http://dx.doi.org/10.5012/bkcs.2013.34.7.2006

Identification of Potent Leukocyte Common Antigen-Related Phosphatase Inhibitors *via* Structure-Based Virtual Screening

Hwangseo Park,[*] Pham Ngoc Chien,[†] Ha-Jung Chun,[‡] and Seong Eon Ryu[†,*]

Department of Bioscience and Biotechnology, Sejong University, Seoul 143-747, Korea. [*]E-mail: hspark@sejong.ac.kr
[†]Department of Bioengineering, Hanyang University, Seoul 133-791, Korea. [*]E-mail: ryuse@hanyang.ac.kr
[‡]Department of Radiation Oncology, College of Medicine, Hanyang University, Seoul 133-791, Korea
Received March 14, 2013, Accepted April 3, 2013

Leukocyte common antigen-related phosphatase (LAR) has been considered a promising target for the development of therapeutics for neurological diseases. Here, we report the first example for a successful application of the structure-based virtual screening to identify the novel small-molecule LAR inhibitors. Five of these inhibitors revealed micromolar inhibitory activities with the associated IC_{50} values ranging from 2 to 6 μM. Because the newly identified inhibitors were also screened for having desirable physicochemical properties as a drug candidate, they may serve as a starting point of the structure-activity relationship study to optimize the medical efficacy. Structural features relevant to the stabilization of the new inhibitors in the active site of LAR are discussed in detail.

Key Words : Virtual screening, Docking, LAR, Inhibitor, Drug design

Introduction

Proteoglycans play a critical role in the regeneration of injured neurons.[1] In particular, the level of chondroitin sulfate proteoglycan (CSPG) increases substantially in the region of neural injury.[2,3] Although the formation of CSPG layer has an effect of protecting the normal regions from the harmful signals that come from the injured neurons, the layer can also serve as a barrier for regenerating the neurons after the disappearance of injury conditions. This hypothesis has become confirmed in a series of experimental evidence that the impairment of neural regeneration by CSPG can be overcome in the presence of bacterial chondroitinase ABC that cleaves the glycoaminoglycan units in CSPGs.[4-6]

The leukocyte common antigen-related phosphatase (LAR) is a functional receptor for CSPGs.[7] The extracellular immunoglobulin-like domains of LAR bind to proteoglycan units of CSPGs and thereby activate the intracellular protein tyrosine phosphatase (PTP) domains.[8] Related with the function of LAR, it was shown that the knockout of LAR gene in mice significantly promoted the neural regeneration.[7] Therefore, the inhibition of LAR activity by small-molecule inhibitors can be a promising therapeutic strategy for the neurological diseases caused by the impairment in neural regeneration. In addition, the possibility of the involvement of LAR in the development of cancer and diabetes was also proposed,[9,10] further motivating the discovery of LAR inhibitors.

LAR is a transmembrane receptor protein tyrosine phosphatase comprising 1907 amino acids. Its intracellular region includes two PTP domains, D1 and D2, both of which contain the conserved CX_5R catalytic loop. Although the X-ray crystal structure of LAR was reported,[8,11] the lack of structural information about the interactions between LAR and the small-molecule ligands has made it difficult to design the potent LAR inhibitors. Indeed, the discovery of LAR inhibitors has lagged behind the neurobiological and biochemical studies. Only a few classes of small-molecule LAR inhibitors have been reported so far including illudalic acid and its derivatives that possess the 5-formyl group and the hemi-acetal lactone as the key pharmacophores.[12,13]

In the present study, we aim to identify the new potent LAR inhibitors by means of a structure-based drug design protocol involving the virtual screening with docking simulations and *in vitro* enzyme assay. The characteristic feature that discriminates our virtual screening approach from the others lies in the implementation of an accurate solvation model in calculating the binding free energy between LAR and the putative ligands, which would have an effect of increasing the hit rate in enzyme assay.[14] We find in this study that the docking simulations with the improved binding free energy function can be a useful computational tool for elucidating the activities of the identified inhibitors, as well as for enriching the chemical library with molecules that are likely to have desired biological activities.

Computational and Experimental Methods

Atomic coordinates of D1 domain in the X-ray crystal structure of human LAR[11] (PDB entry: 1LAR) were selected as the receptor model in the virtual screening with docking simulations. After removing the solvent molecules, hydrogen atoms were added to each protein atom. A special attention was paid to assign the protonation states of the ionizable Asp, Glu, His, and Lys residues in the original X-ray crystal structure. The side chains of Asp and Glu residues were assumed to be neutral if one of their carboxylate oxygens pointed toward a hydrogen-bond accepting group such as the backbone aminocarbonyl oxygen at a distance within 3.5 Å,

a generally accepted distance limit for a hydrogen bond of moderate strength.[15] Similarly, the lysine side chains were assumed to be protonated unless the NZ atom was in proximity of a hydrogen-bond donating group. The same procedure was also applied to determine the protonation states of ND and NE atoms in His residues.

The docking library for LAR comprising about 240,000 compounds was constructed from the latest version of the chemical database distributed by Interbioscreen (http://www.ibscreen.com) containing approximately 477,000 synthetic and natural compounds. Prior to the virtual screening with docking simulations, they were filtrated on the basis of Lipinski's "Rule of Five" to adopt only the compounds with the physicochemical properties of potential drug candidates[16] and without reactive functional group(s). All of the compounds included in the docking library were then processed with the CORINA program to generate their 3D atomic coordinates, followed by the assignment of Gasteiger-Marsilli atomic charges.[17] We used the AutoDock program[18] in virtual screening because the outperformance of its scoring function over those of the others had been shown in several target proteins.[19] AMBER force field parameters were assigned for calculating the van der Waals interactions and the internal energy of a ligand. Docking simulations with AutoDock were then carried out in the active site of LAR to score and rank the compounds in the docking library according to their calculated binding affinities.

In the actual docking simulation between LAR and the putative ligands, we used the empirical AutoDock scoring function improved by the implementation of a new solvation model for a compound. The modified scoring function has the following form:

$$\Delta G_{bind}^{aq} = W_{vdW}\sum_{i=1}\sum_{j=1}\left(\frac{A_{ij}}{r_{ij}^{12}}-\frac{B_{ij}}{r_{ij}^{6}}\right) + W_{hbond}\sum_{i=1}\sum_{j=1} E(t)\left(\frac{C_{ij}}{r_{ij}^{12}}-\frac{D_{ij}}{r_{ij}^{10}}\right) + W_{elec}\sum_{i=1}\sum_{j=1}\frac{q_i q_j}{\varepsilon(r_{ij})r_{ij}} + W_{tor}N_{tor} + W_{sol}\sum_{i=1} S_i\left(Occ_i^{\max}-\sum_{j\neq i} V_j e^{-\frac{r_{ij}^2}{2\sigma^2}}\right), \quad (1)$$

where W_{vdW}, W_{hbond}, W_{elec}, W_{tor}, and W_{sol} are the weighting factors of van der Waals, hydrogen bond, electrostatic interactions, torsional term, and desolvation energy of inhibitors, respectively. r_{ij} represents the interatomic distance, and A_{ij}, B_{ij}, C_{ij}, and D_{ij} are related to the depths of the potential energy well and the equilibrium separations between the two atoms. The hydrogen bond term has an additional weighting factor, $E(t)$, representing the angle-dependent directionality. Cubic equation approach was applied to obtain the dielectric constant required in computing the interatomic electrostatic interactions between LAR and a ligand molecule.[20] In the entropic term, N_{tor} is the number of rotatable bonds in the ligand. In the desolvation term, S_i and V_i are the solvation parameter and the fragmental volume of atom i,[21] respectively, while $Occ_i^{\max}$ stands for the maximum atomic occupancy. In the calculation of the solvation free energy term in Eq. (1), we used the atomic parameters developed by Kang *et al*.[22] The accuracy in virtual screening seems to be enhanced due to the inclusion of the solvation free energy term in the scoring function because the underestimation of ligand solvation often leads to the overestimation of the binding affinity of a ligand with many polar atoms.[14] Indeed, the superiority of this modified scoring function to the previous one was demonstrated in recent studies for virtual screening of kinase and phosphatase inhibitors.[23,24]

The gene for intracellular catalytic domains of LAR (LAR-cat) corresponding to residues 1308-1876 was cloned into BamHI and EcoRI restriction sites of the expression vector pET28a. Cells containing the vector were induced with 0.2 mM IPTG and further grown at 18 ºC for 16 hours. Cell pellets were resuspended in the lysis buffer containing 50 mM Tris-HCl (pH 7.5), 0.5 M NaCl, 1% PMSF, and 5% glycerol. The resuspension was then lysed by sonication on ice. The His-tagged LAR-cat was purified by a Ni-NTA agarose column and dialyzed against 20 mM Tris-HCl, pH 8.0, 5 mM EDTA, 50 mM NaCl, and 5% glycerol. Total 145 compounds selected from the virtual screening were evaluated for *in vitro* inhibitory activity against the purified LAR-cat. These enzyme assays were performed by monitoring the extent of hydrolysis of 6,8-difluoro-4-methyl-umbelliferyl phosphate (DiFMUP) with spectrofluorometric analysis. The purified LAR-cat (0.39 nM), DiFMUP (10 μM), and a candidate inhibitor were incubated in the reaction mixture containing 20 mM Tris-HCl pH 8.0, 0.01% Trition X-100, and 5 mM DTT for 20 minutes. The resulting fluorescence was measured at 460 nm by using the Perkin Elmer 2030 multiplate reader. IC_{50} values of the potent inhibitors were determined from direct regression curve analysis in triplicate.

Results and Discussion

Of the 240,000 compounds screened with docking simulations, 150 top-scored compounds were selected as virtual hits. 145 of them were available from the compound supplier and were tested for inhibitory activity against LAR by *in vitro* enzyme assay. As a result, we identified the five compounds that inhibited the catalytic activity of LAR by more than 50% at the concentration of 10 μM, which were selected to determine the IC_{50} values. The chemical structures and IC_{50} values of the newly identified inhibitors are shown in Figure 1. Also, the dose-response behaviors of the inhibitors measured to obtain their IC_{50} values are shown in Figure 2. We note that compounds **1-5** have the IC_{50} values ranging from 2.7 to 5.9 μM. In particular, the micromolar inhibitory activities of **3-5** are surprising because they are too small to be fully accommodated in the active site of LAR. These low molecular weight inhibitors are therefore expected to be capable of establishing the strong multiple hydrogen bonds with the amino acid residues in the active site. Benzene-1,2-diol in **1** and **3-5** and 5-imino-5,6-dihydro-[1,3,4]thiadiazolo[3,2-*a*]pyrimidin-7-one in **2** seem to serve as an effective surrogate for the substrate phosphotyrosine group.

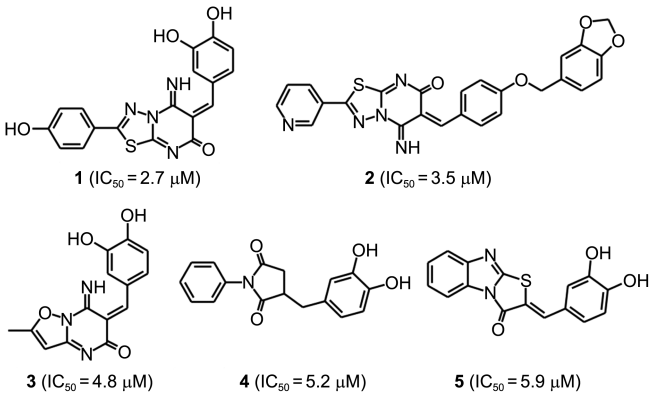

Figure 1. Chemical structures and IC_{50} values of the newly identified LAR inhibitors.

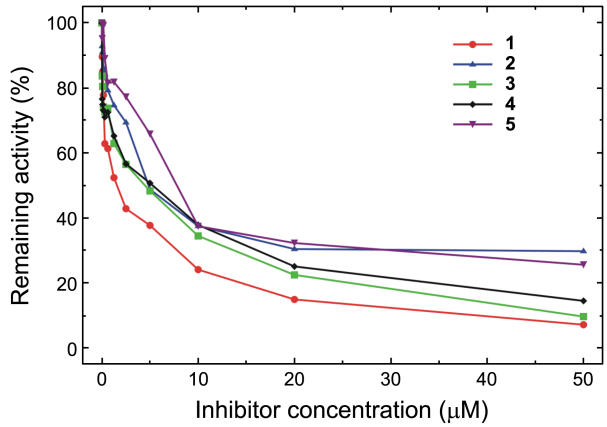

Figure 2. Dose-response curves of **1**, **2**, **3**, **4**, and **5** for the inhibition of LAR.

Besides the micromolar inhibitory activities, **1-5** were also screened for having desirable physicochemical properties as a drug candidate. Thus, they deserve consideration for further development by the structure-activity relationship studies to develop therapeutics for the neurological diseases caused by LAR.

To obtain structural insight into the inhibitory mechanisms of the identified LAR inhibitors, their binding modes in the active site were investigated in a comparative fashion. Figure 3 shows the lowest-energy conformations of **1-5** in the active site gorge of LAR calculated with the modified scoring function. The results of docking simulations are self-consistent because the functional groups of similar chemical character are placed in similar ways with comparable interactions with the protein groups. As revealed by the superposition of the docked structures, for example, the hydrophilic moieties that serve as a surrogate for the substrate phosphotyrosine group are bound in the PTP loop including the catalytic cysteine residue (Cys1522) while the hydrophobic groups point toward the WPD loop located above the active site. These common features in the calculated binding modes indicate that a potent LAR inhibitor should include an effective surrogate for the substrate phosphotyrosine group and simultaneously the hydrophobic groups for binding to the WPD loop. In order to examine the possibility of

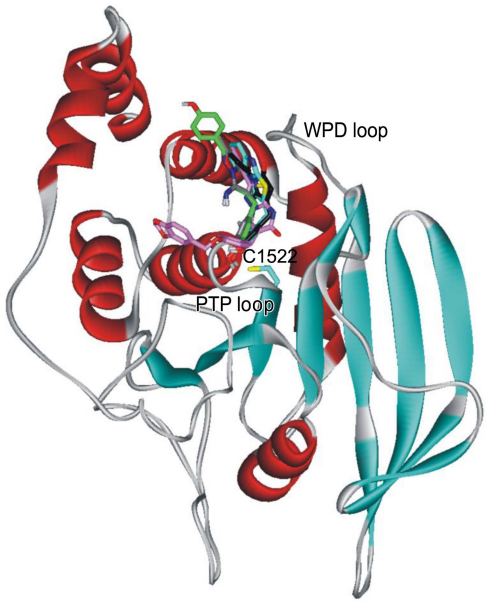

Figure 3. Comparative view of the binding modes of **1-5** in the active site of LAR. The positions of PTP loop, WPD loop, and Cys1522 are also indicated.

the allosteric inhibition of LAR by the identified inhibitors, docking simulations were carried out with the grid maps for the receptor model so as to include the entire D1 domain of LAR. However, the binding configuration in which an inhibitor resides outside the active site was not observed for any of the new inhibitors. These results support the possibility that the inhibitors would impair the catalytic activity of LAR through the specific binding in the active site.

The calculated binding mode of **1** in the active site of LAR

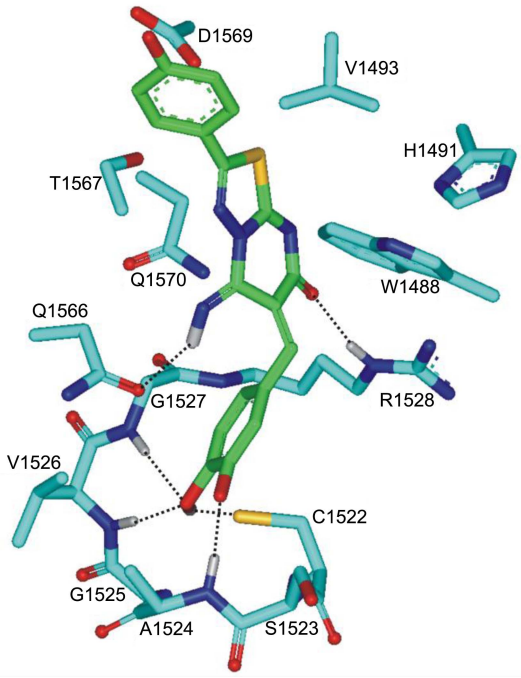

Figure 4. Calculated binding mode of **1** in the active site of LAR. Each dotted line indicates a hydrogen bond.

is shown in Figure 4. We note that one of the two phenolic oxygens receives three hydrogen bonds from the backbone amidic nitrogens of Ala1524, Val1526, and Gly1527, and the other donates a hydrogen bond to the thiolate ion of the catalytic cysteine residue (Cys1522) at the bottom of active site. Apparently, these four hydrogen bonds seem to play a critical role in anchoring the inhibitor in the active site. It is also noted that the benzene-1,2-diol group of **1** resides in the vicinity of the side-chain thiolate ion of Cys1522 with the associated interatomic distances ranging from 2 to 4 Å. Judging from the proximity to Cys1522 and the formation of multiple hydrogen bonds in the active site, the benzene-1,2-diol moiety should be an effective surrogate for the phosphotyrosine group in the substrates of LAR. The aminocarbonyl oxygen and the iminic nitrogen of **1** also appear to form a stable hydrogen bond with the side-chain guanidinium ion of Arg1528 and the side-chain amide moiety of Gln1566, respectively, which should also be a significant binding force in the LAR-**1** complex. The inhibitor **1** can be further stabilized in the active site of LAR by the hydrophobic interactions of its nonpolor groups with the side chains of Trp1488, His1491, Val1493, Ala1524, and Val1526. Thus, the overall structural features derived from docking simulations indicate that the micromolar inhibitory activity of **1** should stem from the multiple hydrogen bonds and hydrophobic interactions established simultaneously in the active site.

Figure 5 shows the lowest-energy binding mode of **2** in the active site of LAR. In this calculated structure of LAR-**2** complex, the roles of hydrogen bond donor with respect to the aminocarbonyl oxygen and the iminic nitrogen of **2** are played by the side-chain ammonium ion of Lys1433 and the backbone amidic group of Arg1528, respectively. Similar to the benzene-1,2-diol moiety in the LAR-**1** complex, 6-imino-5,6-dihydro-1*H*-pyrimidin-4-one ring of **2** stays close to the side-chain thiolate ion of Cys1522 at the interatomic distances within 5 Å. This indicates that it can also serve as an effective surrogate for the substrate phosphotyrosine group that is necessary for the effective inhibition of LAR. An additional hydrogen bond is found between the terminal [1,3]dioxolane ring of **2** and the side-chain amidic moiety of Gln1566, which should also play a significant role in stabilizing the inhibitor in the active site. Hydrophobic interactions in LAR-**2** complex seem to be established in a stronger form than those in LAR-**1** complex: three more hydrophobic residues (Tyr1355, Val1358, and Tyr1563) appear to be involved in the stabilization of hydrophobic groups of **2** in addition to Trp1488, His1491, Val1493, Ala1524, and Val1526 that are also present at the interface of LAR-**1** complex. Because the number of hydrogen bonds decreases from six in LAR-**1** to three in LAR-**2** complex, such a strengthening of hydrophobic interactions can be invoked to explain the similar inhibitory activity of **2** and **1** (Figure 1).

Binding modes of **3-5** are similar to that of **1** in that the benzene-1,2-diol group receives multiple hydrogen bonds from the backbone amidic groups on the PTP loop and donates a hydrogen bond to the side-chain thiolate ion of Cys1522 at the bottom of active site. This confirms the usefulness of the benzene-1,2-diol moiety as a surrogate for the substrate phosphotyrosine group. Hydrophobic interactions with nonpolar residues on the PTP and WPD loops are also found in the calculated structures of LAR-**3**, LAR-**4**, and LAR-**5** complexes. Although the inhibitory activities **3-5** are a little lower than **1** and **2** (Figure 1), they are expected to serve as a good inhibitor scaffold from which much more potent inhibitors can be derivatized because of the low molecular weights (280-310) as compared to 380.4 for **1** and 483.5 for **2**.

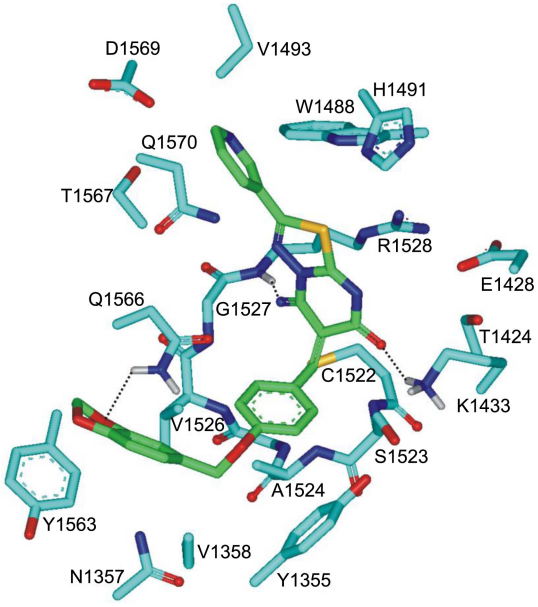

Figure 5. Calculated binding mode of **2** in the active site of LAR. Each dotted line indicates a hydrogen bond.

Conclusions

In summary, we have identified five novel inhibitors of LAR by applying a computer-aided drug design protocol involving the structure-based virtual screening with docking simulations under consideration of the effects of ligand solvation in the scoring function. These inhibitors revealed a high potency with the IC_{50} values ranging from 2.7 to 5.9 mM and were also screened for having desirable physicochemical properties as a drug candidate. Therefore, each of the newly identified inhibitors deserves consideration for further development by SAR studies to optimize the medical efficacy. Detailed binding mode analyses with docking simulations showed that the inhibitors could be stabilized in active site by the simultaneous establishment of multiple hydrogen bonds and van der Waals contacts.

Acknowledgments. This work was supported by Basic Science Research Program through National Research Foundation of Korea (NRF) funded by Ministry of Education, Science and Technology (2012-0008440), and a Hanyang University internal grant.

References

1. Silver, J.; Miller, J. H. *Nat. Rev. Neurosci.* **2004**, *5*, 146.
2. Rhodes, K. E.; Fawcett, J. W. *J. Anat.* **2004**, *204*, 33.
3. Sherman, L. S.; Back, S. A. *Trends Neurosci.* **2008**, *31*, 44.
4. Cafferty, W. B.; Bradbury, E. J.; Lidierth, M.; Jones, M.; Duffy, P. J.; Pezet, S.; McMahon, S. B. *J. Neurosci.* **2008**, *28*, 11998.
5. Houle, J. D.; Tom, V. J.; Mayes, D.; Wagoner, G.; Phillips, N.; Silver, J. *J. Neurosci.* **2006**, *26*, 7405.
6. Tester, N. J.; Howland, D. R. *Exp. Neurol.* **2008**, *209*, 483.
7. Fisher, D.; Xing, B.; Dill, J.; Li, H.; Hoang, H. H.; Zhao, Z.; Yang, X. L.; Bachoo, R.; Cannon, S.; Longo, F. M.; Sheng, M.; Silver, J.; Li, S. *J. Neurosci.* **2011**, *31*, 14051.
8. Biersmith, B. H.; Hammel, M.; Geisbrecht, E. R.; Bouyain, S. *J. Mol. Biol.* **2011**, *408*, 616.
9. Dunah, A. W.; Hueske, E.; Wyszynski, M.; Hoogenraad, C. C.; Jaworski, J.; Pak, D. T.; Simonetta, A.; Liu, G.; Sheng, M. *Nat. Neurosci.* **2005**, *8*, 458.
10. Yang, T.; Massa, S. M.; Longo, F. M. *J. Neurobiol.* **2006**, *66*, 1420.
11. Nam, H.-J.; Poy, F.; Krueger, N. X.; Saito, H.; Frederick, C. A. *Cell* **1999**, *97*, 449.
12. Ling, Q.; Huang, Y.; Zhou, Y.; Cai, Z.; Xiong, B.; Zhang, Y.; Ma, L.; Wang, X.; Li, X.; Li, J.; Shen, J. *Bioorg. Med. Chem.* **2008**, *16*, 7399.
13. Yang, X. N.; Li, J. Y.; Zhou, Y. Y.; Shen, Q.; Chen, J. W.; Li, J. *Biochim. Biophys. Acta* **2005**, *1726*, 34.
14. Shoichet, B. K.; Leach, A. R.; Kuntz, I. D. *Proteins* **1999**, *34*, 4.
15. Jeffrey, G. A. *An Introduction to Hydrogen Bonding*; Oxford University Press: Oxford, 1997.
16. Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. *Adv. Drug Delivery Rev.* **1997**, *23*, 3.
17. Gasteiger, J.; Marsili, M. *Tetrahedron* **1980**, *36*, 3219.
18. Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. *J. Comput. Chem.* **1998**, *19*, 1639.
19. Park, H.; Lee, J.; Lee, S. *Proteins* **2006**, *65*, 549.
20. Park, H.; Jeon, J. H. *Phys. Rev. E* **2007**, *75*, 021916.
21. Stouten, P. F. W.; Frömmel, C.; Nakamura, H.; Sander, C. *Mol. Simul.* **1993**, *10*, 97.
22. Kang, H.; Choi, H.; Park, H. *J. Chem. Inf. Model.* **2007**, *47*, 509.
23. Park, H.; Jeon, J. Y.; Kim, S. Y.; Jeong, D. G.; Ryu, S. E. *J. Comput. Aided Mol. Des.* **2011**, *25*, 469.
24. Park, H.; Chi, O.; Kim, J.; Hong, S. *J. Chem. Inf. Model.* **2011**, *51*, 2986.