Inhibition of Plasminogen Activator Inhibitor-1
Its Mechanism and Effectiveness on Coagulation and Fibrosis

Yuko Izuhara, Satoru Takahashi, Masaomi Nangaku, Shunya Takizawa, Hideyuki Ishida, Kiyoshi Kurokawa, Charles van Ypersele de Strihou, Noriaki Hirayama, Toshio Miyata

Objective—Serine protease inhibitors (serpin) play a central role in various pathological processes including coagulation, fibrinolysis, malignancy, and inflammation. Inhibition of serpins may prove therapeutic. As yet, however, only very few small molecule serpin inhibitors have been reported. For the first time, we apply a new approach of virtual screening to discover novel, orally active, small molecule serpin inhibitors and report their effectiveness.

Methods and Results—We focused on a clinically important serpin, plasminogen activator inhibitor-1 (PAI-1), whose crystal structure has been described. We identify novel, orally active molecules able to enter into the strand 4 position (s4A) of the A β-sheet of PAI-I as a mock compound. In vitro they specifically inhibit the PAI-1 activity and enhance fibrinolysis activity. In vivo the most effective molecule (TM5007) inhibits coagulation in 2 models: a rat arteriovenous (AV) shunt model and a mouse model of ferric chloride–induced testicular artery thrombosis. It also prevents the fibrotic process initiated by bleomycin in mouse lung.

Conclusions—The present study demonstrates beneficial in vitro and in vivo effects of novel PAI-1 inhibitors. Our methodology proves to be a useful tool to obtain effective inhibitors of serpin activity. ***(Arterioscler Thromb Vasc Biol.* 2008;28:672-677)**

Key Words: serpin ■ virtual screening ■ anticoagulation ■ antifibrosis ■ plasminogen activator inhibitor-1 inhibitor

Serine protease inhibitors (serpin) play a central role in the regulation of a variety of pathological processes including coagulation, fibrinolysis, malignancy, and inflammation.[1] Their inhibition may prove therapeutic.

Serpins consist of a β-sheets-rich body. They contain an exposed mobile reactive center loop (RCL)[2] which, once cleaved by a target serine protease, inserts its N terminus part into the strand 4 position (s4A) of the A β-sheet, triggering its antiprotease activity. Small molecule compounds able to enter into the s4A position of the A β-sheet as a mock molecule may thus prevent the biological activity of the serpin.

Up to now, however, only very few small molecule serpin inhibitors have been described, none of which are in clinical use. Most have been discovered by high-throughput random screening (HTS) of a large chemical library,[3–5] a rather inefficient strategy. In the present study, a new approach of virtual screening based on the 3-dimensional structure of a serpin, (PAI-1), was used to discover novel, orally active, small molecule compounds able to inhibit the target molecule.

PAI-1 regulates the plasminogen activation system through inhibition of its target serine proteases, tissue-type and urokinase-type plasminogen activator (tPA and uPA).[1] Studies in humans and animals have demonstrated that PAI-1 expression is enhanced in various disorders such as thrombosis, fibrotic diseases, atherosclerosis, radiation damage, and cancer progression.[6] PAI-1 has been linked with fibrin deposition evolving into organ fibrosis and atherosclerosis, or with striking alterations of cell adhesion and migration mediating cancer progression.[7] The absence of PAI-1 in PAI-1 knockout mice markedly attenuates these pathological processes.[8–12] Inhibition of PAI-1 by a neutralizing antibody[13] provides similar promising results in animal experiments. Small molecule PAI-1 inhibitors, active orally, should prove useful to treat not only thrombotic disorders but also fibrotic processes and cancer.[2,14] They should be more efficient than the thrombolytic agents in present use such as streptokinase and recombinant tPA, both of which are protein-based and require intravenous administration to obtain a rapid onset action.

Original received October 8, 2007; final version accepted January 10, 2008.

From the Institute of Medical Sciences (Y.I., K.K., T.M.), Tokai University, Kanagawa, Japan; the Department of Pathology (S.T.), University of Tsukuba School of Medicine, Tsukuba, Japan; the Division of Nephrology and Endocrinology (M.N.), University of Tokyo School of Medicine, Tokyo, Japan; the Divisions of Neurology (S.T.), Physiology (H.I.), and Basic Medical Science and Molecular Medicine (N.H.), Tokai University School of Medicine, Kanagawa, Japan; the Service de Nephrologie (C.v.Y.d.S.), Universite Catholique de Louvain, Brussels, Belgium; and the Division of Translational Medicine (T.M.), Center for Translational and Advanced Animal Research on Human Disease, Tohoku University School of Medicine, Japan.

Correspondence to Toshio Miyata, MD, PhD, Division of Translational Medicine, Center for Translational and Advanced Animal Research on Human Disease, Tohoku University School of Medicine, 2-1 Seiryo-Machi, Aoba-ku, Sendai, 980-8575, Japan. E-mail t-miyata@mail.tains.tohoku.ac.jp

Arterioscler Thromb Vasc Biol **is available at http://atvb.ahajournals.org** **DOI: 10.1161/ATVBAHA.107.157479**

We relied on the 3-dimensional structure of PAI-1 and on the virtual screening method to identify novel orally bioavailable molecules able to fit into the s4A position of PAI-1 as a mock compound. These compounds inhibit PAI-1 activity in vitro and coagulation in vivo in two rodent models of vascular thrombosis. Furthermore, we demonstrate for the first time that they also prevent the fibrotic process induced in mice lung by bleomycin, and thus that, in this model, PAI-1 is not a surrogate marker of fibrosis but rather its main cause.

Materials and Methods

Virtual Screening

Virtual screening was performed by the software system MOE (Molecular Operating Environment).[15] The initial chemical library, of about 2 240 000 entries, merges various commercially available chemicals and deletes redundant entries. In the first step, 2 different filters based on molecular descriptors were used for coarse screening of compounds. The first filter was calculated from drug molecules that are clinically used in Japan. It was represented by the specific distributions of the molecular descriptors for these molecules and represents their general drug-likeness. The second filter was calculated from distributions of the molecular descriptors common to known reference inhibitors[16] and from the inhibitory reactive-center loop peptide N-Ac-TVASS-NH_2[17] that binds PAI-1. It represents the specific lead-likeness of the inhibitory molecules. Through the 2 filters we efficiently reduced the number of compounds to about 3000.[18] The second step involved docking simulation of the selected compounds to the PAI-1 molecule by the program Ph4Dock.[19] The crystal structure (1A7C) of the complex between PAI-1 and the inhibitory reactive-center loop peptide was obtained from the Protein Data Bank[20] and used as the target structure for the docking study. After removal of water molecules and binding peptide, hydrogen atoms were added in accord with the standard protonation states of acidic and basic residues in proteins and their positions were optimized. Because the cleft that is occupied by β-strand 4A in the latent form of PAI-1 likely plays a critical functional role, we focused on it for docking simulation. This target site was characterized by the alpha site finder function[21] in MOE for the subsequent docking experiments. The MMFF94s[22] force field was used in the docking simulation. In the final step, the positions of hydrogen atoms in PAI-1 and all atoms in small molecules were optimized. Docking results were sorted by the following interaction energy.

$$U_{total} = U_{ele} + U_{vdw} + U_{ligand}$$

U_{ele} and U_{vdw} are the electrostatic and van der Waals interactions between the protein and the small molecule, respectively, and U_{ligand} is the conformation energy of the small molecule. The molecules with the lowest U_{total} values are considered as candidate molecules for biological assay.

In Vitro Assays, Toxicity, and Pharmacokinetics

Please see the supplemental data section, available online at http://atvb.ahajournals.org, for detailed Methods.

Arteriovenous Shunt Model

Animal experiments were performed in accordance with the Animal Experimentation Guidelines of Tokai University School of Medicine. Thrombus formation in arteriovenous (AV) shunts was achieved in male CD rats (Charles River Japan Inc, Kanagawa, Japan) by a previously described method.[23] Before the study, TM5007 (300 mg/kg), warfarin (1.2 mg/kg), or ticlopidine (500 mg/kg), suspended in 0.5% carboxymethyl cellulose sodium salt (CMC) solution, was given by gavage, or tPA (275000 IU/kg) was administered intravenously by a bolus injection (n=7 for each group). Control rats were given 0.5% CMC solution only (n=7). Blood was allowed to circulate through the shunt for 30 min. The wet weight of the thrombus covering the silk thread was eventually measured.

Ferric Chloride–Induced Thrombosis Model and Visualization of Thrombi

Male mice were anesthetized with an intraperitoneal injection of 12 mL/kg ketamine-xylazine. Rhodamine 6G (Sigma; 0.1 mL of 0.1%) was injected intravenously. A testicular artery (100 to 150 μm in diameter) was carefully exposed for ferric chloride ($FeCl_3$) treatment. A cotton thread (0.2 mm in diameter) saturated with 0.25 mol/L $FeCl_3$ was applied to the adventitial surface of the testicular artery. After 5 minutes, the cotton thread was replaced by a saline solution in the wound. Thrombus formation in the testicular artery was subsequently monitored through 3-dimensional imaging using an ultrafast laser confocal microscope equipped with a piezo-electric motor control unit as previously reported.[24] The time from endothelial damage by $FeCl_3$ to occlusion of testicular arteries by large thrombi was measured. Mice were pretreated either by gavage of TM5007 (200 mg/kg), twice a day for 4 days. The antiplatelet glycoprotein (GP) IIb/IIIa agent, tirofiban (0.13 mg/kg, Wako), was single administered in a single intravenous injection before the injury.

Bleomycin-Induced Pulmonary Fibrosis

Male C57BL/6J (CLEA Japan Inc.) mice weighing 19 to 21 g were anesthetized with intraperitoneal pentobarbital and their trachea exposed by a cervical incision. Ten animals served as controls. Ten animals received an intratracheal instillation of bleomycin (Nippon Kayaku) dissolved in saline (1.5 U/kg), and 10 animals received in addition by gavage TM5007 (200 mg/kg) suspended in 0.5% CMC, twice a day for 14 days. Lung tissue was obtained for histological analysis and measurement of hydroxyproline content. Hydroxyproline was measured in tissue hydrolysates by the method of Kivirikko et al.[25] Tissue sections were stained with hematoxylin and eosin and pulmonary fibrosis was scored on a scale of 0 to 8 using a previously described method.[26] Azan stain for collagen was also used.

Statistics

All data are expressed as the mean±SE. Differences among groups were assessed by Kruskal-Wallis test. The statistical significance was determined by 2-tailed Mann Whitney U test. Values are considered significant at $P<0.05$. All statistical analyses were performed on the statistical package SPSS for Windows (Version 14.0, SPSS).

Results

Virtual Screening

Our library, encompassing 2 240 000 chemicals, was virtually screened, as described in methods. The use of 2 different filters reduced the number of compounds to about 3000, the first filter assessing drug-likeness and the second evaluating specific lead-likeness of the inhibitory molecules.

Docking simulation was then undertaken by a program Ph4Dock[19] for the remaining compounds. The crystal structure (1A7C) of the complex of PAI-1 with its inhibitory RCL peptide[20] was used, knowing that the 14-aa peptide corresponding to the N terminus of the RCL of PAI-1 inhibits the in vitro activity of PAI-1.[27] The program evaluated whether the compound is able to fit within the PAI-1 cleft. Virtual screening by a combination of 2 filters and by the docking method identified 95 candidate compounds with high binding affinity to the s4A position of PAI-1. The simulated binding mode of 1 novel candidate compound (TM5001) is illustrated (please see supplemental Figure I). The compound binds tightly within the cleft, in the s4A position.

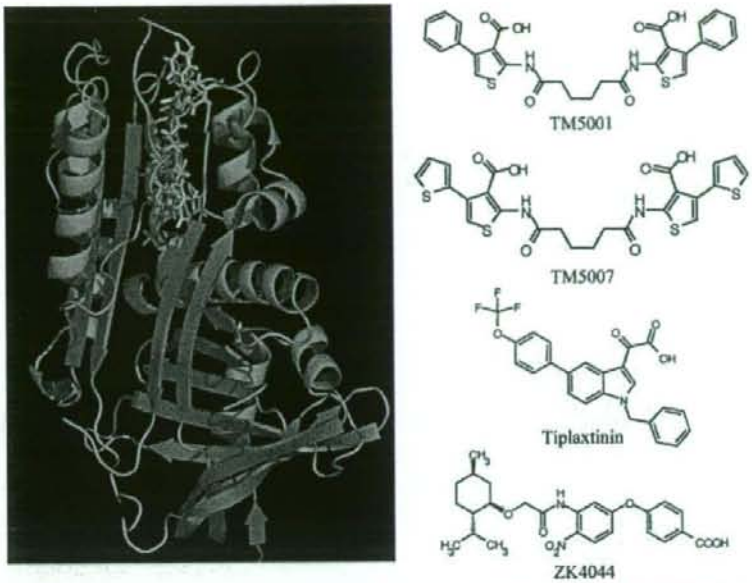

Figure 1. The docking study of small molecules to PAI-1. The binding modes of TM5001 (yellow), TM5007 (red), tiplaxtinin[3] (blue), and ZK4044[5] (green) are shown. Chemical structures of PAI-1 inhibitors are shown. Figures were drawn by a software PyMOL version 0.97 (DeLano Scientific LLC).

Docking simulations were also undertaken for PAI-1 inhibitors previously identified by HTS[3,5] to understand their mechanism of action. Both tiplaxtinin and ZK4044 bind at almost the same site as TM5001 and TM5007 (Figure 1). The similarity of these characteristics suggests that the 4 inhibitors share a common binding region at the s4A position, despite completely different chemical structures.

In Vitro Assessment

We purchased or synthesized 28 of the candidate compounds discovered by virtual screening and tested their biological activities in vitro by three different assays. Inhibition of PAI-1 activity was measured by tPA-dependent hydrolysis of peptide substrate. The 2 most effective candidate compounds, N, N′-bis (3,3′-carboxy-4,4′-phenyl-2,2′-thienyl) hexanedicarboxamide (TM5001) and N, N′-bis [3,3′-carboxy-4,4′-(2,2′-thienyl)-2,2′- thienyl] hexanedicarboxamide (TM5007) had an efficacy comparable to that of tiplaxtinin (IC_{50} for TM5001, TM5007, and tiplaxtinin 28.6±7.3, 29.2±4.2, and 40±7 μmol/L, respectively). TM5001 and TM5007 share a common binding mode with TM5001 as illustrated in Figure 1. It suggests a similar molecular mechanism for their activity. Neither TM5001 nor TM5007 (up to 250 μmol/L) modified other serpin/serine protease systems (ie, α_1-antitrypsin/trypsin and α_2-antiplasmin/plasmin): their PAI-1 inhibitory activity appears thus specific (please see supplemental Figure II).

On SDS-PAGE, PAI-1 formed a covalent complex with tPA whereas no PAI-1/tPA complex formation was observed when PAI-1 was preincubated with our compounds (exemplified for TM5007 in Figure 2).

Finally the inhibition of fibrinolysis was tested on a fibrin plate (exemplified for TM5007). The area of tPA-induced fibrinolysis was decreased by PAI-1. Preincubation of PAI-1 with our compounds prevented this effect (please see supplemental Figure III).

Toxicity and Pharmacokinetics

Cytotoxicity of TM5001 and TM5007 was assessed with HeLa cells as the LDH activity released into the culture medium after 24-h incubation. Results are expressed as percentage of the LDH release induced by the lysis of all cells. TM5001 (100 μmol/L) and TM5007 (250 μmol/L) did

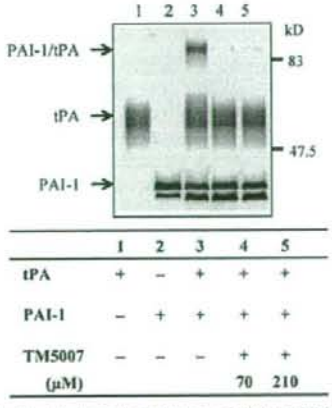

Figure 2. In vitro biological evaluation of PAI-1 inhibitor. The formation of PAI-1/tPA complex was assessed by SDS-PAGE and silver staining.

Table. Effect of TM5007 on Thrombus Weight, and Plasma APTT, PT, and PAI-1 Activity in the Rat Arteriovenous Shunt Model

Treatment		n	Thrombus Weight Subtracting Thread (mg)	APTT (sec)	PT (sec)	PAI-1 Activity (ng/ml)
Vehicle	0.5% CMC, p.o.	7	74.3±3.2	14.8±0.3	14.3±0.3	1.0±0.1
TM5007	300 mg/kg, p.o.	7	54.8±0.8†	14.6±0.5	14.4±0.2	0.8±0.1*
Warfarin	1.2 mg/kg, p.o.	7	51.1±3.5†	29.2±2.8†	72.1±11.6†	1.0±0.1
Ticlopidine	500 mg/kg, p.o.	7	55.3±1.8†	14.9±0.2	14.4±0.1	1.3±0.1*
Tpa	275 000 IU/kg, i.v.	7	64.6±5.8	17.6±0.5†	14.4±0.1	0.6±00†

Data are expressed as mean±SE. * $P<0.05$, † $P<0.01$ vs vehicle by Mann–Whitney U test.

not significantly raise maximum LDH activity above unstimulated controls (30.8±7.1 and 26.6±2.2%, respectively, versus 23.8±1.6%; $P=0.077$ and 0.137).

Acute toxicity of TM5001 and 5007 was evaluated in vivo in mice. Various single doses of up to 2000 mg/kg of both compounds elicited no symptoms up to 2 wks later.

Subacute toxicity of TM5007 was assessed in rats at 2 different doses (300 mg/kg/d for 1 wk or 2000 mg/kg/d for 2 week) given daily. Neither blood pressure nor body weight was modified. No biochemical abnormalities were noted in plasma and urine including bleeding time, APTT, PT, TT, and red blood cell count (please see supplemental Tables).

Plasma Tmax, Cmax, and T1/2 were calculated in rats given orally 50 mg/kg of each of the 2 compounds. They reached 18 h, 32 μmol/L, and 54 h for TM5001, and 18 h, 8.8 μmol/L, and 124 h for TM5007.

In Vivo Assessment

TM5007 was chosen for further investigations, as it proved more effective than TM5001. Its in vivo anticoagulant effectiveness was assessed by weighing the blood clot obtained in a rat AV shunt model (Table). Blood clot weight was significantly lower in rats given 300 mg/kg of TM5007 (54.8±0.8 mg) than in vehicle-treated rats (74.3±3.2 mg) ($P<0.01$). This effect, obtained at a plasma concentration of TM5007 of 5.2±0.7 μmol/L, was equivalent to that of warfarin (1.2 mg/kg) and ticlopidine (500 mg/kg), and superior to that of tPA (275000 IU/kg). PAI-1 activity was significantly reduced by TM5007 (0.8±0.1 versus 1.0±0.1 ng/mL in vehicle treated rat) but not by other agents, whereas APTT and PT were not modified (Table).

The anticoagulant effectiveness of TM5007 was also evaluated in $FeCl_3$-treated mouse testicular arteries. Growth of thrombi, followed by 3-dimensional imaging, led to complete arterial occlusion within 12.7±2.7 min (n=29) in vehicle-treated mice versus only 56.3±8.9 min in TM5007-treated mice (n=15, $P<0.01$ compared to the control group). This effect was obtained at a TM5007 plasma concentration of 4.6±0.6 μmol/L. Tirofiban had a similar effect (58±12.4 min; n=5, $P<0.01$ compared to the control group).

The in vivo antifibrotic effect of TM5007 was tested in a mouse model of bleomycin induced pulmonary fibrosis. Bleomycin increased significantly the lung hydroxyproline content above that of control mice (232.9±8.5 versus 140.2±4.8 μg/lung, $P<0.001$). TM5007 significantly lowered the bleomycin-induced lung hydroxyproline content (204.2±9.5 μg/lung, $P<0.05$). Bleomycin raised plasma PAI-1 activity above that of control mice (1.7±0.2 versus 0.8±0.1 ng/mL, $P<0.001$). TM5007, at a plasma concentration of 9.2±0.2 μmol/L, lowered significantly the bleomycin-induced PAI-1 rise (1.2±0.1 ng/mL, $P<0.05$). Histological evidence of pulmonary fibrosis was markedly improved by TM5007 (Figure 3A through 3C). Azan staining disclosed the accumulation of collagen in bleomycin-treated lungs (Figure 3D through 3F). Bleomycin raised the fibrosis score above control (4.7±0.37 versus 0.5±0.17, $P<0.001$), a rise that was partially prevented by TM5007 (2.9±0.38, $P<0.01$), in good agreement with the results of plasma PAI-1 activity.

Discussion

Relying on virtual screening and the 3-dimensional structure of the complex of PAI-1 with its inhibitory peptide, we have identified 2 novel, orally bioavailable, small molecule PAI-1 inhibitors, TM5001 and TM5007. Both are stable, nontoxic, and devoid of cellular toxicity as demonstrated in vitro by their inability to raise LDH levels in the medium of cultured HeLa cells. The absence of acute and subacute toxicity is confirmed in vivo in mice given a single dose of up to 2000 mg/kg, or in rats fed 300 mg/kg for 1 week or 2000 mg/kg for 2 weeks. The in vivo effectiveness of TM5007 is demonstrated in animal models of either acute vascular thrombosis or of chronic lung fibrosis, without deleterious effects on blood pressure or bleeding, in good agreement with previous results in PAI-1 deficient mice and humans.[8,14]

The specificity of the effect of TM5007 on PAI-1 was further documented in other serpin/serine protease systems (ie, α_1-antitrypsin/trypsin and α_2-antiplasmin/plasmin) by a chromogenic assay with synthetic substrates. TM5007 exhibited no inhibitory activity against any of the closely related serpins or serine proteases at a concentration of 250 μmol/L. This concentration is approximately 10 times above the IC_{50} (29 μmol/L) for TM5007 against PAI-1 in the in vitro study and approximately 25 to 50 times greater than the peak plasma levels of this compound in vivo in mice and rats (5 to 10 μmol/L). The activity of TM5007 appears thus specific for PAI-I.

Inhibition of thrombus formation has been demonstrated for all known PAI-I inhibitors, both in vitro and in vivo in acute thrombotic models.[3–5,28,29] Delayed effects, by contrast, on subsequent tissue remodeling were demonstrated only for tiplaxtinin in a murine model of angiotensin II–induced hypertension.[11] We confirm that TM5007 is a powerful antithrombotic agent, which does not prolong bleeding time, PT, and APTT. Its effect, at a dosage of 300 mg/kg, was

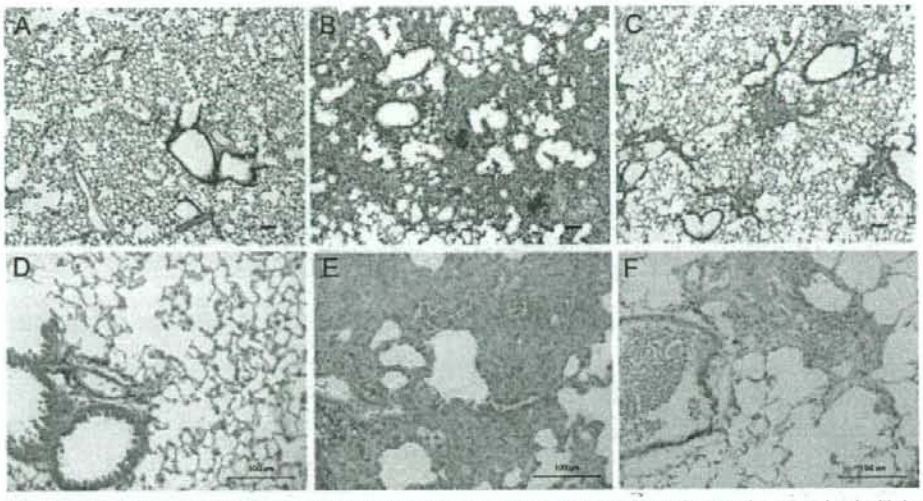

Figure 3. Histological analysis of bleomycin-induced pulmonary fibrosis. Representative photomicrographs are stained with hematoxylin and eosin (A–C) or for collagen with Azan (D–F). Nontreatment (A and D), bleomycin+vehicle (B and E), bleomycin+TM5007 (C and F). Magnification is ×50 (A–C) or ×200 (D–F). Bar shows 100 μm.

equivalent to that of warfarin (1.2 mg/kg) and ticlopidine (500 mg/kg) and superior to that of tPA (275000 IU/kg).

The antithrombotic effect of TM5007 was also demonstrated in a $FeCl_3$-induced testicular artery thrombosis mouse model. Oral pretreatment with TM5007 (200 mg/kg twice daily for 4 days) was as efficacious as a single intravenous injection of an antiplatelet drug, Tirofiban (0.13 mg/kg).

Of note, the IC_{50} of TM5007 against PAI-1 calculated in vitro (29.2 μmol/L) exceeds the peak plasma levels (5 to 10 μmol/L) of this compound observed in vivo in mice and rats. This discrepancy probably reflects the differences in the experimental systems. In vitro, PAI-1 inhibition is measured directly. In vivo, by contrast, its effect on thrombus formation is complex as it involves several factors other than PAI-1/tPA.

We demonstrate for the first time in the present study that PAI-I inhibition prevents the fibrotic process initiated in the lung by bleomycin. Eitzman et al, studying mice overexpressing or lacking the PAI-1 gene, have demonstrated a strong correlation between PAI-1 expression and collagen accumulation in lung tissue.[9] The inhibition of lung fibrosis by TM5007 establishes that PAI-1 is not a mere surrogate marker of fibrosis but rather its main cause. This observation is of potential importance. Fibrotic changes are indeed associated with the failure of several other organs, including the heart, vessels, liver, and kidney. Their prevention might transform the fate of numerous diseases such as cardiovascular disease, liver cirrhosis, renal disease, and radiation injury.

A few small molecule PAI-I inhibitors have been discovered by the rather inefficient HTS screening of a chemical library.[3–5] Gerlatova et al[30] demonstrated that a binding epitope for tiplaxtinin is adjacent to a previously identified interaction site for vitronectin, thereby suggesting that the antiserpin activity of this drug is mediated by an interaction between PAI-I and vitronectin. By contrast, we focused on the s4A position as a target site of PAI-I inhibition. Our docking simulation demonstrated that TM5001 and TM5007 preferentially bind to this site, suggesting that our compounds exert inhibitory activity through blocking the s4A position. On SDS-PAGE, no PAI-1/tPA complex formation was indeed observed when PAI-1 was preincubated with our compounds (Figure 2). It thus appears from the results that the inhibitory mechanism of our compound is not identical with that of tiplaxtinin. Interestingly, despite their completely different chemical structures, tiplaxtinin as well as ZK4044 potentially associate with the s4A position on our docking simulation, in good agreement with a previous assumption by Gerlatova et al[30] that tiplaxtinin inhibits the PAI-1 at multiple mechanisms. Altogether, these findings confirm the critical role of the s4A position as a target site of PAI-I inhibition.

Utilization of the 3-dimensional structure of PAI-1 has allowed not only an understanding of the molecular events leading to PAI-1 inhibition but also the virtual screening of new inhibitors for clinical use. Their potential applications include thrombotic disorders (arterial and venous), fibrotic diseases, amyloidosis, obesity, and type 2 diabetes mellitus.[31,32] The availability of a specific inhibitor should also offer a potentially important pharmacological tool to investigate the role of PAI-1 in these processes.

Finally, elucidation of the 3-dimensional structure of other serpins together with the use of virtual screening, should allow the identification of other small molecule serpin inhibitors, to curtail their harmful effects.

Acknowledgments

We thank H. Saito, K. Okada, O. Matsuo, H. Oishi, and K. Nakazato for their expertise.

Sources of Funding

This study was supported by a grant from the New Energy and Industrial Technology Development Organization in Japan (to T.M., N.H., and K.K.).

Disclosures

None.

References

1. Carrell RW, Boswell DR. Serpins: the superfamily of plasma serine protease inhibitors. In: Barrett, Salvesen, eds. *Proteinase Inhibitors. Elsevier Science.* 1986; 403–420.
2. Gils A, Declerck PJ. The structural basis for the pathophysiological relevance of PAI-1 in cardiovascular diseases and the development of potential PAI-1 inhibitors. *Thromb Haemost.* 2004;91:425–437.
3. Elokdah H, Abou-Gharbia M, Hennan JK, McFarlane G, Mugford CP, Krishnamurthy G, Crandall DL. Tiplaxtinin, a novel, orally efficacious inhibitor of plasminogen activator inhibitor-1: design, synthesis, and preclinical characterization. *J Med Chem.* 2004;47:3491–3494.
4. Crandall DL, Elokdah H, Di L, Hennan JK, Gorlatova NV, Lawrence DA. Characterization and comparative evaluation of a structurally unique PAI-1 inhibitor exhibiting oral in-vivo efficacy. *J Thromb Haemost.* 2004;2:1422–1428.
5. Liang A, Wu F, Tran K, Jones SW, Deng G, Ye B, Zhao Z, Snider RM, Dole WP, Morser J, Wu Q. Characterization of a small molecule PAI-1 inhibitor, ZK4044. *Thromb Res.* 2005;115:341–350.
6. Dellas C, Loskutoff DJ. Historical analysis of PAI-1 from its discovery to its potential role in cell motility and disease. *Thromb Haemost.* 2005;93: 631–640.
7. Bajou K, Noël A, Gerard RD, Masson V, Brunner N, Holst-Hansen C, Skobe M, Fusenig NE, Carmeliet P, Collen D, Foidart JM. Absent of host plasminogen activator inhibitor 1 prevents cancer invasion and vascularization. *Nature Med.* 1998;4:923–928.
8. Carmeliet P, Stassen JM, Schoonjans L, Ream B, van den Oord JJ, De Mol M, Mulligan RC, Collen D. Plasminogen activator inhibitor-1 gene-deficient mice: II Effects on hemostasis, thrombosis, and thrombolysis. *J Clin Invest.* 1993;92:2756–2760.
9. Eitzman DT, McCoy RD, Zheng X, Fay WP, Shen T, Ginsburg D, Simon RH. Bleomycin-induced pulmonary fibrosis in transgenic mice that either lack or overexpress the murine plasminogen activator inhibitor-1 gene. *J Clin Invest.* 1996;97:232–237.
10. Eitzman DT, Westrick RJ, Xu Z, Tyson J, Ginsburg D. Plasminogen activator inhibitor-1 deficiency protects against atherosclerosis progression in the mouse carotid artery. *Blood.* 2000;96:4212–4215.
11. Weisberg AD, Albornoz F, Griffin JP, Crandall DL, Elokdah H, Fogo AB, Vaughan DE, Brown NJ. Pharmacological inhibition and genetic deficiency of plasminogen activator inhibitor-1 attenuates angiotensin II/salt-induced aortic remodeling. *Arterioscler Thromb Vasc Biol.* 2005; 25:365–371.
12. Nicholas SB, Aguiniga E, Ren Y, Kim J, Wong J, Govindarajan N, Noda M, Wang W, Kawano Y, Collins A, Hsueh WA. Plasminogen activator inhibitor-1 deficiency retards diabetic nephropathy. *Kidney Int.* 2005;67: 1297–1307.
13. van Giezen JJJ, Wahlund G, Nerme V, Abrahamsson T. The F_{ab}-fragment of a PAI-1 inhibiting antibody reduces thrombus size and restores blood flow in a rat model of arterial thrombosis. *Thromb Haemost.* 1997;77: 964–969.
14. Wu Q, Zhao Z. Inhibition of PAI-1: A new anti-thrombotic approach. *Cur Drug Targets-Cardiovas & Haemat Dis.* 2002;2:27–42.
15. *MOE (Molecular Operating Environment)*, version 2004.04; Chemical Computing Group Inc.: Montreal, Quebec, Canada, 2004.
16. Egelund R, Einholm AP, Pedersen KE, Nielsen RW, Christensen A, Deinum J, Andreasen PA. A regulatory hydrophobic area in the flexible joint region of plasminogen activator inhibitor-1: defined with fluorescent activity-neutralizing ligands. *J Biol Chem.* 2001;276:13077–13086.
17. Xue Y, Bjorquist P, Inghardt T, Linschoten M, Musil D, Sjolin L, Deinum J. Interfering with the inhibitory mechanism of serpins: crystal structure of a complex formed between cleaved plasminogen activator inhibitor type 1 and a reactive-centre loop peptide. *Structure.* 1998;6:627–636.
18. Horio K, Goto J, Hirayama N. A simple method to improve the odds in finding 'lead-like' compounds from the chemical library. *Chem Pharm Bull.* 2007;55:980–984.
19. Goto J, Kataoka R, Hirayama N. Ph4Dock-Pharmacophore-based protein-ligand docking. *J Med Chem.* 2004;47:6804–6811.
20. Bernstein FC, Koetzle TF, Williams GJ, Meyer EF Jr, Brice MD, Rodgers JR, Kennard O, Shimanouchi T, Tasumi M. The protein data bank: a computer-based archival file for macromolecular structures. *J Mol Biol.* 1997;112:535–542.
21. Edelsbrunner H, Facello M, Fu P, Liang J Measuring Proteins and Voids in Proteins. *Proceedings of the 28th Hawaii International Conference on Systems Science.* 1995; 256–264.
22. Halgren TA. Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. *J Comp Chem.* 1996;17: 490–519.
23. Morishima Y, Tanabe K, Terada Y, Hara T, Kunitada S. Antithrombotic and hemorrhagic effects of DX-9065a, a direct and selective factor Xa inhibitor: comparison with a direct thrombin inhibitor and antithombin III-dependent anticoagulants. *Thromb Haemost.* 1997;78:1366–1371.
24. Goto S, Tamura B. S, Ishida H. Ability of anti-GP IIb/IIIa agents to dissolve platelet thrombi formed on a collagen surface under blood flow conditions. *J Am Coll Cardiol.* 2004;44:316–323.
25. Kivirikko KI, Laitinen O, Prockop D. J. Modifications of a specific assay for hydroxyproline in urine. *Anal Biochem.* 1967;19:249–255.
26. Ashcroft T, Simpson JM, Timbrell V. Simple method of estimating severity of pulmonary fibrosis on a numerical scale. *J Clin Pathol.* 1988;41:467–470.
27. Eitzman DT, Fay WP, Lawrence DA, Francis-Chmura AM, Shore JD, Olson ST, Ginsburg D. Peptide-mediated inactivation of recombinant and platelet plasminogen activator inhibitor-1 in vitro. *J Clin Invest.* 1995; 95:2416–2420.
28. Charlton PA, Faint RW, Bent F, Bryans J, Chicarelli-Robinson I, Mackie I, Machin S, Bevan P. Evaluation of a low molecular weight modulator of human plasminogen activator inhibitor-1 activity. *Thromb Haemost.* 1996;75:808–815.
29. Friederich PW, Levi M, Biemond BJ, Charlton P, Templeton D, van Zonneveld AJ, Bevan P, Pannekoek H, and ten Cate JW. Novel low-molecular-weight inhibitor of PAI-1 (XR5118) promotes endogenous fibrinolysis and reduces postthrombolysis thrombus growth in rabbits. *Circulation.* 1997;96:916–921.
30. Gorlatova NV, Cale JM, Elokdah H, Li D, Fan K, Warnock M, Crandall DL, Lawrence DA. Mechanism of inactivation of plasminogen activator inhibitor-1 by a small molecule inhibitor. *J Biol Chem,* 2007;282: 9288–9296.
31. Smith LH, Dixon JD, Stringham JR, Eren M, Elokdah H, Crandall DL, Washington K, Vaughan DE. Pivotal role of PAI-1 in a murine model of hepatic vein thrombosis. *Blood.* 2006;107:132–134.
32. Vaughan DE, De Taeve BM, Eren M. PAI-1 Antagonists: predictable indications and unconventional applications. *Curr Drug Targets.* 2007;8: 962–970.

ANALYTICAL SCIENCES 2008, VOL. 24

Crystal Structure of (*4SR,5RS*)-5-(4-Cyanophenyl)-4-trimethylsilyl-3-methylisoxazoline

Toshio AKIMOTO,* Hisashi HARA, Taichi NAKANO,** and Noriaki HIRAYAMA*†**

**Basic Medical Science and Molecular Medicine, Tokai University School of Medicine, 143 Shimokasuya, Isehara, Kanagawa 259-1143, Japan*

***Department of Materials Chemistry, School of High-Technology for Human Welfare, Tokai University, 317 Nishino, Numazu, Shizuoka 410-0395, Japan*

The crystal of the title compound, $C_{14}H_{18}N_2OSi$, belongs to space group $P\bar{1}$ with cell dimensions of a = 6.2639(3), b = 10.818(1), c = 12.085(1)Å, α = 69.976(4), β = 76.959(5), and γ = 81.566(3)°. The final R value is 0.0486. The isoxazoline ring takes an envelope conformation. The dihedral angle between the isoxazoline and phenyl rings is 103.9(1)°. The Si-C bond bridging the isoxazoline ring and trimethylsilyl group is significantly longer than the other three Si-C bonds.

(Received, April 15, 2008; Accepted June 23, 2008; Published on web August 1, 2008)

The first organogermanium pharmaceutical propagermanium was launched in 1994. Propagermanium has been introduced into clinics for the treatment of chronic hepatitis. This achievement has stimulated further investigations of organosilicon compounds with biological activities in addition to the organogermanium compounds. In particular, silicon-containing isoxazolines are interesting groups of compounds, since some of them possess biological activity. Aiming to obtain bioactive silicon-containing isoxazolines, we have synthesized a series of silicon-containing isoxazolines. In the present paper we report on the X-ray structure of (*4SR,5RS*)-5-(4-cyanophenyl)-4-trimethylsilyl-3-methylisoxazoline, synthesized by the [2 + 3] cycloaddition[1] of ethanenitrile oxide to (*E*)-1-(4-cyanophenyl)-2-(trimethylsilyl)ethene.[2] The chemical structure of the racemate is shown in Fig. 1.

Single crystals were obtained from 15% ethyl acetate in hexane. A colorless platelet crystal with a size of 0.38 × 0.35 × 0.30 mm was mounted on a glass fiber and used for data collection. The structure was solved by direct methods and non-H atoms were refined by a full-matrix least-squares method with anisotropic temperature factors. The positions of all H-atoms were located by a difference Fourier synthesis and refined isotropically. The crystal and experimental data are given in Table 1. The atomic parameters of the non-H atoms are given in Table 2.

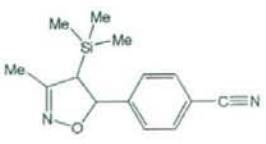

Fig. 1 Chemical structure of (*4SR,5RS*)-5-(4-cyanophenyl)-4-trimethylsilyl-3-methylisoxazoline.

The molecular structure, drawn by ORTEP-III,[5] is shown in Fig. 2. Selected bond lengths, bond angles and torsion angles are given in Table 3. The isoxazoline ring takes an envelope conformation with C6 being deviated by 0.387(4)Å from the least-squares plane defined by the other four non-H atoms in the ring. The dihedral angle between the isoxazoline and phenyl rings is 103.9(1)°. As indicated by the O1-C6-C8-C9 torsion angle of −18.1(3)°, the O1-C6 and C8-C9 bonds are in a *syn* orientation. This conformation may be due to the intramolecular C-H···O hydrogen bond between C9 and O1 [C9···O1 2.779(3), C9-H3 0.89(3), H3···O1 2.42(2)Å, ∠C9-H3···O1 104(2)°]. The

Table 1 Crystal and experimental data

Formula: $C_{14}H_{16}N_2OSi$	
Formula weight = 258.39	
Crystal system: triclinic	
Space group: $P\bar{1}$	$Z = 2$
a = 6.2639(3)Å	α = 69.976(4)°
b = 10.818(1)Å	β = 76.959(5)°
c = 12.085(1)Å	γ = 81.566(3)°
V = 747.4(1)Å^3	
D_x = 1.148 g/cm^3	
No. of observations (*all*) = 2445	
θ_{max} = 68.2° with Cu K_α	
R ($I > 2.00\sigma(I)$) = 0.0486	
$(\Delta/\sigma)_{max}$ = 0.000	
$(\Delta\rho)_{max}$ = 0.21 e/Å^3	
$(\Delta\rho)_{min}$ = −0.20 e/Å^3	
Measurement: Rigaku RAXIS-RAPID	
Program system: CrystalStructure 3.7.0[3]	
Structure determination: SIR92[4]	
Refinement: full-matrix	

CCDC 691367 contains the supplementary crystallographic data for this paper. These data can be obtained free of charge from The Cambridge Crystallographic Data Centre *via* www.ccdc.cam.ac.uk/data_request/cif.

† To whom correspondence should be addressed.
E-mail address: hirayama@is.icc.u-tokai.ac.jp

Table 2 Atomic coordinates and equivalent isotropic thermal parameters (B_{eq})

atom	x	y	z	B_{eq}(Å^2)
Si(1)	0.53606(11)	0.35156(7)	0.22337(6)	4.998(16)
O(1)	0.8373(3)	0.3291(2)	0.44718(16)	6.48(4)
N(1)	0.9304(3)	0.2115(2)	0.41902(19)	6.59(6)
N(2)	0.0930(6)	0.1050(4)	1.0474(2)	11.61(11)
C(1)	0.4562(8)	0.2429(4)	0.1521(4)	7.96(11)
C(2)	0.7958(6)	0.4261(4)	0.1389(3)	7.43(9)
C(3)	0.3158(6)	0.4834(3)	0.2349(3)	6.80(8)
C(4)	0.5743(4)	0.2514(2)	0.3824(2)	4.57(5)
C(5)	0.7852(4)	0.1683(2)	0.3860(2)	5.53(6)
C(6)	0.6041(4)	0.3339(2)	0.4582(2)	4.93(5)
C(7)	0.8343(10)	0.0463(3)	0.3515(3)	8.68(10)
C(8)	0.4918(4)	0.2821(2)	0.5882(2)	4.65(5)
C(9)	0.6068(5)	0.2056(2)	0.6754(2)	5.54(6)
C(10)	0.5028(5)	0.1581(3)	0.7933(2)	6.27(7)
C(11)	0.2820(5)	0.1884(2)	0.8253(2)	6.02(7)
C(12)	0.1653(5)	0.2622(3)	0.7394(2)	6.54(8)
C(13)	0.2682(4)	0.3099(2)	0.6211(2)	5.79(7)
C(14)	0.1752(5)	0.1416(3)	0.9489(2)	8.04(9)

$B_{eq} = 8/3\pi^2(U_{11}(aa^*)^2 + U_{22}(bb^*)^2 + U_{33}(cc^*)^2 + 2U_{12}(aa^*bb^*)\cos\gamma + 2U_{13}(aa^*cc^*)\cos\beta + 2U_{23}(bb^*cc^*)\cos\alpha)$.

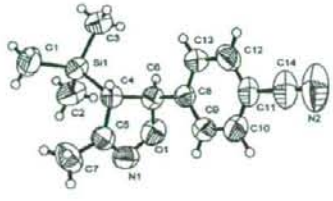

Fig. 2 Molecular structure of (*4SR,5RS*)-5-(4-cyanophenyl)-4-trimethylsilyl-3-methylisoxazoline along with the labeling atoms. Thermal ellipsoids of non-H atoms are drawn at the 40% probability level.

Si1–C4 bond length of 1.905(2)Å is significantly longer the other three Si-C bond lengths. The average distance of the latter three bonds is 1.851 Å. The endocyclic bond angle centered on the C4 atom is 99.1(2)°, whereas the C4-C5-N1 angle is 114.4(2)°. The exocyclic bond angles around the C5 atom are highly asymmetric, and the C4-C5-C7 angle is significantly larger than the counterpart. These geometrical characteristics around the C4 atom imply that a certain steric repulsion might exist between the trimethylsilyl group and the isoxazoline moiety. However, only van der Waals contacts between H6(C3) and H11(C6), and H10(C1) and H13(C7) are observed. The atoms in the parentheses designate the atoms to which the relevant hydrogen atoms are attached. The distances are 2.39(4) and 2.49(5)Å, respectively. The molecules are packed by van der Waals interactions in the crystal.

Table 3 Selected bond lengths (Å), bond angles (°) and torsion angles (°)

Si(1)-C(1)	1.853(6)	Si(1)-C(2)	1.847(3)
Si(1)-C(3)	1.852(3)	Si(1)-C(4)	1.905(2)
O(1)-N(1)	1.436(3)	O(1)-C(6)	1.432(3)
N(1)-C(5)	1.273(4)	N(2)-C(14)	1.140(4)
C(1)-Si(1)-C(2)	111.4(2)	C(1)-Si(1)-C(3)	110.4(2)
C(2)-Si(1)-C(3)	109.7(2)	C(1)-Si(1)-C(4)	109.6(2)
C(2)-Si(1)-C(4)	108.5(1)	C(3)-Si(1)-C(4)	107.2(2)
N(1)-O(1)-C(6)	107.(2)	O(1)-N(1)-C(5)	108.3(2)
Si(1)-C(4)-C(5)	112.9(2)	Si(1)-C(4)-C(6)	114.6(2)
N(1)-C(5)-C(4)	114.4(2)	N(1)-C(5)-C(7)	120.6(3)
O(1)-C(6)-C(4)	104.0(2)	O(1)-C(6)-C(8)	111.2(2)
N(2)-C(14)-C(11)	179.0(4)	C(5)-C(4)-C(6)	99.1(2)
C(4)-C(5)-C(7)	124.9(3)		
C(1)-Si(1)-C(4)-C(5)	76.9(2)	C(1)-Si(1)-C(4)-C(6)	-170.6(2)
C(2)-Si(1)-C(4)-C(5)	-44.9(2)	C(2)-Si(1)-C(4)-C(6)	67.6(2)
C(3)-Si(1)-C(4)-C(5)	-163.2(2)	C(3)-Si(1)-C(4) -C(6)	-50.7(2)
C(6)-O(1)-N(1)-C(5)	14.9(2)	N(1)-O(1)-C(6)-C(4)	-24.5(2)
N(1)-O(1)-C(6)-C(8)	98.0(2)	O(1)-N(1)-C(5)-C(4)	2.0(2)
O(1)-N(1)-C(5)-C(7)	-179.7(2)	Si(1)-C(4)-C(5)-N(1)	105.2(2)
Si(1)-C(4)-C(5)-C(7)	-72.9(2)	C(6)-C(4)-C(5)-N(1)	-16.5(2)
Si(1)-C(4)-C(6)-O(1)	-96.8(2)	Si(1)-C(4)-C(6)-C(8)	142.2(2)
C(5)-C(4)-C(6)-O(1)	23.7(2)	O(1)-C(6)-C(8)-C(9)	-18.1(3)
O(1)-C(6)-C(8)-C(13)	162.8(2)		

Acknowledgements

We thank the Research and Study Program of Tokai University Educational System General Research Organization for financial support.

References

1. T. Mukaiyama and T. Hoshino, *J. Am. Chem. Soc.*, **1960**, *82*, 5339.
2. T. Endo, F. Sasaki, H. Hara, J. Suzuki, S. Tamura, Y. Nagata, T. Iyoshi, A. Saigusa and T. Nakano, *Appl. Organometal. Chem.*, **2007**, *21*, 183.
3. CrystalStructure, verSion 3.5.1, **2000** – **2003** Crystal Structure Analysis Package, Rigaku and Rigaku/MSC.
4. SIR92: A. Altomare, G. Cascarano, C. Giacovazzo, A. Guagliardi, M. Burla, G. Polidori, and M. Camalli, *J. Appl. Cryst.*, **1994**, *27*, 435.
5. ORTEP III, L. J. Farrugia, *J. Appl. Cryst.*, **1997**, *22*, 389.

Chemocavity: Specific Concavity in Protein Reserved for the Binding of Biologically Functional Small Molecules

Shinji Soga,† Hiroki Shirai,† Masato Kobori,† and Noriaki Hirayama*,‡

Molecular Medicine Research Laboratories, Drug Discovery Research, Astellas Pharma Inc., 21 Miyukigaoka, Tsukuba, Ibaraki 305-8585, Japan, and Basic Medical Science and Molecular Medicine, Tokai University School of Medicine, 143 Shimokasuya, Isehara, Kanagawa 259-1143, Japan

Received April 2, 2008

The idea that there should be a specific site on a protein for a particular functional small molecule is widespread. It is, however, usually not so easy to understand what characteristics of the site determine the binding ability of the functional small molecule. We have focused on the concurrence rate of the 20 standard amino acids at such binding sites. In order to correlate the concurrence rate and the specific binding site, we have analyzed high-quality X-ray structures of complexes between proteins and small molecules. A novel index characterizing the binding site based on the concurrency rate has been introduced. Using this index we have identified that there is a specific concavity designated as a chemocavity where a specific group of small molecules, i.e., canonical molecular group, is highly inclined to be bound. This study has demonstrated that a chemocavity is reserved for a specific canonical molecular group, and the prevalent idea has been confirmed.

INTRODUCTION

Protein functions either through interactions with other biological macromolecules, mainly proteins, or small organic molecules such as amino acids, nucleic acids, sugars, and drugs. The interaction between a protein and a small molecule occurs mainly on a surface of the protein. The relevant site on the protein is usually a concavity. The accurate and efficient recognition between the protein and the small molecule is essential for the biological reaction. To realize such highly specific and efficient recognition, the characteristics of the binding site should be well formulated for the small molecule. In other words, a specific concavity must be reserved for the corresponding small molecule. Now we have many quality X-ray structures of complexes between biologically important proteins and small molecules at our disposal for detailed inspection. Not all the structures of biologically important proteins have been determined yet. However, thanks to the rapid growth of the number of protein structures in the Protein Data Bank (PDB)[1] in recent years, a sufficient number of proteins can be employed to extract a general principle regarding specific concavities closely related to biological functions.

The idea that there should be a specific site on a protein for a particular functional small molecule is widespread. Various studies have been undertaken for recognizing common functional sites in proteins. The nucleotide-binding site is one of the typical functional sites observed frequently in proteins. This site is rich in several specific amino acids and shares a common structural pattern.[2,3] As the rate at which three-dimensional structures of proteins are solved increases so rapidly, the number of protein structures without functional annotations is growing. As a result, methods for recognizing protein function from structure are more important now. Most methods currently available are reliant upon the three-dimensional arrangement of amino acids around the relevant site.[4,5] In addition, binding sites for only specific groups of small molecules have been taken into account. Shulman-Peleg *et al.*[4] focused on proteins that bind adenine, estradiol, ATP, and fatty acid. On the other hand, Zhang *et al.*[5] found the most highly connected binding sites for ten specific ligands. Although these studies have demonstrated that a specific site in a protein is recognized by a particular small organic molecule, they were confined to relatively small sets of proteins and ligands. Therefore it is highly required to expand these previous studies to generalize and confirm the prevalent idea that a specific site in a protein is reserved for a particular group of ligands. The shape and the properties of such specific sites in proteins can be determined by the amino acids flanking the cavity. Several studies have been undertaken in order to find the conserved surface structures in a set of similar proteins by translating the cavity-flanking amino acids into a set of pseudocenters.[6,7] Since these methods utilize spatially conserved physicochemical properties, they are relatively sensitive to the predefined interaction patterns and moreover computationally intensive.

Recently we have developed an index named propensity for ligand binding (PLB index) to identify the specific concavity in a protein for binding of druglike small molecules.[8,9] The PLB index is based on the characteristic appearances of the 20 standard amino acids at the concavity. Since this method is not strictly dependent on three-dimensional disposition of amino acids clustering at the concavity, it is computationally light. Despite its simplicity, it has proven very useful in identifying the specific concavity

* Corresponding author phone: +81 463 93 1121; e-mail: hirayama@is.icc.u-tokai.ac.jp.
† Astellas Pharma Inc.
‡ Tokai University School of Medicine.

10.1021/ci800113c CCC: $40.75
Published on Web 07/22/2008

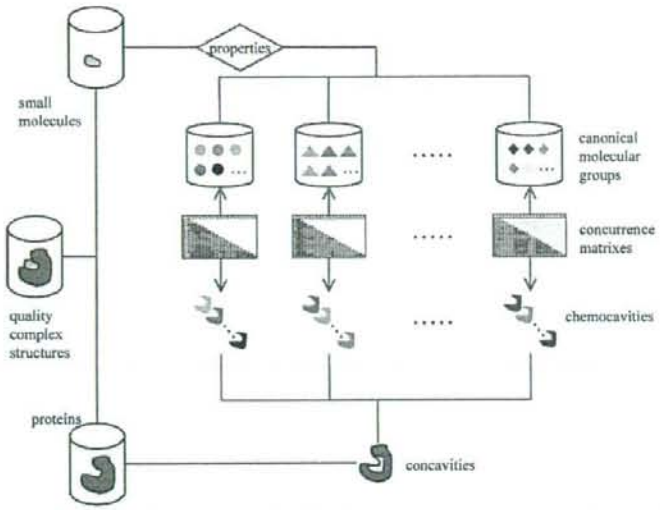

Figure 1. A schematic diagram of classification of canonical molecular groups and corresponding chemocavities. Concurrence matrixes link them.

on the surface of the protein. This concavity can be regarded as a reserved concavity for a promised small molecule. Hence, it is designated as chemocavity in this study. Chemocavity is defined as a concavity in a protein where a particular group of small molecules closely related to the function of the protein should be bound. This group of small molecules is expected to share common characteristics. In this paper, such a group is designated as a canonical molecular group. Suppose a chemocavity is an essential concavity for protein function, the corresponding canonical molecular group can bind the chemocavity so as to regulate the function of the protein. In this study, we have expanded the concept of the PLB index in order to make it sensitive enough to apply in identification of various chemocavities. In the expanded PLB index we introduced the concurrence rate of each of two of the 20 standard amino acids at the concavity. The expanded PLB index named chemocavity index has successfully identified different chemocavities corresponding to different canonical molecular groups. The present study unequivocally has shown that there is a specific chemocavity reserved in a protein structure for binding the corresponding canonical molecule which can control the function of the protein.

COMPUTATIONAL METHODS

The steps taken in this study are illustrated in Figure 1. The quality X-ray structures of complexes between proteins and small organic molecules were selected from PDB. The small organic molecules were classified according to their properties to obtain a set of canonical molecular groups. The proteins were then classified according to the canonical molecular groups that are bound to the proteins. The amino acids concurrence rates in the concavities that share the same canonical molecular group were analyzed. The amino acids concurrence rates can be expressed by concurrence matrixes. The characteristics of the concavities were expressed by the chemocavity indices based on the concurrence rates. The specific concavity for a particular canonical molecular group is designated as chemocavity. The correlations between various chemocavities were finally analyzed to check the independency of chemocavities and corresponding canonical molecular groups.

High Quality and Nonredundant X-ray Structures of Protein-Small Molecule Complexes. Reliable positions of non-hydrogen atoms in the concavities of proteins are highly required in order to accurately identify the amino acids at the concavities. A set of high quality X-ray structures of complexes between proteins and small molecules obtained from PDB on May 18, 2007 was used in this study. This quality data set was selected by use of the following criteria: the R_{free} value of less than or equal to 0.24, the resolution value of less than or equal to 2.5 Å, the occupancy factors of 1.0 for all non-hydrogen atoms of the small molecule, and the atomic displacement parameters of less than 50 Å^2 for all non-hydrogen atoms of the small molecule. The redundant proteins were removed from the data set by consulting the "Non-redundant PDB chain set (NRPDB)" resource [http://www.ncbi.nlm.nih.gov/Structure/VAST/nr-pdb.html] on May 18, 2007. A p-value of 10e-80 was used to judge sequence similarity by BLAST.[10] The homologous proteins above this threshold value were eliminated from the data set. Only the small organic molecules with molecular weight between 100 and 800 were taken into account. Finally, the 3621 complex structures including 4039 small organic molecules were selected and used as a data set for this study.

Selection of Canonical Molecular Groups. Canonical molecular groups were selected by clustering of the small organic molecules. Fingerprint and QuaSAR clustering methods implemented in the software system MOE[11] were used for this purpose. The fingerprint clustering algorithm based on the Jarvis-Patrick method[142] was applied using MACCS structural keys and Tanimoto coefficient. For QuaSAR clustering, molecular weight and SlogP[13] were used as molecular descriptors. By applying these two clustering methods, 4039 small organic molecules were brigaded to

1579 independent groups of small organic molecules. Since the groups with only a small number of member molecules are not appropriate in representing the canonical molecular groups, they are excluded from consideration in this study.

Chemocavity Index Considering Concurrence Rate of Two Standard Amino Acids at Concavity. The PLB index used in the previous studies is based on the frequency of the 20 standard amino acids at a concavity in order to judge its druggability. In this study, however, chemocavities corresponding to 48 canonical molecular groups must be distinguished. Therefore a more informative index is highly required. Gutteridge et al. have reported that combinations of different amino acid residues play important roles for catalysis at enzyme active site.[14] Therefore, based on the assumption that the concurrence rate of each of the two amino acids out of the 20 standard amino acids could augment the information, the concurrence rate of two amino acids was included in the expanded PLB index, now designated as the chemocavity index. The amino acids whose non-hydrogen atoms exist within 4.5 Å from the non-H atoms of the small molecule were used to determine the concurrence frequency of two amino acids at the concavity. Suppose the concurrence frequency of two amino acids x and y at the chemocavities corresponding to a canonical molecular group 'a' is $N_a(x,y)$, the concurrence frequency rate, $CA_a(x,y)$, is given as follows:

$$CA_a(x,y) = \frac{N_a(x,y)}{\sum_{l=1}^{20}\sum_{m=1}^{20} N_a(l,m)} \quad (1)$$

The denominator of the above equation means the total number of the concurrence frequency of every two amino acids at the concavities.

The incidence of every amino acid residue at all chemocavities in the whole data set was also calculated. The occurrence of a particular amino acid x in the chemocavities, $CA(x)$, can be defined as

$$CA(x) = \frac{N(x)}{\sum_{l=1}^{20} N(l)} \quad (2)$$

$N(x)$ denotes the number of amino acid x in the chemocavities. The denominator means the total number of all amino acids in the chemocavities. $CA(x)$ was determined using all structures in the data set.

If the occurrences of amino acids x and y are independent, $PA(x,y)$ defined in the following equation is an expected probability that the amino acid x and y appear concurrently.

$$PA(x,y) = W(x,y)CA(x)CA(y)$$
$$W(x,y) = \begin{cases} 1 (x = y) \\ 2 (x \neq y) \end{cases} \quad (3)$$

If the concurrence rate of the amino acids x and y at the chemocavities for the molecules belonging to canonical molecular group 'a' is more than expected, the following $RA_a(x,y)$ becomes greater than 1.

$$RA_a(x,y) = \frac{CA_a(x,y)}{PA(x,y)} \quad (4)$$

By use of a linear combination of RAs for all combinations of the two amino acids, an index expressing the tendency of a concavity to be a chemocavity 'a' is defined as follows. This chemocavity index means a propensity of a concavity i to be the chemocavity 'a'.

$$CC_{a,i} = \frac{\sum_{x=1}^{20}\sum_{y=1}^{20} N_i(x,y)RA_a(x,y)}{\sum_{x=1}^{20}\sum_{y=1}^{20} N_i(x,y)} \quad (5)$$

$N_i(x,y)$ denotes the concurrence frequency of the two amino acids x and y at a concavity i. If the chemocavity index for a concavity is high, the composition of amino acids at the concavity should be similar to that of the chemocavity. Hence, it is highly expected that the canonical molecules corresponding to this particular chemocavity should bind at this concavity.

Evaluation of Chemocavity Index. In order to assess the tendency of a concavity to be a particular chemocavity, the threshold value of $thCC_a$ was determined. The chemocavity indexes for concavities where small molecules belonging to a particular canonical molecular group were averaged to obtain the average positive CC_a index. On the other hand, chemocavity indices of concavities that do not bind the relevant canonical molecules were averaged to obtain the average negative CC_a index. The threshold value of $thCC_a$ was set to an average of these two values and used as a criterion to assess whether the concavity can be the chemocavity or not. To make the assessment simpler, 0 or 1 was assigned to each concavity as follows:

$$Z_{a,i} = \begin{cases} 1 (CC_{a,i} \geq th(CC_a)) \\ 0 (CC_{a,i} < th(CC_a)) \end{cases} \quad (6)$$

If the CC_a index is larger than or equal to the threshold value of CC_a, the concavity is regarded to be the chemocavity 'a'. Suppose the number of concavities in a set of concavities is N, the tendency of this set of concavities to be a particular chemocavity of 'a' is expressed as $(\sum_{i=1}^{N} Z_{a,i})/N$.

This value, designated as identification index, was used as an index to discriminate the different chemocavities belonging to different canonical molecular groups.

RESULTS AND DISCUSSION

Identification of 48 Canonical Molecular Groups. In this study, a set of the high-quality X-ray structures of complexes between proteins and small molecules was used as a source of information regarding the chemocavity and the canonical molecular group. By use of the initial set of 3621 complex structures, the 1579 independent groups of small molecules were selected based on the combination of chemical topology and physicochemical properties. For selection of molecular groups that well represent canonical molecular groups, only the groups consisting of more than nine molecules were included in the consideration. The total 48 independent groups were picked out and used in this study as canonical molecular groups. These 48 canonical groups represent functional small molecules that are highly inclined to bind their target proteins. These groups are given in Table 1. Typical biological small molecules, such as amino acids, nucleotides, and sugars, are well classified. Although buffer molecules, detergent molecules, and polyethylene glycols are not biological molecular groups, they are tightly bound at

Table 1. 48 Canonical Molecular Groups[a]

group #	molecular category	average molecular weight	average SlogP	# of compounds in the group
1	organic acid	148.07	−4.79	24
2	organic acid	189.32	−5.25	63
3	amino acid	145.18	−3.45	15
4	amino acid	208.01	−0.36	12
5	peptide	301.65	−5.59	21
6	amino sugar	207.17	−2.48	81
7	amino sugar	381.90	−5.19	11
8	amino sugar	410.40	−4.50	27
9	amino sugar	556.63	−6.44	10
10	sugar	150.13	−2.58	11
11	sugar	162.82	−2.14	12
12	sugar	169.16	−2.75	16
13	sugar	179.26	−3.18	58
14	sugar	194.18	−2.57	11
15	sugar	310.28	−4.34	10
16	sugar	341.17	−5.32	62
17	sugar	342.30	−5.40	20
18	sugar	501.19	−7.44	24
19	sugar	666.70	−9.77	17
20	sugar phosphate	225.90	−4.94	10
21	sugar phosphate	260.26	−5.48	14
22	nucleoside	267.25	−1.88	14
23	nucleoside	384.16	−3.75	42
24	nucleoside	397.39	−3.89	22
25	nucleotide	322.35	−4.71	11
26	nucleotide	343.98	−4.11	35
27	nucleotide	358.55	−6.68	35
28	nucleotide	401.49	−6.37	17
29	nucleotide	423.84	−5.77	72
30	nucleotide	441.15	−8.44	57
31	nucleotide	455.63	−3.53	55
32	nucleotide	470.83	−6.80	10
33	nucleotide	507.45	−7.29	75
34	nucleotide	518.67	−9.97	55
35	nucleotide	662.21	−6.27	99
36	nucleotide	733.99	−8.24	90
37	nucleotide	740.93	−11.27	11
38	nucleotide	759.50	−7.83	23
39	nucleotide	785.26	−5.48	89
40	thiamin	426.22	−2.73	13
41	pyridoxal phosphate	232.58	−1.16	52
42	porphyrin	563.46	1.64	198
43	higher fatty acid	265.14	4.56	10
44	buffer	195.58	−2.44	53
45	buffer	239.32	−4.69	32
46	detergent	282.55	−0.11	10
47	polyethylene glycol	150.17	−1.00	12
48	polyethylene glycol	207.56	−0.89	44

[a] The molecular group number corresponds to the chemocavity number.

specific concavities in proteins. Therefore they remain in the canonical molecular groups. It is possible that these molecules can be hints for some biological molecules. The average molecular weight and estimated octanol/water partition coefficient of SlogP of the member compounds in the 48 canonical groups are also given in Table 1. The list of all compounds included in the 48 canonical molecular groups is available in the Supporting Information. The nucleotides and sugars are the major groups in these 48 canonical groups. It is clearly due to the historical reason of protein crystallography. Many proteins with these groups of small molecules have been crystallized and their structures were determined.

Chemocavity. In the PLB index originally used for detecting the drug-binding site in the protein, only the frequency of each 20 standard amino acids was taken into account. The original PLB index was good enough for detecting the drug-binding site. In this study, however, we must characterize the concavities corresponding to the 48 canonical groups. Since the higher identification ability is required for such characterization, the expanded PLB index, termed chemocavity index, considering the concurrence rate of the two amino acids was formulated and applied in this

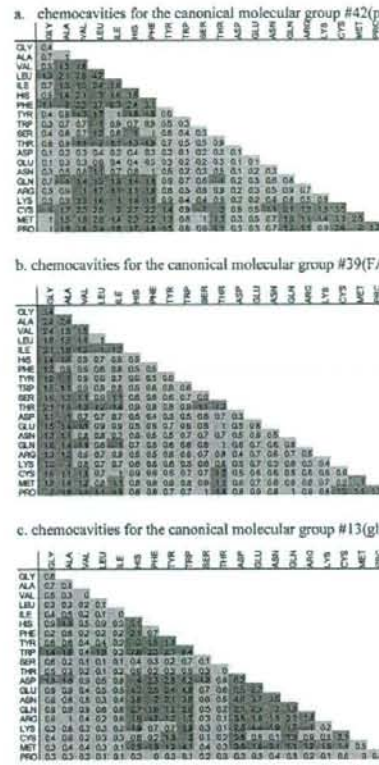

Figure 2. Concurrence matrixes of chemocavites for three canonical molecular groups including porphyrin, FAD, and glucose.

study. The concurrence rate of the two amino acids can be expressed by the concurrence matrix as shown in Figure 2. Each matrix element stands for a ratio of the concurrence frequency of a pair of two amino acids in a chemocavity to the frequency observed in other concavities on the surface of the proteins in the data set. To illustrate the characteristics of the concurrence matrixes of different chemocavities, the matrixes for three different chemocavities are compared in Figure 2. These chemocavities correspond to the canonical chemical groups of porphyrin, flavin adenine dinucleotide (FAD), and glucose. The concurrence frequencies greater than 1.0 are colored in red, and the concurrent pairs with the concurrence frequency being greater than 3.0 feature the most outstanding aspect of the chemocavity. The concurrence frequencies of Leu-Leu, Leu-Phe, Phe-Phe, Cys-Cys, and Cys-Met pairs are greater than 3.0 in the concavity of porphyrin. In the chemocavity of glucose, the concurrence frequencies of amino acids with charged and polar side chains are significantly high as expected. It is particularly interesting that the self-concurrence rate of Trp is exceptionally high. In addition, the concurrence rates of Trp with polar and charged amino acids such as Asn, Asp, Gln, Glu, Lys, and His are also markedly high. Although the self-concurrence rate of Gly is markedly high, it seems that the concurrence matrix of FAD is relatively featureless. As shown in this example, a concurrence matrix for a concavity on a protein where a specific small molecule is bound well represents the characteristics of each chemocavity.

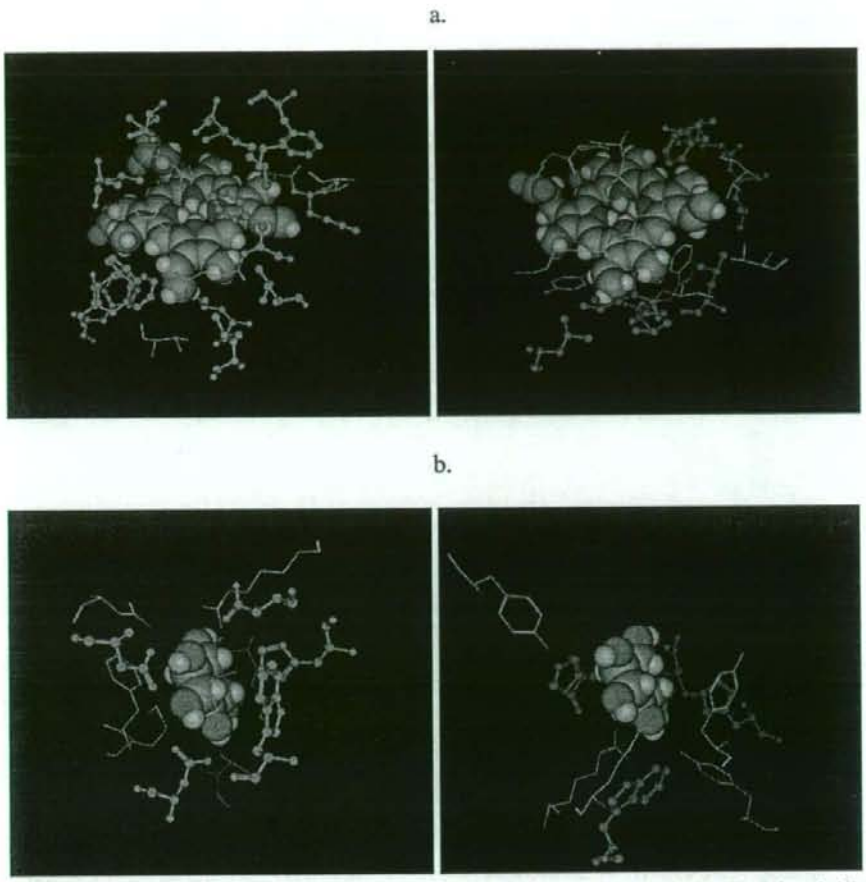

Figure 3. Concurrently existing amino acids at chemocavities of porphyrin and glucose. a. Porphyrin: The left and right structures are 1J3Y(2.57) and 1SOX(2.00), respectively. The values in the parentheses indicate the chemocavity indexes. Leu, Cys, Met, and Phe are depicted in red or blue. b. Glucose. The left and right structures are 2HPH(3.047) and 2BYO(2.759), respectively. The values in the parentheses indicate the chemocavity indexes. Trp, His, Gln, and Asp are depicted in red or blue.

It is particularly worth noting that the concurrence matrix does not directly reflect the three-dimensional structures of the concavities where small molecules are bound. Nevertheless, the concurrence matrix can characterize the chemocavity. This important feature of the concurrency of amino acids at the chemocavity is further illustrated in Figure 3. The chemocavities for porphyrin groups are rich in concurrence of Leu, Cys, Met, and Phe. In Figure 3a, these amino acids are colored in red and blue in the two different proteins. Although the three-dimensional arrangements of these amino acids around the porphyrin group are markedly different, the chemocavity indexes calculated from the concurrence rates of these amino acids clustered around the porphyrin group unequivocally suggest that these proteins share a common chemocavity for the porphyrin group. Another example of glucose is shown in Figure 3b. The chemocavity is rich in concurrence of Trp, His, Gln, and Asp. These amino acids shown in red and blue are clustered around glucose. The clustered amino acids adopt totally different three-dimensional arrangements. The chemocavity indexes, however, tell us that these concavities should belong to a chemocavity inherent to glucose.

Evaluation of Cross Interactions between Different Chemocavities. If each chemocavity is specific to a particular canonical molecular group, chemocavities corresponding to different canonical groups could be distinguished. We made an identification index based on chemocavity index as described in the method section. By use of this index, cross interactions between different chemocavities are evaluated. The results are shown in Figure 4. The chemocavity a is evaluated using the identification index based on the chemocavity b and vice versa. Therefore the (a,b) and (b,a) elements are not always symmetrical with respect to the diagonal line. The chemocavities are clustered according to the conventional classification of the corresponding canonical molecular groups. The main clusters are distinguished by different colors. The cross-element colored in red indicates the high identification index.

It is noteworthy that most of the diagonal elements are colored in red and off-diagonal elements are mostly white. Since the proteins that bind to small molecules belonging to a particular canonical molecular group are not homologous, the results clearly demonstrate that a chemocavity is highly reserved for the corresponding canonical molecular group.

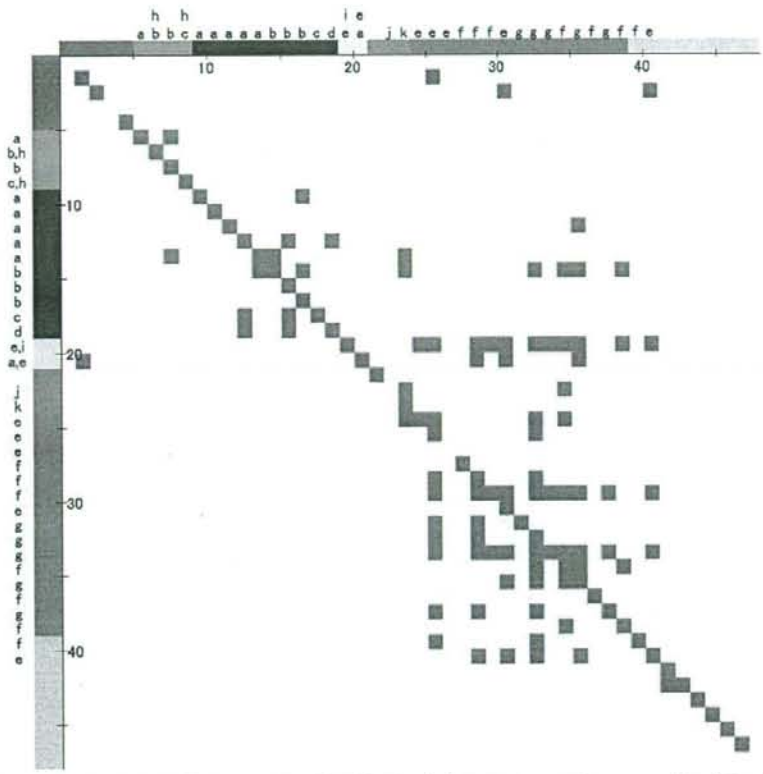

Figure 4. Chemocavity cross-interactions. Elements whose identification indexes are greater than or equal to 0.6 are indicated in red. The main clusters are distinguished by different colors. Color codes for the canonical molecular groups are as follows: orange, organic acid and amino acid (1 to 5); purple, amino sugar (6 to 9); blue, sugar (10 to 19); yellow, sugar phosphate (20 to 21); aqua, nucleoside (22 to 24); green, nucleotide (25 to 39); gray, others (40 to 48). Symbols *a* to *l* denote specific molecular groups included in each canonical molecular group: (*a*) monosaccharide, (*b*) disaccharide, (*c*) trisaccharide, (*d*) tetrasaccharide, (*e*) nucleotide monophosphate, (*f*) nucleotide diphosphate, (*g*) nucleotide triphosphate, (*h*) N-acetyl aminosugar, (*i*) open-chain form, (*j*) S-adenosyl-L-homocysteine, (*k*) S-adenosylmethionine, and (*l*) coenzyme.

In particular, the diagonal elements indicate the relevance of the chomocavity index. On the other hand, the off-diagonal elements are of interest to understand the independency of different chemocavities. Since most of the values of the off-diagonal elements are below the thresholds, major chemocavites are well discriminated to each other. Some cross-interactions are observed between different chemocavities for nucleotides. Significant cross-interactions are seen primarily in chemocavities corresponding to nucleotide triphosphates. Meanwhile, chemocavities for nucleotide monophosphates and diphosphates are well discriminated. The cross-interactions between nucleotide chemocavities and sugar phosphate chemocavities indicate that phosphate groups play important roles to gather specific amino acids into the chemocavities. On the contrary, the cross-interactions between chemocavities for nucleotides and sugars are marginal. Virtually no cross-interactions are observed between the nucleotide chemocavities and other chemocavities of non-nucleotides such as amino sugar, porphyrins, higher fatty acids, amino acids, peptides, and organic acids. Chemocavities regarding sugar groups also deserve special mention. The chemocavities corresponding to monosaccharide, disaccharide, and trisaccharide are well discriminated. It is interesting to note that the cross interactions of these chemocavities and those for amino sugar are not virtually observed.

In this study, we have demonstrated that a chemocavity is highly reserved for the corresponding canonical molecular group. By use of the identification index, we can differentiate a specific chemocavity from trivial concavities. Therefore, this method can be practically applicable to deduce a specific chemocavity in a protein whose ligand is still unknown. It may also be useful in finding a novel binding site for a particular small-molecule in a ligand-known protein. Since the concept of chemocavity is empirical-based, its accuracy and versatility solely depend on the completeness of the database of protein structures. With the increase of the number of proteins in PDB, the identification ability of this method will certainly improve.

CONCLUSIONS

We have introduced a chemocavity index based on the concurrence rate of the 20 standard amino acids at the concavity in proteins in order to identify a site in a protein inherently bound by a specific group of small organic molecules named as canonical molecular group. The chemocavity index has successfully correlated a particular

chemocavity to the corresponding canonical molecular group. But, we think it is still difficult to apply this method to any protein structures with no ligands, since accurate identification of small molecule binding sites are needed when searching against whole protein structures. That is to say, it is difficult to differentiate between small molecule interacting surface and noninteracting surface. The idea that there should be a specific site on a protein for a particular functional small molecule has been recognized widely. This study has confirmed that there should be a chemocavity in the protein to which a particular canonical molecule is highly inclined to be bound.

ACKNOWLEDGMENT

This work was partly supported by Grant-in-Aid for Scientific Research (A) (19200024) from MEXT (Ministry of Education, Culture, Sports, Science and Technology) for N.H.

Supporting Information Available: List of all compounds included in the 48 canonical molecular groups. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES AND NOTES

(1) Berman, H. M.; Westbrook, J.; Fenz, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The protein data bank. *Nucleic Acids. Res.* **2000**, *28*, 235–242.

(2) Walker, J. E.; Saraste, M.; Runswick, M. J.; Gay, N. J. Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J.* **1982**, *1*, 945–951.

(3) Brakoulias, A.; Jackson, R. M. Towards a structural classification of phosphate binding sites in protein-nucleotide complexes: an automated all-against-all structural comparison using geometric matching. *Proteins* **2004**, *56*, 250–260.

(4) Shulman-Peleg, A.; Nussinov, R.; Wolfson, H. J. Recognition of functional sites in protein structures. *J. Mol. Biol.* **2004**, *339*, 607–633.

(5) Zhang, Z.; Grigorov, M. G. Similarity networks of protein binding sites. *Proteins* **2006**, *62*, 470–478.

(6) Davies, J. R.; Jackson, R. M.; Mardia, K. V.; Taylor, C. C. The Poisson Index: a new probabilistic model for protein ligand binding site similarity. *Bioinformatics* **2007**, *23*, 3001–3008.

(7) Konc, J.; Janezic, D. Protein-protein binding-sites prediction by protein surface structure conservation. *J. Chem. Inf. Model.* **2007**, *47*, 940–944.

(8) Soga, S.; Shirai, H.; Kobori, M.; Hirayama, N. Use of amino acid composition to predict ligand-binding sites. *J. Chem. Inf. Model.* **2007**, *47*, 400–406.

(9) Soga, S.; Shirai, H.; Kobori, M.; Hirayama, N. Identification of the druggable concavity in homology models using the PLB index. *J. Chem. Inf. Model.* **2007**, *47*, 2287–2292.

(10) Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W.; Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **1990**, *215*, 403–410.

(11) MOE (Molecular Operating Environment), version 2006.0801; Chemical Computing Group Inc.: Montreal, Quebec, Canada, 2006.

(12) Jarvis, R. A.; Patrick, E. A. Clustering using a similarity measure based on shared near neighbors. *IEEE Trans. Comput.* **1973**, *C22*, 1025–1034.

(13) Wildman, S. A.; Crippen, G. M. Prediction of Physiochemical Parameters by Atomic Contributions. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 868–873.

(14) Gutteridge, A.; Thornton, J. M. Understanding nature's catalytic toolkit. *Trends Biochem. Sci.* **2005**, *30*, 622–629.

CI800113C

ANALYTICAL SCIENCES 2008, VOL. 24

X-ray Structure Analysis Online

Crystal Structure of Benzethonium Chloride Monohydrate

Rumiko TANAKA and Noriaki HIRAYAMA*

Basic Medical Science and Molecular Medicine, Tokai University School of Medicine, 143 Shimokasuya, Isehara, Kanagawa 259-1143, Japan

The crystal of the title compound, $C_{27}H_{42}ClNO_2 \cdot H_2O$, belongs to space group $P2_1/a$ with cell dimensions of $a = 9.37(6)$, $b = 13.23(7)$, $c = 22.5(2)$Å, and $\beta = 92.03(1)^\circ$. The final R value is 0.076. The molecule takes a folded conformation with the distance between the centroids of the two phenyl rings being 8.93 Å. The dihedral angle between the two phenyl rings is 48.3°.

(Received, April 7, 2008; Accepted June 23, 2008; Published on web August 1, 2008)

Benzethonium chloride, *N,N*-dimethyl-*N*-[2-[2-[4-(1,1,3,3-tetramethylbutyl)phenoxy]ethoxy]ethyl]benzenemethanaminium chloride, is a synthetic quaternary ammonium. The chemical structure is shown in Fig. 1. The compound has been used as antiseptic and anti-infective agents. In the food industry it is applied as a disinfectant and preservative. Yip *et al.* recently identified that benzethonium chloride is a novel cancer-specific compound by using a cell-based small-molecule screen.[1] An X-ray analysis of the title compound was undertaken to disclose its inherent three-dimensional structure, which is essentially important information for accelerating the development of novel anticancer agents based on this compound.

Benzethonium chloride was purchased from Prestwick Chemical Inc. Single crystals of the molecule were grown from an ethanol solution. A colorless platelet crystal with a size of $0.4 \times 0.3 \times 0.1$ mm was mounted on a glass fiber and used for data collection. The large standard deviations of the cell constants are obviously due to soft and highly fragile properties of the crystal. The structure was solved by direct methods, and non-H atoms were refined by a full-matrix least-squares method with anisotropic temperature factors. Three methyl groups attached to C22 are disordered at two positions with the occupancy factors being 0.7 and 0.3. A water molecule of crystallization was found during the refinement process. The positions of all H-atoms except those of the water molecule were calculated geometrically and fixed in the final refinement cycle. The H-atoms of the water molecule were located by a difference Fourier synthesis and fixed in the final refinement cycle. The crystal and experimental data are given in Table 1.

The atomic parameters of non-H atoms are given in Table 2.

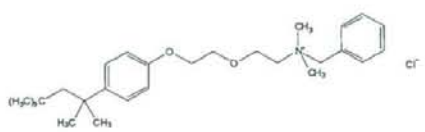

Fig. 1 Chemical structure of benzethonium chloride.

† To whom correspondence should be addressed.
E-mail address: hirayama@is.icc.u-tokai.ac.jp

The molecular structure, drawn by ORTEP-III[4] is shown in Fig. 2. In this figure, only the terminal methyl groups with an occupancy of 0.7 are depicted. Selected bond lengths, bond angles and torsion angles are given in Table 3. The bond lengths, angles and torsion angles in the molecule are within the expected ranges. The crystal structure of benzethonium nitrate has been reported recently.[5] In the following explanation, the geometrical parameters observed in benzethonium nitrate are shown in the square brackets. The dihedral angle between the two phenyl rings is 48.3(2)° [78.5(3)°]. The conformation of the chain moiety linking the two phenyl rings determines the characteristic shape of the molecule. The relevant torsion angles of C2-C1-N1-C10, C1-N1-C10-C11, N1-C10-C11-O1, C10-C11-O1-C12, C11-O1-C12-C13, O1-C12-C13-O2, and C12-C13-O2-C14 are –61.3(4)[–176.6(5)], –173.4(4)[175.7(5)], –98.2(6)[–69.4(6)], –171.1(4)[–73.1(5)], –151.5(4)[165.7(4)].

Table 1 Crystal and experimental data

Formula: $C_{27}H_{42}NO_2Cl \cdot H_2O$	
Formula weight = 466.10	
Crystal system: monoclinic	
Space group: $P2_1/a$	$Z = 4$
$a = 9.37(6)$Å	
$b = 13.23(7)$Å	$\beta = 92.0(3)^\circ$
$c = 22.5(1)$Å	
$V = 2790(28)$Å^3	
$D_x = 1.109$ g/cm^3	
No. of observations ($I > 2.00\sigma(I)$) = 2859	
$\theta_{max} = 68.15^\circ$ with Cu K_α	
R ($I > 2.00\sigma(I)$) = 0.076	
$(\Delta/\sigma)_{max} = 0.001$	
$(\Delta\rho)_{max} = 0.48$ e/Å^3	
$(\Delta\rho)_{min} = -0.50$ e/Å^3	
Measurement: Rigaku RAXIS-RAPID	
Program system: CrystalStructure 3.7.0[2]	
Structure determination: SIR92[3]	
Refinement: full-matrix	

CCDC 691260 contains the supplementary crystallographic data for this paper. These data can be obtained free of charge from The Cambridge Crystallographic Data Centre *via* www.ccdc.cam.ac.uk/data_request/cif.

Table 2 Atomic coordinates and equivalent isotropic thermal parameters (B_{eq})

atom	x	y	z	B_{eq}(Å^2)	occupancy
Cl(1)	0.1257(1)	0.12330(9)	0.40520(5)	6.75(3)	
O(1)	0.5716(4)	0.0897(3)	0.3063(2)	10.8(1)	
O(2)	0.3341(3)	0.0667(2)	0.2211(1)	6.38(8)	
O(3)	0.4344(3)	0.1961(3)	0.4560(1)	9.8(1)	
N(1)	0.7454(3)	-0.0213(2)	0.4177(1)	4.49(8)	
C(1)	0.8626(4)	-0.0998(3)	0.4290(2)	5.3(1)	
C(2)	0.8644(4)	-0.1857(3)	0.3855(2)	4.9(1)	
C(3)	0.7933(5)	-0.2759(4)	0.3963(2)	6.6(1)	
C(4)	0.7977(6)	-0.3554(5)	0.3557(4)	9.1(2)	
C(5)	0.8735(8)	-0.3452(6)	0.3059(4)	9.3(2)	
C(6)	0.9456(7)	-0.2575(6)	0.2945(2)	8.3(2)	
C(7)	0.9409(5)	-0.1784(4)	0.3342(2)	5.8(1)	
C(8)	0.7616(5)	0.0568(3)	0.4656(2)	6.3(1)	
C(9)	0.6028(4)	-0.0693(4)	0.4206(2)	6.7(1)	
C(10)	0.7662(5)	0.0259(3)	0.3578(2)	6.3(1)	
C(11)	0.6754(8)	0.1124(4)	0.3394(2)	9.9(2)	
C(12)	0.4905(5)	0.1711(4)	0.2795(2)	7.7(1)	
C(13)	0.3408(5)	0.1370(3)	0.2689(2)	6.2(1)	
C(14)	0.2008(5)	0.0335(3)	0.2031(2)	5.7(1)	
C(15)	0.0789(5)	0.0478(3)	0.2334(2)	5.9(1)	
C(16)	-0.0495(5)	0.0099(4)	0.2106(2)	6.6(1)	
C(17)	-0.0623(5)	-0.0428(4)	0.1584(2)	6.2(1)	
C(18)	0.0633(6)	-0.0576(4)	0.1294(2)	8.6(2)	
C(19)	0.1930(6)	-0.0214(4)	0.1513(2)	8.7(2)	
C(20)	-0.2052(6)	-0.0803(4)	0.1322(2)	7.5(2)	
C(21)	-0.266(1)	-0.0146(8)	0.0868(5)	21.9(5)	
C(22)	-0.2752(8)	0.0952(7)	0.0747(4)	10.4(3)	
C(23)	-0.3099(6)	-0.1005(6)	0.1795(3)	12.2(2)	
C(24)	-0.1873(9)	-0.1801(7)	0.1021(4)	20.8(4)	
C(25)	-0.351(2)	0.116(2)	0.0188(9)	20.6(8)	0.700
C(26)	-0.334(3)	0.140(2)	0.1271(8)	23(1)	0.700
C(27)	-0.126(1)	0.139(1)	0.0709(9)	15.0(5)	0.700
C(28)	-0.238(6)	0.085(5)	0.018(2)	21(1)	0.300
C(29)	-0.432(4)	0.141(4)	0.077(1)	26(1)	0.300
C(30)	-0.203(3)	0.167(2)	0.115(2)	11.4(9)	0.300

$B_{eq} = 8/3\pi^2(U_{11}(aa^*)^2 + U_{22}(bb^*)^2 + U_{33}(cc^*)^2 + 2U_{12}(aa^*bb^*)\cos\gamma + 2U_{13}(aa^*cc^*)\cos\beta + 2U_{23}(bb^*cc^*)\cos\alpha)$.

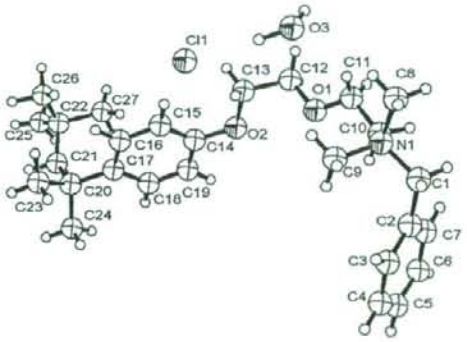

Fig. 2 Molecular structure of benzethonium chloride along with the labeling atoms. Thermal ellipsoids of non-H atoms are drawn at the 40% probability level.

–72.2(5)[77.3(5)] and –175.7(3)[80.7(6)]°, respectively. The distance between the centroids of the two phenyl rings is 8.93[11.45]Å. These geometrical parameters indicate that the benzethonium ion in the chloride salt takes a more folded conformation than that in the nitrate salt. In the chloride salt, the water molecule and the chloride ion are connected by the two hydrogen bonds, forming an infinite chain along the *a*-axis in the crystal structure. The geometrical parameters of these hydrogen bonds are as follows: O3···Cl1 3.22(2), O3-H43 1.03, H43···Cl1 2.21 Å, ∠O3-H43···Cl1 164°; O3···Cl1 (1/2+*x*, 1/2–*y*, *z*) 3.22(2), O3-H44 0.96, H44···Cl1 2.29 Å, ∠O3-H44···Cl1 164°. The symmetry code of the hydrogen acceptor of the second hydrogen bond is given in parentheses. There are no hydrogen bonds between the benzethonium ion and the water molecule.

Table 3 Selected bond lengths (Å), bond angles (°) and torsion angles (°)

O(1)—C(11)	1.242(8)	O(1)—C(12)	1.438(6)
O(2)—C(13)	1.422(5)	O(2)—C(14)	1.372(5)
N(1)—C(1)	1.526(5)	N(1)—C(8)	1.498(5)
N(1)—C(9)	1.483(5)	N(1)—C(10)	1.504(5)
C(11)—O(1)—C(12)	117.5(4)	C(13)—O(2)—C(14)	116.6(3)
C(1)—N(1)—C(8)	107.1(3)	C(1)—N(1)—C(9)	110.2(3)
C(1)—N(1)—C(10)	108.4(3)	C(8)—N(1)—C(9)	109.4(3)
C(8)—N(1)—C(10)	110.2(3)	C(9)—N(1)—C(10)	111.4(3)
C(2)—C(1)—N(1)	115.5(3)	C(11)—C(10)—N(1)	118.8(4)
O(1)—C(11)—C(10)	114.4(4)	C(13)—C(12)—O(1)	108.8(4)
O(2)—C(13)—C(12)	109.7(4)	C(15)—C(14)—O(2)	125.3(4)
C(19)—C(14)—O(2)	116.1(4)		
C(12)—O(1)—C(11)—C(10)	-171.1(4)	C(11)—O(1)—C(12)—C(13)	-151.5(4)
C(14)—O(2)—C(13)—C(12)	-175.7(3)	C(13)—O(2)—C(14)—C(15)	-12.2(6)
C(13)—O(2)—C(14)—C(19)	170.1(4)	C(8)—N(1)—C(1)—C(2)	179.7(3)
C(9)—N(1)—C(1)—C(2)	60.8(4)	C(10)—N(1)—C(1)—C(2)	-61.3(4)
C(1)—N(1)—C(10)—C(11)	-173.4(4)	C(8)—N(1)—C(10)—C(11)	-56.4(5)
C(9)—N(1)—C(10)—C(11)	65.2(5)	N(1)—C(1)—C(2)—C(3)	-93.1(5)
N(1)—C(1)—C(2)—C(7)	88.9(5)	N(1)—C(10)—C(11)—O(1)	-98.2(6)
O(1)—C(12)—C(13)—O(2)	-72.2(5)	O(2)—C(14)—C(15)—C(16)	-179.7(4)
O(2)—C(14)—C(19)—C(18)	-179.5(5)		

Acknowledgements

This work was supported by Grants for the Key Technology Research Promotion Program of New Energy and Industrial Technology Development Organization (NEDO) of Japan, and also by the Research and Study Program of Tokai University Educational System General Research Organization.

References

1. K. W. Yip *et al.*, *Clin. Cancer Res.*, **2006**, *12*, 5557.
2. CrystalStructure, version 3.5.1, **2000 – 2003** Crystal Structure Analysis Package, Rigaku and Rigaku/MSC.
3. SIR92: A. Altomare, G. Cascarano, C. Giacovazzo, A. Guagliardi, M. Burla, G. Polidori, and M. Camalli, *J. Appl. Cryst.*, **1994**, *27*, 435.
4. ORTEP III, L. J. Farrugia, *J. Appl. Cryst.*, **1997**, *22*, 389.
5. J. Pernak (list all authors, pls.), *Green Chem.*, **2006**, *8*, 798.

ANALYTICAL SCIENCES 2008, VOL. 24

Crystal Structure of Linezolid

Rumiko TANAKA and Noriaki HIRAYAMA†

Basic Medical Science and Molecular Medicine, Tokai University School of Medicine, 143 Shimokasuya, Isehara, Kanagawa 259-1143, Japan

The crystal of the title compound, $C_{16}H_{20}FN_3O_4$, belongs to space group $P1$ with cell dimensions of a = 6.5885(9), b = 10.977(1), c = 12.919(1)Å, α = 69.315(9), β = 88.17(1), and γ = 74.23(1)°. The final R value is 0.050. The key structural feature of the molecule, *N*-aryloxazolidinone, is not planar. The dihedral angles between the least-squares planes of the oxazolidine and phenyl rings are 52.7(5) and 39.0(6)° in two crystallographically independent molecules.

(Received November 9, 2007; Accepted January 28, 2008; Published on web March 19, 2008)

Linezolid, *N*-[[(S)-3-(3-fluoro-4-morpholinophenyl)-2-oxo-5-oxazolidinyl]methyl] acetamide, is a synthetic antibiotic belonging to a unique class of antimicrobials, called oxazolidinones. Linezolid disrupts bacterial growth by inhibiting the initiation process in protein synthesis.[1] The chemical structure is shown in Fig. 1. Linezolid has been approved primarily to fight resistant gram-positive cocci, such as vancomycin-resistant enterococcus, methicillin-resistant *Staphylococcus aureus*, and penicillin-resistant pneumococci. Although Linezolid has been successfully used to treat a number of resistant strains of bacteria, the emergence of linezolid resistance had been already reported. An X-ray analysis of the title compound was undertaken to disclose its inherent three-dimensional structure, which is essentially important information in designing the next generation of linezolid to fight resistant bacteria.

Linezolid is a product of Pharmacia. Single crystals of the molecule were grown from an acetonitrile solution. A colorless platelet crystal with a size of 0.5 × 0.3 × 0.1 mm was mounted on a glass fiber and used for data collection. The structure was solved by direct methods and non-H atoms were refined by a full-matrix least squares method with anisotropic temperature factors. The positions of H-atoms bonded to the amide nitrogen atoms were located from a difference Fourier synthesis and refined isotropically. Other H-atoms were calculated geometrically and refined by the riding model. The positions of all H-atoms were fixed in the last refinement cycle. The absolute configuration was assigned based on the known configuration of the asymmetric carbon in the molecule. The non-zero maximum shift/error value in the last refinement cycle is due to the two crystallographically independent molecules related by a pseudo-symmetry. The crystal and experimental data are given in Table 1. Atomic parameters of non-H atoms are given in Table 2.

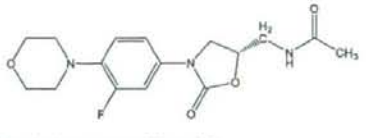

Fig. 1 Chemical structure of linezolid.

The molecular structure, drawn by ORTEP-III,[4] is shown in Fig. 2. There are two crystallographically independent molecules in the crystal. The molecules whose labels are 100s and 200s are designated as molecules I and II, respectively. The two independent molecules adopt significantly different conformations. The oxazolidinone rings take a twist form on C105-C106 and an envelope form on C206 in molecules I and II, respectively. These oxazolidinone rings are twisted with respect to the phenyl rings. The magnitudes of the twisting are different in the two molecules. The torsion angles around the interconnecting N-C bonds are significantly different. C104-N102-C107-C108 and C105-N102-C107-C112 in molecule I are 21(1) and 8(1)°, respectively. The corresponding angles in molecule II of C204-N202-C207-C208 and

Table 1 Crystal and experimental data

Formula: $C_{16}H_{20}FN_3O_4$	
Formula weight = 337.35	
Crystal system: triclinic	
Space group: $P1$	$Z = 2$
a = 6.5885(9)Å	α = 69.315(9)°
b = 10.977(1)A	β = 88.17(1)°
c = 12.919(1)Å	γ = 74.23(1)°
V = 839.0(2)Å^3	
D_x = 1.335 g/cm^3	
No. of observations (I>2.00σ(I)) = 1760	
θ_{max} = 68.2° with Cu $K\alpha$	
R (I>2.00σ(I)) = 0.050	
$(\Delta/\sigma)_{max}$ = 0.142	
$(\Delta\rho)_{max}$ = 0.20 e/Å^3	
$(\Delta\rho)_{min}$ = –0.23 e/Å^3	
Measurement: Rigaku RAXIS-RAPID	
Program system: CrystalStructure 3.7.0[2]	
Structure determination: SIR92[3]	
Refinement: full-matrix	

CCDC 672482 contains the supplementary crystallographic data for this paper. These data can be obtained free of charge from The Cambridge Crystallographic Data Centre *via* www.ccdc.cam.ac.uk/data_request/cif.

† To whom correspondence should be addressed.
E-mail address: hirayama@is.icc.u-tokai.ac.jp

Table 2 Atomic coordinates and equivalent isotropic thermal parameters (B_{eq})

Atom	x	y	z	B_{eq}(Å^2)
Molecule I				
F(101)	-0.500(1)	0.2503(7)	0.3206(6)	11.6(2)
O(101)	0.9656(9)	-0.0736(6)	0.8020(5)	7.2(2)
O(102)	0.286(1)	0.1710(6)	0.6471(5)	6.4(2)
O(103)	-0.039(1)	0.2847(7)	0.5697(6)	9.4(2)
O(104)	-0.682(1)	-0.0695(8)	0.1518(6)	8.9(2)
N(101)	0.635(1)	0.0453(7)	0.8143(6)	6.2(2)
N(102)	0.1309(9)	0.0826(7)	0.5545(5)	5.0(2)
N(103)	-0.480(1)	-0.0013(7)	0.3095(6)	6.0(2)
C(101)	0.844(2)	0.0176(9)	0.8272(7)	5.5(2)
C(102)	0.927(2)	0.099(1)	0.8778(9)	7.0(2)
C(103)	0.532(1)	-0.0285(8)	0.7689(7)	6.8(2)
C(104)	0.111(2)	0.1876(9)	0.5880(7)	5.7(2)
C(105)	0.328(1)	-0.025(1)	0.6039(8)	6.0(2)
C(106)	0.445(1)	0.0493(8)	0.6523(7)	5.8(2)
C(107)	-0.025(1)	0.0608(9)	0.4944(6)	4.6(2)
C(108)	-0.195(2)	0.167(1)	0.4326(8)	6.7(2)
C(109)	-0.333(2)	0.1413(9)	0.3730(8)	6.2(2)
C(110)	-0.322(2)	0.015(1)	0.3714(7)	5.4(2)
C(111)	-0.156(2)	-0.0875(9)	0.4330(8)	5.9(2)
C(112)	-0.007(1)	-0.0671(9)	0.4938(7)	5.4(2)
C(113)	-0.685(2)	0.061(1)	0.134(1)	10.8(4)
C(114)	-0.499(2)	0.075(1)	0.189(1)	9.7(3)
C(115)	-0.478(2)	-0.141(1)	0.3271(8)	6.8(2)
C(116)	-0.664(2)	-0.144(1)	0.2664(9)	8.3(3)
Molecule II				
F(201)	0.865(1)	0.3301(8)	1.0953(7)	14.2(2)
O(201)	-0.6235(9)	0.6535(6)	0.5936(5)	6.8(2)
O(202)	0.1034(9)	0.4124(6)	0.7529(5)	6.8(2)
O(203)	0.388(1)	0.2794(7)	0.8657(7)	9.2(2)
O(204)	1.032(1)	0.6428(9)	1.2709(6)	8.9(2)
N(201)	-0.302(1)	0.5088(7)	0.6349(6)	6.3(2)
N(202)	0.239(1)	0.5029(7)	0.8519(5)	4.8(2)
N(203)	0.838(1)	0.5742(9)	1.1154(7)	6.6(2)
C(201)	-0.496(1)	0.5568(9)	0.5836(7)	5.1(2)
C(202)	-0.544(1)	0.482(1)	0.5147(8)	6.9(2)
C(203)	-0.242(1)	0.567(1)	0.7087(8)	7.2(2)
C(204)	0.258(2)	0.389(1)	0.8287(8)	6.3(2)
C(205)	0.062(1)	0.6094(9)	0.7911(8)	5.8(2)
C(206)	-0.010(1)	0.5535(7)	0.7128(7)	5.4(2)
C(207)	0.384(1)	0.5168(8)	0.9225(7)	5.0(2)
C(208)	0.551(2)	0.4128(9)	0.9825(8)	6.5(2)
C(209)	0.698(2)	0.434(1)	1.0415(9)	7.8(3)
C(210)	0.681(1)	0.5565(9)	1.0545(7)	5.2(2)
C(211)	0.506(1)	0.6605(9)	0.9953(7)	5.1(2)
C(212)	0.367(1)	0.6406(8)	0.9296(7)	5.3(2)
C(213)	1.052(2)	0.501(2)	1.288(1)	12.1(5)
C(214)	0.861(2)	0.493(1)	1.233(1)	9.7(4)
C(215)	0.836(2)	0.712(1)	1.0953(8)	7.2(2)
C(216)	1.023(2)	0.719(2)	1.155(1)	9.4(4)

$B_{eq} = 8/3\,\pi^2 U_{11}(aa^*)^2 + U_{22}(bb^*)^2 + U_{33}(cc^*)^2 + 2U_{12}(aa^*bb^*)\cos\gamma + 2U_{13}(aa^*cc^*)\cos\beta + 2U_{23}(bb^*cc^*)\cos\alpha)$.

C205-N202-C207-C212 are –6(1) and –5(1)°, respectively. Morpholine rings in both molecules take chair conformations. Morpholine rings are twisted with respect to the phenyl rings around the interconnecting N-C bond. The conformations around the N-C bonds are significantly different in molecules I and II. C115-N103-C110-C111 and C114-N103-C110-C109 angles in molecule I are 9(1) and 61(1)°, respectively. They are

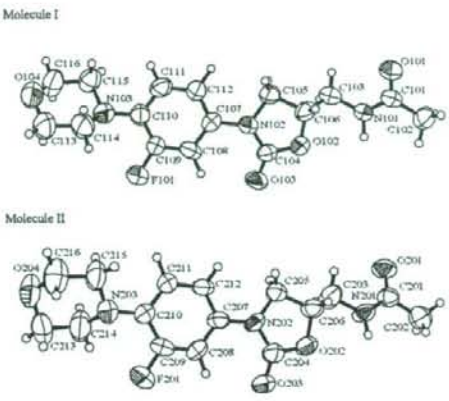

Fig. 2 Molecular structure of linezolid along with the labeling atoms. Thermal ellipsoids of non-H atoms are drawn at the 40% probability level.

significantly different from the corresponding values of –13(1) and –65(1)°, respectively, in molecule II. The conformations of terminal methyl acetamide chains are also significantly different, as represented by the torsion angles of C101-N101-C103-C106 and C201-N201-C203-C206 being –104.1(8) and –151.5(7)°, respectively. The two crystallographically independent molecules are connected by intermolecular hydrogen bonds, and form a dimer in the crystal. The geometry of the hydrogen bonds is as follows: N101-H···O203[N101···O203 2.93(1), N101-H 0.98, H···O203 2.04 Å, ∠N101-H···O203 149°] and N201-H···O103 [N201···O103 2.97(1), N201-H 0.92 H···O103 2.07 Å, ∠N201-H···O103 167°].

Acknowledgements

This work was supported by Grants for the Key Technology Research Promotion Program of New Energy and Industrial Technology Development Organization (NEDO) of Japan, and also by the Research and Study Program of Tokai University Educational System General Research Organization.

References

1. S. M. Swaney, H. Aoki, M. C. Ganoza, and D. L. Shinabarger, *Antimicrob. Agents Chemother.*, **1998**, *42*, 3251.
2. CrystalStructure, version 3.5.1, **2000-2003** Crystal Structure Analysis Package, Rigaku and Rigaku/MSC.
3. SIR92: A. Altomare, G. Cascarano, C. Giacovazzo, A. Guagliardi, M. Burla, G. Polidori, and M. Camalli, *J. Appl. Cryst.*, **1994**, *27*, 435.
4. ORTEP III, L. J. Farrugia, *J. Appl. Cryst.*, **1997**, *22*, 389.

ANALYTICAL SCIENCES 2008, VOL. 24

Crystal Structure of Guanabenz Acetate

Rumiko TANAKA and Noriaki HIRAYAMA†

Basic Medical Science and Molecular Medicine, Tokai University School of Medicine, 143 Shimokasuya, Isehara, Kanagawa 259-1143, Japan

The crystal of the title antihypertension agent, $C_8H_9Cl_2N_4{\cdot}C_2H_3O_2$, belongs to space group *Pbca* with the cell dimensions a = 13.389(5), b = 11.007(6), and c = 17.342(5)Å. The final R value is 0.060. The planar hydrazinecarboximidamide moiety is slightly inclined to the phenyl group. The acetate ion plays an important role to connect guanabenz molecules together by hydrogen bonds in the crystal.

(Received September 28, 2007; Accepted December 10, 2007; Published on web January 24, 2008)

Guanabenz acetate (2-[(2,6-dichlorophenyl)methylene]-hydrazinecarboximidamide acetate) is an antihypertensive whose mechanism of action appears to involve both central alpha adrenergic stimulation leading to a suppression of the sympathetic nervous system activity from bulbar vasoconstriction centers as well as a peripheral adrenergic neuron blockade. This drug is used in the treatment of high blood pressure.[1] It is effectively used alone or in combination with a thiazide type of diuretic. Guanabenz begins to lower the blood pressure within 60 min after taking a single dose, and may slow the pulse rate slightly. The chemical structure is shown in Fig. 1. An X-ray analysis of the title compound was undertaken to disclose its inherent three-dimensional structure.

Guanabenz acetate was purchased from Sigma-Aldrich Co. Single crystals of the molecule were grown from an ethanol solution. A colorless needle crystal with a size of 0.3 × 0.01 × 0.01 mm was mounted on a glass fiber and used for data collection. The structure was solved by direct methods, and non-H atoms were refined by the full-matrix least-squares method with anisotropic temperature factors. The positions of all H-atoms were calculated geometrically, refined by the riding model, and fixed in the final refinement cycle. The crystal and experimental data are given in Table 1. The atomic parameters of non-H atoms are given in Table 2.

The molecular structure drawn by ORTEP-III[4] is shown in Fig. 2. Selected bond lengths, bond angles and torsion angles are given in Table 3. Although the overall structure of the molecule is very planar, the planar moiety, consisting of N1, N2, N3, N4 and C1, is twisted with respect to the phenyl ring with the dihedral angle between the least-squares planes being 9.3(6)°. The bond lengths were within the expected ranges. The exocyclic bond angles around C3 and C4 are highly asymmetric. The bond angles of C3-C4-Cl2 and C2-C3-C4 are significantly larger than the corresponding counterparts. This asymmetry is presumably due to intramolecular contact between N4 and Cl2 and the intramolecular hydrogen bond between C2-H and Cl1. The N4···Cl2 distance of 2.89(1)Å is appreciably shorter than the sum of the van der Waals radii. The geometry of the intramolecular hydrogen bond is as follows: C2···Cl1 2.98(1), C2-H1 1.10, H1···Cl1 2.35 Å, ∠C2-H1···Cl1 114°. N1, N2, N3, O1 and O2 form an extensive intermolecular hydrogen network in the crystal. The O1 and O2 atoms are involved in the bifurcated hydrogen bonds, and an acetate ion connect three guanabenz molecules by the hydrogen bonds. The geometry of the intermolecular hydrogen bonds are as follows: N1···O2(3/2–x, –1/2+y, z) 2.73(1), N1-H5 0.95, H5···O2 1.79 Å, ∠N1-H5···O2 172°; N1···O1(2–x, –1/2+y, 3/2–z) 3.07(1), N1-H6

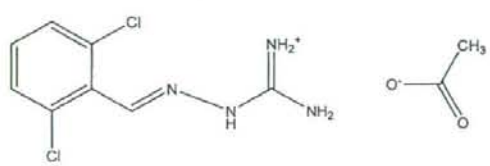

Fig. 1 Chemical structure of guanabenz acetate.

Table 1 Crystal and experimental data

Formula: $C_8H_9Cl_2N_4{\cdot}C_2H_3O_2$
Formula weight = 291.14
Crystal system: orthorhombic
Space group: *Pbca* $Z = 8$
a = 13.389(5)Å
b = 11.007(6)Å
c = 17.342(5)Å
V = 2555(1)Å^3
D_x = 1.513 g/cm^3
No. of observations ($I > 1.80\sigma(I)$) = 652
θ_{max} = 68.25° with Cu K_α
R ($I > 2.00\sigma(I)$) = 0.060
$(\Delta/\sigma)_{max}$ = 0.000
$(\Delta\rho)_{max}$ = 0.41 e/Å^3
$(\Delta\rho)_{min}$ = –0.44 e/Å^3
Measurement: Rigaku RAXIS-RAPID
Program system: CrystalStructure 3.6.0[2]
Structure determination: SIR92[3]
Refinement: full-matrix

CCDC 668117 contains the supplementary crystallographic data for this paper. These data can be obtained free of charge from The Cambridge Crystallographic Data Centre *via* www.ccdc.cam.ac.uk/data_request/cif.

† To whom correspondence should be addressed.
E-mail: hirayama@is.icc.u-tokai.ac.jp