Journal of Computer-Aided Molecular Design, **15:** 553–559, 2001.
KLUWER/ESCOM

Discovery of a novel serine protease inhibitor utilizing a structure-based and experimental selection of fragments technique

Shingo Makino*, Takashi Kayahara, Kazumi Tashiro, Mitsuo Takahashi, Takashi Tsuji and Masataka Shoji
Pharmaceutical Research Laboratories, Ajinomoto Co., Inc. 1-1, Suzuki-cho, Kawasaki-ku, Kawasaki-shi, 210-8681 Japan

Received 28 July 2000; accepted 30 March 2001

Key words: database, docking simulation, fragments, ligands, thrombin

Summary

We report a set of strategies to develop novel ligands (Structure Based and Experimental Selection of Fragments: SbE-SF). First, a docking simulation utilizing DOCK3.5 is performed in order to screen the fragment database, which was generated with the in-house program FRAGMENT++ specifically for docking simulation purposes. Although the affinity of these small molecules (fragments) is expected to be low, the affinity of fragments selected by computation is assayed by experiment to determine which ones can be potent inhibitors. After determining such key fragments, additional fragments are attached to the key ones in order to increase the binding affinity,taking into account the binding modes predicted by computation. This method has been applied to a thrombin inhibitor study, resulting in the discovery of a novel inhibitor exhibiting $pIC_{50} = 7.9$.

Introduction

There have been numerous reported programs for docking molecules into binding pockets of proteins [1–4]. Based on the success of some of these programs in re-discovering known ligands by screening a large chemical database [3, 4], it was expected that the same technique would allow us to identify novel ligands. However, there are problems. One of the issues is whether or not a chemical database such as Available Chemical Directory (ACD) [5] contains novel ligands with high binding affinity to a target protein. In the case of re-discovering ligands [3, 4], these ligands had been registered in the database after the affinity had been intentionally optimized by medicinal chemists. However, considering the number and diversity of molecules in commercially available database [6], it is most likely that novel ligands with high affinity would not exist in these databases, especially in the case of proteases where the ligands tend to be large. Therefore, novel molecules need to be designed and synthesized in order to achieve high affinity for a specific target protein.

Programs aimed at the design of novel ligands have been reported; these generate virtual ligands complementary to a three-dimensional structure of a binding pocket [7–11]. However, synthesis of these generated molecules is often difficult. Thus, it is important to consider synthetic accessibility in the design or generation of ligands.

The evaluation of ligand affinity to a protein is also an issue, as calculations of molecule-protein interaction energies are rough estimations of changes in free energy resulting from a ligand binding to a protein. In the case of database screening documented in previous reports [3,4], most of the molecules in ACD are not complementary to binding pockets, so it was not difficult to distinguish known ligands from other molecules. Nevertheless, if molecules are designed to be complementary to binding pockets, it is extremely difficult to distinguish molecules with high affinity from the others [11]. Therefore, the inclusion of the experi-

*To whom correspondence should be addressed.
E-mail: shingo_makino@ajinomoto.com

mental evaluation of affinity is preferred in regards to the process of designing ligands. The NMR technique to identify small molecules (fragments) that interact with specific sites of a target protein was reported [12, 13]. Their NMR technique is attractive, because such key fragments can be utilized as an anchor and may be developed into ligands with high affinity. However, this method requires acquisition of compounds, and it does not provide precise information about the binding modes of fragments. Contrary to their NMR technique, docking simulation provides possible binding modes, and does not require acquisition of compounds or labeling of proteins. Thus, docking simulation also has some advantages over experimental methods. We believe that it is important to balance computational evaluations of molecular structures with experimental evaluations for the successful discovery of novel ligands. Considering these facts, we have developed practical methods that are described in the following sections.

Methods

Computational resources

The interactive modeling and visualization of molecules was performed using the program SYBYL [14]. Programs were written in SYBYL Programming Language (SPL) [15] or in C++ with the Standard Template Library (GNU gcc-2.7.2) [16]. All calculations were performed on Silicon Graphics Indigo2 workstations with 200 MHz R4400 processors and 128 MB RAM.

A target protein selection and strategy for designing inhibitors

The development of inhibitors with oral activity for thrombin has been interesting in terms of pharmaceuticals, because the serine proteases such as thrombin and factor Xa (FXa) play key roles in the cascade-like activation of the coagulation system [17]. In addition, X-ray elucidated structures of thrombin were available from the Protein Data Bank (PDB) [18]. Therefore, we selected thrombin as a target protein for testing our strategy. There have been numerous reports of thrombin inhibitors such as argatoroban (**1**) and NAPAP (**2**) as shown in Figure 1. However, they showed poor oral activity which was likely due to the components guanidine (pKa = 13) or benzamidine (pKa = 11) both of which have high levels of basicity [19]. Therefore, replacements of those components with some of lower basicity are expected to improve oral activity [20].

The X-ray structure of thrombin with NAPAP (PDB: 1DWD) shows a benzamidine component occupying the S1 pocket, which is the deepest binding pocket of thrombin (Figure 2). Thus, we should be able to identify alternative fragments for benzamidine by docking fragments into the S1 pocket. Novel ligands can be designed based on accessible chemistry and the presumed binding modes of fragments as described in the following sections.

Generation of the fragment database

Since fragment databases with appropriate molecular structures for docking simulations were not commercially available, we decided to develop a program named FRAGMENT++ to generate a fragment database for the use of docking simulations [21]. Because the identification of rigid fragments was preferred over that of flexible fragments for the purposes of facilitating additions of fragments during the optimization of affinity, only rigid fragments were generated in the following algorithm. First, the flexible bonds of the molecules in the chemical database are identified and each molecule is decomposed into rigid fragments (Figure 3). Then, fragments with ring systems are selected and the remnants of flexible bonds are trimmed.

Since different molecules in a chemical database can possess the same fragment components, a significant number of the same fragments can be generated by such decomposition of molecules. Thus, the generated fragments are tagged with 'unique SLN' [22] to identify the same fragments and all of the duplicate fragments are then removed by SNL tagging. Then, those molecular structures are appropriately modified for docking simulation. As shown in Figure 4a–c, amidine, nitro and carboxylic acid components are modified both in terms of atom type [23] and geometry. In the case of amines including pyridines, it is improper to fix states of protonation, as the local pH environment of binding pockets may vary from buffer conditions. Thus, both protonated and unprotonated states of fragments are included in the fragment database. (Figure 4d, e).

We applied FRAGMENT++ to ACD to generate a fragment database with 7050 molecules. Then, charges were calculated utilizing the Gasteiger–Marsili method [24]. The whole process was performed automatically and completed within 4 hours.

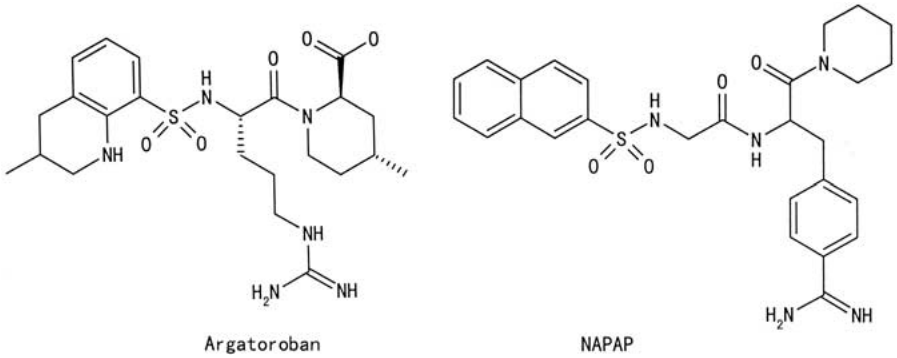

Figure 1. Thrombin inhibitors such as argatoroban (**1**) and NAPAP (**2**) as shown showed poor oral activity which was likely due to the components guanidine (pKa = 13) or benzamidine (pKa = 11) both of which have high levels of basicity.

Docking simulation with DOCK3.5 and the selection of fragments

The X-ray crystallographic structure of thrombin (PDB: 1DWD) [25] was selected for the target of docking simulation. After removing the ligand (NAPAP) and water molecules from the thrombin structure, hydrogen atoms were added and charges were calculated using the Gasteiger–Marsili method [26]. Finally, the energy of the protein structure was minimized by MAXMIN2 [27], with the non-hydrogen atoms fixed. The rigid docking program DOCK3.5 [28] was employed for this simulation, as all the fragments in the fragment database were rigid molecules. Sphere points were generated within the S1 pocket using SPHGEN [29] so that the fragments could be specifically docked into the S1 pocket. Then, the docking simulation of the fragment database was performed with DOCK3.5. (CPU time 25 h) The 200 molecules exhibiting the best score were examined visually to select fragments based on binding modes and synthetic accessibility of the fragment derivatives. Although a significant number of amidine derivatives was found to be complementary to the pocket, they were excluded because of their high basicity. Finally, 8 commercially available compounds were purchased for the experimental assay (Figure 5).

Experimental evaluation of selected fragments

Docking simulations can exclude molecules that are not complementary to a binding pocket. This is especially effective when a binding pocket is narrow such as the S1 pocket of thrombin. However, the computational evaluation of binding affinity is still in the experimental stage. Therefore, we chose to experimentally evaluate selected fragments before conducting the time-consuming synthesis of fragment derivatives. This evaluation of fragment affinity was measured as the negative logarithm of fragment concentration when cleavage of chromogenic substrate was inhibited by 50% (pIC_{50}). Initially, the binding affinity of benzamidine (1) (Figure 5) was measured both for thrombin and FXa [30], as according to X-ray structures, the S1 pocket of thrombin (PDB: 1DWD) is almost identical to that of FXa. (PDB: 1FAX) [31] The experimental result showed that benzamidine has a higher affinity for FXa ($pIC_{50} = 3.4$) than thrombin ($pIC_{50} = 2.5$). Since we wished to order fragments according to their affinity determined by experiment, it was preferable to employ the assay system with higher sensitivity. Therefore, we decided to perform such an assay by experiment of the fragments with FXa instead of thrombin. The result of this assay showed that 1-aminoisoquinoline (**4**) and 2-aminoquinoline (**5**) have a moderate affinity for FXa (Figure 5). Considering the similarity of the S1 pockets between FXa and thrombin [32], we interpreted this result as an indication that these molecules should also have an affinity for thrombin. Since 1-aminoisoquinoline had a higher affinity than that of 2-aminoquinoline, we decided to synthesize derivatives of 1-aminoisoquinoline derivatives, taking the synthetic accessibility into account.

Design and synthesis of the fragment derivatives and the discovery of a novel inhibitor

Binding modes of benzamidine (X-ray) and 1-aminoisoquinoline (calculation) were compared. As shown in Figure 6, hydrogen bonds in a benzamidine – thrombin complex are well preserved in a 1-aminoisoquinoline – thrombin complex (GLY219 & ASP189, areas labeled A & B). Furthermore, 1-aminoisoquinoline has additional hydrophobic inter-

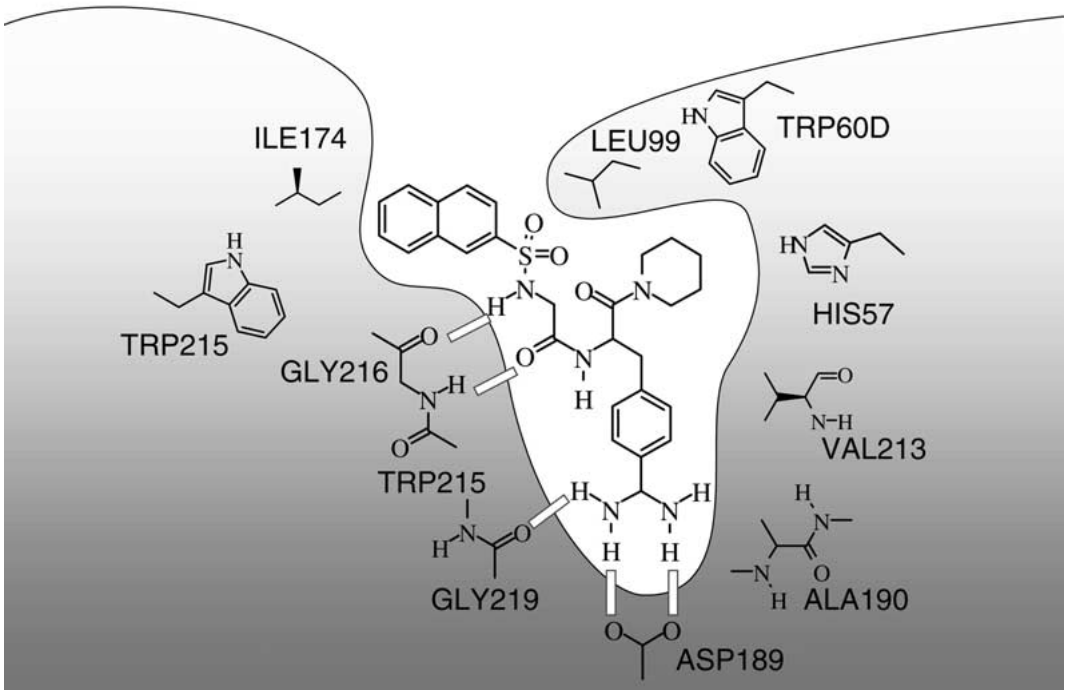

Figure 2. X-ray crystallographic structure of thrombin/NAPAP complex (PDB: 1DWD).

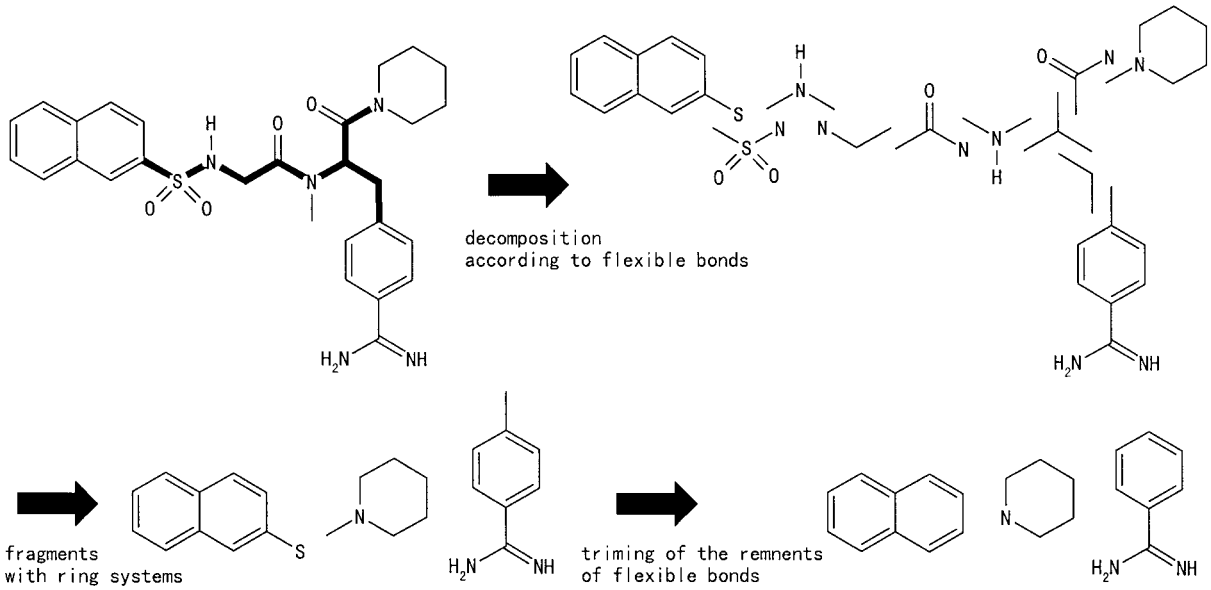

Figure 3. Fragments are generated by decomposition of molecules in a database. Flexible bonds are indicated by bold lines.

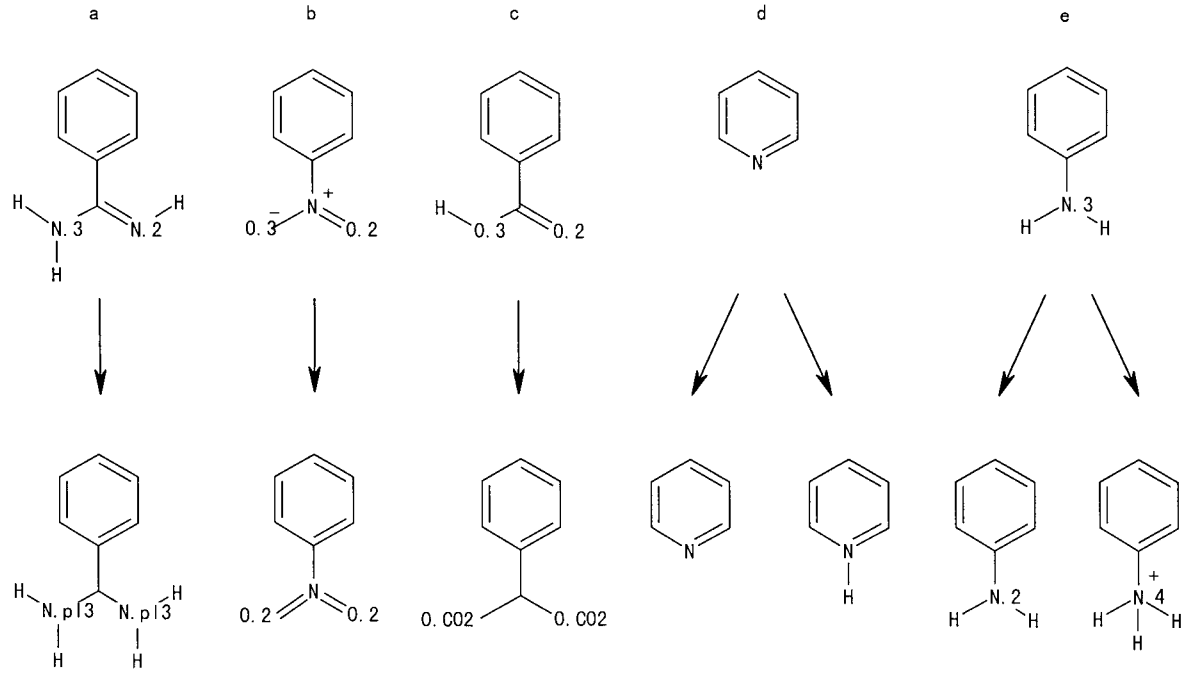

Figure 4. Molecular structures are modified both in terms of atom type and geometry.

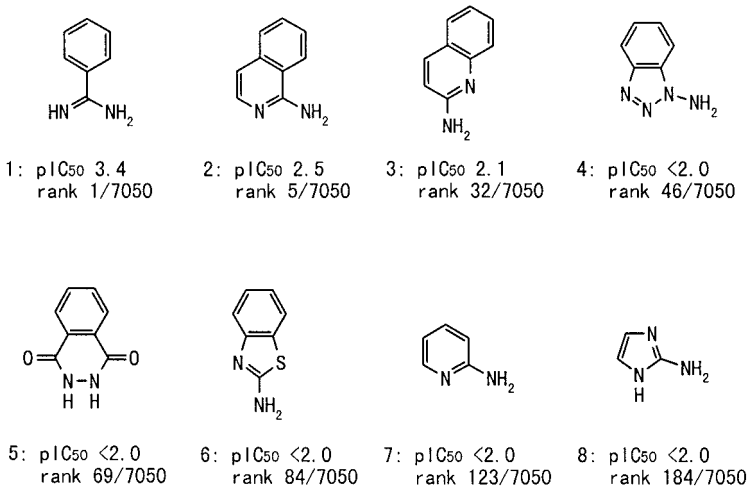

Figure 5. Eight commercially available compounds were purchased for experimental assay. The affinities were determined experimentally.

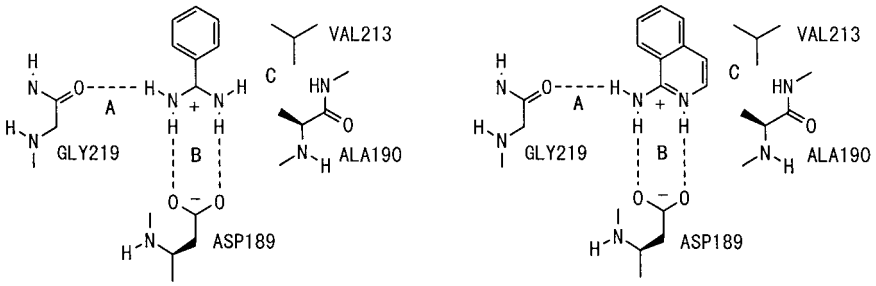

Figure 6. The binding modes of benzamidine (X-ray) and 1-aminoisoquinoline (calculation) complex with thrombin.

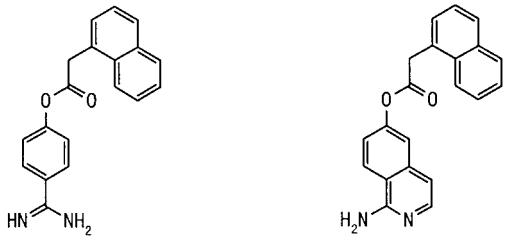

Figure 7. A known thrombin inhibitor and the designed 1-aminoisoquinoline derivative.

action with ALA190 and VAL213 (labeled C). Although 1-aminoisoquionoline has a lower affinity than benzamidine, which can probably be attributed to its lower basicity, such a chemical feature was preferred from the viewpoint of oral activity.

Molecule **1** in Figure 7 has been known as a thrombin inhibitor with a binding affinity of $pIC_{50} = 7.2$. If the presumed binding mode of 1-aminoisoquinoline in Figure 6 was correct, molecule **2** in Figure 7 would be expected to have high affinity. Accordingly, molecule **2** was synthesized and the affinity to thrombin was examined. The result of the chromogenic substrate assay showed an excellent affinity for thrombin ($pIC_{50} = 7.9$), indicating that the presumed binding mode in Figure 6 was correct.

Discussion

Our designing and screening strategy was found to be highly effective for finding novel ligands, taking advantage of both computational and experimental methods. However, in order to increase the possibility of discovering novel ligands, the following points have to be considered carefully.

Selection of a target protein

We selected a deep binding pocket as the target for a docking simulation, because simulations with such pockets are most effective in excluding molecules that are not complementary to the pocket. If a shallow binding pocket was selected as a target, the efficacy of screening chemical databases by computation would be expected to be significantly diminished. In addition, thrombin is one of the most intensively studied proteins in the history of structure-based drug design and thus, the research facilitates the assignment of the binding modes.

Solubility of fragments

The solubility of fragments is also an important factor in the evaluation of affinity by experiments. Small fragments usually have low affinity to a target protein and need to be dissolved in high concentrations of aqueous solutions for measurement of their affinities. Hydrophobic fragments often aggregate in water, hence their binding affinities may not be assayed correctly. The fragments selected for evaluation by experiment in this research have a hydrophilic moiety, so we did not encounter this problem. Our experimental evaluation of commercially available fragments was speedy, and significantly reduced the time to discover potent ligands.

Synthetic accessibility

It is important to factor in both computational evaluation and synthetic accessibility when searching for non-virtual inhibitors. The involvement of organic chemists in the early stages of designing ligands was the key for our successful discovery of the novel inhibitor. If synthetic accessibility was not properly considered, we would have certainly ended up identifying only virtual ligands that might bind to the target protein. In this case, the evaluation of the synthetic accessibility was achieved solely by human inspection, although a few computational programs exist [33, 34].

Conclusion

We have developed a method that is both practical and effective for discovering novel ligands. By applying this strategy to thrombin, the 1-aminoisoquinoline derivative was successfully discovered as a potent inhibitor ($pIC_{50} = 7.9$). Since the 1-aminoisoquinoline component has lower basicity than that of benzamidine, 1-aminoisoquinoline derivatives should be appropriate from the standpoint of oral activity. Considering the structural similarity of S1 pockets between thrombin and FXa, the 1-aminoisoquinoline component is expected to be utilized in inhibitors for both thrombin and FXa. In fact, 1-aminoisoquinoline derivatives have been extensively synthesized for these serine protease inhibitors in our research center [35] and our hypothesis of the binding mode for 1-aminoisoquinoline was supported by X-ray analysis. These results will be published shortly [36].

Acknowledgements

SM would like to thank Dr Iyer Vijay at Boston University and Dr Tim Mitchell at Cambridge Discovery Chemistry for their critical reading of this manuscript.

References

1. Mizutani, M.Y., Tomioka, N. and Itai, A., J. Mol. Biol., 243 (1994) 310.
2. Rarey, M., Kramer, B., Lengauer, T., Klebe and G., J. Mol. Biol., 261 (1996) 470.
3. Welch, W., Ruppert, J. and Jain, A., Chem. & Biol., 3 (1996) 449.
4. Makino, S. and Kuntz, I.D., J. Comp. Chem., 18 (1997) 1812.
5. Available Chemical Directory (ACD) available from Center for Scientific Computing (CSC), Tekniikantie 15 a D,Otaniemi, Espoo, Finland.
6. e.g. Available Chemical Directory (ACD), MDL Drug Data Report (MDDR), National Cancer Institute (NCI) are available from Center for Scientific Computing (CSC), Tekniikantie 15 a D,Otaniemi, Espoo, Finland.
7. Nishibata, Y. and Itai, A., J. Med. Chem., 36 (1989) 2921.
8. Moon, J.B. and Howe, W.J., Proteins, 11, (1991) 314.
9. Roe, D.C. and Kuntz, I.D., J. Comp. Aid. Mol. Des., 9 (1995) 269.
10. Böhm, H.-J., J. Comp. Aid. Mol. Des., 10 (1996) 265.
11. Kick, E.K., Roe, D.C., Skillman, A.G., Liu, G.C., Ewing, T.J.A., Sun, Y., Kuntz, I.D. and Ellman, J.A., Chem. Biol., 4 (1997) 297.
12. Shuker, S.B., Hajduk, P.J., Meadows, R.P. and Fesik, S.W., Science, 274, (1996) 1531.
13. Fejzo, J., Lepre, C.A., Peng, J.W., Bemis, G.W., Ajay, Murcko, M.A. and Moore, J.M., Chem. Biol., 6 (1999) 755.
14. SYBYL, Version 6.0.2, Tripos Associates, St. Louis, MO, 1993.
15. SYBYL Programming Language (SPL), Tripos Associates, St. Louis, MO, 1993.
16. Available from ftp://ftp.gnu.org/pub/gnu/gcc.
17. Kaiser, B., Drugs Future, 23 (1998) 423.
18. Protein Data Bank, Chemistry Department, Building 555, Brookhaven National Laboratory, Upton, NY 11973.
19. Dyke, S. F. and Kinsman, R. G., in Weissberger, A. and Yaylor, E.C. (eds), The Chemistry of Heterocyclic Compounds, Vol.
20. 38 John Wiley & Sons, New York, NY, 1981; p. 8.
21. Sturzebecher, J., Vieweg, H., Wikstrom, P., Turk, D., and Bode, W., BiolChem. Hoppe–Seyler, 373 (1992) 491.
22. Because this program is still under active development, interested readers should contact the authors concerning its availability.
23. SLN Manual, Tripos Associates, St. Louis, MO, 1993.
24. SYBYL, Version 6.02, Tripos Associates, St. Louis, MO, 1993.
25. Gasteiger, J. and Marsili, M., Tetrahedron, 36 (1980) 3210.
26. Banner, D.W. and Hadvary, P., J. Biol. Chem., 266 (1991) 20085.
27. Docking simulations with various charge models were performed without finding significant difference in the results.
28. SYBYL Command Manual, Tailor:MAXIMIN2, Tripos Associates, St. Louis, MO, 1993.
29. DOCK3.5 manual, Edited by D. A. Gschwend, Department of Pharmaceutical Chemistry, University of California, San Francisco.
30. DOCK3.5 manual, p. 13.
31. Hara, T., Yokoyama, A., Ishihara, H., Yokoyama, Y., Nagahara, T. and Iwamoto, M., Thrombosis Haemostasis, 71 (1994) 314.
32. Brandstetter, H., Kuhne, A., Bode, W., Huber, R., Saar, W.V.D., Wirthernsohn, K. and Engh, R.A., J. Biol. Chem., 271 (1996) 29988.
33. Although affinities of several fragments for thrombin were determined experimentally, the affinity of all the fragments other than benzamidine were not determined for FXa ($pIC_{50} < 2.0$).
34. Makino, S., Ewing, T.J.A. and Kuntz, I.D., J. Comp. Aid. Mol. Des., 13 (1999) 513.
35. CAESA by Peter Johnson, http:// www.chem.leeds.ac.uk/ICAMS/new_web/CAESA/caesa.htm.
36. Nakagawa, T., Makino, S., Sagi, K., Takayanagi, M. and Kayahara, T., New Inhibitors of FXa, Patent WO99-47503.
37. Nakagawa, T, Makino, S., Sagi, K., Takayanagi, M., Yamanashi, M., Yoshida, K., Yukuda, Y., Ishikawa, K., Kayahara, T., Takehana, S., Takahashi, M., Tsuji, T. and Shoji, M., J. Med. Chem. (submitted).