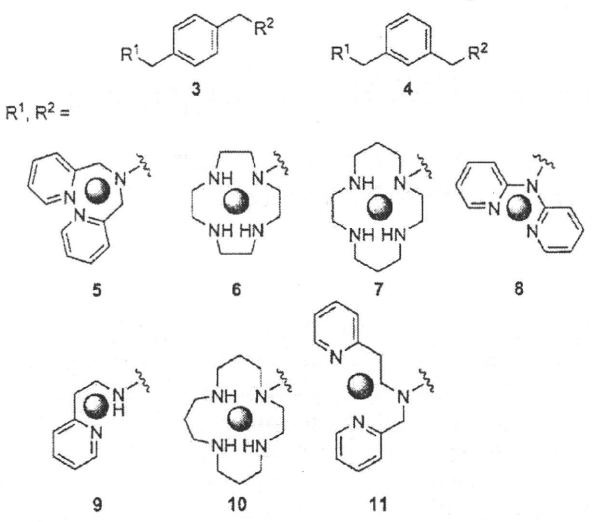


Figure 1. The structures of aromatic spacers (upper) and cationic moieties (R^1 and R^2). The shaded circle represents the position of the metal cation (Zn^{II} or Cu^{II}) in the chelate.

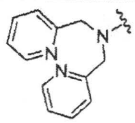
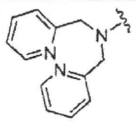
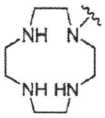
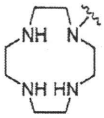
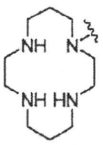
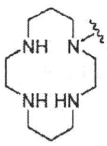
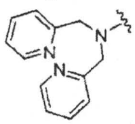
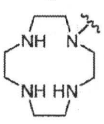
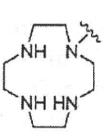
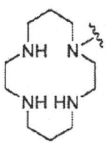
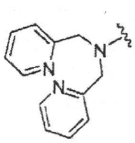
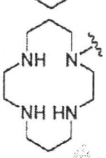
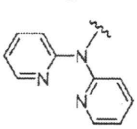
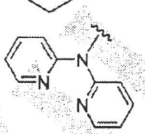
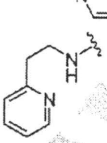
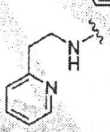
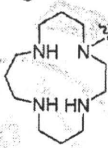
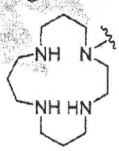
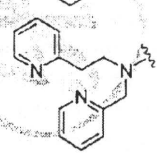
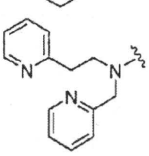
Jurkat cells, which express CXCR4.[38] The percent inhibition of all compounds at 1 μM is shown in Table 1. Seven compounds (**16**, **17**, **20–22**, **28**, and **29**, Table 1) resulted in greater than 87% inhibition. The high activity of **16** is consistent with results reported previously.[20,22] The anti-HIV activities of **17** and **29**, which contain only cyclam or cyclal rings, were reported by De Clercq et al.[39,40] Compounds with only pyridine and/or cyclen rings did not show any high binding activity. The presence of azamacrocyclic rings is presumably indispensable to the interaction of these compounds with CXCR4, and the size of rings appears to be important because not only compounds **16** and **17**, with two cyclam rings in the molecule, but also compounds **28** and **29**, with two cyclal rings, have remarkably more potent CXCR4 binding activity than compounds **14** and **15**, which have two cyclen rings. Compound **22**, with a *p*-xylene moiety, exhibited higher activity than compound **23**, which has an *m*-xylene moiety, indicating that *p*-xylene is more suitable than *m*-xylene as a spacer for approximate positioning of cationic moieties. At 0.1 μM, compound **22** resulted in 86% inhibition of [^{125}I]CXCL12 binding, while the other six compounds exhibited 37–66% inhibition. The IC_{50} value of compound **22** was estimated to be 37 nM.

$ZnCl_2$ was added to phosphate-buffered saline (PBS) solutions of these 20 compounds, **12–31**, to form zinc(II) complexes. The percent inhibition for each compound at 1 μM against [^{125}I]CXCL12 binding was determined and is given in Table 1. Zinc complexation of **12–15**, **18**, **19**, and **23** resulted in a remarkable increase in CXCR4 binding activity compared to the corresponding zinc-free compounds. These molecules contain dipicolylamine and/or cyclen moieties, suggesting that chelation of the nitrogen atoms with the zinc(II) ion significantly affects their interactions with CXCR4. The high activity of the zinc chelates of **12** and **13** is consistent with results provided in our previous paper.[37] Additionally, the anti-HIV activity of zinc complexes of **14** and **15** was reported by Kimura et al.[41] For compounds with only dipicolylamine and/or cyclen macrocycles as cationic moieties (**12–15**, **18**, and **19**), zinc complexation is critical to achieve high binding activity; the corresponding zinc-free compounds exhibit no significant activity. Compounds **16**, **17**, **20–22**, **28**, and **29** demonstrated high binding affinity in metal-free states as well as in zinc complexation states, indicating that zinc complexation of either of the macrocyclic rings in these compounds is not essential for high activity. The CXCR4 binding activity and anti-HIV activity of the zinc complex of **16** were reported previously.[42,43] Measured inhibition percentages for 0.1 μM of the zinc complexes of **12**, **14–23**, **28**, and **29** are given in Table 1. The zinc complexes of **20–22**, **28**, and **29** at 0.1 μM exhibited greater than 79% inhibition of [^{125}I]CXCL12 binding, and the other eight zinc complexes (of **12**, **14–19**, and **23**) showed less than 55% inhibition. The IC_{50} values of zinc complexes of **20–22**, **28**, and **29** were estimated to be 11, 8.3, 22, 40, and 52 nM, respectively. Zinc complexes of compounds containing a combination of cyclen and cyclam moieties, **20** and **21**, had remarkably small ■■low? potent■■ IC_{50} values.

To form chelates with a copper(II) cation, $CuCl_2$ was added to solutions in PBS of **12-31**. The inhibition percentages of all the compounds at 1 μM against [^{125}I]CXCL12 binding are shown in Table 1. Copper complexes of **14** and **15** exhibited a significant increase in CXCR4 binding activity as compared to the corresponding copper-free compounds, a phenomenon which is also seen in the zinc chelates. These compounds have two cyclen moieties in the molecules, suggesting that zinc or copper complexation is critical for high binding activity. Compounds **16**, **17**, and **20–22** showed high binding affinity in metal-free states and zinc- and copper-complexed states, indicating that metallic complexation of the cyclam rings in these compounds is not necessary for high activity. The CXCR4 binding activity of the copper complex of **16** was previously reported.[42] For compounds **17**, **22**, **23**, **28**, and **29**, copper complexation caused a significant decrease in binding activity compared to the corresponding copper-free compounds, whereas for compounds **14**, **15**, **18**, and **19**, copper complexation caused an increase in binding activity. This phenomenon may be due to the difference in ring sizes and structures of macrocycles, and was not observed upon zinc-complex formation. Inhibition at 0.1 μM of the copper complexes of **16** and **20–22**, which exhibited greater than 85% inhibition of [^{125}I]CXCL12 binding at 1 μM, are given in Table 1. The copper complexes of **16**, **20**, **21**, and **22** at 0.1 μM showed 39, 69, 88, and 39% inhibition, respectively, with the IC_{50} value of the copper complex of **21** estimated to be 16 nM.

Molecular modeling analysis of compound **21** and its zinc(II) and copper(II) complexes predicted that these complexes would form a stable coordinate conformation as shown in Figure 2. In general, zinc(II) complexes ■■are predicted to■■ adopt a tetrahedral conformation, while copper(II) complexes form a planar four coordinate/square conformation. The zinc(II) complex of **21** has■■is predicted to have■■ a tetrahedral conformation and the copper(II) complex a square planar conformation in both the cyclen and cyclam rings. The carboxyl group of either Asp171 or Asp262 in CXCR4 is

Table 1. CXCR4 binding activity of compounds **12–31** in the metal ion-free form, the zinc complex, and the copper complex.

Compd	Spacer	R^1	R^2	Metal free			Zinc complex			Copper complex		
				Inhibition[a] [%]		IC_{50}[b]	Inhibition[a] [%]		IC_{50}[b]	Inhibition[a] [%]		IC_{50}[b]
				1 μM	0.1 μM	[nM]	1 μM	0.1 μM	[nM]	1 μM	0.1 μM	[nM]
12	*p*-xylene			0	n. d.	n. d.	83±2	24±5	n. d.	10±4	n. d.	n. d.
13	*m*-xylene			0	n. d.	n. d.	31±3	n. d.	n. d.	0	n. d.	n. d.
14	*p*-xylene			30±4	n. d.	n. d.	87±4	0	n. d.	60±2	n. d.	n. d.
15	*m*-xylene			33±2	n. d.	n. d.	94±1	13±6	n. d.	80±3	n. d.	n. d.
16	*p*-xylene			94±4	59±6	n. d.	97±5	28±3	n. d.	98±1	39±3	n. d.
17	*m*-xylene			95±3	49±9	n. d.	98±4	55±7	n. d.	75±1	n. d.	n. d.
18	*p*-xylene			32±0.7	n. d.	n. d.	97±6	0	n. d.	52±3	n. d.	n. d.
19	*m*-xylene			17±5	n. d.	n. d.	91±4	0	n. d.	22±6	n. d.	n. d.
20	*p*-xylene			89±3	62±3	n. d.	>100	79±1	11	>100	69±3	n. d.
21	*m*-xylene			89±3	66±3	n. d.	92±3	>100	8.3	>100	88±1	16
22	*p*-xylene			94±3	86±3	37	99±8	79±0.6	22	85±3	39±3	n. d.
23	*m*-xylene			58±8	n. d.	n. d.	90±17	37±0.3	n. d.	48±4	n. d.	n. d.
24	*p*-xylene			3±0.9	n. d.	n. d.	0	n. d.	n. d.	0	n. d.	n. d.
25	*m*-xylene			4±3	n. d.	n. d.	0	n. d.	n. d.	0	n. d.	n. d.
26	*p*-xylene			14±2	n. d.	n. d.	10±3	n. d.	n. d.	0	n. d.	n. d.
27	*m*-xylene			10±3	n. d.	n. d.	10±4	n. d.	n. d.	0	n. d.	n. d.
28	*p*-xylene			91±0.4	37±0.9	n. d.	97±4	>100	40	57±4	n. d.	n. d.
29	*m*-xylene			87±2	50±1	n. d.	>100	91±4	52	55±1	n. d.	n. d.
30	*p*-xylene			0	n. d.	n. d.	14±3	n. d.	n. d.	14±3	n. d.	n. d.
31	*m*-xylene			24±2	n. d.	n. d.	20±3	n. d.	n. d.	0	n. d.	n. d.
FC-131	*cyclo*-[D-Tyr-Arg-Arg-Nal-Gly-]			100	100	1.8	–	–	–	–	–	–

[a] CXCR4 binding activity was assessed based on inhibition of [^{125}I]CXCL12 binding to Jurkat cells. Percent inhibition for all compounds at 1 and 0.1 μM were calculated relative to the percent inhibition by FC131 (100%). [b] IC_{50} values are the concentrations which correspond to 50% inhibition of [^{125}I]CXCL12 binding to Jurkat cells. All data are mean■■±SD? ±SEM?■■ values of at least three independent experiments. n.d = not determined.

thought to coordinate strongly with zinc ions but not copper ions in the complexes,[41–43] and as a consequence, the zinc complex of **21** would bind more strongly than **21** or its copper complex■■ok?■■. This order of binding affinities is commonly seen for these compounds and their zinc(II) or copper(II) complexes.

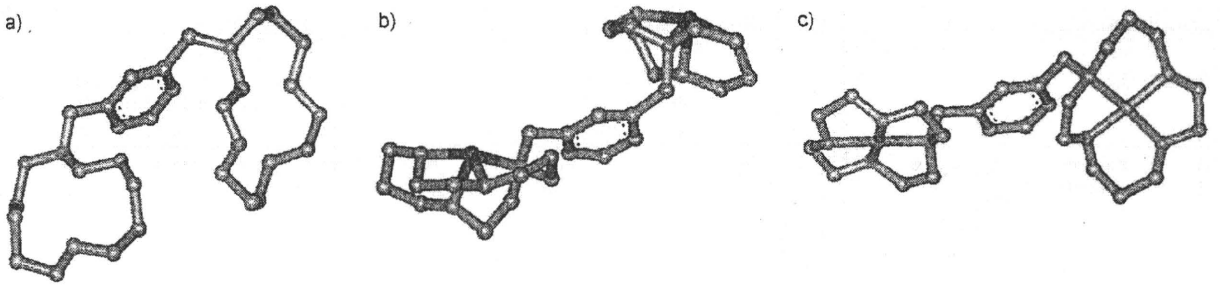


Figure 2. Structures calculated by molecular modeling of a) compound **21**, and its b) zinc and c) copper complexes. Atom color code: nitrogen = blue, carbon = gray, zinc = red, copper = light red.

We investigated the CXCR4 antagonistic activity of compound **22** and the zinc complexes of **20**, **21**, **22**, and **28**, all of which possess strong CXCR4 binding activity. The CXCR4 antagonistic activity was assessed based on the inhibitory activity of the compounds against Ca^{2+} mobilization induced by CXCL12 stimulation through CXCR4 (figure S1 in the Supporting Information). All of the tested compounds showed significant antagonistic activity at 1 μM.

The representative compounds **14**, **16**, **20–23**, **28**, and **29**, as well as their zinc chelates, were evaluated for anti-HIV activity. CXCR4 is the major co-receptor for the entry of T-cell-line-tropic (X4) HIV-1.[10,11] Inhibitory activity against X4-HIV-1 (NL4-3 strain)-induced cytopathogenicity in MT-4 cells was assessed and is shown in Table 2.[38] A correlation between CXCR4 binding activity and anti-HIV activity was observed. For compound **16** and its zinc complex, anti-HIV activity was significantly stronger than CXCR4 binding activity, and for the zinc complexes of compounds **20–22**, the CXCR4 binding activity is two to four-times stronger than the anti-HIV activity. The anti-HIV activity of the zinc complex of **16** was the most potent (EC_{50} = 8.2 nM). This is comparable to the anti-HIV activities of **16** and its zinc complex that were reported previously.[20,22,42,43] The zinc complex of **21**, which was the most active compound in terms of CXCR4 binding activity, also exhibited potent anti-HIV activity (EC_{50} = 36 nM).

Table 2. Anti-HIV activity and cytotoxicity of representative compounds in the metal ion-free and zinc chelates.

Compd	Metal ion-free		Zinc chelate	
	EC_{50}[a] [nM]	CC_{50}[b] [μM]	EC_{50}[a] [nM]	CC_{50}[b] [μM]
14	200	> 10	200	> 10
16	21	> 10	8.2	> 10
20	38	> 10	39	> 10
21	50	> 10	36	> 10
22	93	> 10	48	> 10
23	290	> 10	220	> 10
28	36	> 10	56	> 10
29	130	> 10	42	> 10
FC131	93	> 10		
AZT	69	> 100		

[a] EC_{50} values are the concentrations corresponding to 50% protection from X4-HIV-1 (NL4-3 strain)-induced cytopathogenicity in MT-4 cells. [b] CC_{50} values are the concentrations at which the viability of MT-4 cells is reduced by 50%. All data are mean values from at least three independent experiments.

Taken together, these results show that all of the compounds exhibiting CXCR4 binding activity also showed significant anti-HIV activity (EC_{50} values < 300 nM), and none of the tested compounds exhibited significant cytotoxicity (CC_{50} values > 10 μM; Table 2). Conversely, zinc complexes of **20**, **21**, **22**, and **28** did not exhibit significant anti-HIV activity against macrophage-tropic (R5) HIV-1■■X4-HIV-1?■■ (NL(AD8) strain)-induced cytopathogenicity in PM-1■■MT-4?■■ cells at concentrations below 10 μM. Since R5-HIV-1 strains use CCR5 instead of CXCR4 as the major co-receptor for entry, this suggests that these compounds do not bind CCR5 but rather are highly selective for CXCR4.■■Previously talked of X4-HIV-1 and MT-4 cells? Please clarify.■■

Conclusions

The present study introduces a new class of low-molecular-weight CXCR4 antagonists and their zinc(II) or copper(II) complexes, which contain pyridyl or azamacrocycle moieties with *p*-xylene or *m*-xylene spacers. These compounds demonstrated strong CXCR4 binding activity. Zinc complexes of **20** and **21**, which were the two most active compounds, contain cyclen and cyclam rings with *p*- and *m*-xylene spacers and exhibited remarkably small■■low? potent?■■ IC_{50} values (11 and 8.3 nM, respectively). These compounds showed significant CXCR4 antagonistic activity, based on inhibitory activity against Ca^{2+} mobilization induced by CXCL12 stimulation through CXCR4, as well as potent anti-HIV activity, as assessed by protection from X4-HIV-1-induced cytopathogenicity in MT-4 cells. These results provide useful insights into the future design of novel CXCR4 antagonists, complementing information from other CXCR4 antagonists such as T140, FC131, and KRH-1636. Furthermore, these new compounds are useful for the development of therapeutic strategies for CXCR4-relevant diseases and chemical probes to study the biological activity of CXCR4.

Experimental Section

Chemistry

Compounds **12–17**, **20**, **21**, **24**, **25**, **27–29**, and **31** were synthesized as previously reported.[20,22,37,40,41,44–47] Compounds **18**, **19**, **22**, **23**, **26**, and **30** were synthesized in the present study; details are provided in the Supporting Information. A representative compound, **18**, was synthesized by coupling *p*-dibromoxylene (1,4-bis(-bromomethyl)benzene) with tri-Boc-protected 1,4,7,10-tetraazacyclododecane, followed by treatment with trifluoroacetic acid and subsequent coupling with bis(pyridin-2-ylmethyl)amine. All crude compounds were purified by RP-HPLC and identified by FAB/ESI-HRMS. Zinc(II) or copper(II) complex formation was accomplished by treatment of the above compounds with 10 equiv of $ZnCl_2$ or $CuCl_2$ in PBS. All zinc(II) or copper(II) complexes were characterized by chemical shifts of their methylene protons in ^{1}H NMR analysis. The pyridyl zinc(II) complex was characterized previously,[37] and zinc(II) or copper(II) complex formation with these macrocyclic compounds has been reported elsewhere.[41,42,48,49] Detailed procedures and data are provided in the Supporting Information.

Biological assays

A CXCR4 binding assay for compounds, based on the inhibition of [^{125}I]CXCL12 binding to Jurkat cells, was performed as reported by Tanaka et al.[38] CXCR4 antagonistic activity was evaluated as described by Ichiyama et al[27], measuring inhibitory activity against Ca^{2+} mobilization induced by CXCL12 stimulation in HOS cells expressing CXCR4. Anti-HIV activity was determined by inhibitory activity against X4-HIV-1(NL4–3)-induced cytopathogenicity in MT-4 cells as reported by Tanaka et al.[38] An X4 HIV-1 infectious molecular clone (pNL4–3) was obtained from the AIDS Research and Reference Reagent Program. The virus NL4–3 was obtained from the culture supernatant of 293T cells transfected with pNL4–3.

Molecular modeling

Molecular modeling calculations were performed using Sybyl (version 7.0, Tripos). Energy minimization was performed using the Tripos force field and Gasteiger–Hückel charge parameters. The lowest energy conformation was obtained by random search methods.

Acknowledgements

T.T. and N.O. are supported by research fellowships for young scientists from the Japan Society for the Promotion of Science. This work was supported in part by a Grant-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science, and Technology of Japan, and Health and Labor Sciences Research Grants from the Japanese Ministry of Health, Labor, and Welfare. The authors thank Mr. Wu Honggui (Tokyo University of Science) for his assistance with the anti-HIV assay.

Keywords: azamacrocycles · Ca^{2+} mobilization · CXCR4 · HIV · structure–activity relationships

[1] T. Nagasawa, H. Kikutani, T. Kishimoto, *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 2305–2309.

[2] C. C. Bleul, M. Farzan, H. Choe, C. Parolin, I. Clark-Lewis, J. Sodroski, T. A. Springer, *Nature* **1996**, *382*, 829–833.

[3] E. Oberlin, A. Amara, F. Bachelerie, C. Bessia, J. L. Virelizier, F. Arenzana-Seisdedos, O. Schwartz, J. M. Heard, I. Clark-Lewis, D. L. Legler, M. Loetscher, M. Baggiolini, B. Moser, *Nature* **1996**, *382*, 833–835.

[4] K. Tashiro, H. Tada, R. Heilker, M. Shirozu, T. Nakano, T. Honjo, *Science* **1993**, *261*, 600–603.

[5] C. C. Bleul, R. C. Fuhlbrigge, J. M. Casanovas, A. Aiuti, T. A. Springer, *J. Exp. Med.* **1996**, *184*, 1101–1109.

[6] K. Tachibana, S. Hirota, H. Iizasa, H. Yoshida, K. Kawabata, Y. Kataoka, Y. Kitamura, K. Matsushima, N. Yoshida, S. Nishikawa, T. Kishimoto, T. Nagasawa, *Nature* **1998**, *393*, 591–594.

[7] T. Nagasawa, S. Hirota, K. Tachibana, N. Takakura, S. Nishikawa, Y. Kitamura, N. Yoshida, H. Kikutani, T. Kishimoto, *Nature* **1996**, *382*, 635–638.

[8] Y. Zhu, Y. Yu, X. C. Zhang, T. Nagasawa, J. Y. Wu, Y. Rao, *Nat. Neurosci.* **2002**, *5*, 719–720.

[9] R. K. Stumm, C. Zhou, T. Ara, F. Lazarini, M. Dubois-Dalcq, T. Nagasawa, V. Hollt, S. Schulz, *J. Neurosci.* **2003**, *23*, 5123–5130.

[10] H. K. Deng, R. Liu, W. Ellmeier, S. Choe, D. Unutmaz, M. Burkhart, P. D. Marzio, S. Marmon, R. E. Sutton, C. M. Hill, C. B. Davis, S. C. Peiper, T. J. Schall, D. R. Littman, N. R. Landau, *Nature* **1996**, *381*, 661–666.

[11] Y. Feng, C. C. Broder, P. E. Kennedy, E. A. Berger, *Science* **1996**, *272*, 872–877.

[12] T. Koshiba, R. Hosotani, Y. Miyamoto, J. Ida, S. Tsuji, S. Nakajima, M. Kawaguchi, H. Kobayashi, R. Doi, T. Hori, N. Fujii, M. Imamura, *Clin. Cancer Res.* **2000**, *6*, 3530–3535.

[13] A. Müller, B. Homey, H. Soto, N. Ge, D. Catron, M. E. Buchanan, T. McClanahan, E. Murphy, W. Yuan, S. N. Wagner, J. L. Barrera, A. Mohar, E. Verastegui, A. Zlotnik, *Nature* **2001**, *410*, 50–56.

[14] H. Tamamura, A. Hori, N. Kanzaki, K. Hiramatsu, M. Mizumoto, H. Nakashima, N. Yamamoto, A. Otaka, N. Fujii, *FEBS Lett.* **2003**, *550*, 79–83.

[15] N. Tsukada, J. A. Burger, N. J. Zvaifler, T. J. Kipps, *Blood* **2002**, *99*, 1030–1037.

[16] J. Juarez, K. F. Bradstock, D. J. Gottlieb, L. J. Bendall, *Leukemia* **2003**, *17*, 1294–1300.

[17] T. Nanki, K. Hayashida, H. S. El-Gabalawy, S. Suson, K. Shi, H. J. Girschick, S. Yavuz, P. E. Lipsky, *J. Immunol.* **2000**, *165*, 6590–6598.

[18] H. Tamamura, M. Fujisawa, K. Hiramatsu, M. Mizumoto, H. Nakashima, N. Yamamoto, A. Otaka, N Fujii, *FEBS Lett.* **2004**, *569*, 99–104.

[19] T. Murakami, T. Nakajima, Y. Koyanagi, K. Tachibana, N Fujii, H. Tamamura, N. Toshida, M. Waki, A. Matsumoto, O. Yoshie, T. Kishimoto, N. Yamamoto, T. Nagasawa, *J. Exp. Med.* **1997**, *186*, 1389–1393.

[20] D. Schols, S. Struyf, J. Van Damme, J. A. Este, G. Henson, E. DeClarcq, *J. Exp. Med.* **1997**, *186*, 1383–1388.

[21] B. J. Doranz, K. Grovit-Ferbas, M. P. Sharron, S.-H. Mao, M. Bidwell Goetz, E. S. Daar, R. W. Doms, W. A. O'Brien, *J. Exp. Med.* **1997**, *186*, 1395–1400.

[22] G. A. Donzella, D. Schols, S. W. Lin, J. A. Este, K. A. Nagashima, *Nat. Med.* **1998**, *4*, 72–76.

[23] O. M. Z. Howard, J. J. Oppenheim, M. G. Hollingshead, J. M. Covey, J. Bigelow, J. J. McCormack, R. W. Buckheit, Jr., D. J. Clanton, J. A. Turpin, W. G. Rice, *J. Med. Chem.* **1998**, *41*, 2184–2193.

[24] H. Tamamura, Y. Xu, T. Hattori, X. Zhang, R. Arakaki, K. Kanbara, A. Omagari, A. Otaka, T. Ibuka, N. Yamamoto, H. Nakashima, N. Fujii, *Biochem. Biophys. Res. Commun.* **1998**, *253*, 877–882.

[25] H. Tamamura, A. Omagari, S. Oishi, T. Kanamoto, N. Yamamoto, S. C. Peiper, H. Nakashima, A. Otaka, N. Fujii, *Bioorg. Med. Chem. Lett.* **2000**, *10*, 2633–2637.

[26] N. Fujii, S. Oishi, K. Hiramatsu, T. Araki, S. Ueda, H. Tamamura, A. Otaka, S. Kusano, S. Terakubo, H. Nakashima, J. A. Broach, J. O. Trent, Z. Wang, S. C. Peiper, *Angew. Chem.* **2003**, *115*, 3373–3375; *Angew. Chem. Int. Ed.* **2003**, *42*, 3251–3253.

[27] K. Ichiyama, S. Yokoyama-Kumakura, Y. Tanaka, R. Tanaka, K. Hirose, K. Bannai, T. Edamatsu, M. Yanaka, Y. Niitani, N. Miyano-Kurosaki, H. Takaku, Y. Koyanagi, N. Yamamoto, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 4185–4190.

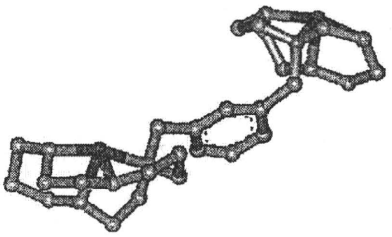
[28] H. Tamamura, N. Fujii, *Curr. Drug Targets-Infectious Disorders* **2004**, *4*, 103–110.

[29] H. Tamamura, K. Hiramatsu, S. Ueda, Z. Wang, S. Kusano, S. Terakubo, J. O. Trent, S. C. Peiper, N. Yamamoto, H. Nakashima, A. Otaka, N. Fujii, *J. Med. Chem.* **2005**, *48*, 380–391.

[30] H. Tamamura, T. Araki, S. Ueda, Z. Wang, S. Oishi, A. Esaka, J. O. Trent, H. Nakashima, N. Yamamoto, S. C. Peiper, A. Otaka, N. Fujii, *J. Med. Chem.* **2005**, *48*, 3280–3289.
[31] G. C. Valks, G. McRobbie, E. A. Lewis, T. J. Hubin, T. M. Hunter, P. J. Sadler, C. Pannecouque, E. De Clercq, S. J. Archibald, *J. Med. Chem.* **2006**, *49*, 6162–6165.
[32] W. Zhan, Z. Liang, A. Zhu, S. Kurtkaya, H. Shim, J. P. Snyder, D. C. Liotta, *J. Med. Chem.* **2007**, *50*, 5655–5664.
[33] A. Khan, G. Nicholson, J. Greenman, L. Madden, G. McRobbie, C. Pannecouque, E. De Clercq, R. Ullom, D. L. Maples, R. D. Maples, J. D. Silversides, T. J. Hubin, S. J. Archibald, *J. Am. Chem. Soc.* **2009**, *131*, 3416–3417.
[34] G. J. Bridger, R. T. Skerlj, P. E. Hernandez-Abad, D. E. Bogucki, Z. Wang, Y. Zhou, S. Nan, E. M. Boehringer, T. Wilson, J. Crawford, M. Metz, S. Hatse, K. Princen, E. De Clercq, D. Schols, *J. Med. Chem.* **2010**, *53*, 1250–1260.
[35] R. T. Skerlj, G. J. Bridger, A. Kaller, E. J. McEachern, J. B. Crawford, Y. Zhou, B. Atsma, J. Langille, S. Nan, D. Veale, T. Wilson, C. Harwig, S. Hatse, K. Princen, E. De Clercq, D. Schols, *J. Med. Chem.* **2010**, *53*, 3376–3388.
[36] M. Takenaga, H. Tamamura, K. Hiramatsu, N. Nakamura, Y. Yamaguchi, A. Kitagawa, S. Kawai, H. Nakashima, N. Fujii, R. Igarashi, *Biochem. Biophys. Res. Commun.* **2004**, *320*, 226–232.
[37] H. Tamamura, A. Ojida, T. Ogawa, H. Tsutsumi, H. Masuno, H. Nakashima, N. Yamamoto, I. Hamachi, N. Fujii, *J. Med. Chem.* **2006**, *49*, 3412–3415.
[38] T. Tanaka, H. Tsutsumi, W. Nomura, Y. Tanabe, N. Ohashi, A. Esaka, C. Ochiai, J. Sato, K. Itotani, T. Murakami, K. Ohba, N. Yamamoto, N. Fujii, H. Tamamura, *Org. Biomol. Chem.* **2008**, *6*, 4374–4377.
[39] G. J. Bridger, R. T. Skerlj, D. Thornton, S. Padmanabhan, S. A. Martellucci, G. W. Henson, M. J. Abrams, N. Yamamoto, K. De Vreese, R. Pauwels, E. De Clercq, *J. Med. Chem.* **1995**, *38*, 366–378.
[40] G. J. Bridger, R. T. Skerlj, S. Padmanabhan, S. A. Martellucci, G. W. Henson, M. J. Abrams, H. C. Joao, M. Witvrouw, K. De Vreese, R. Pauwels, E. De Clercq, *J. Med. Chem.* **1996**, *39*, 109–119.
[41] Y. Inouye, T. Kanamori, T. Yoshida, T. Koike, M. Shionoya, H. Fujioka, E. Kimura, *Biol. Pharm. Bull.* **1996**, *19*, 456–458.
[42] L. O. Gerlach, J. S. Jakobsen, K. P. Jensen, M. R. Rosenkilde, R. T. Skerlj, U. Ryde, G. J. Bridger, T. W. Schwartz, *Biochemistry* **2003**, *42*, 710–717.
[43] H. F. Egberink, E. De Clercq, A. L. Van Vliet, J. Balzarini, G. J. Bridger, G. Henson, M. C. Horzinek, D. Schols, *J. Virol.* **1999**, *73*, 6346–6352.
[44] M. Le Baccon, F. Chuburu, L. Toupet, H. Handel, M. Soibinet, I. Dechamps-Olivier, J.-P. Barbier, M. Aplincourt, *New J. Chem.* **2001**, *25*, 1168–1174.
[45] B. Antonioli, D. J. Bray, J. K. Clegg, K. Gloe, K. Gloe, O. Kataeva, L. F. Lindoy, J. C. McMurtrie, P. J. Steel, C. J. Sumby, M. Wenzel, *Dalton Trans.* **2006**, 4783–4794.
[46] S. P. Foxon, D. Utz, J. Astner, S. Schindler, F. Thaler, F. W. Heinemann, G. Liehr, J. Mukherjee, V. Balamurugan, D. Ghosh, R. Mukherjee, *Dalton Trans.* **2004**, 2321–2328.
[47] S. Mandal, F. Lloret, R. Mukherjee, *Inorg. Chim. Acta* **2009**, *362*, 27–37.
[48] M. Soibinet, I. De'champs-Olivier, E. Guillon, J.-P. Barbier, M. Aplincourt, F. Chuburu, M. Le Baccon, H. Handel, *Eur. J. Inorg. Chem.* **2003**, 1984–1994.
[49] R. W. Hay, M. T. Tarafder, *Transition Met. Chem.* **1990**, *15*, 490–492.

Received: December 19, 2010
Published online on ■■ ■■, 2011

To chelate or not to chelate? Low-molecular-weight nonpeptidic compounds were designed by combining common structural motifs of alkylamino and pyridyl macrocyclic antagonists of the chemokine receptor CXCR4. Several new zinc(II) or copper(II) complexes exhibited potent anti-HIV activity, strong CXCR4 binding activity, and significant inhibitory activity against Ca^{2+} mobilization induced by CXCL12 stimulation. ■■ Text edited for brevity: ok? ■■



T. Tanaka, T. Narumi, T. Ozaki, A. Sohma, N. Ohashi, C. Hashimoto, K. Itotani, W. Nomura, T. Murakami, N. Yamamoto, H. Tamamura**

■■ – ■■

Azamacrocyclic Metal Complexes as CXCR4 Antagonists

Chem-Bio Informatics Journal, Vol.10, pp.32-45(2010)

Comparison of binding affinity evaluations for FKBP ligands with state-of-the-art computational methods: FMO, QM/MM, MM-PB/SA and MP-CAFEE approaches

Hirofumi Watanabe[1,2*], Shigenori Tanaka[1,2*], Noriaki Okimoto[3], Aki Hasegawa[3], Makoto Taiji[3*], Yoshiaki Tanida[4], Takashi Mitsui[5], Mariko Katsuyama[5], Hideaki Fujitani[4]

[1]*Graduate School of Engineering, Kobe University, 3-11 Tsurukabuto, Nada, Kobe, Japan*
[2]*JST-CREST*
[3]*RIKEN Advanced Science Institute, Computational Systems Biology Research Group, 61-1 Ono-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0046, Japan*
[4]*Fujitsu Laboratories Ltd., 10-1 Morinosato-Wakamiya, Atsugi 243-0197, Japan*
[5]*Fujitsu Ltd, 9-3, Nakase 1-chome, Mihama-ku, Chiba City Chiba 261-8588, Japan*

**E-mail: watanabe@radix.h.kobe-u.ac.jp, tanaka2@kobe-u.ac.jp, taiji@riken.jp*

(Received October 9, 2009; accepted March 15, 2010; published online April 20, 2010)

Abstract

We compared binding affinity evaluations for 10 FKBP ligands with such state-of-the-art computational methods as FMO, QM/MM, MM-PB/SA, and MP-CAFEE. For the FKBP ligands, we confirmed that each method could provide good correlations between the experimental and computational binding affinities. From the calculated results, we discussed the importance of solvation effect and structural sampling for these methods in detail. In addition, we addressed the issues of computational time and present arguments on the future perspective of the computational binding affinity evaluations.

Key Words: Protein-ligand binding affinity, Fragment molecular orbital method, QM/MM method, MM-PB/SA method, MP-CAFEE method, FKBP

Area of Interest: Molecular Computing

1. Introduction

Recent advances in computational resources and increasing numbers of registered three dimensional (3D) protein structures have accelerated the application of computer simulations to large biomolecular systems. Further, recent progresses of computational methods of evaluation of binding affinity, which are expected to improve drug design efficiency, manifest various features in each method employed [1][2]. For example, the fragment molecular orbital (FMO) method [3] can deal with whole large biomolecules quantum mechanically. In the FMO method, a large molecular system is divided into small fragments, and the conventional molecular orbital (MO) calculations are performed for each fragment and fragment pair. The FMO method overcomes the size limitation of the conventional MO method while maintaining chemical accuracy for energy evaluation. The quantum mechanics/molecular mechanics (QM/MM) method is another method for overcoming the size limitation of quantum chemical methods [4]. In the QM/MM method, a region that requires accurate analysis is studied quantum-mechanically, and other regions are assigned to be studied by classical force field calculations. The molecular mechanics Poisson-Boltzmann surface area (MM-PB/SA) method [5] is often used for binding affinity evaluation as well. In the MM-PB/SA method, the free energy is calculated by using the snapshots of solute molecules obtained from explicit-solvent MD simulation. At this time, the explicit solvent is replaced with implicit models. The massively parallel computation for absolute binding free energy (MP-CAFEE) method [6][7][8][9][10][11] directly evaluates free energy difference by the Bennet acceptance ratio (BAR) method, which can be interpreted as a maximum likelihood estimate of free energy differences [9] from the non-equilibrium work based on the non-equilibrium identity [10][11]. Auxiliary restraints for keeping the ligand position are not employed, while such restraints are used in the standard scheme for evaluation of absolute binding affinities by thermodynamic integration or the free energy perturbation method [1].

Successful results have been published by employing these methods [6][7][8][12][13][14][15][16] which provides us with the expectation that such methods can be put to practical use. In these reports, however, the target proteins were not identical for the different methods. Even if the target proteins were identical, their structures might be different, thus making it difficult to perform reliable comparisons among the different methods. In addition, unsuccessful results of binding affinity evaluations were rarely reported [17]. Thus, it is difficult to determine which methods are useful for each specific situation. To provide the answer for this question, in this paper, we attempt to perform reliable comparisons of binding affinity evaluations among the state-of-the-art calculation methods. To achieve this, we evaluated protein-ligand binding affinity by different methods for the same target protein and ligands, with the same structures. As a test set for binding affinity evaluation, we employed one of the FK506 binding proteins (FKBPs), FKBP12, which is well known as a target protein of the immune suppressor tacrolimus (FK506), and its 10 ligands because this protein is often used as the test set for computational binding affinity evaluations [6][7][14][16].

The computational methods for evaluation of binding affinities are characterized by several aspects such as energy evaluation based on quantum mechanics (QM) or molecular mechanics (MM) force field, a way of incorporating entropic contributions, and solvation effects. To clarify which contributions are important for obtaining good correlations between computational binding affinities and experimental ones, a systematic comparison of results obtained from various computational methods is important. Thus, we used four state-of-the-art computational methods, FMO, QM/MM, MM-PB/SA, and MP-CAFEE, for evaluating the binding affinity of the

protein-ligand bound state. Features of these methods are given in Table 1.

We believe the results of our comparisons will become a benchmark for choosing appropriate methods for practical application. In addition, we expect that they would also become a basis for the development of a new binding affinity evaluation method.

Table 1. Features for each computational method employed in the present study
(* In the MP-CAFEE method, free energy difference is directly evaluated by the BAR method. The entropic contribution is incorporated as a part of free energy difference.)

Method	FMO	QM/MM	MM-PB/SA	MP-CAFEE
Quantum or Classical	Quantum	Quantum + Classical	Classical	Classical
Entropic effect	No	No	Normal mode analysis	BAR method*
Solvation effect	No	No	Implicit (PB equation)	Explicit (TIP3P)
Selection of atomic coordinate	Single Point	Single Point	Sampling by MD	Sampling by MD

2. Methods

2.1 Preparation of protein-ligand structures

The 10 complex structures of FKBP12 and ligands were constructed based on four X-ray crystallographic structures (PDB ID 1FKG (L8), 1FKH (L9), 1FKI (L13) and 1FKF (L20)), where the numbering of the ligand molecules is in accordance with the literature [6][7][18]. As all ligands have common binding elements, a pipecolate and an alpha-keto amide region, we assumed that bound conformations of the other ligands are similar to the known four complexes. Three-dimensional structural image of complex of FKBP and FK-506 (L20) is shown in Figure 1. The molecular formulas of 10 ligands are shown in Figure 2 and the experimental binding affinities and molecular weights are shown in Table 2 [18]. To obtain the initial structures for FMO and QM/MM methods, we performed MM energy minimizations of respective protein-ligand complexes above with the TIP3P water solvent molecules [19] with AMBER 8 [20]. After that, each initial structure of protein-ligand complex for the FMO and QM/MM calculations was constructed by removing water molecules from energy-minimized structure.

2.2 FMO method

For the FMO calculations we used the ABINIT-MP program [21]. The Schrödinger equation was solved by the Hartree-Fock (HF) and Møller-Plesset second order perturbation (MP2) methods with the 6-31G basis set. Fragmentation was performed by the following rule. For the protein region, each fragment has an amino acid residue, and a ligand is included in the fragment. Thresholds for electrostatic interaction approximations [22] such as L_{aoc}, L_{ptc} and L_{dimer} were 0.0, 2.0 and 2.0, respectively. In FMO calculations, we evaluated binding energies without entropies to reduce computational time. We regarded those values as the approximate values of binding free energies.

2.3 QM/MM method

In this article, we used the QSite program for the QM/MM calculations [23]. This software adopts a frozen-core orbital for the QM/MM boundary [4]. The QM region contained side-chain atoms of 7 residues (Tyr26, Asp37, Arg42, Phe46, Trp59, Tyr82, and Phe99), entire atoms of 3 residues (Glu54, Val55, and Ile56) and ligand atoms. The Schrödinger equation for each structure

was solved by the HF and MP2 methods with the 6-31G basis set and the OPLS force field was used for MM region. In the QM/MM calculations, entropic contributions are neglected as well as in the FMO calculations.

2.4 MM-PB/SA method

We performed MD simulations of protein-ligand complexes in explicit water molecules to calculate MM-PB/SA binding free energy. All of the MD simulations were performed using AMBER 8.0 [20] modified for use on the special-purpose computer MDGRAPE-3 [24][25]. All of the protein-ligand complex structures were solvated in a rectangular box containing TIP3P water molecules (explicit solvent) [19] under periodic boundary conditions. The box dimensions were chosen so that the minimum distance of any protein atoms from the wall of the box was 15 Å. The SHAKE algorithm [26] was applied to bonds involving hydrogen atoms, considering an integration time step of 1.0 fs. Long-range Coulomb interactions were treated by applying the particle mesh Ewald (PME) method [27]. The real-space component of the PME method was calculated by using MDGRAPE-3, while the wave number-space component for this method and bonded interactions were calculated by the host computers. To optimize the balance between the calculation times for these components, a cutoff distance of 14 Å was used for the real-space component. Each system was gradually heated to 298 K for the first 50 ps. The Berendsen's temperature and pressure control methods [28] were used to maintain the temperature and pressure constant at 298 K and 1 atm, respectively. The force-field parameters and charges for the protein and ligands which were the same ones as in the MP-CAFEE method, were determined by the force field formulator for organic molecules (FF-FOM) [7]. The production MD trajectory (500 snapshots) was collected for the last period of 5 ns. In the calculation of the binding free energies by the MM-PB/SA method, the water molecules were replaced with implicit solvation models, that is, the PB equation and the SA term, while TIP3P explicit solvent model was employed for obtaining the trajectory. The entropic contribution was evaluated by the normal mode analysis. The evaluations of binding free energy using the MM-PB/SA method have been reported in previous studies [28][29][30][31].

2.5 MP-CAFEE method

The results we show in this paper for the MP-CAFEE method were published by Fujitani et al. [7]. We again show the results by this method in this paper for the comparison with the results by the other methods. Thus, the calculation procedures are to be found in the literature [7]. Here, we summarize the calculation procedure and the setting of the study. We used an in-house modified version of the GROMACS package to perform the MP-CAFEE calculation [32]. MD calculations were performed with the following conditions. A Nose-Hoover thermostat was used for temperature control at 298K and with a time constant of 0.3ps. We used Berendsen's algorithm was employed with time constant of 1.0ps and the pressure of 1.0 atm. In addition, we calculated Coulomb interaction by the PME method. The equilibration process was as follows. First a conjugate gradient minimization was performed, followed by an MD simulation for 200ps with the solute position being restrained. After that, long time equilibration at 298K between the ligand and surrounding system was performed. The equilibration times were 5ns for the solvated ligand and from 10ns to 50ns for solvated protein-ligand complex. After the equilibration, structural samplings for work measurements were performed. In the annihilation process, the Coulomb interaction and van der Waals interaction were turned off by using 32 intermediate λ_i points. For each λ_i point, 12 MD simulations with different initial momenta were performed, where the sampling times for each λ_i point were 2.5 ns for the solvated protein-ligand complex and 1.0 ns for the solvated ligand. The

MD calculations were performed by on a FUJITSU Bioserver test machine with 1920 VLIW architecture processors in a rack. The force fields for both proteins and ligands were determined by the FF-FOM [7], which properly assigns atom types of the general amber force field (GAFF) [33], and RESP charge [34] was used for ligand charges. A more detailed explanation is described in the literature [7].

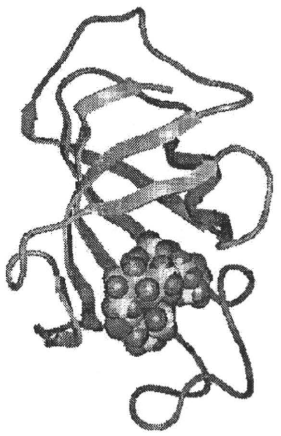


Figure 1. Complex structure of FKBP12, which is composed of 107 residues, and FK506 (L20)

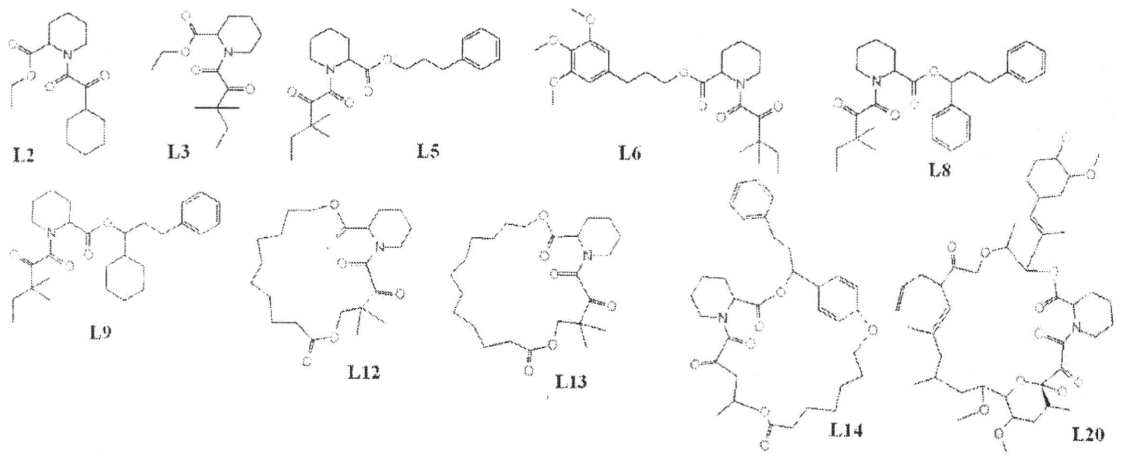


Figure 2. Molecular structures of 10 FKBP ligands, where the numbering of the ligand molecules is in accordance with the literature [6][7][18]

Table 2. Experimental binding free energy of each ligand

ΔG_{exp} and MW indicate the experimental binding affinities and molecular weights for 10 ligands respectively.

Ligand	L2	L3	L5	L6	L8	L9	L12	L13	L14	L20
ΔG_{exp}	-7.8	-8.4	-9.5	-10.8	-10.9	-11.1	-10.3	-9.5	-12.3	-12.8
MW	295.4	283.4	373.5	463.6	449.6	455.6	411.5	437.6	581.4	804.0

3. Results

3.1 FMO method

The correlation between the binding energies evaluated by the FMO method and experimental binding affinities is shown in Figure 3. The results by the FMO-HF and FMO-MP2 [36][37] methods are depicted in Figure 3(a) and (b), respectively, where the 6-31G basis set was used. These figures indicate that the binding energies by the FMO-HF methods were hardly correlated with the experimental values, while the result by the FMO-MP2 method showed better correlation. The bound ligands were surrounded by several hydrophobic residues in the binding pocket. Thus, a better result by the FMO-MP2 method would thus be due to the inclusion of the dispersion force between the ligands and hydrophobic residues. The regression slope was far from unity in the case of the FMO-MP2 (slope=4.13). This result is reasonable, however, because both the entropic contribution and the solvation effect are ignored in the present FMO calculation.

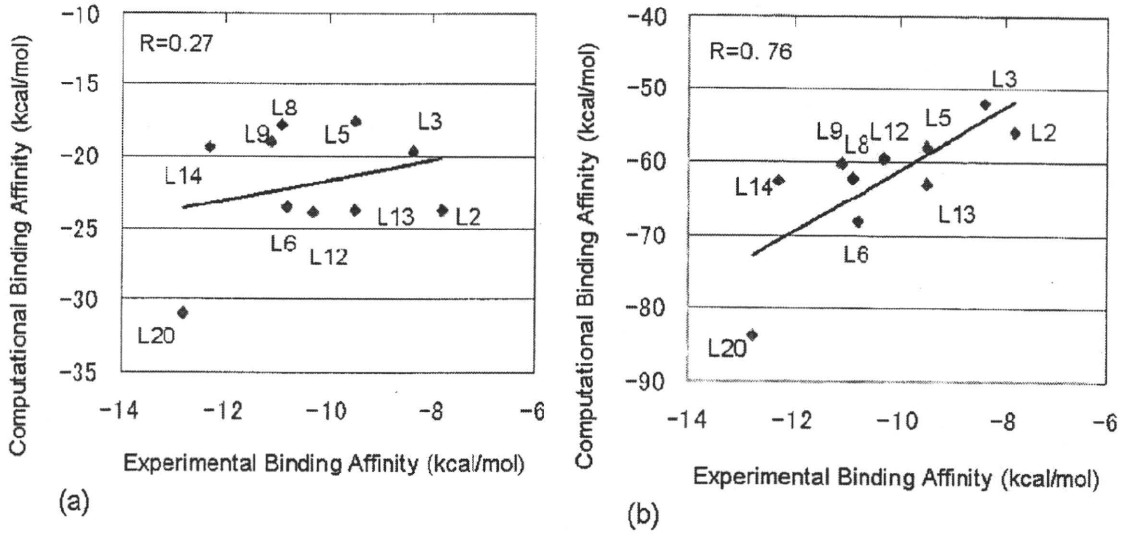


Figure 3. Correlations between the experimental binding affinities and calculated ones by the FMO method, (a) FMO-HF, and (b) FMO-MP2
R is the correlation coefficient.

3.2 QM/MM method

The correlation between binding energies by the QM/MM method and the experimental binding affinities is shown in Figure 4. The QM/MM (HF) and QM/MM (MP2) results are shown in Figure 4(a) and (b), respectively, where the 6-31G basis set was used and the OPLS force field was used for the MM region. The QM/MM (MP2) result (correlation coefficient R=0.86) showed a slightly better correlation than the QM/MM (HF) result (R=0.82). The slopes of regressions, however, were far from unity for both QM/MM (HF) and QM/MM (MP2). In the same way as FMO method, this would be due to the fact the entropic and solvation effects are not included in the present calculations.

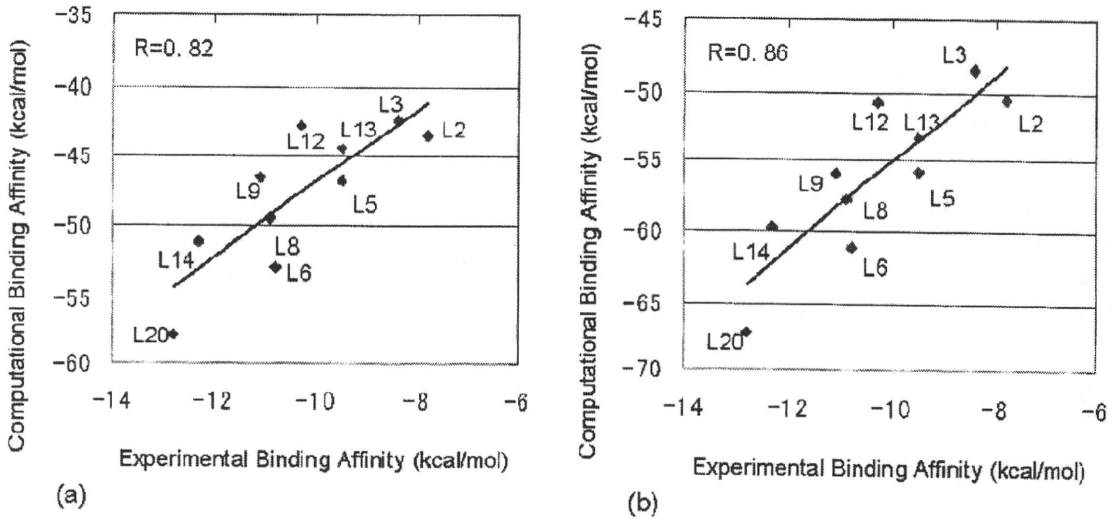


Figure 4. Correlations between the experimental and calculated binding affinities by the QM/MM method, (a) QM/MM HF/6-31G, and (b) QM/MM MP2/6-31G

3.3 MM-PB/SA method

The correlation between binding free energies evaluated by the MM-PB/SA method and experimental binding affinities is shown in Figure 5. In this work, we compared the MM-PB/SA energies with and without entropic term (Figure 5(a) and (b)). This comparison indicated the introduction of the entopic effect improves the absolute binding energies in the MM-PB/SA method, but the MM-PB/SA energies with entropic term show a worse correlation (R=0.76) as compared with the MM-PB/SA energies without entropic term (R=0.83). In addition, for the MM-PB/SA energies with entropic contribution, the slope of regression is far from unity although all the energy components of free energy are incorporated. It is noticeable that these results are in contrast to an earlier binding affinity evaluation for FKBP ligands by the MM-PB/SA method performed by Xu et al. [16]. In their results, the binding affinity evaluation with the entropic contribution provided a better correlation than did the evaluation without the entropic contribution. This difference might be caused by the difference in ligands used for evaluations. In the present evaluation, 4 compounds out of 10 are macrocyclic compounds (see Figure 2). It would be difficult to predict accurate entropic values for these macrocyclic compounds by the normal mode analysis.

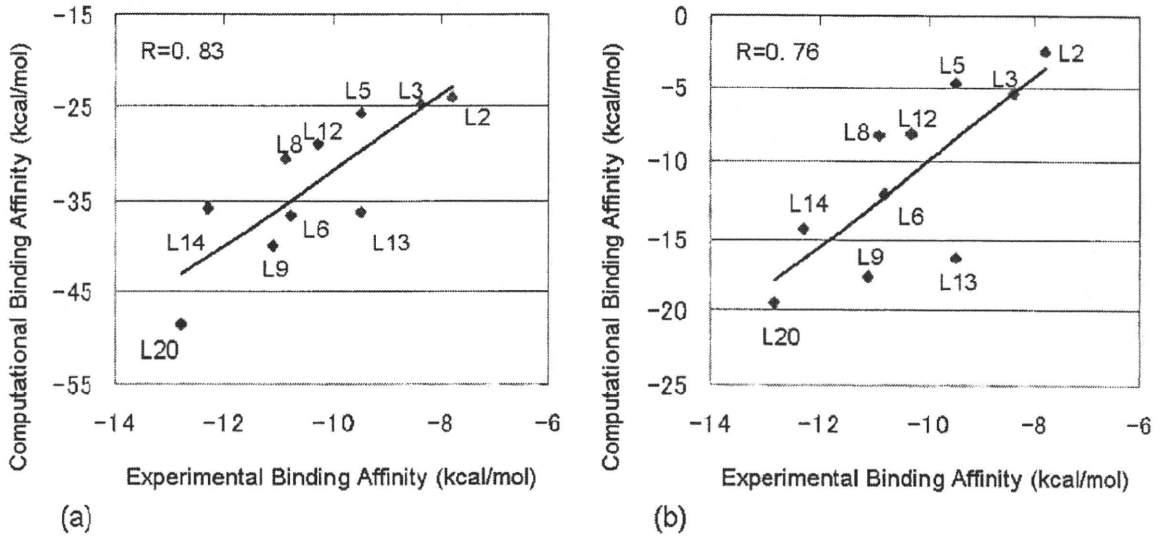


Figure 5. Correlations between the experimental binding affinities and calculated ones by the MM-PB/SA method, (a) without entropy term, and (b) with entropy term

3.4 MP-CAFEE method

The correlation between binding free energies evaluated by the MP-CAFEE method and the experimental binding affinities is shown in Figure 6. This result was previously published in the literature [7] and is shown here for comparison with the other methods. Not only was the correlation very good (R=0.98), but also the slope of regression was 1.26, which is close to unity. In addition, it is remarkable that the calculated absolute binding free energies calculated by this method reproduced the absolute values of the experimental binding affinities.

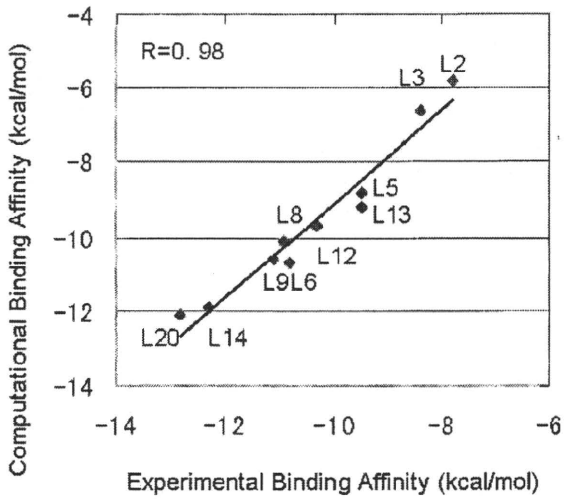


Figure 6. Correlation between the experimental binding affinities and calculated ones by the MP-CAFEE method

4. Discussion

To analyze the relationship between experimental and computational binding affinities, the correlation coefficients, the regressing slopes and intercepts are summarized in Table 3. From this table, it is obvious that MP-CAFEE method showed extremely good correlation with experimental data. The slope of the regression is so close to unity; moreover, the absolute values of the computational binding free energy are so close to those of experimental binding free energies. On the other hand, the other methods show relatively good correlations (correlation coefficients R=0.76-0.86) with experimental results. For the FMO and QM/MM calculations, the correlation coefficients by MP2 level calculations are better than those by the HF calculations. Considering that FKBP ligands interact with the hydrophobic residues in the binding pocket, this improvement is due to the effect of the dispersion force incorporated in the MP2 calculation [38][39]. However, the values of regression slopes deviate from unity since the entropic contribution and the solvation effect are not taken into account in the present FMO and QM/MM calculations. The MM-PB/SA (with entropic term) calculations take into account all contributions approximately, but the slope is far from unity in contrast with that of MP-CAFEE method. To achieve improvement of the correlation coefficient and the regression slope, the precise evaluation of entropic term and incorporation of more accurate solvation energy are essential. In the following, we discuss the importance of solvation effect and structural sampling for these methods in detail.

Table 3. Summary of calculation results

Regression equation is $\Delta G(E)_{comput}=a\,\Delta G_{exp}+b$, where ΔG_{exp}, $\Delta G(E)_{comput}$, a and b are experimental binding affinities, calculated binding free energies (or binding energies), regression slopes and intercepts, respectively. Values in parentheses are standard errors.

Method	FMO (HF)	FMO (MP2)	QM/MM (HF)	QM/MM (MP2)	MM-PB/SA (with entropy)	MM-PB/SA (without entropy)	MP-CAFEE
Correlation	0.27	0.76	0.82	0.86	0.76	0.83	0.98
Intercept: b	-14.98(9.09)	-19.90(13.10)	-20.92(6.69)	-24.28(6.84)	19.13(8.94)	8.67(10.10)	3.56(1.06)
Slope: a	0.69(0.87)	4.13(1.25)	2.60(0.64)	3.07(0.65)	2.90(0.86)	4.04(0.97)	1.26(0.10)

4.1 Structural sampling

Structural sampling is important for two reasons. One is to find the appropriate stable structures, and the other is to cover enough phase space for entropy estimations. Although the structural sampling was not performed into the FMO and QM/MM methods, the fairly good correlations were obtained by these methods. These results may suggest the structural sampling is not important to obtain good correlations for the current target. However, Ishikawa et al. [39] pointed out that appropriate selection of atomic coordinate is crucial to obtain good correlations. Since the structural sampling may contribute to select appropriate atomic coordinates, this would lead to good correlations found in other cases. Further, these methods produce fair correlations in spite of the neglect of entropies. This may be a result of specific structural properties of the current system, or entropy-enthalpy compensation. Since the enthalpy changes are roughly proportional to the entropy changes, the free energy changes, which consist of the sum of both contributions to energy changes, are proportional to enthalpy changes.

In the MM/PB-SA method, the inclusion of the entropic contribution by normal mode analysis made the correlation somewhat worse. In the normal mode analysis, it is known that conformations at different local energy minima provide rather similar entropy values even though there are

differences in the case of finite temperature [40]. One of alternative methods is principal component analysis. The values are sensitive to the data sampling frequency [40][41] and are likely to be overestimated [42]. Therefore, in order to predict more accurate binding free energy by MM/PB-SA method, it is necessary to improve the calculations for the entropy terms. In the MP-CAFFE method, entropic contribution can be treated exactly. If the sampling is enough, the entropy will converge to the theoretical value. The very precise results from the method confirm that the entropic contributions can be estimated precisely after enough structural sampling.

4.2 Solvation effect

Although solvation effect was not incorporated into the FMO and QM/MM methods, relative good correlations were obtained with both of the methods for the current system. Retegan et al. [43] reported that incorporation of solvation effects in semi-empirical QM/MM method significantly improved the correlation between experimental binding affinities and computational binding energies for the CK2 proteins and its ligands. Therefore, introduction of solvation effect into the FMO and QM/MM methods would be one of the improvement means of the correlation coefficients. Furthermore, we discuss the influence of different strategies to treat solvation by MM-based free-energy evaluations. In the MM-PB/SA method, the implicit solvent model described by the PB equation is used for energy evaluation. Many reports suggest that the MM-PB/SA method can provide the fairly good correlation like in the case, and the implicit solvent model seems to be a good and easy way to approximate solvent effects in single-point or short-trajectory calculations [44][45]. In addition, since it was suggested that the implicit solvent model provided slightly worse evaluation of solvation energies than the explicit solvent model did [46], improvement of the implicit solvent model would lead to a better binding affinity evaluation. Also for the solvation energies, the MP-CAFFE method provides the accurate estimations within the framework of classical force fields. It is difficult to distinguish the origin of the difference between the MP-CAFFE and the MM/PB-SA methods, since in the MP-CAFFE method energetic contributions are not estimated separately. As we have already discussed partly, there are several possibilities – solvation effect, entropy estimation, sampling length – and these all will be related to the differences in the results.

4.3 Computational time

Computational time of each method is summarized in Table 4, where the calculation time for the MP-CAFEE method is represented by the converted value in the Pentium4 case for convenience, whereas this calculation was performed with the FUJITSU Bioserver. Although the methods based on the quantum mechanics generally require extensive computational time, the present FMO and QM/MM calculations required less computational time because in the present treatment the structural sampling was not performed in the cases of both QM based methods. If structural sampling were to be performed with the quantum mechanical method, much more computational time would be required. Concerning the classical MD based free energy evaluation method, the MP-CAFEE method required much more computational time than the MM-PB/SA method. This difference is caused by the fact that MP-CAFEE method requires a number of MD runs ($12 \times 32 \times 2$ runs for a FKBP ligand) for obtaining binding affinity of one protein-ligand complex.

Table 4. Typical computational time for the evaluations of binding affinity and of computational resources

Method	FMO (HF)	FMO (MP2)	QM/MM (HF)	QM/MM (MP2)	MM-PB/SA (with entropy)	MM-PB/SA (without entropy)	MP-CAFEE
Typical time	7hours	22hours	2.1hours	1.25days	14days	15days	4.5days
Computational resource	Opteron246 x 16 cores		Xeon5450 3.0GHZ		Xeon5150+MD-GRAPE3		Pentium4 x 300cores

4.4 Further Perspectives

We will discuss the future perspective of computational binding affinity evaluation for drug designs. Although the four state-of-the-art methods discussed here require large computational resources, such computation will become daily in near future by continuous growth of computer performance. We believe our present analysis would provide an important basis for making practical use of the computational binding affinity evaluation. To provide the significant information for practical use, we should continue to work on the comparison of binding affinity evaluation. Current results with only one target protein are not enough to clarify merits and demerits of each method in different situations. This paper is just the first report of our attempt. When many target proteins will have been studied for the comparison of the state-of-the-art methods, our aim will truly be achieved. We will report our next comparison results based on other target proteins.

In this paper we have mainly discussed the correlation coefficients between experimental and computational binding affinities. Although this focused point is suitable for judging which effect is important for reproducing experimental binding affinity, other points would also be needed to satisfy the aim of practical use in drug design because there are many kinds of demands in the pharmaceutical industry. For example, throughput of the screening should be at least more than 100 compounds per week for practical use. Prediction in weak affinity region is important for the fragment-based drug design, since the affinity of a small fragment, which will form the skeleton of lead compound, is very weak. To answer these kinds of questions, the way to assess the computational methods should be reconsidered.

5. Conclusion

To clarify the merits and demerits of various computational approaches for binding free-energy estimation in drug design, we compared the results of the binding energy evaluations by the four state-of-the-art computational methods, such as the FMO, QM/MM, MM-PB/SA and MP-CAFEE methods. We assessed the computational binding affinity evaluation methods from several aspects such as correlation with experimental binding affinity, slope of the regression and computational costs, and discussed which contributions are related to the good correlation.

Acknowledgments

The authors are highly grateful to the CBI society for the accomplishment of this work. A plan of this study was born at the site of the CBI annual meeting in 2007. In the symposium on the last day of this meeting, successful results of the FMO, MM-PB/SA and MP-CAFEE methods were reported. At that time, to clarify more detailed applicability of each method, this comparative study was proposed. In addition, this work is based on presentations in a workshop "Present Status and

Problems of *in silico* Technology in Structure Based Drug Design", which was held as a parallel session of the annual meeting of the CBI society in 2008. The authors also acknowledge all the participants of this workshop. We thank Drs. Yoshio Okiyama and Ikuo Kurisaki for useful comments on the manuscript. This work was partially supported by the CREST project of the Japan Science and Technology Agency (JST).

References

[1] M. K. Gilson, J. A. Given, B. Bush, and J. A. McCammon, The statistical-thermodynamic basis for computation of binding affinities, A critical review, *Biophys. J.*, **72**.1047-1069 (1997).

[2] M. K. Gilson and H. Zhou, Calculation of Protein-Ligand Binding Affinities, *Annu. Rev. Biophys. Biomol. Struct.*, **36**, 21-42 (2007).

[3] K. Kitaura K, E. Ikeo E, T. Asada, T. Nakano and M. Uebayasi, Fragment molecular orbital method: an approximate computational method for large molecule, *Chem. Phys. Lett.*, **313**, 701-706 (1999).

[4] D. M. Philipp and R. A. Friesner, Mixed ab initio QM/MM modeling using frozen orbitals and tests with alanine dipeptide and tetrapeptide, *J. Comput. Chem.*, **20**, 1468-1494 (1999).

[5] J. Srinivasan, J. Miller J, P. A. Kollman, D. A. Case, Continuum solvent studies of the stability of RNA hairpin loops and helices, *J. Biomol. Struct. Dyn.*, **16**, 671-682 (1998).

[6] H. Fujitani, Y. Tanida, M. Ito, G. Jayachandran, C. D. Snow, M. R. Shirts, E. J. Sorin and V. S. Pande, Direct calculation of the binding energies of FKBP ligands, *J. Chem. Phys.*, **123**, 084108 (2005).

[7] H. Fujitani, Y. Tanida and A. Matsuura, Massively parallel computation of absolute binding free energy with well-equilibrated states, *Phys. Rev. E*, **79**, 021914 (2009).

[8] Y. Tanida, M. Ito and H. Fujitani., Calculation of absolute free energy of binding for theophylline and its analog to aptamer using nonequilibrium work values, *Chem. Phys.*, **337**, 135-143 (2007).

[9] M. R. Shirts, E. Bair, G. Hooker and V. S. Pande, Equilibrium free energies from nonequilibrium measurements using maximum likelihood methods, *Phys. Rev. Lett.*, **91** 140630 (2003).

[10] C. Jarzynski, Nonequilibrium Equality for Free Energy Differences, *Phys. Rev. Lett.*, **78**, 2690 (1997).

[11] G. E. Crooks, Path-ensemble averages in systems driven far from equilibrium, *Phys. Rev. E*, 61, 2361 (2000).

[12] K. Fukuzawa, K. Kitaura, K. Uebayasi, K. Nakata, T. Kaminuma and T. Nakano, Ab initio quantum mechanical study of the binding energies of human estrogen receptor α with its ligands: an application of fragment molecular orbital method, *J. Comput. Chem.*, **26**, 1-10 (2005).

[13] T. Harada, K. Yamagishi, T. Nakano, K. Kitaura and H. Tokiwa, Ab initio fragment molecular orbital theory of ligand binding to human progesterone receptor ligand binding domain, *Naunyn-Schmiedeberg's Arch Pharmacol*, **377**, 607-615 (2008).

[14] I. Nakanishi, D. G. Fedorov and K. Kitaura, Molecular Recognition Mechanism of FK506 Binding Protein: An All-Electron Fragment Molecular Orbital Study, *Proteins*, **68**, 145-158 (2007).

[15] P. A. Kollman, I. Massova, C. Reyes, B. Kuhn, S. Huo, L. Chong, M. Lee, T. Lee, Y. Duan, W. Wan, O. Donini, P. Cieplak, J. Srivivasan, D. A. Case and T. E. Cheatham III,

Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models, *Acc. Chem. Rev.*, **33**, 889-897 (2000).

[16] Y. Xu and R. Wang, A Computational of Analysis of the Binding Affinities of FKBP12 Inhibitors Using the MM-PB/SA Method, *Proteins*, **64**, 1058-1068 (2006).

[17] D. A. Pearlman, Evaluating the molecular mechanics Poisson-Boltzmann surface area free energy method using a congeneric series of ligands to p38 MAP kinase, *J. Med. Chem.*, **48**, 7796-7807 (2005).

[18] D. A. Holt, J. I. Luengo, D. S. Yamashita, H. J. Oh, A. L. Konialian, H. K. Yen, L. W. Rozamus, M. Brandt, M. J. Bossard, M. A. Levy, D. S. Eggleston, J. Liang, L. W. Schultz, T. J. Stout and J. Clardy, Design, synthesis, and kinetic evaluation of high-affinity FKBP ligands and the X-ray crystal structures of their complexes with FKBP12, *J. Am. Chem. Soc.*, **115**, 9925-9938 (1993).

[19] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, Comparison of Simple Potential Functions for Simulating Liquid Water, *J. Chem. Phys.*, **79**. 926-935 (1983).

[20] D. A. Case, T. A. Darden, T. E. Cheatham, III, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, K. M. Merz, B. Wang, D. A. Pearlman, M. Crowley, S. Brozell, V. Tsui, H. Gohlke, J. Mongan, V. Hornak, G. Cui, P. Beroza, C. Schafmeister, J. W. Caldwell, W. S. Ross, and P. A. Kollman. AMBER 8, *University of California San Francisco* (2004).

[21] T. Nakano, T. Kaminuma, T. Sato, Y. Akiyama, M. Uebayasi and K. Kitaura, Fragment molecular orbital method: application to polypeptides, *Chem. Phys. Lett.*, **318**, 614-618 (2000).

[22] T. Nakano, T. Kaminuma, T. Sato, K. Fukuzawa, Y. Akiyama, M. Uebayasi and K. Kitaura, Fragment molecular orbital method: use of approximate electrostatic potential, *Chem. Phys. Lett.*, **351**, 475-480 (2002).

[23] Schrödinger Inc. Portland, OR. http://www.schrodinger.com/

[24] M. Taiji, MDGRAPE-3 chip: a 165 Gflops application specific LSI for molecular dynamics simulations, *Proceedings of Hot Chips 16, IEEE Computer Society, in CD-ROM.* (2004).

[25] T. Narumi, Y. Ohno, N. Okimoto, T. Koishi, A. Suenaga, N. Futatsugi, R. Yanai., R. Himeno, S. Fujikawa, M. Ikei and M. Taiji . A 185 Tflops simulation of amyloid-forming peptides from Yeast Prion Sup35 with the special-purpose computer System MD-GRAPE3, *Proc Supercomputing 2006, in CD-ROM* (2006).

[26] J. P. Ryckaert, G. Ciccotti and H. J. C. Berendsen, Numerical-Integration of Cartesian Equations of Motion of a System with Constraints - Molecular-Dynamics of N-Alkanes, *J. Comput. Phys.*, **23**, 327-341 (1977).

[27] T. Darden, D. York and L. Pedersen, Particle Mesh Ewald - an N.Log(N) Method for Ewald Sums in Large Systems, *J. Chem. Phys.*, **98**, 10089-10092 (1993).

[28] H. J. C. Berendsen and J. P. M. Postma, W. F. Vangunsteren, A. Dinola, J. R. Haak, Molecular-Dynamics with Coupling to an External Bath, *J. Chem. Phys.*, **81**. 3684-3690 (1984).

[29] A. Suenaga, M. Hatakeyama, M. Ichikawa, X. Yu, N. Futatsugi, T. Narumi, K. Fukui, T. Terada and M. Taiji, M. Shirouzu, S. Yokoyama, A. Konagaya, Molecular dynamics, free energy, and SPR analyses of the interactions between the SH2 domain of Grb2 and ErbB phosphotyrosyl peptides, *Biochemistry*, **42**, 5195-5200 (2003).

[30] A. Suenaga, N. Takada, M. Hatakeyama, M. Ichikawa, X. Yu, K. Tomii, N. Okimoto, N. Futatsugi, T. Narumi, M. Shirouzu, S. Yokoyama, A. Konagaya and M. Taiji, Novel mechanism of interaction of p85 subunit of phosphatidylinositol 3-kinase and ErbB3 receptor-derived phosphotyrosyl peptides, *J. Biol. Chem.*, **280**, 1321-1326 (2005).

[31] N. Okimoto, T. Nakamura, A. Suenaga, N. Futatsugi, Y. Hirano, Y. Yamaguchi and T. Ebisuzaki, Cooperative motions of protein and hydration water molecules: Molecular dynamics study of scytalone dehydratase, *J. Am. Chem. Soc.*, **126**, 13132-13139 (2004).
[32] http://www.gromacs.org/
[33] J. Wang, R. M. Wolf, Caldwell, J. W.;P. A. Kollman and D. A. Case, Development and testing of a general AMBER force field, *J. Comput. Chem.*, 25, 2004, 1157-1174.
[34] C. I. Bayly, P. Chieplak, W. D. Cornell and P. A. Kollman, A Well Behaved Electrostatic Potential Based Method Using Charge Restraint deriving Atomic Charges: The RESP model, *J. Phys. Chem.*, **97**, 10268-10280 (1993).
[35] Y. Mochizuki, S. Koikegami, T. Nakano, S. Amari, and K. Kitaura, Large scale MP2 calculations with fragment molecular orbital scheme, *Chem. Phys. Lett.*, **396**, 473-479 (2004).
[36] Y. Mochizuki, T. Nakano, S. Koikegami, S. Tanimori, Y. Abe, U. Nagashima, and K. Kitaura, A parallelized integral-direct second-order Moeller-Plesset perturbation theory method with a fragment molecular orbital scheme, *Theor. Chem. Acc.*, **112**, 442-452 (2004).
[37] K. Fukuzawa, Y. Mochizuki, S. Tanaka, K. Kitaura, and T. Nakano, Molecular Interactions between Estrogen Receptor and Its Ligand Studied by the Ab Initio Fragment Molecular Orbital Method, *J. Phys. Chem. B*, **110**, 16102-16110 (2006).
[38] K. Fukuzawa, Y. Komeiji, Y. Mochizuki, A. Kato, T. Nakano and S. Tanaka, Intra- and Inter-molecular Interactions between Cyclic-AMP Receptor Protein and DNA: Ab initio Fragment Molecular Orbital Study, *J. Comput. Chem.*, **27**, 948-960 (2006).
[39] T. Ishikawa, T. Ishikura and K. Kuwata, Theoretical study of the prion protein based on the fragment molecular orbital method, *J. Comput. Chem.*, **30**, 2594-2601, (2009).
[40] H. Gohlke and D. A. Case, Converging free energy estimates: MM-PB(GB)SA studies on the protein-protein complex Ras-Raf, *J. Comput. Chem.*, **25**, 238-250 (2004).
[41] J. Numata, M. Wan and E. W. Knapp, Conformational entropy of biomolecules:beyond the quasi-harmonic approximation, *Genome Inform.*, **18**, 192-205 (2007).
[42] C. E. Chang, W. Chen and M. K. Gilson, Evaluating the accuracy of the quasiharmonic approximation, *J. Chem. Theor. Comput*, **1**, 1017-1028 (2005).
[43] M. Retegan, A. Milet and H. Jamet, Exploring the Binding of Inhibitors Derived from Tetrabromobenzimidazole to the CK2 Protein Using a QM/MM-PB/SA Approach, *J. Chem. Inf. Model*, **49**, 963-871 (2009).
[44] B. Kuhn, P. Gerber, T. Schulz-Gasch and M. Stahl, Validation and use of the MM-PBSA approach for drug discovery, *J. Med. Chem.*, **48**, 4040-4048 (2005).
[45] S. Huo, J. Wang, P. Cieplak, P. A. Kollman and I. D. Kuntz, Molecular dynamics and free energy analyses of cathepsin D-inhibitor interactions: insight into structure-based ligand design, *J. Med. Chem.*, **45**, 1412-1419 (2002).
[46] A. Nicholls, D. L. Mobley, J. P. Guthrie, J. D. Chodera, D. I. Bayly, M. D. Cooper and V. J Pande, Predicting small-molecule solvation free energies: an informal blind test for computational chemistry, *J. Med. Chem.*, **51**, 769-779 (2008).