elution with 0.2 M of imidazole and 50 mM Tris–HCl (pH 7.5). Fusion proteins eluted from the column (2 mL) were dialyzed against 1 L buffer containing 0.5 M NaCl and 50 mM Tris–HCl (pH 7.5) at 4 °C for 16 h.

2.2. NMR spectroscopy

NMR experiments were performed on a Bruker AvanceDRX (500 MHz) or an Bruker Avance III (600 MHz) NMR spectrometer, the latter being equipped with a cryogenic triple-resonance probe. To determine the structure in MeOH, ~400 μg of ^{13}C/^{15}N-VIP-G was dissolved in 0.25 mL of 50%–50% H_2O–d_3-MeOH containing 20 mM Tris–HCl (pH 7.4). For structural determination in a DPC micelle, ~500 μg of ^{15}N- or ^{13}C/^{15}N-VIP-G was dissolved in 0.25 mL of 90%–10% H_2O–D_2O containing 50 mM potassium phosphate buffer at pH 7.2, with and without 50 mM DPC micelle. Sodium 2,2-dimethyl-2-silapentane-5-sulfonate was used as a reference for the chemical shift. Heteronuclear single quantum coherence (HSQC) spectra [33] adapted with a gradient sensitivity enhancement [34] were acquired with eight transients and 256 increments at 288 or 298 K, and zero filling during spectral processing. All NMR spectra were recorded at 288 (in MeOH) or 298 K (in DPC micelle). All spectra were processed using NMRPipe [35] and analyzed using the program nmrDraw [35]. Interproton distances were obtained from 3D ^{13}C- and ^{15}N-edited HSQC NOESY spectra recorded with a 200-ms (in MeOH) and 150-ms (in DPC micelle) mixing time. Structures were calculated using a standard seven iteration cycle protocol of the program CYANA version 2.0.17 [36,37]. All NOE cross-peaks were selected manually using SPARKY [38]. Dihedral angle restraints were calculated using the TALOS program based on backbone atom chemical shifts [39]. The structural coordinates were deposited to PDB (PDB ID: 2RRH for the structure in MeOH), (PDB ID: 2RRI for the micelle-bound VIP), and the assignments were deposited to BMRB (accession numbers = 11419 and 11420, respectively).

3. Results and discussion

3.1. Structure of VIP-G in 50% MeOH

Organic solvents such as MeOH or TFE have been used in conformational studies of VIP and its derivatives by CD and/or NMR analysis to determine their SARs. These solvents, either neat or mixed with water, mimic hydrophobic environments, such as the cell membrane or N-terminal ectodomain (N-ted) of VIP receptors. Wild-type VIP is a 28-amino acid peptide having a C-terminal amide moiety. However, Fahrenkrug and co-workers clearly showed that the amide group of the C-terminus of VIP was not very effective for receptor binding or activation [40]. Therefore, glycine-extended VIP (VIP-G) (Fig. 1A) was used for the NMR analysis in this study. We employed ^{13}C/^{15}N isotopically labeled VIP-G produced by *E. coli*.

Assignments for the backbone HN, N, Cα, Cβ, and Co resonances of VIP-G in 50% MeOH were achieved by analysis of the triple-resonance HNCACB, CBCA(CO)NH, and HNCO spectra according to the standard method [41]. Hα, the side-chain proton, and carbon resonances were assigned from the HCC(CO)NH and CC(CO)NH spectra. Fig. 1B shows the assigned ^{1}H-^{15}N HSQC spectra of VIP-G in 50% MeOH. All backbone resonances, except for His^1 and Ser^2, and 93% of the nonexchangeable protons of the side-chain signals were assigned. Backbone dihedral angle restraints were determined using the TALOS program [39]. TALOS showed an α-helix in VIP-G between Ala^4 and Ser^{25} in 50% MeOH.

The final ensemble of 20 low-energy structures of VIP-G in 50% MeOH was generated from a total of 395 experimental constraints derived by NMR. Many medium-range NOEs showing helix, $d_{\alpha N}(i, i+3)$, $d_{\alpha\beta}(i, i+3)$, and $d_{\alpha N}(i, i+4)$ were observed among residues Val^5–Gly^{29} (Supplementary Fig. 1A). The RMSD values of the backbone atoms evaluated for Ala^4–Asn^{28} of the 20 structures were 0.35 Å for the

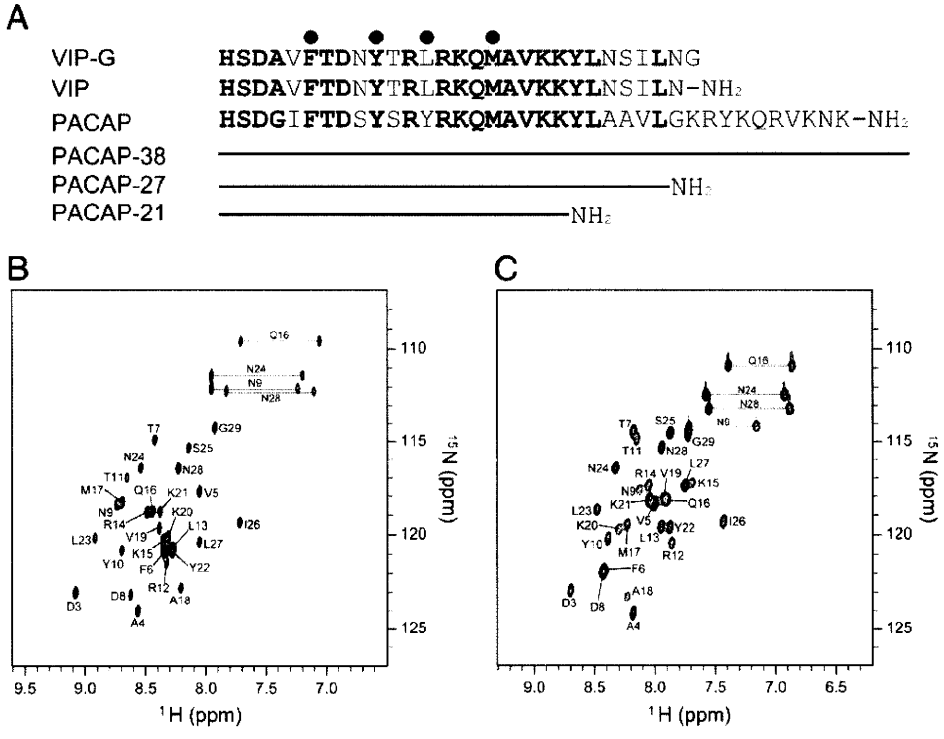


Fig. 1. Amino acid sequence and ^{1}H-^{15}N HSQC spectra. (A) Amino acid sequence of VIP and PACAP. Conserved residues are shown in bold. The filled circles indicate the residues forming a hydrophobic surface facing to DPC micelle. (B) and (C) ^{1}H-^{15}N HSQC spectra in 50% MeOH containing 20 mM Tris-HCl (pH 4.4) at 288 K (B) and in the presence of the 50 mM DPC micelle containing 50 mM potassium phosphate buffer (pH 7.2) at 298 K (C). Spectral cross-peaks are labeled by a 1-letter amino acid code and residue number.

backbone atoms and 0.95 Å for all the heavy atoms. The Ramachandran plot for all refined 20 structures for VIP-G shows that the backbone dihedral angles occupy the most favored or additionally allowed regions (Table 1). Superposition of the backbones of the 20 lowest target function structures in 50% MeOH is shown in Fig. 2. The structure of VIP-G in MeOH contains a central α-helix region comprising residues 4–29 and unstructured N-terminal regions (residues 1–3) (Fig. 2B).

Tan et al. previously reported the structure of VIP in 30% TFE [31]. The α-helical contents in 50% MeOH are similar to that in 30% TFE, which is also consistent with CD data. In contrast, the side-chain conformation is different in 50% MeOH and 30% TFE. In 30% TFE, several side-chain–side-chain interactions have been reported (Asn^{9}–Arg^{12}, Leu^{13}–Met^{17}, Val^{19}–Leu^{23}, and Lys^{20}–Asn^{24}). However, there were no side-chain interactions in this experiment, although the side-chain structure is well converged (RMSD values <1.0 Å).

Table 1
Structural statistics for the final 20 structures of VIP in DPC and 50% MeOH.

	MeOH	DPC
Distance restraints		
Total number of NOE restraints	395	376
Intraresidue	Unused	Unused
Sequential restraints ($\|i-j\|=1$)	246	246
Medium-range restraints ($1<\|i-j\|\leq 4$)	149	129
Long-range restraints ($\|i-j\|>5$)	0	1
Dihedral angle restraints		
$\varphi/\psi/\chi$	23/23/0	22/22/0
Hydrogen bond restraints	0	0
Final statistics		
Maximum target function	0.05	0.47
RMSD for experimental structure (Å)		
All backbone atoms (4–28 in MeOH, 6–26 in DPC)	0.35	0.55
All heavy atoms (4–28 in MeOH, 6–26 in DPC)	0.93	0.93
Ramachandran plot statistics (%)		
Most favored region	97.0	89.1
Additionally allowed region	3.0	10.9
Generously allowed region	0.0	0.0
Disallowed region	0.0	0.0

3.2. Structure of VIP-G bound to DPC micelle

We then determined the solution structure of VIP-G in the presence of the 50 mM DPC micelle (peptide/DPC ratio, 1:84). DPC is a well-known membrane-mimicking compound and is often used in solution NMR studies of peptide–membrane interactions. A DPC micelle is composed of ~60 DPC molecules in water. DPC has a neutral head group and mimics the nonisotropic environment of a lipid membrane. DPC forms stable micelles, resulting in reasonable correlation time and manageable line width for solution NMR studies. A remarkable difference in the DPC micelle compared with planar lipid bilayers is that the micelles possess a highly curved surface.

The ^{1}H-^{15}N HSQC spectra of VIP-G in the presence of DPC micelles are shown in Fig. 1C. We observed broader cross-peaks on the association of VIP-G to the DPC micelles (molecular mass of the DPC micelle is approximately 20 kDa). When more DPC was added to this sample (peptide/micelle ratio, 1:3), NMR signals were unchanged. Therefore, these NMR signals were from only the bound form within this experimental condition. We assigned all backbone resonances, except for His^{1} and Ser^{2}, and 93% of the nonexchangeable protons of the side-chain signals. An ensemble of 20 structures with low CYANA target functions was constructed from 376 experimental NMR constraints. The RMSD values after superimposing backbone atoms of Phe^{6}–Ile^{26} were 0.55 Å and 0.93 Å for backbone and heavy atoms, respectively. The secondary structure of the micelle-bound VIP-G consisted of an α-helix around Phe^{6}–Ile^{26} (Fig. 3). The N-terminal

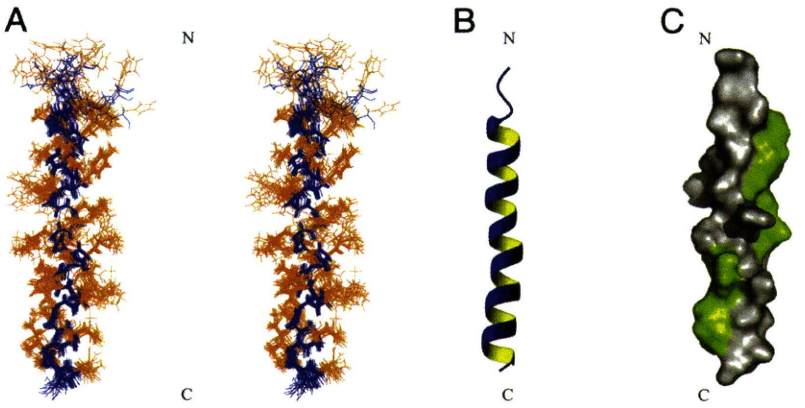


Fig. 2. Three-dimensional structures of VIP-G in 50% MeOH. (A) Ensemble of the 20 lowest energy structures of VIP-G in 50% MeOH. The backbone atoms of residues 4–28 were used for the superimposition. (B) and (C) Side views of the ribbon diagram (B) and surface model (C) of VIP-G of a representative conformer. The hydrophobic residues (V, I, L, Y, F, M) are colored green.

region (residues 1–5) was largely disordered, whereas the C-terminal region was more restricted and was close to an α-helical structure.

3.3. Comparison of 50% MeOH- and DPC micelle-bound forms of the VIP-G structure

To determine the interaction at the peptide–membrane interface, we first compared the chemical shift of VIP-G in DPC micelle to that in MeOH. Although the chemical shifts of several peaks were different in DPC and in MeOH, we could not determine the interaction at the VIP-G–micelle interface specifically from the chemical shift data. However, we determined the 3D structure of VIP-G in two different membrane-mimicking environments allowing us to compare two structures. Mapping of the hydrophobic residues on the molecular surface of VIP-G reveals that two hydrophobic patches are located at the opposite side of molecule (Figs. 2C and 3C). Compared with the structure of VIP-G in 50% MeOH, the α-helix of the micelle-bound form displayed a slight curvature (Fig. 4A and B). The diameter of this curvature was 21.84 ± 2.01 Å, which is similar to the reported size of a DPC micelle (Supplementary Fig. 2) [42]. The side chains of residues Phe^{6}, Tyr^{10}, Leu^{13}, and Met^{17} found at the concave face formed a hydrophobic patch in the micelle-bound state. VIP-G localized at the water–micelle boundary with these side chains toward the micelle interior. Thus, we inferred that this structural difference between the MeOH form and the micelle-bound form was caused by the peptide–micelle interaction on the curved micelle surfaces. Note that these two hydrophobic clusters were largely conserved within VIP and PACAP, which suggests the biological importance of the residues. The hydrophobic cluster consists of Ala^{18}, Val^{19}, Tyr^{22}, Leu^{23}, Ile^{26}, and Leu^{27} located at the peptide–N-ted interface on the structure of PACAP–PAC1R complex [25]. It is party consistent with an extensive SAR study on VIP, in which Ala-substituted mutants for each residue were examined [43]. Of course, direct proof of the peptide–lipid interaction of these residues requires further investigation by NMR experiments using spin-labeled DPC analogs.

3.4. Comparison of the structure of VIP-G and PACAP

The micelle- and receptor-bound structure of PACAP, which has high sequence homology with VIP, has been reported previously [24]. In this section, we compare our MeOH and micelle-bound forms of the VIP structure with other reported structures. Inooka et al. published a preliminary analysis of the micelle-bound structure of PACAP27, which adopts an α-helix (data not deposited on PDB). In addition, the micelle-bound structure of PACAP38 (PDB ID 2D2P), which comprises the flexible N-terminus (residues 1–4) followed by the

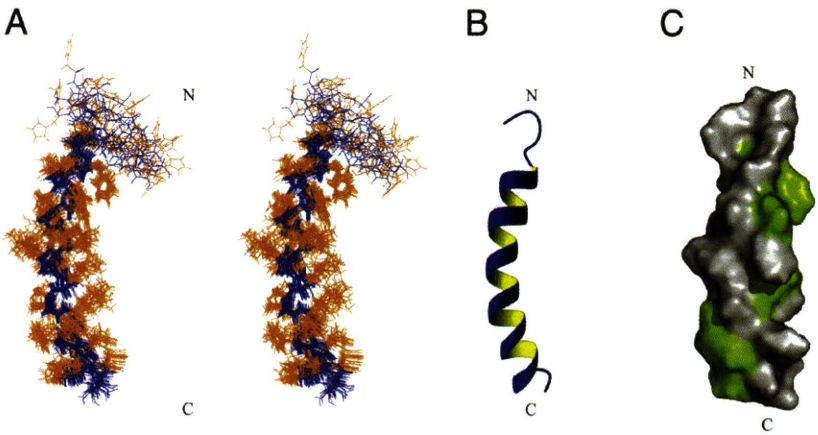


Fig. 3. Three-dimensional structures of VIP-G bound to a DPC micelle. (A) Ensemble of the 20 lowest energy structures of VIP-G in the DPC micelle. The backbone atoms of residues 6–26 were used for the superimposition. (B) and (C) Side views of the ribbon diagram (B) and surface model (C) of micelle-bound VIP-G of a representative conformer. The hydrophobic residues (V, I, L, Y, F, M) are colored green.

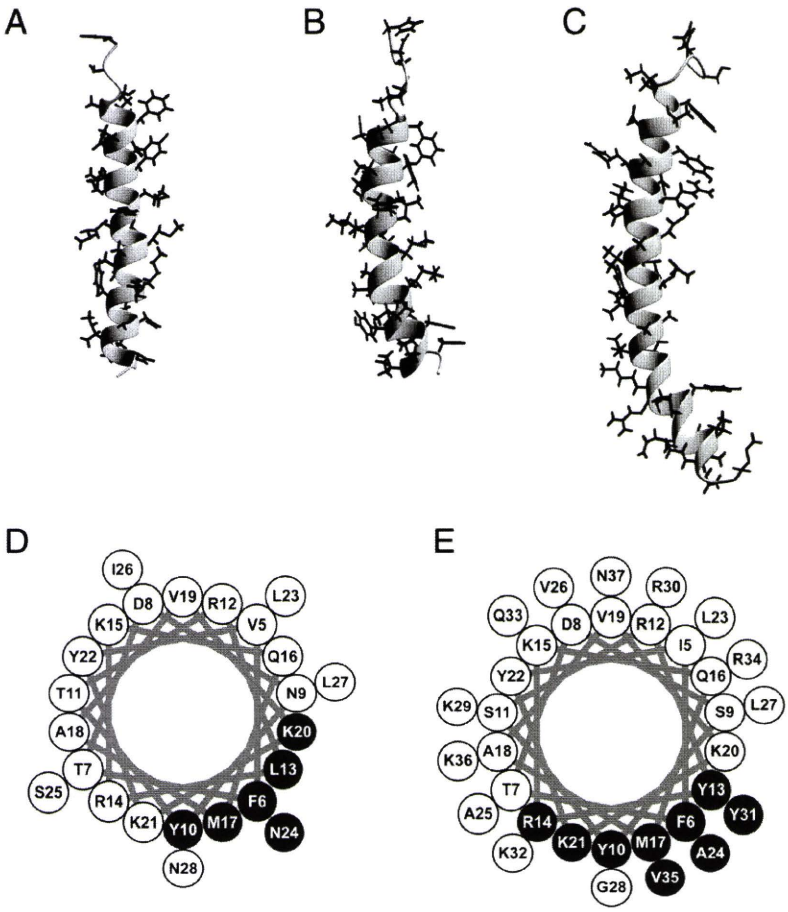


Fig. 4. Comparison of the structure of VIP-G and PACAP-38. (A–C) Side views of the ribbon diagram of VIP-G (in MeOH (A) and in the DPC micelle (B)) and PACAP-38 bound to the DPC micelle (C). (D and E) A helical wheel projections of the membrane-bound states of VIP (D) and PACAP (E). Amino acid residues located the peptide–membrane interface are shown in black.

stable α-helical region (5–38), is substantially similar to that of VIP-G (Fig. 4B and C). In particular, the C-terminal region of PACAP38 adopts a curved helix, which may reflect the curvature of the micelle. Moreover, the identical hydrophobic side chains (Phe^{6}, Tyr^{10}, Leu^{13}, and Met^{17}) form a micelle-binding interface inserted into the hydrophobic interior of the DPC micelle (Fig. 4D and E). This result indicates that VIP-G and PACAP have the same micelle-anchoring topology.

The conformations of VIP and PACAP in the lipid bilayer have also been studied, providing a better model for mimicking a cell membrane. A dimyristoylphosphatidylcholine/dimyristoylphosphatidylglycerol-based lipid bilayer was used as the peptide carrier to determine the structure of PACAP21 and PACAP27 using solid state NMR [44]. Interestingly, PACAP21 and PACAP27 adopted extended structures in the lipid bilayer-bound state. Gololobov et al. studied the conformation of VIP in liposomes using CD spectra and showed that VIP forms an α-helix in the presence of anionic phospholipid phosphatidylglycerol, whereas VIP remains unstructured in the presence of the neutral phospholipid phosphatidylcholine [30]. Although class II GPCR ligands usually adopt α-helices in the membrane-mimicking environment, VIP and PACAP vary their conformation depending on the environment, which might be unique. Considering the stable α-helical formation with an anionic lipid bilayer, some conserved cationic residues (Arg^{14}, Lys^{20}, and Lys^{21}) located near the peptide–micelle interface may play an important part in the conformational transition and α-helical stabilization.

3.5. Two-step ligand transportation mechanism of VIP in GPCR activation

As mentioned previously, the biological action of class II GPCR ligands is because of the ligation of peptides to the N-ted of the receptors. All the examples showed that these ligand peptides adopt an α-helical conformation on binding to the N-ted. A two-step ligand transportation model of GPCR activation was suggested [24,45]. Their model proposed that membrane interaction is an important event preceding receptor binding (Fig. 5). Peptide–membrane interaction induces an α-helical conformation closely homologous to the active one, leading to reduced losses in entropy upon receptor binding.

SAR was investigated by analyzing the biological activities of several variants of VIP. In particular, alanine-scanning studies revealed the key residues for VIP receptor binding or activation [43,46–48]. Modeling of the VIP structure in membrane-mimicking environments or with the receptor was also challenged [31,49]. Unfortunately, the interpretations of the existing pharmacological studies seemed to be controversial and did not succeed in perfectly explaining the mode of VIP action. Therefore, more structural information on VIP is required. In this study, we determined the micelle-bound VIP structure and found that four residues form a hydrophobic interface with the DPC micelle. Mutation on some of these residues showed substantial reduction in the activity of VIP, probably because they destabilized the α-helix. Thus, the residues in the peptide–membrane interface and those in the peptide–receptor interface must be separately considered in a SAR study.

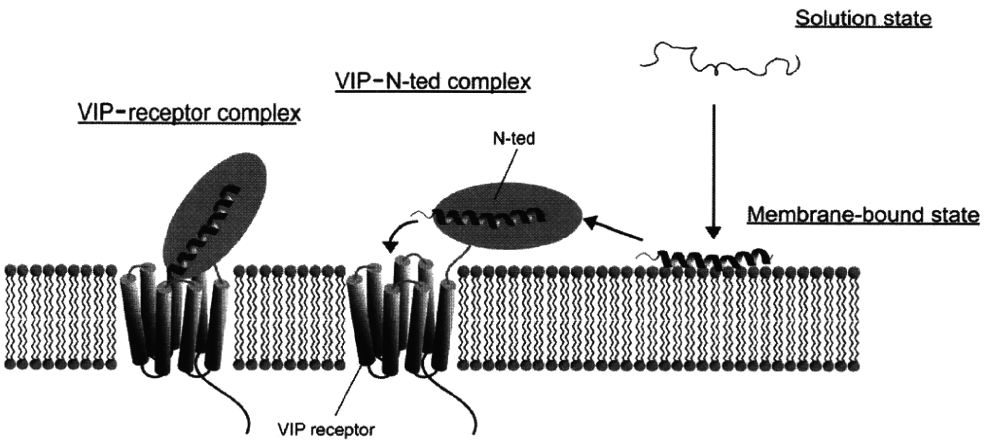


Fig. 5. Schematic diagram of the two-step ligand transportation model for VIP. The peptide nonspecifically binds to cell membranes and forms an α-helix at the C-terminal region. Then, VIP undergoes 2D diffusion to the receptor and binds to the N-ted of the VIP receptor. Finally, VIP is sandwiched between the N-ted and core of the receptor.

In conclusion, we have determined the solution structure of VIP in two distinct membrane-mimicking environments, both of which induced an α-helix in VIP. The data presented here will provide a rational basis for the development of VIP analogs, which will be useful in therapy for inflammatory diseases as well as cancer diagnoses.

Supplementary materials related to this article can be found online at doi:10.1016/j.bbapap.2011.03.009.

References

[1] B.H. Du, J. Eng, J.D. Hulmes, M. Chang, Y.C. Pan, R.S. Yalow, Guinea pig has a unique mammalian VIP, Biochem. Biophys. Res. Commun. 128 (1985) 1093–1098.

[2] S. Onoue, S. Misaka, S. Yamada, Structure–activity relationship of vasoactive intestinal peptide (VIP): potent agonists and potential clinical applications, Naunyn Schmiedebergs Arch. Pharmacol. 377 (2008) 579–590.

[3] M. Laburthe, A. Couvineau, V. Tan, Class II G protein-coupled receptors for VIP and PACAP: structure, models of activation and pharmacology, Peptides 28 (2007) 1631–1639.

[4] R.M. Campbell, C.G. Scanes, Evolution of the growth hormone-releasing factor (GRF) family of peptides, Growth Regul. 2 (1992) 175–191.

[5] G.V. Segre, S.R. Goldring, Receptors for secretin, calcitonin, parathyroid hormone (PTH)/PTH-related peptide, vasoactive intestinal peptide, glucagonlike peptide 1, growth hormone-releasing hormone, and glucagon belong to a newly discovered G-protein-linked receptor family, Trends Endocrinol. Metab. 4 (1993) 309–314.

[6] S.R. Rawlings, M. Hezareh, Pituitary adenylate cyclase-activating polypeptide (PACAP) and PACAP/vasoactive intestinal polypeptide receptors: actions on the anterior pituitary gland, Endocr. Rev. 17 (1996) 4–29.

[7] M. Laburthe, M. Rousset, C. Boissard, G. Chevalier, A. Zweibaum, G. Rosselin, Vasoactive intestinal peptide: a potent stimulator of adenosine 3':5'-cyclic monophosphate accumulation in gut carcinoma cell lines in culture, Proc. Natl Acad. Sci. U.S.A. 75 (1978) 2772–2775.

[8] R. Salomon, A. Couvineau, C. Rouyer-Fessard, T. Voisin, D. Lavallee, A. Blais, D. Darmoul, M. Laburthe, Characterization of a common VIP-PACAP receptor in human small intestinal epithelium, Am. J. Physiol. 264 (1993) 294–300.

[9] C.J. MacKenzie, E.M. Lutz, D.A. McCulloch, R. Mitchell, A.J. Harmar, Phospholipase C activation by VIP1 and VIP2 receptors expressed in COS 7 cells involves a pertussis toxin-sensitive mechanism, Ann. NY Acad. Sci. 805 (1996) 579–584.

[10] D.L. Bellinger, D. Lorton, S. Brouxhon, S. Felten, D.L. Felten, The significance of vasoactive intestinal polypeptide (VIP) in immunomodulation, Adv. Neuroimmunol. 6 (1996) 5–27.

[11] S.G. Smalley, P.A. Barrow, N. Foster, Immunomodulation of innate immune responses by vasoactive intestinal peptide (VIP): its therapeutic potential in inflammatory disease, Clin. Exp. Immunol. 157 (2009) 225–234.

[12] M. Kojima, T. Ito, T. Oono, T. Hisano, H. Igarashi, Y. Arita, K. Kawabe, D.H. Coy, R.T. Jensen, H. Nawata, VIP attenuation of the severity of experimental pancreatitis is due to VPAC1 receptor-mediated inhibition of cytokine production, Pancreas 30 (2005) 62–70.

[13] A. Arranz, C. Abad, Y. Juarranz, J. Leceta, C. Martinez, R.P. Gomariz, Vasoactive intestinal peptide as a healing mediator in Crohn's disease, Neuroimmunomodulation 15 (2008) 46–53.

[14] E. Gonzalez-Rey, M. Delgado, Therapeutic treatment of experimental colitis with regulatory dendritic cells generated with vasoactive intestinal peptide, Gastroenterology 131 (2006) 1799–1811.

[15] M. Delgado, C. Martinez, D. Pozo, J.R. Calvo, J. Leceta, D. Ganea, R.P. Gomariz, Vasoactive intestinal peptide (VIP) and pituitary adenylate cyclase-activation polypeptide (PACAP) protect mice from lethal endotoxemia through the inhibition of TNF-alpha and IL-6, J. Immunol. 162 (1999) 1200–1205.

[16] M. Delgado, C. Abad, C. Martinez, J. Leceta, R.P. Gomariz, Vasoactive intestinal peptide prevents experimental arthritis by downregulating both autoimmune and inflammatory components of the disease, Nat. Med. 7 (2001) 563–568.

[17] M.L. Thakur, C.S. Marcus, S. Saeed, V. Pallela, C. Minami, L. Diggles, H. Le Pham, R. Ahdoot, E.A. Kalinowski, 99mTc-labeled vasoactive intestinal peptide analog for rapid localization of tumors in humans, J. Nucl. Med. 41 (2000) 107–110.

[18] M.L. Thakur, M.R. Aruva, J. Gariepy, P. Acton, S. Rattan, S. Prasad, E. Wickstrom, A. Alavi, PET imaging of oncogene overexpression using 64Cu-vasoactive intestinal peptide (VIP) analog: comparison with 99mTc-VIP analog, J. Nucl. Med. 45 (2004) 1381–1389.

[19] J.C. Reubi, In vitro identification of vasoactive intestinal peptide receptors in human tumors: implications for tumor imaging, J. Nucl. Med. 36 (1995) 1846–1853.

[20] C. Boissard, J.C. Marie, G. Hejblum, C. Gespach, G. Rosselin, Vasoactive intestinal peptide receptor regulation and reversible desensitization in human colonic carcinoma cells in culture, Cancer Res. 46 (1986) 4406–4413.

[21] I. Rubinstein, Human VIP-alpha: an emerging biologic response modifier to treat primary pulmonary hypertension, Expert Rev. Cardiovasc. Ther. 3 (2005) 565–569.

[22] V. Sethi, H. Onyuksel, I. Rubinstein, Liposomal vasoactive intestinal peptide, Meth. Enzymol. 391 (2005) 377–395.

[23] F. Hajos, B. Stark, S. Hensler, R. Prassl, W. Mosgoeller, Inhalable liposomal formulation for vasoactive intestinal peptide, Int. J. Pharm. 357 (2008) 286–294.

[24] H. Inooka, T. Ohtaki, O. Kitahara, T. Ikegami, S. Endo, C. Kitada, K. Ogi, H. Onda, M. Fujino, M. Shirakawa, Conformation of a peptide ligand bound to its G-protein coupled receptor, Nat. Struct. Biol. 8 (2001) 161–165.

[25] C. Sun, D. Song, R.A. Davis-Taber, L.W. Barrett, V.E. Scott, P.L. Richardson, A. Pereda-Lopez, M.E. Uchic, L.R. Solomon, M.R. Lake, K.A. Walter, P.J. Hajduk, E.T. Olejniczak, Solution structure and mutational analysis of pituitary adenylate cyclase-activating polypeptide binding to the extracellular domain of PAC1-RS, Proc. Natl Acad. Sci. U.S.A. 104 (2007) 7875–7880.

[26] C.R. Grace, M.H. Perrin, J. Gulyas, M.R. Digruccio, J.P. Cantle, J.E. Rivier, W.W. Vale, R. Riek, Structure of the N-terminal domain of a type B1 G protein-coupled receptor in complex with a peptide ligand, Proc. Natl Acad. Sci. U.S.A. 104 (2007) 4858–4863.

[27] S. Runge, H. Thogersen, K. Madsen, J. Lau, R. Rudolph, Crystal structure of the ligand-bound glucagon-like peptide-1 receptor extracellular domain, J. Biol. Chem. 283 (2008) 11340–11347.

[28] C. Parthier, M. Kleinschmidt, P. Neumann, R. Rudolph, S. Manhart, D. Schlenzig, J. Fanghanel, J.U. Rahfeld, H.U. Demuth, M.T. Stubbs, Crystal structure of the incretin-bound extracellular domain of a G protein-coupled receptor, Proc. Natl Acad. Sci. U.S.A. 104 (2007) 13942–13947.

[29] A.A. Pioszak, H.E. Xu, Molecular recognition of parathyroid hormone by its G protein-coupled receptor, Proc. Natl Acad. Sci. U.S.A. 105 (2008) 5034–5039.

[30] G. Gololobov, Y. Noda, S. Sherman, I. Rubinstein, J. Baranowska-Kortylewicz, S. Paul, Stabilization of vasoactive intestinal peptide by lipids, J. Pharmacol. Exp. Ther. 285 (1998) 753–758.

[31] Y.V. Tan, A. Couvineau, S. Murail, E. Ceraudo, J.M. Neumann, J.J. Lacapere, M. Laburthe, Peptide agonist docking in the N-terminal ectodomain of a class II G protein-coupled receptor, the VPAC1 receptor. Photoaffinity, NMR, and molecular modeling, J. Biol. Chem. 281 (2006) 12792–12798.

[32] T. Tenno, N. Goda, Y. Tateishi, H. Tochio, M. Mishima, H. Hayashi, M. Shirakawa, H. Hiroaki, High-throughput construction method for expression vector of peptides for NMR study suited for isotopic labeling, Protein Eng. Des. Sel. 17 (2004) 305–314.

[33] D. Neri, T. Szyperski, G. Otting, H. Senn, K. Wuthrich, Stereospecific nuclear magnetic resonance assignments of the methyl groups of valine and leucine in the DNA-binding domain of the 434 repressor by biosynthetically directed fractional 13C labeling, Biochemistry 28 (1989) 7510–7516.

[34] L.E. Kay, P. Keifer, T. Saarinen, Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity, J. Am. Chem. Soc. 114 (1992) 10663–10665.
[35] F. Delaglio, S. Grzesiek, G.W. Vuister, G. Zhu, J. Pfeifer, A. Bax, NMRPipe: a multidimensional spectral processing system based on UNIX pipes, J. Biomol. NMR 6 (1995) 277–293.
[36] T. Herrmann, P. Guntert, K. Wuthrich, Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA, J. Mol. Biol. 319 (2002) 209–227.
[37] P. Guntert, Automated NMR protein structure calculation, Prog. Nucl. Magn. Reson. Spectrosc. 43 (2003) 105–125.
[38] T. D. Goddard, D. G. Kneller, Sparky 3. San Francisco: University of California.
[39] G. Cornilescu, F. Delaglio, A. Bax, Protein backbone angle restraints from searching a database for chemical shift and sequence homology, J. Biomol. NMR 13 (1999) 289–302.
[40] J. Fahrenkrug, B. Ottesen, C. Palle, Non-amidated forms of VIP (glycine-extended VIP and VIP-free acid) have full bioactivity on smooth muscle, Regul. Pept. 26 (1989) 235–239.
[41] J. Cavanagh, W.J. Fairbrother III, A.G. Palmer, M. Rance, N.J. Skelton, Protein NMR Spectroscopy: Principles and Practice, 2nd Ed, Academic Press, San Diego, 2007, pp. 535–673.
[42] F. Chevalier, J. Lopez-Prados, P. Groves, S. Perez, M. Martin-Lomas, P.M. Nieto, Structure and dynamics of the conserved protein GPI anchor core inserted into detergent micelles, Glycobiology 16 (2006) 969–980.
[43] P. Nicole, L. Lins, C. Rouyer-Fessard, C. Drouot, P. Fulcrand, A. Thomas, A. Couvineau, J. Martinez, R. Brasseur, M. Laburthe, Identification of key residues for interaction of vasoactive intestinal peptide with human VPAC1 and VPAC2 receptors and development of a highly selective VPAC1 receptor agonist. Alanine scanning and molecular modeling of the peptide, J. Biol. Chem. 275 (2000) 24003–24012.
[44] N. Komi, K. Okawa, Y. Tateishi, M. Shirakawa, T. Fujiwara, H. Akutsu, Structural analysis of pituitary adenylate cyclase-activating polypeptides bound to phospholipid membranes by magic angle spinning solid-state NMR, Biochim. Biophys. Acta 1768 (2007) 3001–3011.
[45] S.R.J. Hoare, Mechanisms of peptide and nonpeptide ligand binding to Class B G-protein-coupled receptors, Drug Discov. Today 10 (2005) 417–427.
[46] M. O'Donnell, R.J. Garippa, N.C. O'Neill, D.R. Bolin, J.M. Cottrell, Structure–activity studies of vasoactive intestinal polypeptide, J. Biol. Chem. 266 (1991) 6389–6392.
[47] H. Igarashi, T. Ito, W. Hou, S.A. Mantey, T.K. Pradhan, C.D. Ulrich, S.J. Hocart, D.H. Coy, R.T. Jensen, Elucidation of vasoactive intestinal peptide pharmacophore for VPAC(1) receptors in human, rat, and guinea pig, J. Pharmacol. Exp. Ther. 301 (2002) 37–50.
[48] H. Igarashi, T. Ito, S.A. Mantey, T.K. Pradhan, W. Hou, D.H. Coy, R.T. Jensen, Development of simplified vasoactive intestinal peptide analogs with receptor selectivity and stability for human vasoactive intestinal peptide/pituitary adenylate cyclase-activating polypeptide receptors, J. Pharmacol. Exp. Ther. 315 (2005) 370–381.
[49] E. Ceraudo, S. Murail, Y.V. Tan, J.J. Lacapere, J.M. Neumann, A. Couvineau, M. Laburthe, The vasoactive intestinal peptide (VIP) alpha-helix up to C terminus interacts with the N-terminal ectodomain of the human VIP/Pituitary adenylate cyclase-activating peptide 1 receptor: photoaffinity, molecular modeling, and dynamics, Mol. Endocrinol. 22 (2008) 147–155.

Published online 3 November 2010 *Nucleic Acids Research, 2011, Vol. 39, Database issue* ***D487–D493***
doi:10.1093/nar/gkq1057

SAHG, a comprehensive database of predicted structures of all human proteins

Chie Motono[1,2,*], Junichi Nakata[1,2], Ryotaro Koike[2,3], Kana Shimizu[1,2], Matsuyuki Shirota[2,4], Takayuki Amemiya[2,5], Kentaro Tomii[1,2], Nozomi Nagano[1,2], Naofumi Sakaya[1,2,6], Kiyotaka Misoo[1,2,6], Miwa Sato[1,2,5,7], Akinori Kidera[2,5,8], Hidekazu Hiroaki[2,9], Tsuyoshi Shirai[2,10], Kengo Kinoshita[2,4], Tamotsu Noguchi[1,2] and Motonori Ota[2,3,*]

[1]Computational Biology Research Center (CBRC), National Institute of Advanced Industrial Science and Technology (AIST), Tokyo 135-0064, [2]Institute for Bioinformatics Research and Development (BIRD), Japan Science and Technology Agency (JST), Tokyo 102-0081, [3]Graduate School of Information Science, Nagoya University, Nagoya 464-8601, [4]Graduate School of Information Science, Tohoku University, Sendai 980-8579, [5]Department of Supramolecular Biology, Yokohama City University, Yokohama 230-0045, [6]Information and Mathematical Science Laboratory Inc., Tokyo 112-0012, [7]Mitsui Knowledge Industry Co., Ltd, Tokyo 105-6215, [8]Department of Computational Science Research Program, RIKEN, Wako 351-0198, [9]Graduate School of Medicine, Kobe University, Kobe 650-0017 and [10]Department of Bioscience, Nagahama Institute of Bioscience and Technology, Nagahama 526-0829, Japan

Received August 15, 2010; Revised October 2, 2010; Accepted October 13, 2010

ABSTRACT

Most proteins from higher organisms are known to be multi-domain proteins and contain substantial numbers of intrinsically disordered (ID) regions. To analyse such protein sequences, those from human for instance, we developed a special protein-structure-prediction pipeline and accumulated the products in the Structure Atlas of Human Genome (SAHG) database at http://bird.cbrc.jp/sahg. With the pipeline, human proteins were examined by local alignment methods (BLAST, PSI-BLAST and Smith–Waterman profile–profile alignment), global–local alignment methods (FORTE) and prediction tools for ID regions (POODLE-S) and homology modeling (MODELLER). Conformational changes of protein models upon ligand-binding were predicted by simultaneous modeling using templates of apo and holo forms. When there were no suitable templates for holo forms and the apo models were accurate, we prepared holo models using prediction methods for ligand-binding (eF-seek) and conformational change (the elastic network model and the linear response theory). Models are displayed as animated images. As of July 2010, SAHG contains 42 581 protein-domain models in approximately 24 900 unique human protein sequences from the RefSeq database. Annotation of models with functional information and links to other databases such as EzCatDB, InterPro or HPRD are also provided to facilitate understanding the protein structure-function relationships.

INTRODUCTION

Nowadays, genome sequencing projects are producing complete genome sequences at an extremely high rate (1,2). With the rise of next-gen sequencers (3–5), this is the continuous trend for the future without a doubt. Consequently, the number of known protein sequences (6) grows more rapidly than the number of known protein structures experimentally determined (7). However, to make full use of genome sequences, proteins encoded in genomes should be analysed and for this purpose, protein three-dimensional (3D) structures provide much information (8,9). Computational methods for protein 3D structure prediction are anticipated to

*To whom correspondence should be addressed. Tel: +81 3 3599 8067; Fax: +81 3 3599 8081; Email: c-motono@aist.go.jp
Correspondence may also be addressed to Motonori Ota. Tel: +81 52 789 4782; Fax: +81 52 789 4782; Email: mota@is.nagoya-u.ac.jp

The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

bridge the gap between the number of known protein sequences and the number of known protein structures. According to assessments of the accuracy of those methods, e.g. recent Critical Assessment of Techniques for Protein Structure Prediction (CASP) experiments (10,11), template-based protein structure prediction often produced 3D models accurate enough for functional annotations, modification of protein functions or even for structure-based drug design (12,13). In addition, in the CASP7 and 8 experiments, fully automated structure prediction methods had reached a comparable level to the best prediction performance by methods with human intervention (14).

In the CASP experiments, target protein sequences are ones whose 3D structures will be determined. It means that such protein structures are expected to be single domains or a couple of domains and suitable for the experimental structure determination. Therefore, sometimes protein sequences are truncated from their full-length forms. On the other hand, most protein sequences coded in genomes from higher organisms are known to be long and should be multi-domain proteins (15), and contain a significant portion of intrinsically disordered (ID) regions (16–19). Clearly, these proteins are unsuitable for experimental structure determination in the full-length form and distinct from the target protein sequences of CASPs. To analyse such proteins, we have developed a special protein-structure-prediction pipeline, by integrating and arranging various computational tools, either developed by us or widely used as global standards. This pipeline was applied to all proteins coded in the human genome. The resulting 3D models as well as other annotations for protein functions were accumulated in the Structural Atlas of Human Genome (SAHG) database and presented through the web interface at http://bird.cbrc.jp/sahg.

There are other databases of protein structure models, e.g. SWISS-MODEL Repository (20) or ModBase (21). Both databases contain annotated protein structure models generated by original automated modeling pipelines. They also allow the users to build models on demand. Compared with them, the SAHG database is distinct mainly in the following points: (i) The 3D models in SAHG were generated by an original pipeline, specific for multi-domain proteins with substantial ID regions; (ii) Conformational changes of proteins upon ligand-binding are predicted by simultaneous modeling using templates of the ligand-bound state (holo form) and the unbound state (apo form) and displayed as animated images; and (iii) Functional annotations for protein interactions, e.g. ligand-binding and protein–protein interactions, are available. All these features are suitable for analysing eukaryotic proteins toward a deep understanding of their functions and interactions.

PREDICTION SCHEME AND CONTENTS

Overview

Schematically, two types of prediction systems were used to analyse protein sequences [RefSeq sequence (22)] automatically. One is the 'Structure prediction pipeline' (right pink regions in Figure 1) in which several homology search and protein structure prediction tools, conducting sequence–sequence, sequence–profile and profile–profile alignments, are combined sequentially, and it processes protein sequences, assigns them with 3D templates and finally produces 3D models. If available, 3D models of apo and holo forms were generated. The other components are 'Other structure and function predictors' (bottom light blue regions in Figure 1). They are an ensemble of independent prediction tools, which analyse protein sequences. All the results from these systems were accumulated in SAHG in XML formats.

Structure prediction pipeline

Construction of 3D models. Protein structure prediction consists of the following procedures: template searches and selection, alignment of target sequence and template, building 3D models and evaluation of model quality.

The template searches and their assignments to a target protein are the 'step-wise-multi-methods' approach. In the first step, a BLAST (23) search against all the latest Protein Data Bank (PDB) (7) and Structure Classification of Proteins (SCOP) (24,25) sequences is performed with 10^{-5} *E*-value cut-off. We selected templates, at least 90% of whose sequence could be aligned with the target, to ensure that the 3D models corresponded to stable domains or proteins. The resulting target sequence-template alignments were ranked based on their *E*-values. The best combination of templates for each domain was determined using an original algorithm to maximize the coverage of the target sequence (label I in Figure 1). In the second step, a PSI-BLAST (23) search with the same parameters was conducted for the remaining regions of the target sequence, where no models had been assigned and the best templates were assigned onto the target sequence (II in Figure 1). Protein sequence profiles were prepared using the latest NCBI-nr database. In the third step, a Smith–Waterman profile–profile alignment method (SWPPA) (26) was applied to the remaining regions against restricted templates (SCOP and PDB subsets with less than 40% sequence identity) with a cut-off of Z-score > 10, the comparable threshold to E-value $< 10^{-5}$ in PSI-BLAST (III in Figure 1). Finally, the FORTE (27) search, a profile–profile comparison method, was performed for the remaining regions, with a strict cut-off of Z-score > 20, to detect distantly related templates (V in Figure 1). FORTE is based on the global–local alignment method and was adjusted to perform best (28) when the target proteins were almost the same length as the PDB entries (around 400 aa) (29). However, more than half of human proteins (53%) are larger than 400 amino acids and even the remaining regions are sometimes over 2000 amino acids. Thus, prior to the FORTE search, potential domains were carved out from the remaining regions using an algorithm based on the prediction of ID regions (IV in Figure 1) and fed into FORTE (see 'Prediction of potential domains' section for details).

Once the target sequence-template alignments were obtained, all templates were checked against our 'apo