stability is strongly affected by the presence of water molecules, is difficult to be predicted in a computer only from its amino sequence. We have presented a modified force field that is applicable even for these proteins. The advantage of our new force field was demonstrated in terms of conformational stability of amino acids, stability of the secondary structure, and predictability of the tertiary structure. The accuracy in predicting protein folding has been improved by this force field. A distinction between the modified force field and ff03 force field is the solvent effect. The effective charges have been calculated by QC calculations in water phase to deduce a good valance between the α-helix and β-sheet structures.

References

1. Tama, F.; Miyashita, O.; Kitao, A.; Go, N. Eur Biophys J 2000, 29, 472.
2. Samatey, F. A.; Matsunami, H.; Imada, K.; Nagashima, S.; Shaikh, T. R.; Thomas, D. R.; Chen, J. Z.; Derosier, D. J.; Kitao, A.; Namba, K. Nature 2004, 431, 1062.
3. Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; Caldwell, J.; Wang, J.; Kollman, P. J Comput Chem 2003, 24, 1999.
4. Lee, M. C.; Duan, Y. Proteins 2004, 55, 620.
5. Pedersen, A. K.; Peters, G. G.; Moller, K. B.; Iversen, L. F.; Kastrup, J. S. Acta Crystallogr D Biol Crystallogr 2004, 60, 1527.
6. Chen, J. M.; Xu, S. L.; Wawrzak, Z.; Basarab, G. S.; Jordan, D. B. Biochemistry 1998, 37, 17735.
7. Zhao, B.; Guengerich, F. P.; Voehler, M.; Waterman, M. R. J Biol Chem 2005, 280, 42188.
8. Cherbavaz, D. B.; Lee, M. E.; Stroud, R. M.; Koshland, D. E., Jr. J Mol Biol 2000, 295, 377.
9. Katagiri, D.; Tsuchiya, T.; Tsuda, M.; Hata, M.; Hoshino, T. J Phys Chem B 2002, 106, 9151.
10. Tsuchiya, T.; Katagiri, D.; Hata, M.; Hoshino, T.; Tsuda, M. J Mol Struct (Theochem) 2002, 589–590, 413.
11. Wang, Z. X.; Duan, Y. J Comput Chem 2004, 25, 1699.
12. Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. Chem Phys Lett 1995, 246, 122.
13. Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. J Phys Chem 1996, 100, 19824.
14. Tsui, V.; Case, D. A. Biopolymers (Nucl Acid Sci) 2001, 56, 275.
15. Tsui., V.; Case., D. A. J Am Chem Soc, 2000, 122, 2489.
16. Onufriev, A.; Bashford, D.; Case, D. A. Proteins 2004, 55, 383.
17. Feig, M.; Onufriev, A.; Lee, M. S.; Im, W.; Case, D. A.; Brooks, C. L., 3rd. J Comput Chem 2004, 25, 265.
18. Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. J Phys Chem 1993, 97, 10269.
19. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A.; Jr., T. V.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. Gaussian, Inc, Wallingford, CT; 2004.
20. Tomasi, J.; Mennucci, B.; Cancès, E. J Mol Struct (Theochem) 1999, 464, 211.
21. Pomelli, C. S.; Tomasi, J.; Barone, V. Theor Chem Acc 2001, 105, 446.
22. Baronea, V.; Cossi, M.; Tomasi, J. J Chem Phys 1997, 107, 3210.
23. Cieplak, P.; Cornell, W. D.; Bayly, C.; Kollman, P. A. J Comput Chem 1995, 16, 1357.
24. Pastor, M. T.; Lopez de la Paz, M.; Lacroix, E.; Serrano, L.; Perez-Paya, E. Proc Natl Acad Sci USA 2002, 99, 614.
25. Cochran, A. G.; Skelton, N. J.; Starovasnik, M. A. Proc Natl Acad Sci USA 2001, 98, 5578.
26. Neidigh, J. W.; Fesinmeyer, R. M.; Andersen, N. H. Nat Struct Biol 2002, 9, 425.
27. McKnight, C. J.; Matsudaira, P. T.; Kim, P. S. Nat Struct Biol 1997, 4, 180.
28. Case, D. A.; Darden, T. A.; T.E. Cheatham, I.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Merz, K. M.; Wang, B.; Pearlman, D. A.; Crowley, M.; Brozell, S.; Tsui, V.; Gohlke, H.; Mongan, J.; Hornak, V.; Cui, G.; Beroza, P.; Schafmeister, C.; Caldwell, J. W.; Ross, W. S.; Kollman, AMBER8, P. A. University of California, San Francisco; 2004.
29. Pastor, R. W.; Brooks, B. R.; Szabo, A. Mol Phys 1988, 65, 1409.
30. Loncharich, R. J.; Brooks, B. R.; Pastor, R. W. Biopolymers (Nucl Acid Sci) 1992, 32, 523.
31. Izaguirre, J. A.; Catarello, D. P.; Wozniak, J. M.; Skeel, R. D. J Chem Phys 2001, 114, 2090.
32. Tugarinov, V.; Zvi, A.; Levy, R.; Anglister, J. Nat Struct Biol 1999, 6, 331.
33. Sharon, M.; Kessler, N.; Levy, R.; Zolla-Pazner, S.; Gorlach, M.; Anglister, J. Structure 2003, 11, 225.
34. Berendsen, H. J. C.; Postma, J. P. M.; Gunsteren, W. F. V.; DiNola, A.; Haak, J. R. J Chem Phys 1984, 81, 3684.
35. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J.; Klein, M. L. J Chem Phys 1983, 79, 926.
36. Darden, T.; York, D.; Pedersen, L. J Chem Phys 1993, 98, 10089.
37. Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. J Chem Phys 1995, 103, 8577.
38. Crowley, M. F.; Darden, T. A.; Cheatham, T. E. III.; Deerfield, D. W. II. J Supercomput 1997, 11, 255.
39. Toukmaji, A.; Sagui, C.; Board, J.; Darden, T. J Chem Phys 2000, 113, 10913.
40. Vranken, W. F.; Budesinsky, M.; Fant, F.; Boulez, K.; Borremans, F. A. FEBS Lett 1995, 374, 117.
41. Yuki, H.; Tanaka, Y.; Hata, M.; Ishikawa, H.; Neya, S.; Hoshino. T. J Comp Chem 2007, 28, 1091.
42. Chowdhury, S.; Lee, M. C.; Xiong, G.; Duan, Y. J Mol Biol 2003, 327, 711.
43. Zhou, R. J Mol Graph Model 2004, 22, 451.
44. Pitera, J. W.; Swope, W. Proc Natl Acad Sci USA 2003, 100, 7587.
45. Zhou, R. Proc Natl Acad Sci USA 2003, 100, 13280.
46. Herzberg, O. J Mol Biol 1991, 217, 701.
47. Fujii, Y.; Okimoto, Y.; Hata, M.; Narumi, T.; Yasuoke, K.; Susukita, R.; Suenaga, R.; Futatsugi, N.; Koishi, T.; Furusawa, H.; Kawai, A.; Ebisuzaki, T.; Neya, S.; Hoshino, T. J Phys Chem B 2003, 107, 10274.
48. Kaldor, S.W.; Kalish, V. J.; Davies, J. F., II; Shetty, B. V.; Fritz, J. E.; Appelt, K.; Burgess, J. A.; Campanale, K. M.; Chirgadze, N. Y.; Clawson, D.K.; Dressman, B. A.; Hatch, S. D.; Khalil, D. A.; Kosa,

M. B.; Lubbehusen, P. P.; Muesing, M. A.; Patick, A. K.; Reich, S. H.; Su, K. S.; Tatlock, J. H. J Med Chem 1997, 40, 3979.
49. Ode, H.; Neya, S.; Hata, M.; Sugiura, W.; Hoshino, T. J Am Chem Soc 2006, 128, 7887.
50. Walker, R. *Amber Tutorials*. Available at: http://amber.scripps.edu/tutorials/. Last accessed 3 March 2008.
51. Hata, M.; Fujii, Y.; Tanaka, Y.; Ishikawa, H.; Ishii, M.; Neya, S.; Tsuda, M.; Hoshino, T. Biol Pharm Bull 2006, 29, 2151.
52. Ode, H.; Neya, S.; Hata, M.; Sugiura, W.; Hoshino, T. J Am Chem Soc 2006, 128, 7887.
53. Ode, H.; Matsuyama, S.; Hata, M.; Hoshino, T.; Kakizawa, J.; Sugiura, W. J Med Chem 2007, 50, 1768.
54. Ode, H.; Matsuyama, S.; Hata, M.; Neya, S.; Kakizawa, J.; Sugiura, W.; Hoshino, T. J Mol Biol 2007, 370, 598.
55. Hata, M.; Tanaka, Y.; Fujii, Y.; Neya, S.; Hoshino, T. J Phys Chem B 2005, 109, 16153.
56. Wang, Z.-X.; Duan, Y. J Comput Chem 2004, 25, 1699.
57. Hirota, N.; Mizuno, K.; Goto, Y. J Mol Biol 1998, 275, 365.
58. Luo, P.; Baldwin, R. L. Biochemistry 1997, 36, 8413.
59. Kurpin, S.; Graslund, A.; Ehrenberg, A.; Koch, M. H. J. Biochem Biophys Res Commun 1995, 217, 1151.
60. Fraczkiewicz, R.; Braun, W. J. Comp. Chem. 1998, 19, 319.
61. Jorgensen, W. L.; Madura, J. D. Mol Phys 1985, 56, 1381.
62. Mahoney, M. W.; Jorgensen, W. L. J Chem Phys 2000, 112, 8910.
63. Caldwell, J. W.; Kollman, P. A. J Phys Chem 1995, 99, 6208.
64. Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. J Phys Chem 1987, 91, 6269.

Japanese Journal of Applied Physics
Vol. 47, No. 5, 2008, pp. 3719–3725

Accurate Evaluation Method of Molecular Binding Affinity from Fluctuation Frequency

Tyuji HOSHINO[1,2*], Koji IWAMOTO[1], Hirotaka ODE[1], and Iwao OHDOMARI[3]

[1]*Graduate School of Pharmaceutical Sciences, Chiba University, Chiba 263-8522, Japan*
[2]*PRESTO, JST, Kawaguchi, Saitama 332-0012, Japan*
[3]*School of Science and Engineering, Waseda University, Tokyo 169-8555, Japan*

(Received November 22, 2007; accepted January 9, 2008; published online May 16, 2008)

Exact estimation of the molecular binding affinity is significantly important for drug discovery. The energy calculation is a direct method to compute the strength of the interaction between two molecules. This energetic approach is, however, not accurate enough to evaluate a slight difference in binding affinity when distinguishing a prospective substance from dozens of candidates for medicine. Hence more accurate estimation of drug efficacy in a computer is currently demanded. Previously we proposed a concept of estimating molecular binding affinity, focusing on the fluctuation at an interface between two molecules. The aim of this paper is to demonstrate the compatibility between the proposed computational technique and experimental measurements, through several examples for computer simulations of an association of human immunodeficiency virus type-1 (HIV-1) protease and its inhibitor (an example for a drug–enzyme binding), a complexation of an antigen and its antibody (an example for a protein–protein binding), and a combination of estrogen receptor and its ligand chemicals (an example for a ligand–receptor binding). The proposed affinity estimation has proven to be a promising technique in the advanced stage of the discovery and the design of drugs. [DOI: 10.1143/JJAP.47.3719]

KEYWORDS: binding energy, drug design, computation, affinity, interfacial fluctuation

1. Introduction

Computers play a significant role in the drug discovery and development of today. One of the major applications of computers in drug discovery is the virtual *in silico* screening to find hit chemicals interacting with a target protein, which reduces experimental efforts in the real *in vitro* screening. In the *in silico* screening, hit chemicals are searched from a database in terms of the similarity to several bioactive molecules which were already known to be bound to the target protein. Unless the information on these bioactive molecules is available in advance, we cannot perform the *in silico* screening. Further the screening cannot find hit chemicals, if the database has no entry on the chemicals. In order to overcome the difficulty in the *in silico* screening, the structure-based drug design has attracted much attention recently. Since the structure-based drug design takes both the chemical compound and the target protein into account, it demands heavy computational costs. Nevertheless, the structure-based drug design currently becomes a practical technology due to the rapid progress in the performance of computers.

The structure-based drug design requires an accurate evaluation of the binding affinity of chemicals to the target protein. One of the approaches to estimate the binding affinity in a computer is the energy calculations based on molecular dynamics (MD) simulations, in which molecular mechanics generalized Born surface area (MMGBSA)[1] or molecular mechanics Poisson Boltzmann surface area (MMPBSA) method[2] is a standard technique and frequently used in the field of computational biology. The energy computed by MMGBSA or MMPBSA method, however, usually show differences of at least more than several kcal/mol among snapshot structures during MD simulation. For the purpose of estimating more accurate binding energy, some advanced techniques like the free energy perturbation[3,4] or the thermodynamic integration[5] have been developed so far. The free energy perturbation method was shown to be effective[3] when an accurate binding free energy for one chemical to the target protein was already available and simulation was performed to evaluate the affinity of the designed compounds having a similar but another chemical structure. The other computational approaches for affinity estimation are surface complementarity,[6] B-factor analysis,[7] and molecular orbital approach.[8] While these techniques provide good estimation in some biomolecules, they have proved unsatisfactory in our trials for the estimation of drug efficacy of inhibitors against human immunodeficiency virus type-1 (HIV-1) protease. Hence, a more precise estimation method of molecular binding affinity is required for the success of the structure-based drug design and for the advance of drug discovery. In this study, HIV-1 protease and its inhibitor are firstly instanced to explain the relationship between the binding affinity and the fluctuation property at their interface. Next a combination of an hen egg-white lysozyme and its antibody HyHEL-10 is examined to demonstrate a compatibility between the estimation from the interfacial fluctuation and the experimentally measured binding energy. Furthermore, an association of human estrogen receptor α-subtype ligand binding domain (hERαLBD) and its ligand chemicals is picked as an example for the ligand–receptor complex.

2. Methods

2.1 *Construction of the model structures for the drug–enzyme, antigen–antibody, and ligand–receptor complexes*

All initial structures have been constructed from the X-ray crystal structures (PDB code: 1OHR for NFV-bound HIV-1 protease, 1C08, 1IC4, 1IC5, 1IC7, 1J1O, 1J1P, 1J1X for the wild type and HD32A, HD96A, HD32AD96A, LY50F, LS91A, LS93A mutants of HyHEL-10 antibody, 1ERE, 3ERD, 1ERR, 3ERT, 1X7R for EST, DES, RAL, OHT, or GEN-associated hERαLBD). Mutated residues in 1OHR crystal data have been converted to be identical with the wild-type sequence: HXB2cv.[9] Missing residues in the hERαLBD crystal data have been added manually to complete the protein structure. Model structures for inhib-

*E-mail address: hoshino@faculty.chiba-u.jp

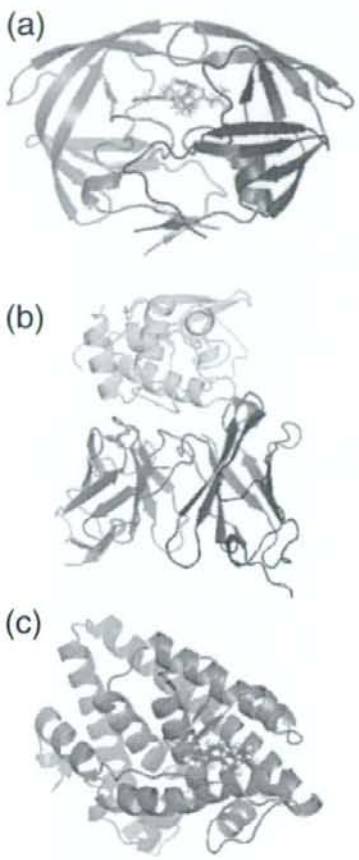

Fig. 1. (Color online) Structures of (a) NFV-bound wild type HIV-1 protease, (b) the complex of an antigen lysozyme and its antibody HyHEL-10, and (c) ligand-associated hERαLBD. Drug is represented by sticks in (a). The antigen is represented in yellow, and the heavy and light chains of the antibody are blue and green in (b). Ligands are represented by sticks in (c).

itor-bound HIV-1 protease, antigen-associated HyHEL-10 antibody, and ligand-bound hERαLBD are shown in Fig. 1.

2.2 *Minimization and equilibration*

MD simulations are carried out using the AMBER8 program package with the Amber ff03 force field.[10] Complexes are solvated in a box and the periodic boundary condition is applied. An integration time step of the simulation is 1 fs. The particle mesh Ewald method and a cut-off distance of 12 Å for non-bonded interactions are employed, and the pressure is kept constant at 1 atm. First, potential energy minimizations are performed starting from the respective crystal structure. The earlier 3000 steps in minimization are executed by the steepest descent method and the later 7000 steps are by the conjugated gradient method. Next, MD simulation is performed from the energy minimized structure. The temperature of the complex is gradually increased by heating up to 300 K for 60 ps and then kept at 300 K for over next 2 ns. Judging from the changes in root mean square deviation and total potential energy and from the principal component analysis, the complex is confirmed to be in equilibration. Equilibration is very important for correct affinity evaluation. The trajectory at 300 K for the subsequent 128 ps simulation is considered to be the most probable structure under the physiological conditions and the energies are collected for affinity evaluation. The principal component analysis is again executed to confirm that the complex is in a single conformation during the acquired trajectory. MMGBSA and MMPBSA calculations are carried out by the mm-pbsa module[11] of AMBER8 using the acquired trajectory for the 128 ps MD simulation.

2.3 *Data acquisition, Fourier transformation, and affinity evaluation*

During the 128 ps MD simulation for data acquisition, Coulomb and van der Waals potential energies between the drug and enzyme, the antigen and antibody, or the ligand and receptor are calculated every integration time step (1 fs). No cut-off method is applied for this energy calculation. In the present work, only van der Waals energy is employed for the evaluation. Not only interfacial energy but also some specific inter-atomic distances can be a barometer to reflect the binding state at the interface.[12,13] Nevertheless, the interfacial van der Waals potential energy has been selected for data acquisition, because van der Waals energy collects the contributions from not only specific local bonds but all area of the interface. The energies are averaged every 1000 steps (1 ps), which generates a sequence of 128 energies along the simulation time. For normalization, 128 energies are scaled by the average. The deviation from the average, this average is equivalent to 1.0 at this stage, is calculated for the respective time point and 128 deviations are again scaled by the maximum one. These 128 sequential values are converted into the frequency components through discrete Fourier transformation. A sequence of 128 energies generates 2 non-periodic and 63 cosine and 63 sine wave components, each of which is characterized by the integer corresponding to the wave number; in other word, frequency. The cosine and sine waves with the same frequency can be combined into a single form. Two wave functions are combined by $A_i \cos(\omega_i t) + B_i \sin(\omega_i t) = q_i \cos(\omega_i t + \phi)$, where $q_i = \sqrt{A_i^2 + B_i^2}$ and $\tan\phi = -B_i/A_i$. A_i and B_i are Fourier coefficients and q_i is called Fourier amplitude. Each of the Fourier components is labelled by a wave number i, which is one of the integers from 0 to 63.

2.4 *Theory*

MD simulations were performed in our recent studies to investigate the action of HIV-1 protease inhibitors,[14,15] the effectiveness of antibiotics,[16] and the interaction of membrane proteins.[17–19] In the present work, these simulation results have been closely examined by converting the obtained dynamic atom motions into animations. The examination has suggested that the atoms locating at the interface between a drug and its target enzyme or between a protein and its conjugating protein rapidly fluctuate with the decrease in binding affinity. This is an important clue to finding a characteristic index of binding affinity. Figures 2(a) and 2(b) depict the situation that a drug is bound to its target enzyme, where several structures are super-

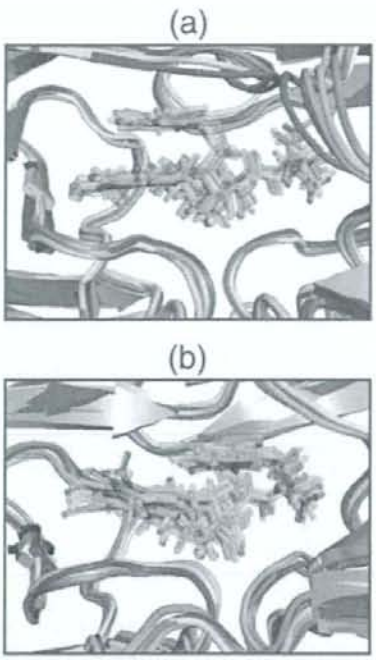

Fig. 2. (Color online) Structure of the binding pocket of HIV-1 protease associated with an inhibitor; NFV. Each image is obtained by the superimposition of six snapshot structures acquired every 10 ps for 50 ps MD simulation. The protease is depicted by a tube, gradated from dark blue to light green with the time-evolution. The inhibitor is represented by sticks with a gradation from light pink to bright red. (a) The firmly fit situation with the wild type protease. (b) The loosely bound situation with the D30N mutant. In the D30N mutant (b), NFV loses the hydrogen bond with the 30th residue of the protease and its side part starts to flicker.

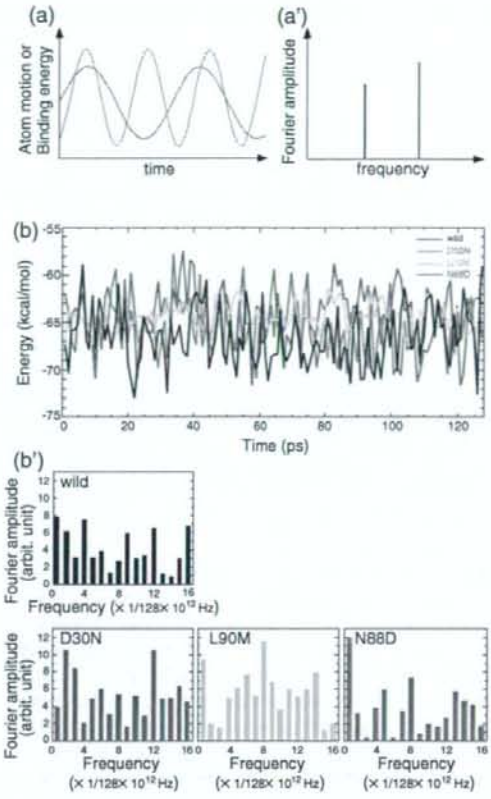

Fig. 3. (Color online) (a) Schematic representation of periodic fluctuations in atom motion and/or binding energy along the simulation time. Blue and red curves indicate the firmly fit and loosely bound cases. (a') Fourier components derived from Fourier transformation of (a). This is a core concept of binding affinity estimation from interfacial fluctuation. (b) Change of van der Waals potential energy at the interface between HIV-1 protease and its inhibitor; NFV, acquired through 128 ps MD simulation at 300 K after the equilibration of the drug–enzyme complex. (b') Fourier spectrum derived from the Fourier transformation of the 128 sampling energy data of (b). D30N, L90M, and N88D mutants are contrasted with the wild type protease.

imposed to represent the time-evolution of atom geometry. In the firmly fit case, the drug is steadily held inside the binding pocket and chemical groups of the drug seldom alter their orientation [Fig. 2(a)]. In contrast, the drug rapidly fluctuates and frequently changes its conformation in the loosely bound case [Fig. 2(b)].

Fourier transformation is one of the most promising approaches to analyze the dynamic motion of atoms. A concept of the presently proposing method of the binding affinity evaluation is explained in Fig. 3. Figure 3(a) shows schematic atom motion and/or binding energy seen in MD simulations, where the blue and red curves represent the firmly fit and the loosely bound cases. If the time-evolution of atom motion and/or binding energy shows the periodic change like Fig. 3(a), Fourier transformation gives a prominent difference in the Fourier components with respect to frequency. Namely, Fourier transformation will suggest the tendency that Fourier component appears at the higher frequency region in the loosely bound case, while at the lower frequency region in the firmly fit case [Fig. 3(a')]. The fluctuation in atom motion or binding energy is not so simple as Fig. 3(a) in the actual drug–enzyme complex. That is, the fluctuation consists of more motions with different vibration frequencies. Accordingly contributions from every motion with different frequencies should be taken into account to extract an index for the molecular binding affinity. The summation of the Fourier components weighted by a function of frequency will be one of the promising means.

Since the fluctuation frequency at the interface is closely correlated with the stability of the complex, an index for binding affinity will be derived from the time-evolved data on the interfacial energy.[20] Fluctuation can be explained from a collective set of vibrations, each of which is quantized by an oscillation frequency of ω_i and has an energy of $\hbar\omega_i$. The total vibration energy at a temperature T is given as the sum of the energies over all vibrational frequencies; $\sum_i \hbar\omega_i \langle n_i \rangle$. Here $\langle n_i \rangle$ is the thermal equilibrium occupancy of vibrations that is expressed by the distribution function $1/[\exp(\hbar\omega_i/kT) - 1]$ called the Planck formula. Since the interfacial fluctuation contains many vibrational modes even for one frequency ω_i, the number of possible modes of vibrations; in other words, the density of vibrational modes, should be taken into account in the estimation.

The number of possible vibrational modes is reflected in the respective Fourier amplitude. Hence, Fourier amplitude (q_i) contributes as a weighting factor for a frequency (ω_i) in the summation over i;

$$\sum_i q_i \frac{\hbar\omega_i}{\exp(\hbar\omega_i/kT) - 1} \quad (1)$$

It should be emphasized that the estimated value from eq. (1) becomes small with the increase of binding affinity, because this quantitative function reflects the instability of the interface of a complex. In eq. (1), $\omega_i = 2\pi/(128 \times 10^{-12}) \times i(s^{-1})$ and $T = 300\,\mathrm{K}$. k is Boltzmann's constant and $\hbar$ is Planck's constant divided by 2π. Since high frequency components may contain local libration of residues, they should be excluded in the evaluation to avoid an unfavourable error. Hence, the upper bound of the summation of eq. (1) is set to 16, which corresponds to the periodicity of 8 ps.

3. Results

HIV-1 protease is an enzyme essential for the reproduction of the virus. Nelfinavir (NFV) is one of the protease inhibitors currently used in the treatment of HIV infection. HIV-1 protease frequently causes point mutations and, for example, the D30N and L90M mutations are known to decrease the inhibitory ability of NFV; in contrast, the N88D mutant becomes more susceptible to NFV. These changes in drug efficacy of NFV were already evaluated in experiments. The effective concentration 90% (EC_{90}), which is one of the experimental measurements of drug efficacy, was reported to increase 6 and 5 times for the D30N and L90M mutants compared to the wild type and to decrease to two-thirds for the N88D mutant.[21] MD simulations were performed for the NFV-bound HIV-1 protease in the wild type and the D30N, L90M, and N88D mutants. After the complex structure of the NFV-bound HIV-1 protease became equilibrated in simulation, van der Waals energy at the interface between NFV and the protease was acquired as a function of time [Fig. 3(b)]. The number of the acquired sequential energies is 128 for the respective complex. The sequential data on the interfacial energy have been converted into the frequency components through the Fourier transformation. Figure 3(b′) shows the Fourier spectrum which represents the magnitude of the respective periodic components up to the wave number 16. It is obvious that larger Fourier amplitudes appear in the higher frequency regions in the D30N and L90M mutants compared to the wild type, while the large amplitudes in the higher frequency regions seem less prominent in the N88D mutant.

The binding affinity between NFV and HIV-1 protease has been estimated from eq. (1), based on the Fourier components shown in Fig. 3(b′). The estimations for the D30N, L90M, and N88D mutants are presented relative to the wild type protease. Figure 4(a) shows a correlation between the estimated binding affinity and the experimental virus resistivity. The resistivity is given in logarithmic transform of the ratio of drug concentrations in the same inhibitory effect relative to the wild type; $\log(EC_{90\_mutant}/EC_{90\_wild})$, which has the dimension of energy and is usually defined as the drug potency. This drug potency shows a clear linearity with the computational binding affinity in Fig. 4(a). The correlation is excellent between the computational affinity and the experimental measurement ($R = 0.99$).

A combination of an antigen (lysozyme) and its antibody (HyHEL_10) has been well studied experimentally on the wild type antibody and several mutants.[22] This is a good example for testing the validity of our affinity estimation and a typical instance for protein–protein interaction. In order to execute the computational affinity evaluation, MD simulations have been performed for the wild type HyHEL_10 and its mutants (HD32A, HD96A, and HD32AD96A). Through these MD simulation, van der Waals potential energies have been acquired as sequential data. The binding affinities between the antigen and the wild type or mutated antibodies have been estimated using eq. (1) and compared with the experimentally measured binding free energy [Fig. 4(b)]. Our affinity evaluation again shows excellent consistency with the experimental measurement ($R = 0.99$). Another experimental measurement was reported on the mutations in the light chain of the HyHEL_10 antibody; LS91A, LS93A, LY50F.[23] The binding affinities between the antigen and the wild type and mutated antibodies have been evaluated by eq. (1) [Fig. 4(c)]. Our affinity evaluation also shows a linear correlation with experimentally measured binding energy ($R = 0.91$).

Due to the therapeutic interest and the concern as an environmental hormone, estrogens have been extensively studied from the viewpoint of the affinity with their receptor. Because an estrogen receptor is capable of binding a variety of ligands and because experimental measurement of their binding affinity is available, the association of an estrogen receptor with its ligand chemicals has been a suitable subject for the trial of the computational binding assay.[8,24] The computational evaluation has been carried out, acquiring the sequential data on the interfacial energy through the MD simulations on the human estrogen receptor α-subtype ligand binding domain (hERαLBD) complexed with several kinds of ligand chemicals; 17β-estradiol (EST), diethylstilbestrol (DES), raloxifene (RAL), 4-hydroxytamoxifen (OHT), genistein (GEN). Our affinity evaluation for the above ligand chemicals provides a sound compatibility with the experimental ligand activity[25] ($R = 0.91$) [Fig. 4(d)]. In all present trials with HIV-1 protease, HyHEL_10 antibody, and estrogen receptor, the approach proposed in this study shows better compatibility with the experimental measurements compared to the conventional MMGBSA and MMPBSA methods (Fig. 5).

4. Discussion

The proposed affinity evaluation from interfacial fluctuation shows a keen analogy to the thermal energy of solids. A solid contains a lot of lattice vibrations, each of which is quantized by a frequency ν associating with an energy $h\nu$. The thermal energy of the solid at a temperature T is given as the total sum of the energies over all vibrational modes;

$$\int_0^{\nu_D} Z(\nu) \frac{h\nu}{\exp(h\nu/kT) - 1}\, d\nu, \quad (2)$$

where $1/[\exp(h\nu/kT) - 1]$ is the Planck distribution function, $Z(\nu)\,d\nu$ is the number of possible modes of vibrations in the frequency interval between ν and $\nu + d\nu$, and ν_D is the upper bound for the vibration frequency. Debye derived his

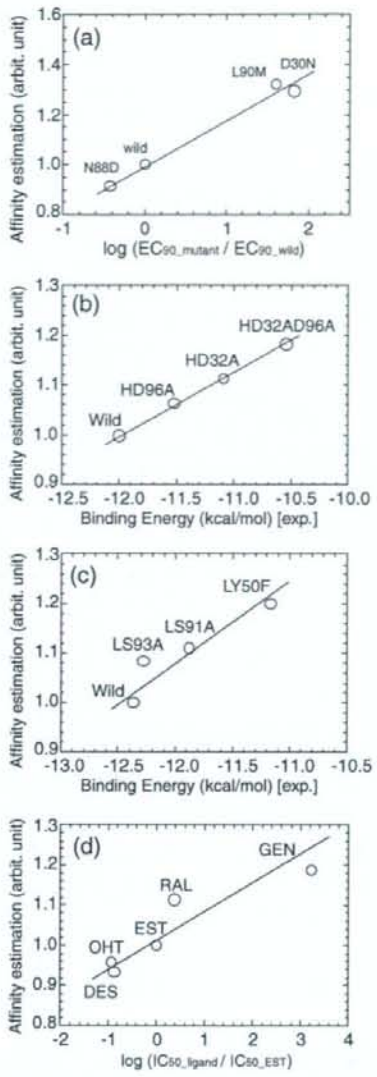

Fig. 4. Comparison of the presently proposed affinity evaluation with the experimentally measured effective concentration 90% (EC_{90}), binding energy, or inhibitory concentration 50% (IC_{50}). (a) Affinity of NFV with the wild type HIV-1 protease and its mutants. (b) Affinity of antigen–antibody complexes. HyHEL_10, which is a typical antibody for lysozyme, is exemplified with its variants mutated in the heavy chain. The affinity estimation is presented relative to that for the wild type. (c) Affinity of the lysozyme-bound HyHEL_10 antibody and its variants mutated in the light chain. (d) Affinity of ligand–receptor complex, exemplified by human estrogen receptor α ligand binding domain and its ligand chemicals. The binding affinities are presented relative to the estimation for 17β-estradiol (EST).

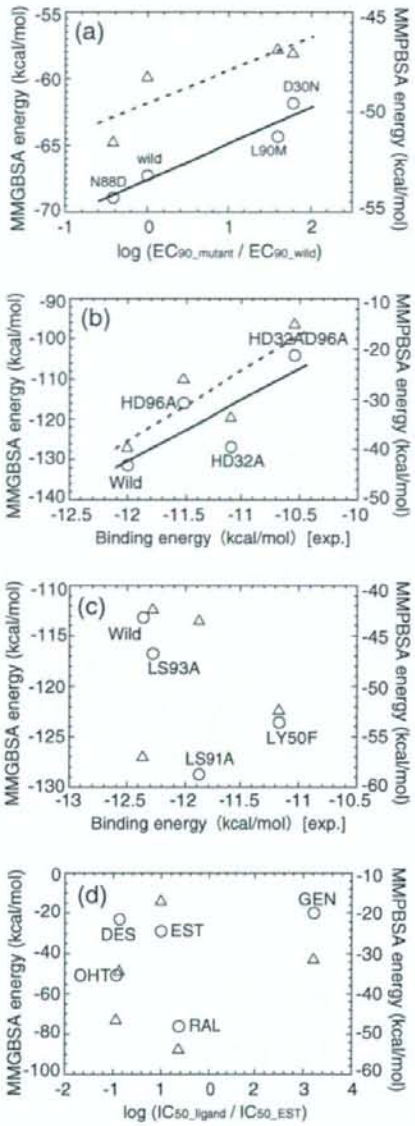

Fig. 5. Comparison of the computational energy with the experimental binding affinity. The binding energies computed by MMGBSA method are presented by triangles and the energies by MMPBSA method are by circles. (a) NFV-bound HIV-1 protease. Both MMGBSA and MMPBSA methods show good correlation with experimentally measured virus resistivity ($R = 0.87$ and 0.96). (b) Complex of an antigen lysozyme and its antibody HyHEL_10, with the variants mutated in the heavy chain. A compatibility with the experimental binding energy is good both in MMGBSA ($R = 0.83$) and MMPBSA ($R = 0.78$) methods. (c) Antigen-complexed HyHEL_10 antibody, with the variants mutated in the light chain. Computations are inconsistent with the experimental binding energy in both methods. (d) Ligand-associated hERαLBD. Neither method gives a clear compatibility with the experimental measurement.

T^3 law on the heat capacity from eq. (2) and achieved a perfect explanation on the experimental thermal property of a solid crystal.[26] Equation (1) proposed in this study is equivalent to eq. (2) when the summation is displaced by integration. The value estimated from eq. (2) is in a dimension of energy, which is stored in a solid as lattice vibration. Likewise, eq. (1) is interpreted to represent a value in a dimension of energy, which is accumulated at an interface as atom fluctuation.

The high frequency components, whose wave number is over 16; i.e., whose periodicity is below 8 ps, have been excluded in the affinity estimation of our present trials, which gives a better compatibility with the experimental binding energy compared to the inclusion of all frequency components. This exclusion is related to the atom motions detected by the nuclear magnetic resonance (NMR). NMR is an experimental technique essential for analyzing the internal motions of proteins and is sensitive not only to the magnitude but also to the time scale of the motions. A typical NMR time-correlation function for internal motion in proteins shows two major time regions.[27] One is a large peak at the region below several picoseconds and the other is a moderate plateau at the region from tens to a few hundred picoseconds. The former corresponds to the librational motion of residues and the latter to the dynamic motion of domains. The latter dynamics is responsible for the hold of molecules as a whole, while the former libration scarcely characterizes the binding state at the interface. This means that the upper bound of the frequency in the summation of eq. (1) is effective to focus only on the fluctuations relevant to the molecular binding.

Our approach using the Fourier transformation provides two important pieces of information. One is the summation of Fourier components weighted by a function of frequency like eq. (1). The other is the composition of frequency spectrum. For example, the character of the inter-molecular interaction appears in the frequency spectra for the D30N L90M, and N88D mutants of Fig. 2(b′). The 30th residue of HIV-1 protease makes a direct hydrogen bond with NFV. This hydrogen bond disappears due to the D30N mutation and the binding ability of NFV is seriously reduced. In contrast, the L90M mutation causes serious deformation of the binding pocket.[14] Due to this deformation, the orientation of NFV alters inside the binding pocket to avoid the fatal collision with the protease. The N88D mutant also causes noticeable deformation of the binding pocket. Interestingly the fitness of NFV for the deformed pocket is improved in the N88D mutant. We notice that the frequency spectrum of the D30N mutant resembles that of the wild type. On the other hand, the frequency spectra of the L90M and N88D mutants are quite different from that of the wild type. This means that the L90M and N88D mutants change their binding constraints on NFV compared to the wild type protease.

As exemplified in the estimation of binding affinity between HIV-1 protease and its inhibitor, the D30N or L90M mutation causes the decrease of drug efficacy. In case of an approved anti-HIV-1 drug; NFV, EC_{90} was reported to increase 6 and 5 times for the D30N and L90M mutants compared to the wild type.[21] Hence more than 5-fold dosage is required to suppress the enzymatic activity of the mutants at the same level as the suppression of the wild type virus. This means that NFV cannot be prescribed for the patient who was infected with the D30N or L90M HIV-1 variants. Since the virus easily causes amino acid mutation and the diversity of variants is high, it is preferable to assay the efficacy of drugs for individual patient before prescription. The computational assay will be one of the helpful techniques for analyzing the efficacy of drugs for individuals and will contribute the progress of the tailor-made medicine because the process can be automated to treat a large number of samples and the cost for assay will be reduced due to the increase of the performance of computer. For the purpose of drug screening, rapidness should be weighted at the cost of accuracy, since subsequent experimental assay will cover more precise screening of potent candidates. In contrast, accuracy is critically important for the choice of drugs for patient. Therefore, an accurate computational technique is indispensable for the therapy in future and the approach proposed in this study will be one of the promising methodologies.

In our present approach to estimate the binding affinity, Fourier transformation is executed to convert the data on the fluctuation of van der Waals energy into frequency components. Discrete Fourier transformation is usually employed when applied to the sampling data acquired by monitoring some signal. Since Fourier transformation can be fundamentally applied only to the periodic function, discontinuity between the values at the first and last sampling points sometimes causes an unfavourable error. Similar difficulty will also arises when a complex of some ligand and its receptor has two stable configurations and the binding configuration frequently alters. While confirming the structural equilibration is one of the fundamental notes to avoid such accidental errors, some technique to solve this difficulty should be developed in our future study.

5. Conclusions

In order to solve the problem that conventional energy calculations to estimate the ligand–receptor binding affinity are not accurate enough to distinguish a slight difference in drug efficacy, we proposed a concept of estimating molecular binding affinity, focusing on the fluctuation at a ligand–receptor interface. The proposed affinity estimation coincides well with the available experimental results such as an association of a virus protease and its inhibitor, a complexation of an antigen and its antibody, and a combination of estrogen receptor and its ligand chemicals, and has proven to be one of the promising techniques in drug design and discovery.

Acknowledgment

We thank Professor Y. Shinoda of Waseda University for his critical reading of the manuscript. This work was supported by a Grant-in-Aid for COE research from the Ministry of Education, Culture, Sports, Science, and Technology of Japan and the Health and Labour Science Research Grant for Research on HIV/AIDS from the Ministry of Health and Labour of Japan. A part of this work is also supported by the grants from Japan Science and Technology Agency.

1) D. Bashford and D. A. Case: Annu. Rev. Phys. Chem. **51** (2000) 129.
2) P. Kollman: Chem. Rev. **93** (1993) 2395.
3) K. M. Merz and P. A. Kollman: J. Am. Chem. Soc. **111** (1989) 5649.
4) S. Hirono and P. A. Kollman: J. Mol. Biol. **212** (1990) 197.
5) W. L. Jorgensen: Acc. Chem. Res. **22** (1989) 184.
6) M. C. Lawrence and P. M. Colman: J. Mol. Biol. **234** (1993) 946.
7) C. B. Post, C. M. Dobson, and M. Karplus: Proteins: Struct. Funct. Genet. **5** (1989) 337.
8) K. Fukuzawa, K. Kitaura, M. Uebayasi, K. Nakata, T. Kaminuma, and T. Nakano: J. Comput. Chem. **26** (2005) 1.

9) W. Sugiura, Z. Matsuda, Y. Yokomaku, K. Hertogs, B. Larder, T. Oishi, A. Okano, T. Shiino, M. Tatsumi, M. Matsuda, H. Abumi, N. Takata, S. Shirahata, K. Yamada, H. Yoshikura, and Y. Nagai: Antimicrob. Agents Chemother. **46** (2002) 708.
10) D. A. Case, T. A. Darden, T. E. Cheatham, 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: AMBER8 (University of California, San Francisco, CA, 2004).
11) A. Onufriev, D. A. Case, and D. Bashford: J. Comput. Chem. **23** (2002) 1297.
12) S. Y. Sheu, D. Y. Yang, H. L. Selzle, and E. W. Schlag: Proc. Natl. Acad. Sci. U.S.A. **100** (2003) 12683.
13) H. Nojima, M. Takeda-Shikata, Y. Kurihara, M. Adachi, S. Yoneda, K. Kamiya, and H. Umeyama: Chem. Pharm. Bull. **50** (2002) 1209.
14) H. Ode, M. Ota, S. Neya, M. Hata, W. Sugiura, and T. Hoshino: J. Phys. Chem. B **109** (2005) 565.
15) H. Ode, M. Hata, S. Neya, W. Sugiura, and T. Hoshino: J. Am. Chem. Soc. **128** (2006) 7887.
16) Y. Fujii, N. Okimoto, M. Hata, T. Narumi, K. Yasuoka, R. Susukita, A. Suenaga, N. Futatsugi, T. Koishi, H. Furusawa, A. Kawai, T. Ebisuzaki, S. Neya, and T. Hoshino: J. Phys. Chem. B **107** (2003) 10274.
17) Y. Sato, M. Hata, S. Neya, and T. Hoshino: Protein Sci. **14** (2005) 183.
18) K. Mori, M. Hata, S. Neya, and T. Hoshino: Chem-Bio Inf. J. **4** (2004) 15.
19) K. Mori, M. Hata, S. Neya, and T. Hoshino: J. Am. Chem. Soc. **127** (2005) 15127.
20) K. Iwamoto, H. Ode, M. Ohta, T. Misu, M. Hata, S. Neya, and T. Hoshino: Jpn. J. Appl. Phys. **44** (2005) L1370.
21) A. K. Patick, M. Duran, Y. Cao, D. Shugarts, M. R. Keller, E. Mazabel, M. Knowles, S. Chapman, D. R. Kuritzkes, and M. Markowitz: Antimicrob. Agents Chemother. **42** (1998) 2637.
22) M. Shiroishi, A. Yokota, K. Tsumoto, H. Kondo, Y. Nishimiya, K. Horii, M. Matsushima, K. Ogasawara, K. Yutani, and I. Kumagai: J. Biol. Chem. **276** (2001) 23042.
23) A. Yokota, K. Tsumoto, M. Shiroishi, H. Kondo, and I. Kumagai: J. Biol. Chem. **278** (2003) 5410.
24) P. D. Kirchhoff, R. Brown, S. Kahn, M. Waldman, and C. M. Venkatachalam: J. Comput. Chem. **22** (2001) 993.
25) G. G. J. M. Kuiper, J. G. Lemmen, B. Carlsson, J. C. Corton, S. H. Safe, P. T. van der Saag, B. van der Burg, and J. Gustafsson: Endocrinology **139** (1998) 4252.
26) A. J. Dekker: *Solid State Physics* (Prentice-Hall, New York, 1957).
27) C. L. Brooks III, M. Karplus, and B. M. Pettitt: in *Proteins: A Theoretical Perspective of Dynamics, Structure, and Thermodynamics* (John Wiley & Sons, New York, 1988) p. 201.

Journal of General Virology (2008), **89**, 3144–3149 DOI 10.1099/vir.0.2008/004820-0

Short Communication

Substitution of the myristoylation signal of human immunodeficiency virus type 1 $Pr55^{Gag}$ with the phospholipase C-δ1 pleckstrin homology domain results in infectious pseudovirion production

Emiko Urano,[1,2] Toru Aoki,[1] Yuko Futahashi,[1] Tsutomu Murakami,[1] Yuko Morikawa,[2] Naoki Yamamoto[1] and Jun Komano[1]

Correspondence
Jun Komano
ajkomano@nih.go.jp

[1]AIDS Research Center, National Institute of Infectious Diseases, 1-23-1 Toyama, Shinjuku-ku, Tokyo 162-8640, Japan

[2]Kitasato Institute of Life Sciences, Kitasato University, Shirokane 5-9-1, Minato-ku, Tokyo 108-8641, Japan

Received 11 June 2008
Accepted 20 August 2008

The matrix domain (MA) of human immunodeficiency virus type 1 $Pr55^{Gag}$ is covalently modified with a myristoyl group that mediates efficient viral production. However, the role of myristoylation, particularly in the viral entry process, remains uninvestigated. This study replaced the myristoylation signal of MA with a well-studied phosphatidylinositol 4,5-biphosphate-binding plasma membrane (PM) targeting motif, the phospholipase C-δ1 pleckstrin homology (PH) domain. PH–Gag–Pol PM targeting and viral production efficiencies were improved compared with Gag–Pol, consistent with the estimated increases in Gag–PM affinity. Both virions were recovered in similar sucrose density-gradient fractions and had similar mature virion morphologies. Importantly, PH–Gag–Pol and Gag–Pol pseudovirions had almost identical infectivity, suggesting a dispensable role for myristoylation in the virus life cycle. PH–Gag–Pol might be useful in separating the myristoylation-dependent processes from the myristoylation-independent processes. This the first report demonstrating infectious pseudovirion production without myristoylated $Pr55^{Gag}$.

The N-terminal region [$p17^{MA}$, matrix (MA) domain] of human immunodeficiency virus type 1 (HIV-1) $Pr55^{Gag}$ (Gag), a structural protein with multiple roles in the virus life cycle (Swanstrom & Wills, 1997), is covalently modified with a myristyol group that aids in plasma membrane (PM) targeting. Removal of this region leads to inefficient Gag targeting to the PM, resulting in dramatically reduced virus production (Bryant & Ratner, 1990; Göttlinger *et al.*, 1989; Pal *et al.*, 1990; Zhou *et al.*, 1994). Although viral particles can be produced by substituting MA with heterologous PM-targeting motifs, such substitution mutants show markedly reduced infectivity (Jouvenet *et al.*, 2006; Scholz *et al.*, 2008), probably due to an active role of MA in viral entry (Kiernan *et al.*, 1998; Wang *et al.*, 1993). However, direct experimental evidence of a viral entry-specific role for MA myristoylation is lacking. Such specific roles of Gag myristoylation can only be determined by separating the myristoylation-dependent PM-targeting function from other MA-associated functions.

We constructed a mutant *gag* expression plasmid where the myristoylated region of Gag was replaced with the N-terminal pleckstrin homology (PH) domain of phospholipase C-δ1 (PLCδ1), a well-studied cellular PM-targeting motif that functions similarly to the myristoyl moiety. PLCδ1 is a member of a family of inositol phospholipid-specific PLC isozymes involved in transducer-mediated intracellular responses (Berridge, 1993). The ~120 aa PH domain can bind to phosphatidylinositol 4,5-biphosphate [PI(4,5)P(2)] and localize to the PM with high affinity and specificity (Ferguson *et al.*, 1995b; Fiorentini *et al.*, 2006; Harlan *et al.*, 1994, 1995; Rhee, 2001; Yagisawa *et al.*, 1994), and green fluorescent protein-bound PLCδ1 PH domains have been used to visualize the PM in living cells (Stauffer *et al.*, 1998; Tall *et al.*, 2000).

We used a codon-optimized HIV-1 *gag–pol* expression vector (p*gag–pol*) for genetic modification of *gag*, as p*gag–pol* increases Gag expression and facilitates protein analyses (Wagner *et al.*, 2000). The substituted mutant retained an intact MA, with the exception of two N-terminal amino acid mutations (ATG→CTG and GGC→GCG), resulting in an MG→LA mutation to knock out the myristoylation signal of Gag and prevent internal translational initiation (Fig. 1a). The PLCδ1 PH domain residues 1–175 (Stauffer *et al.*, 1998) were linked to the LA–Gag N terminus by the amino acids PRAEFT, creating a PH–*gag–pol* expression vector (pPH-*gag-pol*, Fig. 1a). A control PH domain

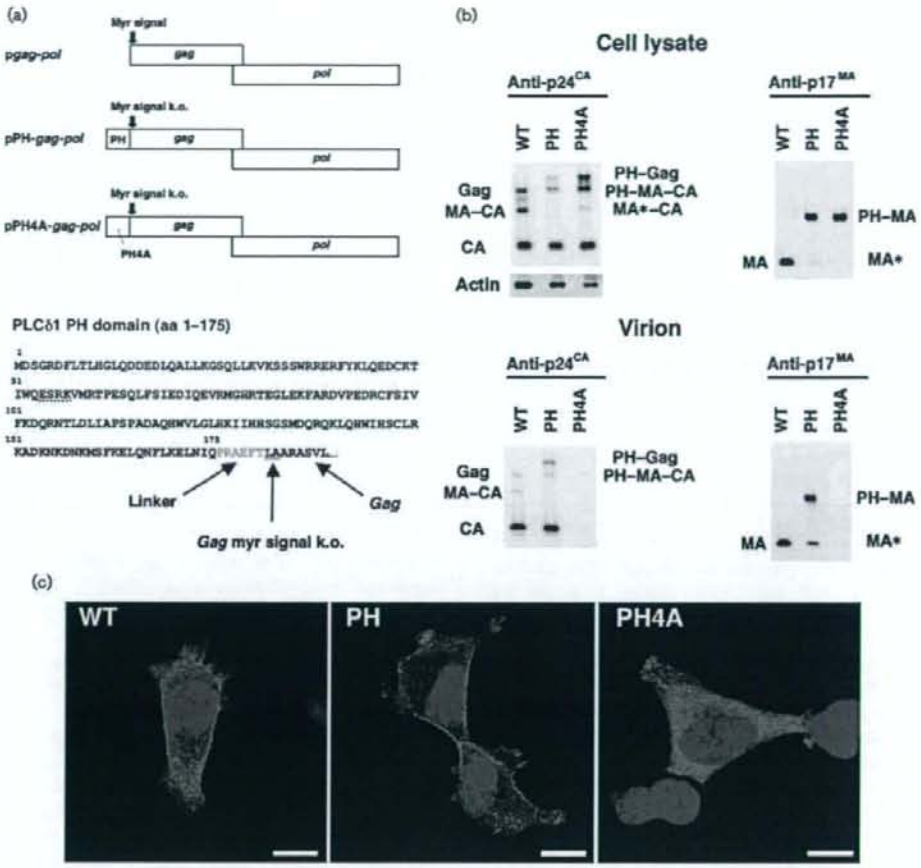

Fig. 1. Viral production by the *gag–pol* expression vectors. (a) The genetic structure of the p*gag-pol*, pPH-*gag-pol* and pPH4A-*gag-pol* expression vectors, and the amino acid sequence of the PH–Gag junction are shown. The N-terminal PH domain of PLCδ1 (aa 1–175) was fused to LA–Gag, linked by a 5 aa spacer (shown in grey). The MG→LA mutation to knock out the myristoylation signal of Gag (myr signal k.o.) is underlined. Four alanine mutations were introduced to replace the ESRK sequence (dotted line) to create the PH4A mutant. (b) Protein expression from p*gag-pol* (WT), pPH-*gag-pol* (PH) and pPH4A-*gag-pol* (PH4A) in transfected 293T cell lysates and Gag cleavage in the virions were examined by Western blot analysis using anti-$p24^{CA}$ or anti-$p17^{MA}$ antibodies. Note that the anti-$p17^{MA}$ antibody recognizes the cleaved $p17^{MA}$ protein only. The band denoted as PH–MA–CA in the virion detected by the anti-$p24^{CA}$ antibody (lower left panel) possibly overlaps with a faint Gag signal derived from PH–Gag and PH4A–Gag from which the PH and PH4A domains have been cleaved. (c) Immunofluorescence assay showing the distribution of Gag, PH–Gag and PH4A–Gag in 293T cells transfected with the respective expression plasmid. Red and blue represent $p24^{CA}$ and the Hoechst 33258-stained nucleus, respectively. Bars, 10 μm.

mutant (PH4A) had mutations at aa 54–57 (ESRK→AAAA; Fig. 1a); these residues are responsible for the PH domain–PI(4,5)P(2) interaction (Ferguson *et al.*, 1995a). PH–Gag, PH4A–Gag and their cleaved products were detected in transfected 293T cell lysates with mouse monoclonal antibodies specific for the $p24^{CA}$ (capsid) domain (anti-$p24^{CA}$; NIH AIDS Research and Reference Reagent Program) and MA domain (anti-$p17^{MA}$; Advanced Biotechnologies) (Fig. 1b). PH–Gag cleavage was more efficient than that of Gag, suggesting efficient PM targeting of PH–Gag (Fig. 1b). The Gag protein levels in the pPH4A-*gag-pol*-transfected cell lysate were higher than those in p*gag-pol*- and pPH-*gag-pol*-transfected cell lysates when adjusted for the amount of protein loaded, indicating the low virus-like particle (VLP) production efficiency by PH4A–Gag (Fig. 1b).

The intracellular distribution of Gag, PH–Gag and PH4A–Gag was analysed by immunofluorescence microscopy of transfected 293T cells (Fig. 1c). Transfected cells were grown for 24 h, fixed (4 % formaldehyde), permeabilized (0.1 % Triton X-100 for 5–30 min) and incubated with mouse anti-$p24^{CA}$ and goat anti-mouse antibodies (GE Healthcare Bio-Sciences) conjugated to streptavidin–Alexa Fluor 555 (Invitrogen). Cells were stained with Hoechst 33258, mounted and analysed using confocal microscopy as described previously (Futahashi *et al.*, 2007). Gag was found to be distributed throughout the cytoplasm and at the cell periphery. In contrast, PH–Gag signals were mostly detected at the cell periphery and PH4A–Gag was distributed homogeneously in the cytoplasm. These data clearly showed that PH–Gag targeted the PM more efficiently than Gag, consistent with the Western blot analysis (Fig. 1b). These results also suggested that the efficient PM targeting of PH–Gag depends on the ability of the PH domain to bind PI(4,5)P(2). Similar observations were made in NP2 and COS7 cells.

VLP production was also examined. Tissue culture supernatants of p*gag-pol*-, pPH-*gag-pol*- or pPH4A-*gag-pol*-transfected 293T cells were passed through nitrocellulose filters (0.45 μm) and the virions were collected by centrifugation (541 000 ***g*** for 1 h). Viral antigens, except for PH4A, were detected with anti-$p24^{CA}$ and anti-$p17^{MA}$ antibodies (Fig. 1b). Gag and PH–Gag were further processed by the viral proteases in the virions compared with the cell lysates, as indicated by the increased signals for CA and MA relative to Gag. Interestingly, approximately one-fifth of the PH–Gag in the virion was cleaved close to the PH–MA junction. Presumably, the amino acid sequence at the C end of the PH domain ELQN/FLKE (aa 164–171, where the protease cleaves at the N–F junction) served as a viral protease recognition site as it matched the substrate consensus sequence and resembled the NC–*p1* junction, RQAN/FLGK (de Oliveira *et al.*, 2003; Swanstrom & Wills, 1997). Alternatively, the N terminus of LA–Gag (EFTL/AADS) might be targeted by the viral protease. Thus, the MA released from PH–MA, designated MA*, possibly has 10 aa attached to its N terminus.

The VLP production efficiency was quantified as the concentration of CA in transfected 293T cell culture supernatants relative to that in cell lysates using a p24 ELISA (Zeptometrics). When the CA concentrations of the virion fractions were normalized to those of the cell lysates, the pPH-*gag-pol* viral production efficiency was 3.2-fold higher than that of p*gag-pol* (3.2 ± 2.0-fold, $n=14$, $P<0.001$ by Wilcoxon's matched pairs rank test; representative experiments are shown in Table 1). In contrast, pPH4A-*gag-pol* produced viral particles less efficiently than p*gag-pol* (0.09 ± 0.07-fold, $n=6$, $P<0.05$ by Wilcoxon's matched pairs rank test; representative experiments are shown in Table 1). These data were consistent with the Western blot analysis (Fig. 1b).

To characterize the physical properties of PH–Gag–Pol VLPs, we measured the specific density of virions and examined the virion morphology. Firstly, the VLPs were subjected to 20–70 % (w/w) equilibrium sucrose gradient centrifugation (120 000 ***g*** for 16 h) and fractions were recovered from the bottom to the top. The peak fraction containing the viral CA antigen was determined by p24 ELISA. Gag–Pol and PH–Gag–Pol VLPs were detected in fractions with densities of 1.15 ± 0.01 ($n=5$) and 1.16 ± 0.01 ($n=4$) g ml^{-1}, respectively (not statistically significant; representative experiments are shown in Fig. 2a). Secondly, ultrathin sections of fixed 293T cells (2 % glutaraldehyde, 2 % osmium tetroxide) transfected with p*gag-pol* or pPH-*gag-pol* were imaged by transmission electron microscopy (JEM1200EX at 80 kV or JEM2000EX at 100 kV; JEOL). The PH–Gag–Pol and Gag–Pol VLP

Table 1. Efficiency of virus production from 293T cells transfected with p*gag-pol*, pPH-*gag-pol* or pPH4A-*gag-pol* expression vector

Experiment	Plasmid	$p24^{CA}$ (ng per well)*		Virus production efficiency (B/A)	Fold increase relative to p*gag-pol*
		Cell lysate (A)	Culture supernatant (B)		
1	p*gag-pol*	4 043	4 869	1.204	–
	pPH-*gag-pol*	1 989	8 363	4.206	3.49
	pPH4A-*gag-pol*	3 175	103	0.033	0.03
2	p*gag-pol*	3 521	3 887	1.104	–
	pPH-*gag-pol*	2 638	7 688	2.914	2.64
	pPH4A-*gag-pol*	5 913	1 125	0.190	0.17
3	p*gag-pol*	3 359	4 160	1.239	–
	pPH-*gag-pol*	1 454	5 172	3.558	2.87
	pPH4A-*gag-pol*	4 226	75	0.018	0.01
4	p*gag-pol*	9 666	8 996	0.931	–
	pPH-*gag-pol*	4 699	18 273	3.889	4.18
	pPH4A-*gag-pol*	5 527	534	0.097	0.10

*Cells grown in six-well plates were transfected using Lipofectamine 2000 according to the manufacturer's protocol (Invitrogen).

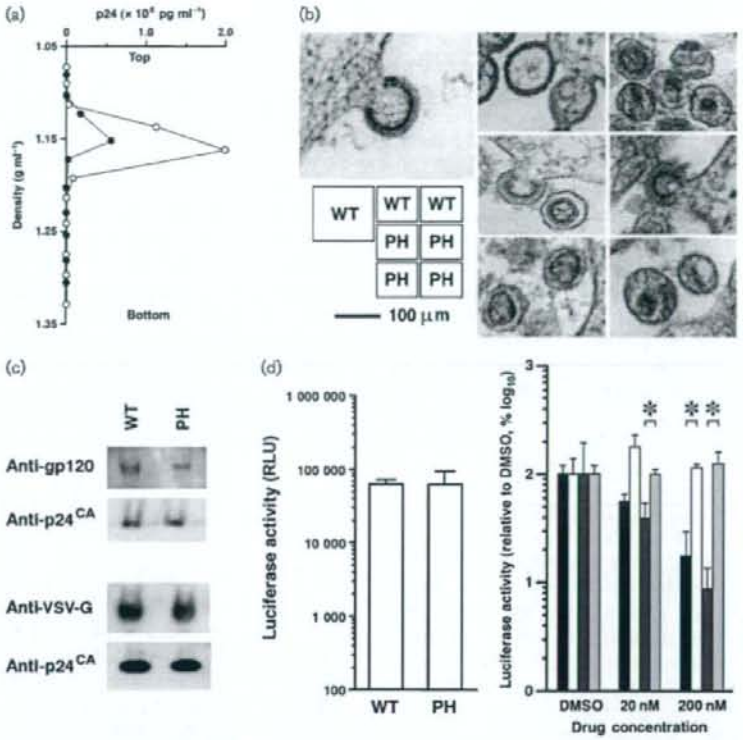

Fig. 2. Physical and biological properties of PH–Gag–Pol VLPs. (a) The virions produced by the p*gag-pol* (WT) and pPH-*gag-pol* (PH) expression vectors were analysed by equilibrium sucrose density-gradient centrifugation. The virion-containing fraction was determined by an ELISA detecting $p24^{CA}$. Representative data from four to five independent experiments are shown. In this experiment, WT (filled circles) and PH (open circles) VLPs migrated in the 1.152 or 1.162 g ml^{-1} density fractions, respectively. (b) Transmission electron microscopy images of 293T cells transfected with p*gag-pol* (WT) or pPH-*gag-pol* (PH). Representative images are shown. (c) Incorporation of HIV-1 Env (upper panel) and VSV-G (lower panel) into Gag–Pol (WT) and PH–Gag–Pol (PH) virions. The virion fractions were subjected to Western blot analysis detecting gp120, VSV-G and $p24^{CA}$. (d) The early phase of the HIV-1 life cycle is supported by PH–Gag–Pol. 293T cells were exposed to virus-containing culture supernatants with similar CA concentrations (270 and 220 ng ml^{-1} for Gag–Pol and PH–Gag–Pol, respectively), and luciferase activities were measured at 2–3 days post-infection as relative light units (RLU). The luciferase activities of PH–Gag–Pol (PH) and Gag–Pol (WT) virus-infected cells were almost identical (left graph). The luciferase transduction by WT (bars 1 and 2) and PH (bars 3 and 4) pseudovirions was performed in the presence of nevirapine (NVP, bars 1 and 3) or TAK-779 (bars 2 and 4). The luciferase signals decreased in the presence of NVP for both WT (bar 1) and PH (bar 3) but not in the presence of TAK-779 for both WT (bar 2) and PH (bar 4), respectively. Representative data from several independent experiments are shown. Asterisks indicate statistical significance ($P<0.01$, $n=3$, Student's *t*-test).

diameters were almost identical (Fig. 2b). Virion budding structures showed that the electron-dense layer, which represented multimerized Gag, of the PH–Gag–Pol VLP was slightly separated from the viral envelope compared with that of the Gag–Pol VLP (Fig. 2b). This indicated that the PH domain was positioned between the viral envelope and the electron-dense layer. In contrast, the morphologies of the mature PH–Gag–Pol and Gag–Pol virions were similar, suggesting that myristoylation is dispensable for mature virion morphology and that the PH–Gag–Pol virion may be infectious.

We examined HIV-1 Env incorporation into the PH–Gag–Pol virion. To do this, we used codon-optimized gp160 (p96ZM651gp160-opt; NIH AIDS Research and Reference Reagent Program). The PH–Gag–Pol virion incorporated HIV-1 Env less efficiently than the Gag–Pol virion as demonstrated by Western blot analysis detecting CA and

gp120 (anti-gp120 antibodies from Santa Cruz Biotechnology; Fig. 2c). This was presumably because PH interfered with the MA–Env interaction. Alternatively, PH may actively incorporate cellular proteins that block efficient Env incorporation into virions. We were unable to evaluate the entry efficacy of PH–Gag–Pol virions pseudotyped by HIV-1 Env because of the limit of detection. PH–Gag–Pol virion infectivity was re-examined by virions pseudotyped with vesicular stomatitis virus G glycoprotein (VSV-G). The incorporation efficiencies of VSV-G into Gag–Pol and PH–Gag–Pol virions were similar (anti-VSV-G antibody from Sigma; Fig. 2c). The lentiviral vector system was used to test this, as HIV-1 provirus gene modifications often fail to produce infectious virions, probably due to viral gene dysregulation. 293T cells were transfected with expression plasmids for Gag–Pol, VSV-G (Komano *et al.*, 2004), Rev and Vpu (a generous gift from Dr H. Göttlinger, University of Massachusetts Medical School, MA, USA), and with a packaging vector encoding a luciferase expression cassette. The HIV-1-based vector expressing firefly luciferase upon infection was recovered 2 days post-transfection. 293T cells were exposed to virus-containing culture supernatants with similar CA concentrations, and luciferase activities were measured at 2–3 days post-infection. When viral preparations with similar p24 concentrations were used, the luciferase activities of PH–Gag–Pol and Gag–Pol virus-infected cells were almost identical to each other (Fig. 2d, left graph). Luciferase expression was blocked by the non-nucleoside reverse transcriptase inhibitor nevirapine (NVP; Boehringer Ingelheim) but not by the CCR5 inhibitor (TAK-779; NIH AIDS Research and Reference Reagent Program), suggesting that gene transduction was mediated by viral infection (Fig. 2d, right panel). Similar results were obtained in several independent experiments. Thus, the PLCδ1 PH domain can functionally replace the HIV-1 Gag myristoylation signal to support both viral production and entry processes, and this myristoylation is dispensable for MA function in the early phase of the virus life cycle. This is the first report describing an infectious pseudovirion without myristoylated Gag. Given that PH–Gag can enhance virus production, HIV-1 with PH–Gag might have been expected to be selected in nature. This is not the case, presumably because the addition of PH to the HIV-1 genome would increase its genome size close to the upper limit that can be incorporated into the retroviral particle, leading to a decrease in genome uptake efficiency, which is clearly a growth disadvantage, despite the enhanced virus production with PH-Gag. More importantly, PH–Gag is unable to incorporate HIV-1 Env efficiently enough to support the production of fully infectious virions. Our data point to the selective advantage of myristoylated Gag in viral evolution.

The myristoylation-dependent Gag–PM association [maximal dissociation constant (K_d) of ~0.5–1.0 $\times 10^{-5}$ M] is presumably important for Gag multimerization at the PM (Provitera *et al.*, 2006). After the first contact of Gag with the PM, the membrane binding of Gag is assumed to be stabilized by the Gag–PI(4,5)P(2) interaction (Ono *et al.*, 2004; Saad *et al.*, 2006). The multimerization of Gag appears to induce a conformational change in MA to expose myristoyl groups to enhance the PM targeting of Gag. The higher-order Gag multimerization is probably facilitated by the increased local concentrations of Gag at the PM. Although Gag and PH–Gag are similar to the extent that PI(4,5)P(2) is involved in their PM association, Gag binds to one of the acyl chains of PI(4,5)P(2), as modelled previously (Saad *et al.*, 2006), whilst the PH domain binds the phosphorylated inositol group (Lemmon *et al.*, 1995). The K_d of binding between the PLCδ1 PH domain and PI(4,5)P(2) (~1–2 $\times 10^{-6}$ M; Lemmon *et al.*, 1995) suggests that the primary force driving PH–Gag to the PM is at least 2.5-fold stronger than that of myristoylation-mediated PM targeting of Gag. This might be one reason why PH–Gag–Pol was 3.2-fold more efficient at virion production than Gag–Pol. Our data suggest that the myristoyl group-dependent Gag–PM affinity is not a prerequisite for efficient Gag assembly at the PM or for viral production.

The MA has multiple functions throughout the virus life cycle (reviewed by Bukrinskaya, 2007; Fiorentini *et al.*, 2006; Hearps & Jans, 2007; Klein *et al.*, 2007). In the PH–Gag–Pol virion, approximately one-fifth of the PH–MA was unanchored from the PM as MA* (Fig. 1b), which might accompany the pre-integration complex to support nuclear targeting. Using PH–Gag–Pol might enable separation of myristoylation-dependent and -independent MA functions, particularly during the entry phase. PH–Gag–Pol might also be useful for producing high-titre lentiviral vectors or for studying Gag trafficking in cells that poorly support PM targeting of myristoylated Gag, such as rodent cells. Furthermore, functional assays comparing the virus production of Gag–Pol and PH–Gag–Pol might enable the identification of chemical inhibitors or cellular factors specifically targeting myristoylated Gag.

Acknowledgements

We thank Dr H. Göttlinger for critically reading the manuscript. This work was supported by the Japan Health Science Foundation, the Japanese Ministry of Health, Labor and Welfare (H18-AIDS-W-003) and the Japanese Ministry of Education, Culture, Sports, Science and Technology (18689014 and 18659136).

References

Berridge, M. J. (1993). Inositol trisphosphate and calcium signalling. *Nature* **361**, 315–325.

Bryant, M. & Ratner, L. (1990). Myristoylation-dependent replication and assembly of human immunodeficiency virus 1. *Proc Natl Acad Sci U S A* **87**, 523–527.

Bukrinskaya, A. (2007). HIV-1 matrix protein: a mysterious regulator of the viral life cycle. *Virus Res* **124**, 1–11.

de Oliveira, T., Engelbrecht, S., Janse van Rensburg, E., Gordon, M., Bishop, K., zur Megede, J., Barnett, S. W. & Cassol, S. (2003). Variability at human immunodeficiency virus type 1 subtype C protease cleavage sites: an indication of viral fitness? *J Virol* **77**, 9422–9430.

Ferguson, K. M., Lemmon, M. A., Schlessinger, J. & Sigler, P. B. (1995a). Structure of the high affinity complex of inositol trisphosphate with a phospholipase C pleckstrin homology domain. *Cell* **83**, 1037–1046.

Ferguson, K. M., Lemmon, M. A., Sigler, P. B. & Schlessinger, J. (1995b). Scratching the surface with the PH domain. *Nat Struct Biol* **2**, 715–718.

Fiorentini, S., Marini, E., Caracciolo, S. & Caruso, A. (2006). Functions of the HIV-1 matrix protein p17. *New Microbiol* **29**, 1–10.

Futahashi, Y., Komano, J., Urano, E., Aoki, T., Hamatake, M., Miyauchi, K., Yoshida, T., Koyanagi, Y., Matsuda, Z. & Yamamoto, N. (2007). Separate elements are required for ligand-dependent and -independent internalization of metastatic potentiator CXCR4. *Cancer Sci* **98**, 373–379.

Göttlinger, H. G., Sodroski, J. G. & Haseltine, W. A. (1989). Role of capsid precursor processing and myristoylation in morphogenesis and infectivity of human immunodeficiency virus type 1. *Proc Natl Acad Sci U S A* **86**, 5781–5785.

Harlan, J. E., Hajduk, P. J., Yoon, H. S. & Fesik, S. W. (1994). Pleckstrin homology domains bind to phosphatidylinositol-4,5-bisphosphate. *Nature* **371**, 168–170.

Harlan, J. E., Yoon, H. S., Hajduk, P. J. & Fesik, S. W. (1995). Structural characterization of the interaction between a pleckstrin homology domain and phosphatidylinositol 4,5-bisphosphate. *Biochemistry* **34**, 9859–9864.

Hearps, A. C. & Jans, D. A. (2007). Regulating the functions of the HIV-1 matrix protein. *AIDS Res Hum Retroviruses* **23**, 341–346.

Jouvenet, N., Neil, S. J., Bess, C., Johnson, M. C., Virgen, C. A., Simon, S. M. & Bieniasz, P. D. (2006). Plasma membrane is the site of productive HIV-1 particle assembly. *PLoS Biol* **4**, e435.

Kiernan, R. E., Ono, A., Englund, G. & Freed, E. O. (1998). Role of matrix in an early postentry step in the human immunodeficiency virus type 1 life cycle. *J Virol* **72**, 4116–4126.

Klein, K. C., Reed, J. C. & Lingappa, J. R. (2007). Intracellular destinies: degradation, targeting, assembly, and endocytosis of HIV Gag. *AIDS Rev* **9**, 150–161.

Komano, J., Miyauchi, K., Matsuda, Z. & Yamamoto, N. (2004). Inhibiting the Arp2/3 complex limits infection of both intracellular mature vaccinia virus and primate lentiviruses. *Mol Biol Cell* **15**, 5197–5207.

Lemmon, M. A., Ferguson, K. M., O'Brien, R., Sigler, P. B. & Schlessinger, J. (1995). Specific and high-affinity binding of inositol phosphates to an isolated pleckstrin homology domain. *Proc Natl Acad Sci U S A* **92**, 10472–10476.

Ono, A., Ablan, S. D., Lockett, S. J., Nagashima, K. & Freed, E. O. (2004). Phosphatidylinositol (4,5) bisphosphate regulates HIV-1 Gag targeting to the plasma membrane. *Proc Natl Acad Sci U S A* **101**, 14889–14894.

Pal, R., Reitz, M. S., Jr, Tschachler, E., Gallo, R. C., Sarngadharan, M. G. & Veronese, F. D. (1990). Myristoylation of gag proteins of HIV-1 plays an important role in virus assembly. *AIDS Res Hum Retroviruses* **6**, 721–730.

Provitera, P., El-Maghrabi, R. & Scarlata, S. (2006). The effect of HIV-1 Gag myristoylation on membrane binding. *Biophys Chem* **119**, 23–32.

Rhee, S. G. (2001). Regulation of phosphoinositide-specific phospholipase C. *Annu Rev Biochem* **70**, 281–312.

Saad, J. S., Miller, J., Tai, J., Kim, A., Ghanam, R. H. & Summers, M. F. (2006). Structural basis for targeting HIV-1 Gag proteins to the plasma membrane for virus assembly. *Proc Natl Acad Sci U S A* **103**, 11364–11369.

Scholz, I., Still, A., Dhenub, T. C., Coday, K., Webb, M. & Barklis, E. (2008). Analysis of human immunodeficiency virus matrix domain replacements. *Virology* **371**, 322–335.

Stauffer, T. P., Ahn, S. & Meyer, T. (1998). Receptor-induced transient reduction in plasma membrane PtdIns(4,5)P2 concentration monitored in living cells. *Curr Biol* **8**, 343–346.

Swanstrom, R. & Wills, J. W. (1997). Synthesis, assembly, and processing of viral proteins. In *Retroviruses*, pp. 263–334. Edited by J. M. Coffin, S. H. Hughes & H. Varmus. Cold Spring Harbor: Cold Spring Harbor Laboratory Press.

Tall, E. G., Spector, I., Pentyala, S. N., Bitter, I. & Rebecchi, M. J. (2000). Dynamics of phosphatidylinositol 4,5-bisphosphate in actin-rich structures. *Curr Biol* **10**, 743–746.

Wagner, R., Graf, M., Bieler, K., Wolf, H., Grunwald, T., Foley, P. & Uberla, K. (2000). Rev-independent expression of synthetic *gag-pol* genes of human immunodeficiency virus type 1 and simian immunodeficiency virus: implications for the safety of lentiviral vectors. *Hum Gene Ther* **11**, 2403–2413.

Wang, C. T., Zhang, Y., McDermott, J. & Barklis, E. (1993). Conditional infectivity of a human immunodeficiency virus matrix domain deletion mutant. *J Virol* **67**, 7067–7076.

Yagisawa, H., Hirata, M., Kanematsu, T., Watanabe, Y., Ozaki, S., Sakuma, K., Tanaka, H., Yabuta, N., Kamata, H. & other authors (1994). Expression and characterization of an inositol 1,4,5-trisphosphate binding domain of phosphatidylinositol-specific phospholipase C-δ1. *J Biol Chem* **269**, 20179–20188.

Zhou, W., Parent, L. J., Wills, J. W. & Resh, M. D. (1994). Identification of a membrane-binding domain within the amino-terminal region of human immunodeficiency virus type 1 Gag protein which interacts with acidic phospholipids. *J Virol* **68**, 2556–2569.

Ligand-independent higher-order multimerization of CXCR4, a G-protein-coupled chemokine receptor involved in targeted metastasis

Makiko Hamatake, Toru Aoki, Yuko Futahashi, Emiko Urano, Naoki Yamamoto and Jun Komano[1]

AIDS Research Center, National Institute of Infectious Diseases, 1-23-1 Toyama, Shinjuku-ku, Tokyo 162-8640, Japan

(Received July 19, 2008/Revised September 1, 2008/Accepted September 7, 2008/Online publication October 31, 2008)

CXCR4, a G-protein-coupled receptor of CXCL12/stromal cell-derived factor-1α, mediates a wide range of physiological and pathological processes, including the targeted metastasis of cancer cells. CXCR4 has been shown to homo-oligomerize in several experimental systems. However, it remains unclear with which domains CXCR4 interacts homotypically, and whether it dimerizes or forms a higher-order complex. To address these issues, we used bioluminescent resonance energy transfer and bimolecular fluorescence complementation analyses to measure the homotypic interactions of CXCR4 in living cells. Both assays indicated that CXCR4 interacts homotypically, which is consistent with previous studies. By studying CXCR4 mutants lacking various domains, we found that multiple transmembrane domains probably serve as potential molecular interaction surfaces for oligomerization. The relative contribution of the amino- or carboxy-termini to oligomerization was small. To differentiate between a dimer and a multimer consisting of more than two molecules, bioluminescent resonance energy transfer–bimolecular fluorescence complementation analysis was conducted. It revealed that CXCR4 engages in higher-order oligomerization in a ligand-independent fashion. This is the first report providing direct experimental evidence for the higher-order multimerization of CXCR4 *in vivo*. We hypothesize that CXCR4 distributes to the cell surface as a multimer, in order to effectively sense, with increased avidity, the chemotaxis-inducing ligand in the microenvironment. Studying the structure and function of the oligomeric state of CXCR4 may lead us to develop novel CXCR4 inhibitors that disassemble the molecular cluster of CXCR4. (*Cancer Sci* 2009; 100: 95–102)

CXCR4, a widely expressed chemokine receptor of CXCL12/stromal cell-derived factor (SDF)-1α, plays a role in various physiological and pathological processes, including neuronal network development, normal and malignant cell migration, inflammatory reactions, the genetic immunodeficiency syndrome WHIM (warts, hypogammaglobulinemia, infections, myelok-athexis), and human immunodeficiency virus (HIV)-1 infection.[1–7] The CXCR4–CXCL12/SDF-1α axis is reportedly involved in the tumor progression of breast cancer and more recently of pancreatic, esophageal, prostate, thyroid, colorectal, and cutaneous cancers.[8–16] Thus, it has been emphasized that the CXCR4–CXCL12/SDF-1α axis may be an important therapeutic target.[9,14] Understanding the regulatory mechanisms of CXCR4 functions may provide clues to develop therapeutic approaches for such disorders.

CXCR4 was shown to homo-oligomerize by several experimental systems. It is possible that CXCR4 dimerizes because CXCL12/SDF-1α forms a dimer, as indicated by structural analyses.[17–19] However, past biophysical analyses did not critically distinguish between dimers and complexes consisting of more than two molecules (the higher-order oligomer).[20–23] The oligomerization of G protein-coupled recepters (GPCR) has been suggested to play a role in ligand-initiated signaling and protein trafficking, including egress from the endoplasmic reticulum (ER) or internalization from the plasma membrane.[24,25] It appears that steady-state oligomerization and its functional significance vary among GPCR species. Although it is known to homo-oligomerize, the functional significance of CXCR4 oligomerization is not well understood. Isolating a monomeric CXCR4 mutant is one straightforward approach toward understanding the functional significance of CXCR4 oligomerization. However, to achieve this goal, it remains to be clarified whether CXCR4 dimerizes or forms a higher-order complex, and with which domains CXCR4 interacts homotypically. Additionally, it is unknown how selfoligomerization is regulated.

It is hypothesized that, for most mammalian GPCR, the stable homotypic interaction is mediated by a specific transmembrane domain (TMD), as has been shown for some GPCR (e.g. dopamine D2 receptor, adenosine A2A receptor, and CCR5).[26–28] If CXCR4 forms a dimer, it is postulated that two CXCR4 molecules should face each other in a symmetrical configuration on the two-dimensional plane of the biological lipid bilayer (Fig. 1a). If this were the case, it might not be difficult to isolate a CXCR4 mutant lacking TMD that fails to interact homotypically. However, if CXCR4 has multiple interaction facets that multimerize, monomeric CXCR4 mutants may be hard to isolate by the simple mutagenesis approach (Fig. 1b).

To measure TMD–TMD interactions, we used bioluminescent resonance energy transfer (BRET) and bimolecular fluorescence complementation (BiFC) assay systems[29,30] that allowed us to measure the specific homotypic and heterotypic interactions of membrane proteins bearing multiple TMD in living mammalian cells. By analyzing a series of CXCR4 mutants using these two techniques, we found that CXCR4 forms a steady-state molecular cluster consisting of more than two molecules, and that this cluster formation is not mediated by the amino- or carboxy-termini. This is the first report to provide direct experimental evidence that CXCR4 engages in higher-order multimerization *in vivo*.

Materials and Methods

Plasmids. The plasmids pCXCR4, pCXCR4-GFP, and d-31-GFP were described previously.[31] The following oligonucleotides encoding the myristoylation and palmitoylation signal region of *lyn* were annealed and cloned into the *Eco*RI and *Bam*HI sites of pDsRed2-N1 (Clontech, Palo Alto, CA, USA), generating pLynDsRed2: sense, 5′-AATTGCCACCATGGGATGTATTAAA-TCAAAAAGGAAAGAC-3′ and antisense, 5′-GATCGTCTTTC-CTTTTTGATTTAATACATCCCATGGTGGC-3′. The *Sna*BI–*Age*I fragment from pLynDsRed2 was cloned into the corresponding sites of pEGFP-C2, generating pMEM-GFP. The 5′ *Nhe*I end of

[1]To whom correspondence should be addressed. E-mail: ajkomano@nih.go.jp

doi: 10.1111/j.1349-7006.2008.00997.x

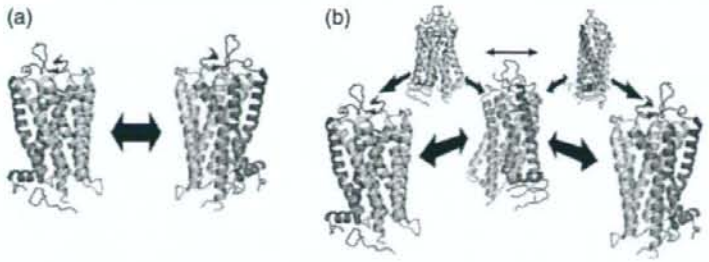

Fig. 1. Hypothetical models of CXCR4 oligomerization. (a) If CXCR4 forms a dimer, it is postulated that two CXCR4 molecules would physically associate with each other via the specific side of the molecule where the same key transmembrane helices are situated. (b) If CXCR4 forms a higher-order multimerization consisting of more than two molecules, CXCR4 should have multiple interaction domains. These models were drawn based on the bovine rhodopsin structure.

the *Nhe*I–*Xba*I fragment from phRL/CMV (Promega, Madison, WI, USA) encoding renilla luciferase (hRL) was ligated to the annealed oligonucleotides of 5′-AGATCTGGTTACCCAATT-3′ and 5′-CTAGAATTGGGTAACCAGATCT-3′. The 3′ end of the same fragment was ligated to the annealed oligonucleotides of 5′-CTAGGATCTGAATTCAGATCT-3′ and 5′-AGATCTGAATTCAGATC-3′. The *Bgl*II fragment was cloned into the *Bgl*II and *Bam*HI sites of pEGFP-C3 (Clontech), generating pEGFP-hRluc. The plasmid pCXCR4 FL-Rluc was constructed by cloning the *Mfe*I fragment from pEGFP-hRLuc BRET into the *Eco*RI and *Mfe*I sites of pCXCR4.

5TMDΔ23 was amplified by a two-step polymerase chain reaction (PCR). In the first step, PCR products were amplified by two separate PCR using the following two primer pairs targeting full-length CXCR4: R4 forward 5′-ACCGGTGCCACCATGGAGGGGATCAGTATATACACTTCAG-3′ and 5TMDΔ23 reverse 5′-GGGATCCAGACGCCAACATAGACCACCTTGTCCGTCATGCTTCTCAGTTTCTTC-3′; 5TMDΔ23 forward 5′-GAAGAAACTGAGAAGCATGACGGACAAGGTGGTCTATGTTGGCGTCTGGATCCC3′ and R4 reverse 5′-AGATCTCGCTGGAGTGAAAACTTGAAGACTCAGACTC-3′. The second PCR step was carried out by mixing the two first-step PCR products and amplifying the DNA with the R4 forward and R4 reverse primers. The final PCR product was cloned into pCRblunt2-TOPO (Invitrogen, Tokyo, Japan) and sequenced. The *Age*I–*Bgl*II fragment, encoding 5TMDΔ23, was cloned into the *Age*I and *Bgl*II sites of pCXCR4-GFP and pCXCR4 FL-Rluc, generating pCXCR4 5TMDΔ23-GFP and pCXCR4 5TMDΔ23-Rluc, respectively. The 5TMDΔ12, 5TMDΔ34, 5TMDΔ34, and 5TMDΔ34 derivatives were generated similarly by using the following primers: 5TMDΔ12 forward 5′-CCGTGAAGAAAATGCTAATTTCAATAAAGCAGTCCATGTCATCTACACAGTTAACCTC-3′, 5TMDΔ12 reverse 5′-GAGGTTAACTGTGTAGATGACATGGACTGCTTTATTGAAATTAGCATTTTCTTCACGG-3′, 5TMDΔ34 forward 5′-CTGGTACTTTGGGAACTTCCTATGCAAGCACATCATGGTTGGCCTTATCCTGCC-3′, 5TMDΔ34 reverse 5′-GGCAGGATAAGGCCAACCATGATGTGCTTGCATAGGAAGTTCCCAAAGTACCAG-3′, 5TMDΔ45 forward 5′-GGCCAAGGAAGCTGTTGGCTGAAAAGACCACAATCATCCCCATCCTGGCTTTC-3′, 5TMDΔ45 reverse 5′-GAAAGCCAGGATGGGGATGATTGTGGTCTTTTCAGCCAACAGCTTCCTTGGCC-3′, 5TMDΔ67 forward 5′-GTCCTGCTATTGCATTATCATCTCCAAATTTAAAACCTCTGCCCAGCACGCACTC-3′, and 5TMDΔ67 reverse 5′-GAGTGCGTGCTGGGCAGAGGTTTTAAATTTGGAGATGATAATGCAATAGCAGGAC-3′. The 3TMD125 and 3TMΔ367 derivatives were constructed by conducting a two-step PCR using the 5TMDΔ34 forward and reverse primers with 5TMDΔ67 and 5TMDΔ12 as templates, respectively.

The plasmid pCXCR4 d-31-Rluc was constructed by replacing the *Age*I–*Bgl*II region of pCXCR4-Rluc with the *Age*I–*Bgl*II fragment from pCXCR4 d-31-GFP. The plasmid pCXCR2/444 has been described previously,[32] and the *Sna*BI–*Bam*HI fragment from pCXCR2/444 replaced the *Sna*BI–*Bam*HI region of pCXCR4-GFP, generating pCXCR2/4-GFP.

The following oligonucleotides encoding the amino-terminal myristoylation signal region of *lyn* were annealed and cloned into the *Eco*RI and *Bam*HI sites of pEGFP-N1, generating pLyn-GFP: sense 5′-AATTGCCACCATGGGAGCTATTAAATCAAAAAGGAAAGAC-3′ and antisense 5′-GATCGTCTTTCCTTTTTGATTTAATAGCTCCCATGGTGGC-3′. The plasmid pLyn-R4cyt-GFP was constructed by cloning the *Xma*I fragment of a PCR product amplified by 5′-CCCGGGGAAATTTAAAACCTCTGCCC-3′ and 5′-CCCGGGCTGGAGTGAAAACTTGAAG-3′ using pCXCR4-GFP as a template into the *Age*I site of pLynGFP.

CXCR3 was amplified from RNA extracted from human peripheral blood mononuclear cells by reverse transcription-PCR using the following primers; 5′-ACCGGTGCCACCATGGTCCTTGAGGTGAGTGACC-3′ and 5′-GAGCTCGAGATCTCCAAGCCCGAGTAGGAGGCCTCTG-3′. CXCR2 was amplified from a plasmid carrying the CXCR2 open reading frame by PCR using the following primers: 5′-GCCACCGGTGCCACCATGGAAGATTTTAACATGG-3′ and 5′-CCTCGAGCCGAGAGTAGTGGAAGTGTGCCCTG-3′. The *Age*I and *Xho*I fragments encoding CXCR2 and CXCR3 were further cloned into the corresponding sites of pCXCR4-GFP and pCXCR4-Rluc, generating pCXCR2-GFP and pCXCR3-GFP, and pCXCR2-Rluc and pCXCR3-Rluc, respectively. The *Bgl*II and *Mfe*I fragments from phmKGN-MN or phmKGC-MN encoding the fragments of Kusabira-Green (mKG) (MBL, Nagoya, Japan) were cloned into the corresponding sites of pCXCR4-GFP, generating pCXCR4-mKGN and pCXCR4-mKGC, respectively. The *Xho*I and *Mfe*I fragments from phmKGN-MN and phmKGC-MN were cloned into the corresponding sites of pCXCR3-GFP, generating pCXCR3-mKGN and pCXCR3-mKGC, respectively.

Cells, transfection, and imaging. All mammalian cells were maintained in RPMI-1640 (Sigma, St Louis, MO, USA) supplemented with 10% fetal bovine serum (Japan Bioserum, Tokyo, Japan), 50 U/mL penicillin, and 50 μg/mL streptomycin (Invitrogen). Cells were incubated at 37°C in a humidified 5% CO_2 atmosphere. Cells were transfected with Lipofectamine 2000 according to the manufacturer's protocol (Invitrogen). Cells were imaged by confocal microscopy as described previously.[31]

Western blotting. Western blotting was carried out according to techniques described previously,[33] except that the cells were processed with an M-PEK kit according to the manufacturer's protocol (Calbiochem, Darmstadt, Germany). The antibodies used are as follows: a monoclonal anti-FLAG epitope (M2; Sigma), a monoclonal anti-GFP antibody (MAB3580; Chemicon International, Temecula, CA, USA), a monoclonal anti-Rluc antibody (MAB4400 and MAB4100; Chemicon), and EnVision⁺ (Dako, Glostrup, Denmark).

Bioluminescent resonance energy transfer assay. 293T cells grown in a six-well plate were transfected with 0.05–0.2 μg renilla luciferase (Rluc) derivatives along with 1–2 μg of the green fluorescent protein (GFP) fusion derivative expression plasmids. Varying amounts of plasmids were used to control the expression levels of GFP and Rluc in the transfected cells. At 2 days after transfection, cells were collected and incubated with the Rluc

doi: 10.1111/j.1349-7006.2008.00997.x

substrate according to the manufacturer's protocol except that a 10-fold higher substrate concentration was used (ViviRen Live Cell Substrate; Promega). The fluorescent and bioluminescent signals were measured using a Tristar LB941 (Berthold Technologies, Bad Wildbad, Germany).

Bimolecular fluorescence complementation–BRET analysis. 293T cells grown in a six-well plate were cotransfected with equal amounts of mKGN and mKGC (31.3–1000 ng) along with the Rluc constructs for CXCR4 or CXCR3 (20–200 ng). The BiFC–BRET signal was measured according to the BRET protocol described above.

Results

Establishment of the BRET system to measure molecular interactions. We used the BRET system to measure the homotypic interaction of CXCR4 (Fig. 2a). Wild-type CXCR4 was fused to either enhanced GFP or Rluc at its carboxy-terminus. Previous studies indicate that CXCR4 function is not affected by this modification.[31,34,35] The expression of CXCR4-GFP and CXCR4-Rluc were verified by western blot analysis using anti-GFP and anti-Rluc antibodies, confocal microscopy, flow cytometric analysis, and immunofluorescence analysis (Fig. 2b,c; data not shown). Two days after cotransfection of the CXCR4-GFP and CXCR4-Rluc expression plasmids into 293T cells, we incubated cells with the cell-permeable Rluc substrate and measured the GFP fluorescence, the net Rluc bioluminescence, and the bioluminescence through 485-nm (window A) and 525-nm (window B) emission filters. The BRET efficiency was defined as the ratio of the window B and window A values. As a positive control, we used the GFP-Rluc fusion protein that should theoretically have the best BRET efficiency. As a negative control, we cotransfected cells with GFP and Rluc expression plasmids. The average and standard deviation of BRET values between CXCR4-GFP and CXCR4-Rluc was 1.103 ± 0.160 where the negative and positive controls yielded 0.667 ± 0.084 and 1.444 ± 0.109, respectively (Table 1). The CXCR4 BRET signals were significantly higher than those of the negative control ($P < 0.001$, Student's t-test). Even the transfected cells expressed ~5–10-fold higher levels of GFP and Rluc in the negative control than did cells transfected with pCXCR4-GFP and pCXCR4-Rluc (data not shown).

The specificity of CXCR4-to-CXCR4 BRET signals has been a concern, as GPCR may yield non-specific BRET signals. To address this, we carried out out two experiments: a competition assay and a BRET measurement between heterologous chemokine receptors. The tripartite transfection of cells with pCXCR4, pCXCR4-GFP, and pCXCR4-Rluc resulted in a reduction of the BRET signal to 0.812 ± 0.058, which is significantly less than the BRET levels between CXCR4-GFP and CXCR4-Rluc ($P < 0.001$; Table 1). We defined the BRET level between GFP and Rluc as 0%, and that between CXCR4-GFP and CXCR4-Rluc as 100%, and used these normalized values to determine the relative BRET efficiency of each interaction. The presence of pCXCR4 reduced the BRET level to 39.3%, most likely because of the interference of the BRET-inducing homotypic CXCR4 interaction by untagged CXCR4.

We next examined various chemokine receptors, including CXCR2, CXCR3, and CCR5. Membrane targeting of these chemokine receptors was visualized by confocal microscopy (Fig. 2c), fluorescence activated cell sorting (FACS) analysis, and an immunofluorescence assay (data not shown). CXCR4-Rluc yielded low BRET signals (Table 1) with either membrane-targeted GFP (MEM-GFP; 0.609 ± 0.041, 0.0%), CXCR2-GFP (0.791 ± 0.093, 16.4%), CXCR3-GFP (0.718 ± 0.083, 7.9%), or CCR5-GFP (0.737 ± 0.057, 2.0%). Similar observations were made for the BRET signals between CXCR4-GFP and CXCR2-Rluc (0.762 ± 0.093, 10.2%) or CXCR4-GFP and CXCR3-Rluc (0.709 ± 0.064, 6.2%). These data suggest that the BRET assay detects specific homotypic interactions of CXCR4. Consistent with this, we also detected the homotypic interaction of CXCR4, but not the CXCR4–CXCR3 interaction, using a BiFC assay. Although some have suggested a possible homotypic interaction of CXCR2,[36] this interaction did not yield BRET signals as high as those with CXCR4 in our experimental set up, suggesting that the homotypic interaction of CXCR4 is substantially stronger than that of other chemokine receptors. Throughout this study, the BRET and BiFC assays were carried out under conditions where the transfected cells expressed similar levels of GFP and Rluc.

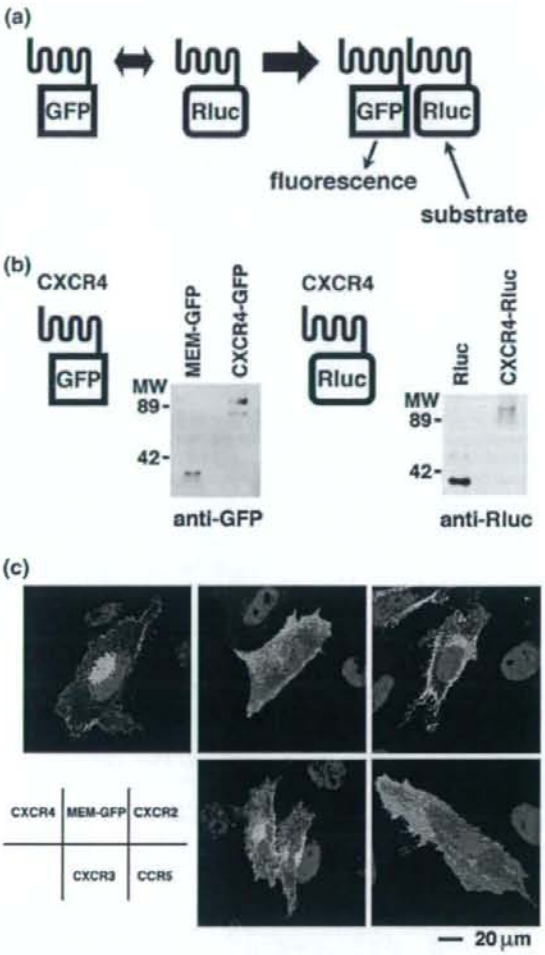

Fig. 2. Bioluminescent resonance energy transfer (BRET) assay to detect CXCR4 homotypic interactions. (a) Schematic drawing of the BRET assay system. The bioluminescence energy from renilla luciferase (Rluc) was absorbed by green fluorescent protein (GFP) to emit green fluorescence. (b) Detection of CXCR4-GFP and CXCR4-Rluc expression by western blotting. The GFP and Rluc fusion proteins were detected by anti-GFP and anti-Rluc antibodies, respectively. The molecular-weight markers are indicated on the left (kDa). (c) The intracellular distribution of CXCR4-GFP (top, left), MEM-GFP (top, middle), CXCR2-GFP (top, right), CXCR3-GFP (bottom, left), and CCR5-GFP (bottom, right). Green and blue represent GFP and the Hoechst 33258-stained nucleus, respectively. Magnification, ×630; scale bar = 20 μm.

Table 1. Detection of specific CXCR4–CXCR4 interactions by bioluminescent resonance energy transfer (BRET) analysis

Sample	BRET level†	*n*	*P*-value	Relative BRET efficiency (%)‡
GFP-Rluc fusion protein	1.444 ± 0.109	20	Not tested	–
GFP + Rluc	0.667 ± 0.084	21	Not tested	0.0
CXCR4-GFP + CXCR4-Rluc	1.103 ± 0.160	21	Not tested	100.0
CXCR4 FL-GFP + CXCR4-Rluc + CXCR4	0.812 ± 0.058	4	<0.001	39.3
MEM-GFP + CXCR4-Rluc	0.609 ± 0.041	5	<0.001	0.0
CXCR2-GFP + CXCR2-Rluc	0.784 ± 0.096	6	<0.001	13.0
CXCR2-GFP + CXCR4-Rluc	0.791 ± 0.093	6	<0.001	16.4
CXCR4-GFP + CXCR2-Rluc	0.762 ± 0.093	6	<0.001	10.2
CXCR3-GFP + CXCR3-Rluc	0.688 ± 0.060	6	<0.001	5.1
CXCR3-GFP + CXCR4-Rluc	0.718 ± 0.083	6	<0.001	7.9
CXCR4-GFP + CXCR3-Rluc	0.709 ± 0.064	6	<0.001	6.2
CCR5-GFP + CXCR4-Rluc	0.737 ± 0.057	4	<0.001	2.0

†The average and standard deviation of the indicated number of independent experiments.
The statistical significance of each BRET efficiency was tested compared to the BRET signal of CXCR4-GFP and CXCR4-Rluc (Student's *t*-test).
‡The relative BRET efficiency for each experiment was averaged.

Fig. 3. Analyses of the transmembrane domain (TMD) mutants of CXCR4. (a) Schematic drawing of CXCR4 5TMD mutants. The deleted portions are indicated in gray. The intracellular distribution of 5TMDΔ23-GFP is shown as a representative. Green and blue represent green fluorescent protein (GFP) and the Hoechst 33258-stained nucleus, respectively. Magnification, ×630; scale bar = 20 μm. (b) Detection of CXCR4 5TMD-GFP and 5TMD-Rluc derivative expression by western blotting. The GFP and renilla luciferase (Rluc) fusion proteins were detected by anti-GFP and anti-Rluc antibodies, respectively.

Measuring the interaction between CXCR4 mutants carrying five TMD. The dimerization model predicts the presence of a CXCR4 mutant defective in the homotypic interaction. To map the domains required for the homotypic interaction of CXCR4, we constructed a series of deletion mutants. It has been reported that removing two transmembrane domains alters neither the membrane topology nor the functions of CXCR4.[37] Thus, we constructed four mutants, 5TMDΔ23, 5TMDΔ34, 5TMDΔ45, and 5TMDΔ67, and fused these mutants to either GFP or Rluc at their carboxy-termini (Fig. 3a). The expression of the 5TMD derivatives was verified by western blot analysis (Fig. 3b). The intracellular distribution of 5TMDD23-GFP was mostly cytoplasmic with reticular patterns (Fig. 3a), and all 5TMD-GFP derivatives showed similar distributions. The BRET efficiencies of homotypic pairs of GFP- and Rluc-tagged 5TMD derivatives were comparable to those of full-length CXCR4 (109.4–133.8%; Table 2). Furthermore, any pair of 5TMD derivatives yielded BRET signal levels similar to those of CXCR4-GFP and CXCR4-Rluc (81.8–161.0%; Table 2). None of the combinations yielded significantly lower BRET signals than the full-length CXCR4

Table 2. Detection of homotypic and heterotypic interactions between five transmembrane domain (5TMD) derivatives by bioluminescent resonance energy transfer (BRET) analysis

Sample	BRET level†	*n*	Relative BRET efficiency (%)‡
Homotypic interaction			
5TMDΔ23-GFP + 5TMDΔ23-Rluc	1.130 ± 0.091	7	115.7
5TMDΔ34-GFP + 5TMDΔ34-Rluc	1.209 ± 0.097	10	133.8
5TMDΔ45-GFP + 5TMDΔ45-Rluc	1.117 ± 0.095	7	112.6
5TMDΔ67-GFP + 5TMDΔ67-Rluc	1.102 ± 0.087	7	109.4
Heterotypic interaction			
5TMDΔ23-GFP + 5TMDΔ34-Rluc	1.060 ± 0.085	4	104.5
5TMDΔ23-GFP + 5TMDΔ45-Rluc	1.078 ± 0.093	3	112.7
5TMDΔ23-GFP + 5TMDΔ67-Rluc	0.978 ± 0.078	7	81.8
5TMDΔ34-GFP + 5TMDΔ23-Rluc	1.299 ± 0.077	3	161.0
5TMDΔ34-GFP + 5TMDΔ45-Rluc	1.177 ± 0.091	3	129.7
5TMDΔ34-GFP + 5TMDΔ67-Rluc	1.225 ± 0.105	3	141.3
5TMDΔ45-GFP + 5TMDΔ23-Rluc	1.080 ± 0.108	8	105.5
5TMDΔ45-GFP + 5TMDΔ34-Rluc	1.131 ± 0.050	4	121.5
5TMDΔ45-GFP + 5TMDΔ67-Rluc	1.161 ± 0.069	3	133.7
5TMDΔ67-GFP + 5TMDΔ23-Rluc	1.072 ± 0.096	3	111.3
5TMDΔ67-GFP + 5TMDΔ34-Rluc	1.065 ± 0.088	4	104.8
5TMDΔ67-GFP + 5TMDΔ45-Rluc	1.042 ± 0.094	3	104.0

†The average and standard deviation for the indicated number of independent experiments.
The statistical significance of each BRET efficiency was tested compared to the BRET signal of CXCR4-GFP and CXCR4-Rluc (Student's *t*-test). All were found to be not significant.
‡The relative BRET efficiency was calculated for each experiment and the average is shown.

combination. These data suggest that either: (1) the interactions between the 5TMD mutants are mediated by the amino- or caroboxy-termini, or the first TMD, as these are shared by all of the 5TMD derivatives; or (2) there are more than two interaction facets in CXCR4, which would make interactions between any two of the 5TMD mutants possible.

Contribution of the amino- or carboxy-terminal domains and the first TMD to CXCR4 oligomerization. To differentiate between the above possibilities, we assessed the potential contribution of the extracellular amino-terminal and the cytoplasmic carboxy-terminal domains to CXCR4 oligomerization. We used a CXCR2–CXCR4 chimeric construct, CXCR2/4, which has the amino-terminus of CXCR2 fused to the first TMD of CXCR4 (Fig. 4a). CXCR2/4 is

doi: 10.1111/j.1349-7006.2008.00997.x

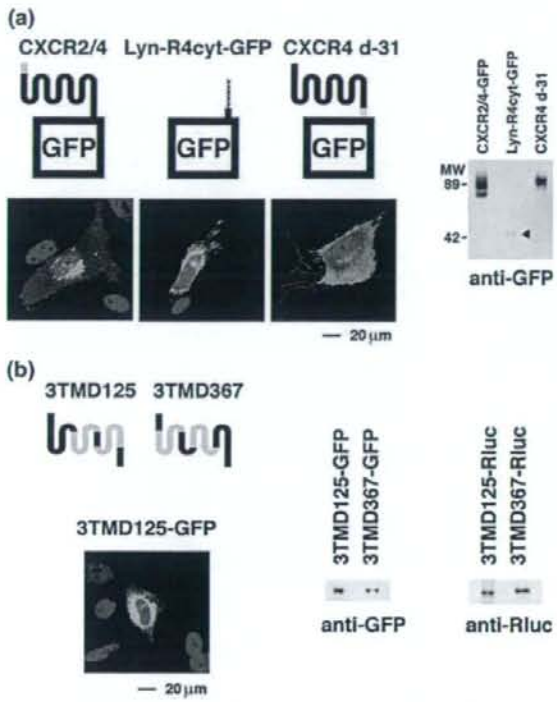

Fig. 4. Assessment of the relative contribution of the CXCR4 amino- and carboxy-termini and the first transmembrane domain (TMD) to oligomerization. (a) Schematic drawing of CXCR2/4-GFP (left), Lyn-R4cyt-GFP (middle), and CXCR4 d-31 (right). The substituted or deleted regions are indicated in gray. The intracellular distribution of each construct is shown. Protein expression was detected by anti-GFP antibody (Lyn-R4cyt-GFP, arrowhead). The molecular-weight markers are indicated on the left (kDa). (b) Schematic drawing of the CXCR4 3TMD mutants. The deleted portions are indicated in gray. The green fluorescent protein (GFP) and renilla luciferase (Rluc) fusion proteins were detected by anti-GFP and anti-Rluc antibodies, respectively. The intracellular distribution of 3TMD125-GFP is shown as a representative. Green and blue represent GFP and the Hoechst 33258-stained nucleus, respectively. Magnification, ×630; scale bar = 20 μm.

known to support HIV-1 entry and CXCL12/SDF-1α signaling.[32,38] The intracellular distribution of CXCR2/4-GFP was similar to that of CXCR4-GFP (Fig. 4a). The interaction between CXCR2/4-GFP and CXCR4-Rluc yielded the BRET signal (to 1.163 ± 0.106, 71.9%; Table 3), which was not significantly different from the BRET levels of CXCR4-GFP and CXCR4-Rluc. Given that only a low BRET signal was detected between CXCR2 and CXCR4 in these experimental conditions (Table 1), this suggests that the amino-terminus of CXCR4 contributes little to its oligomerization.

We next investigated the cytoplasmic tail of CXCR4. The amino-terminus of the CXCR4 cytoplasmic domain was fused to the myristoylation signal motif of *lyn* and the carboxy-terminus was fused to GFP (Lyn-R4cyt-GFP), thereby allowing the cytoplasmic domain of CXCR4 to target the plasma membrane in a native membrane topology (Fig. 4a). As expected, Lyn-R4cyt-GFP was distributed to the plasma membrane, as indicated by microscopic observation (Fig. 4a). However, the BRET signal between Lyn-R4cyt-GFP and CXCR4-Rluc was not detected (0.758 ± 0.057, 4.8%; Table 3). We also tested the cytoplasmic tail-deleted mutant of CXCR4 (Fig. 4a). By removing the cytoplasmic tail of CXCR4, it lost its ability to signal upon ligand binding, but retained its HIV-1 coreceptor function.[32,38] When the carboxy-terminal 31 amino acids were removed (CXCR4 d-31), the BRET signal between CXCR4 d-31-GFP and CXCR4-Rluc was 1.111 ± 0.083 (98.9%), which is comparable to levels seen with the homotypic interaction of full-length CXCR4 (Table 3). These data suggest that CXCR4 oligomerization is not mediated by the amino- or carboxy-termini, but mainly by TMD.

Finally, to examine the possible role of the first TMD in CXCR4 oligomerization, we constructed two CXCR4 mutants carrying three TMD. The 3TMD125 and 3TMD367 mutants have the first, second, and fifth, or the third, sixth, and seventh TMD, respectively, and both were fused to either GFP or Rluc (Fig. 4b). The intracellular distribution of 3TMD-GFP derivatives was similar to those of the 5TMD-GFP derivatives (3TMD125-GFP in Fig. 4b). If the first TMD plays a major role in CXCR4 oligomerization, the BRET signal between 3TMD125-GFP and 3TMD125-Rluc derivatives should be higher than that between 3TMD367 derivatives. Furthermore, if the dimerization model is correct, then the 3TMD125 and 3TMD367 derivatives should not yield detectable BRET signals as they do not share any TMD. We found that the BRET signal representing the homotypic interaction of 3TMD125 was 0.806 ± 0.071 (30.8%), which was significantly higher than the negative control ($P < 0.01$) but lower than that of 3TMD367 (0.917 ± 0.061, 64.0%, $P < 0.05$; Table 3). These data are not consistent with the dimerization model, suggesting that the contribution of the first TMD to the homotypic CXCR4 interaction is relatively small.

The heterotypic interactions between 3TMD and 5TMD derivatives were also tested by the BRET assay. The 3TMD and 5TMD combinations yielded BRET signals that were significantly higher than the 3TMD combinations (72.0 ± 17.5 vs 51.5 ± 16.2%, $P < 0.05$; Table 3) but lower than the 5TMD pairs (72.0 ± 17.5 vs 117.7 ± 18.7%, $P < 0.001$; Tables 2,3). The observations that any pair of 5TMD mutants yielded BRET signals that were comparable to those of CXCR4, and that the BRET signals increased in parallel to the number of TMD, do not agree with a dimerization model and instead favor a multimerization model.

Detecting a higher-order oligomerization of CXCR4 with the BiFC–BRET system. We used the BiFC–BRET technique to directly demonstrate the presence of a trimolecular complex *in vivo*.[39] BiFC uses a 'split' GFP that fluoresces only when the amino- and carboxy-terminal fragments are forced into close proximity. When the BiFC assay is built into the BRET system, the BRET signal should be detected only when three molecules assemble, that is, proteins carrying Rluc and the amino- or carboxy-terminal fragments of GFP (Fig. 5a). We used monomeric mKG as a GFP.[40] The expression vectors of CXCR4 fused to the mKG amino- and carboxy-terminal fragments (mKGN and mKGC, respectively) were constructed. For controls, we constructed CXCR3-mKGN and CXCR3-mKGC. The expression of fusion proteins was verified by immunofluorescence analysis (Fig. 5b) and flow cytometric analysis (data not shown).

We detected BiFC signals when CXCR4-mKGN and CXCR4-mKGC were coexpressed, but not when either was expressed alone or when a pair of CXCR4 and CXCR3 derivatives were coexpressed, confirming the specificity of the BiFC analysis and the BRET results (Fig. 5b; Table 4). We considered the combination of CXCR4-mKGN, CXCR4-mKGC, and Rluc as a negative control. When increasing amounts of pCXCR4-mKGN and pCXCR4-mKGC were transfected into 293T cells along with pCXCR4-hRluc, we detected increasing BiFC–BRET signals that were significantly higher than the negative control (Table 4). Such high BiFC–BRET signals were not observed by tripartite transfection of plasmids expressing CXCR4-mKGN, CXCR4-mKGC, and CXCR3-Rluc, CXCR3-mKGN, CXCR3-mKGC, and CXCR4-Rluc, or CXCR4-mKGN, CXCR4-mKGC, and Rluc