in 72% yield from (+)-**5**. The amino group of hydroxylamine in **8** was protected with a carbobenzoxy (Cbz) group using Cbz chloride in the presence of potassium carbonate (83%), then the two N–O bonds in **9** were reduced with $TiCl_3$–NaOAc and zinc powder in aqueous MeOH–HCl to give diamine **10**. Next, a *tert*-butoxycarbonyl (Boc)-protected guanidine group was introduced with bis(Boc)-2-methyl-2-thiopseudourea (**16**) in the presence of mercury(II) chloride and triethyl amine to give guanidine **11** in 82% yield from **8**. Cyclization of guanidine in **11** was carried out by treatment with monochloromethanesulfonyl chloride and diisopropylethylamine to give **12** in 87% yield, and this product was subsequently converted into **13** in three steps: 1) deprotection of the triisopropylsilyl (TIPS) group with tetra-*n*-butylammonium fluoride (TBAF) in tetrahydrofuran (THF) at 0°C, 2) deprotection of the Cbz group with hydrogen in the presence of Pearlman's catalyst,[17] and 3) guanylation with bis(Cbz)-2-methyl-2-thiopseudourea (**17**) in the presence of mercury(II) chloride in 79%.[18] One of the Cbz groups of **13** was removed with sodium methoxide in THF/methanol to give **14** in 65% yield.[19] Construction of the tricyclic fused-type guanidine was carried out by oxidation of alcohol **14** with IBX (4 equiv) in DMSO at 70°C to give **15** in 24% yield. In this reaction, oxidation of the alcohol at C12 takes place first, then the C4 position is oxidized through the enamine–IBX complex.[15] Finally, (−)-FD-dcSTX (**4b**) was obtained by deprotection of the Cbz and Boc groups on guanidines by hydrogenolysis in the presence of Pearlman's catalyst, followed by treatment with 3N HCl in 78% yield from **15** (i.e., two steps).

The synthesis of (−)-FD-STX (**4a**) is illustrated in Scheme 5. Synthesis of (−)-FD-STX (**4a**) was attempted firstly from the tricyclic compound **15** by deprotection of the *p*-methoxybenzyl (MPM) group with 2,3-dichloro-5,6-dicyano-1,4-benzoquinone (DDQ). However, the desired reaction did not occur, and decomposition of the starting material was observed. The difficulty of this deprotection reaction presumably arises from steric effects at C13. Thus, functionalization at C13 was carried out before construction of the tricyclic FD-STX ring system. After protection of the hydroxyl group of bis-guanidine **13** at C12 with acetic anhydride, the MPM group in the resulting compound was successfully removed with DDQ to give alcohol **19** in 90% yield (two steps). A carbamoyl group was introduced at C13 by treating the alcohol **19** with trichloroacetyl isocyanate, followed by hydrolysis with potassium carbonate in MeOH, thereby affording carbamate **20** in 34% yield. Then, the fused-type tricyclic skeleton was constructed under IBX oxidation conditions to give **21** in 20% yield. Finally, deprotection of the Cbz and Boc groups was carried out under the same conditions as described for the synthesis of **4b** to give (−)-FD-STX (**4a**) in 82% yield (two steps).

(−)-FD-doSTX (**4c**) was synthesized by deprotection of fused-type tricyclic guanidine (−)-**3**, which was obtained from methyl crotonate (**22**) and the chiral nitrone (+)-**5** as previously reported (Scheme 5).[15a,b]

Evaluation of Na_VCh-inhibitory activity of (−)-FD-STXs 4a–c: Next, the Na_VCh-inhibitory activity of **4a–c** was evaluated in terms of inhibition of the cytotoxicity of ouabain and veratridine in mouse neuroblastoma cell line Neuro-2a.[20] In this cell-culture assay, Neuro-2a is treated with a sodium channel activator, veratridine, in the presence of ouabain, an inhibitor of Na^+/K^+ ATPase. This serves to block sodium ion efflux, and decreases the cell viability. Na_VCh inhibition by TTX, STX, and related compounds antagonizes this effect, and rescues the cells in a dose-dependent manner. The results obtained with **4a–c** are summarized in Table 1. (−)-FD-STX (**4a**) and (−)-FD-dcSTX (**4b**) showed a concentration-dependent inhibitory effect on the cytotoxicity of ouabain and veratridine, and their IC_{50} values were determined to be (7.3±2.7) and (15±5.4) μM, respectively (synthetic (+)-dcSTX (**1b**),[15c] used as a control, showed an IC_{50} value of (82±17) nM). On the other hand, (−)-FD-doSTX (**4c**) showed no inhibitory activity even at concentrations over 500 μM. Since (−)-FD-STX (**4a**) and (−)-FD-dcSTX (**4b**) were inhibitory, these compounds were further evaluated by means of the patch-clamp method with cells that express $Na_V1.4$ and $Na_V1.5$, which are representative members of the TTX-s and TTX-r channel subtypes, respectively.

Table 1. Na_VCh-inhibitory activity of FD-STXs **4a–c** in Neuro-2a cells.[21]

Compounds	IC_{50} (mean ±SD)	*n*
(−)-FD-STX (**4a**)	(7.3±2.7) μM	3
(−)-FD-dcSTX (**4b**)	(15±5.4) μM	3
(−)-FD-doSTX (**4c**)	>500 μM	3
(+)-dc-STX (**1b**)	(82±17) nM	3

The sodium ion channel α subunits of the rat skeletal muscle channel $rNa_V1.4$ and rat heart channel $rNa_V1.5$ were each transiently transfected into HEK293 cells, and electrophysiological recording was performed with a whole-cell patch clamp. In this experiment, the sodium ion channels were activated by applying 10-millisecond voltage steps from a holding potential of −120 mV every 5 s, which induces peak sodium currents (I_{Na}).[22,23] Figure 1 shows time courses of changes in I_{Na}, the effects of inhibitors **4a** and **4b**, and the recovery upon washout of these compounds. In the case of $Na_V1.4$, sodium ion current was decreased by the addition of 10 μM (−)-FD-STX (**4a**) or (−)-FD-dcSTX (**4b**), and reached an equilibrium peak current ($I_{inhibitor}$ **4a** or **4b**). Recovery of sodium ion current was observed upon washout of the inhibitor, which indicates that the ligands were removed from the channel (Figure 1a and b). This is a typical reversible inhibition, like that of TTX and STX.[24] The IC_{50} values of **4a** and **4b** were (3.8±0.9) and (16±8.6) μM, respectively (Table 2),[25] and paralleled the results of the cell-based assay. In the case of $Na_V1.5$, the sodium ion current was decreased by the addition of 1 mM **4a** or **4b**, but recovery of the current following washout was observed only in the case of (−)-FD-STX (**4a**) (Figure 1c); the IC_{50} value of **4a** was (118±8.1) μM (Table 2).[25] On the other hand, recov-

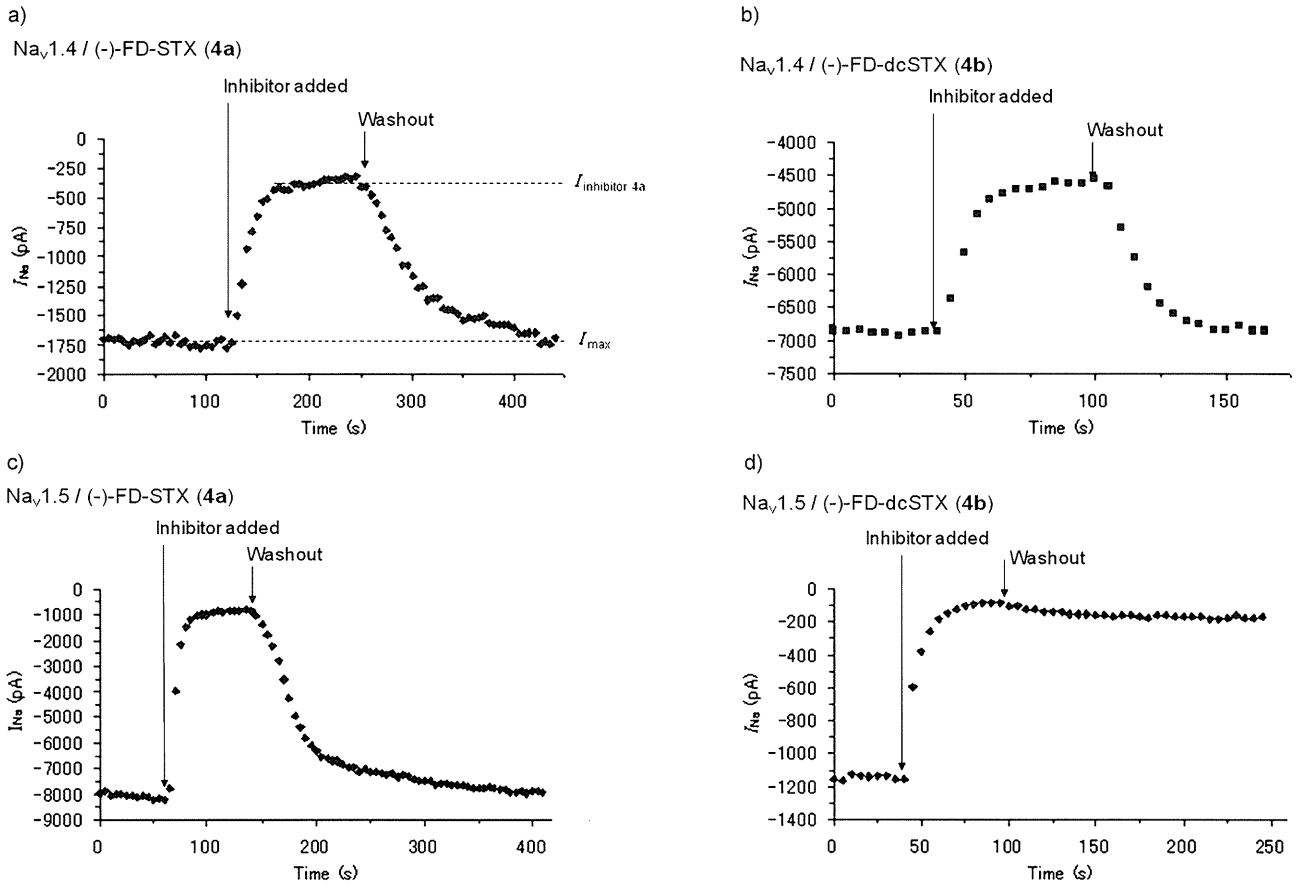


Figure 1. Time courses of inhibition by (−)-FD-STXs and recovery following washout. a) Effect of (−)-FD-STX (**4a**) (10 μM) on rNa$_v$1.4. b) Effect of (−)-FD-dcSTX (**4b**) (10 μM) on rNa$_v$1.4. c) Effect of (−)-FD-STX (**4a**) (1 mM) on rNa$_v$1.5. d) Effect of (−)-FD-dcSTX (**4b**) (1 mM) on rNa$_v$1.5.

Table 2. Inhibitory activity of (−)-FD-STXs **4a** and **4b** towards Na$_v$1.4 and Na$_v$1.5.

Compounds	IC_{50} (Na$_v$1.4) (mean ±SD)	n	IC_{50} (Na$_v$1.5) (mean ±SD)	n
(−)-FD-STX (**4a**)	(3.8±0.9) μM	3	(118±8.1) μM	3
(−)-FD-dcSTX (**4b**)	(16±8.6) μM	4	(182±16) μM	4

ery of the sodium ion current did not occur when (−)-FD-dcSTX (**4b**) was washed out (Figure 1d), thus indicating that the inhibition of Na$_v$1.5 by (−)-FD-dcSTX (**4b**) was "irreversible." Although the inhibitory action of **4b** against Na$_v$1.5 was quite weak, with an IC_{50} value of (182±16) μM (Table 2), this is the first observation of "irreversible" inhibition of Na$_v$1.5 by any STX- or TTX-related molecule. The interaction mechanism of **4b** and Na$_v$1.5 has not yet been clarified, but **4b** may be a good lead compound for the development of Na$_v$Ch-subtype-selective modulators.

Conclusion

In summary, we have synthesized a novel skeletal analogue of saxitoxin, FD-STX. Three FD-STX derivatives **4a–c**, which differ in substitution at C13, were synthesized, and their inhibitory effect on sodium ion channels was evaluated by means of cell-based assay and the patch-clamp method. (−)-FD-STX (**4a**) and (−)-FD-dcSTX (**4b**) showed IC_{50} values of (7.3±2.7) and (15±5.4) μM, respectively, in cell-based assay. These compounds inhibited Na$_v$1.4 (a representative tetrodotoxin-sensitive Na$_v$Ch subtype) and Na$_v$1.5 (a representative tetrodotoxin-resistant Na$_v$Ch subtype) in cells that express those channels. Uniquely, (−)-FD-dcSTX (**4b**) showed "irreversible" inhibition of Na$_v$1.5; it is the first STX- or TTX-related molecule to do so. Further work to elucidate the mechanism of the "irreversible" interaction between (−)-FD-dcSTX (**4b**) and Na$_v$1.5 is in progress.

Experimental Section

General: Flash chromatography was performed on silica gel 60 (spherical, particle size 40–100 μm; Kanto), Chromatorex NH (particle size 75–150 μm; Fuji Silysia). Optical rotations were measured using a JASCO P-2200 polarimeter with the sodium D line. ^{1}H and ^{13}C NMR spectra were recorded using a JEOL JNM ECX 400 or JNM-ECA 500 spectrometer. The spectra are referenced internally according to residual solvent signals of $CDCl_3$ (^{1}H NMR, δ=7.26 ppm; ^{13}C NMR, δ=77.0 ppm); CD_3OD (^{1}H NMR, δ=3.30 ppm; ^{13}C NMR, δ=49.0 ppm); or 4% CD_3CO_2D–D_2O (^{1}H NMR, δ=4.79 ppm (D_2O); ^{13}C NMR, δ=179.0 ppm

(CD_3CO_2D)). Mass spectra were recorded using a JEOL JMS-T100X spectrometer in ESI-MS mode with methanol as solvent.

Isoxazolidine 8: A mixture of nitrone (+)-**5** (10.4 g, 40.5 mmol) and nitro-olefin **6** (9.49 g, 42.5 mmol) was stirred for 30 min at 40 °C without any solvents. The reaction mixture was diluted with CH_2Cl_2 (150 mL) and cooled to −40 °C. DBU (6.28 mL, 44.6 mmol) was added to the solution, and the resulting mixture was stirred for 1 h at the same temperature. AcOH (11.6 mL, 203 mmol) and Zn powder (freshly activated with diluted aqueous HCl, 10.6 g, 162 mmol) were added to the reaction mixture, and the resulting mixture was warmed to 0 °C over 5 h. The reaction mixture was then filtered through a pad of Celite, and the filtrates were diluted with CH_2Cl_2 (150 mL) and washed with H_2O and brine. The combined organic layer was dried over $MgSO_4$, filtered, and concentrated in vacuo. The residue was purified by column chromatography on silica gel (hexane/EtOAc = 15:1 to 5:1) to give hydroxylamine **8** (13.5 g, 72%). Spectral data for **8**: $[\alpha]_D^{16} = +1.3$ ($c = 2.3$ in $CHCl_3$); ^{1}H NMR (400 MHz, $CDCl_3$): $\delta = 7.25$ (d, $J = 8.2$ Hz, 2H), 6.86 (d, $J = 8.2$ Hz, 2H), 5.47 (br, 1H), 4.48 (s, 2H), 4.23 (dt, $J = 2.3$, 5.0 Hz, 1H), 4.12 (q, $J = 6.0$ Hz, 1H), 3.79 (s, 3H), 3.72 (m, 2H), 3.58 (m, 2H), 3.3, (m, 2H), 2.06 (m, 1H), 1.73 (m, 1H), 1.04 ppm (m, 21H); ^{13}C NMR (100 MHz, $CDCl_3$) : $\delta = 159.3$, 129.6, 129.5, 113.8, 77.9, 77.8, 77.4, 73.4, 71.9, 67.9, 55.8, 55.2, 34.4, 17.9, 12.0 ppm; HRMS (ESI): *m/z* calcd for $C_{24}H_{42}N_2O_5SiNa$: 489.2761 [M+Na$^+$]; found: 489.2758.

Isoxazolidine 9: CbzCl (4.01 mL, 28.1 mmol) was added to a solution of 5α-hydroxylamine **8** (10.1 g, 21.6 mmol) and K_2CO_3 (15.0 g, 108 mmol) in THF/H_2O (1:1, 60 mL) at 0 °C. After stirring for 10 min at 0 °C, methanol (150 mL) was added to the solution, and resulting mixture was warmed to room temperature. The reaction mixture was then concentrated in vacuo, and the resulting material was diluted with EtOAc/H_2O and extracted with EtOAc. The combined organic layer was dried over $MgSO_4$, filtered, and concentrated in vacuo. The residue was purified by column chromatography on silica gel (hexane/EtOAc = 8:1 to 4:1) to give *N*-Cbz-hydroxylamine **9** (10.7 g, 83%). Spectral data for isoxazolidine **9**: $[\alpha]_D^{16} = +4.5$ ($c = 2.9$ in $CHCl_3$); ^{1}H NMR (400 MHz, $CDCl_3$): $\delta = 7.35$ (m, 5H), 7.19 (d, $J = 8.7$ Hz, 2H), 6.79 (d, $J = 8.7$ Hz, 2H), 5.20 (d, $J = 12.4$ Hz, 1H), 5.15 (d, $J = 12.4$ Hz, 1H), 4.90 (dd, $J = 4.6$, 5.0 Hz, 1H), 4.44 (d, $J = 11.0$ Hz, 1H), 4.40 (d, $J = 11.0$ Hz, 1H), 4.34 (m, 1H), 4.12 (dt, $J = 3.2$, 6.0 Hz, 1H), 3.97 (dd, $J = 2.8$, 3.2 Hz, 1H), 3.83 (dd, $J = 5.5$, 11.0 Hz, 1H), 3.77 (s, 3H), 3.63 (dd, $J = 3.2$, 11.0 Hz, 1H), 3.40 (ddd, $J = 6.0$, 12.6, 9.6 Hz, 1H), 3.16 (ddd, $J = 4.1$, 6.9, 12.6 Hz, 1H) 2.02 (m, 1H), 1.75 (m, 1H), 1.01 ppm (m, 21H); ^{13}C NMR (100 MHz, $CDCl_3$): $\delta = 159.5$, 156.0, 136.0, 129.8, 128.5, 128.2, 113.9, 77.4, 76.3, 75.2, 73.9, 67.8, 67.1, 66.7, 55.2, 33.9, 17.9, 17.8, 11.9 ppm; HRMS (ESI): *m/z* calcd for $C_{32}H_{48}N_2O_7SiNa$: 623.3129 [M+Na$^+$]; found: 623.3137.

Guanidine 11: $TiCl_3$/HCl (20% aq, 14.0 mL, 21.5 mmol; Kanto Chemical) was added to a solution of *N*-Cbz hydroxylamine **9** (10.7 g, 17.9 mmol), NaOAc (4.40 g, 53.7 mmol), and freshly activated Zn powder (3.51 g, 53.7 mmol) in CH_2Cl_2/MeOH (2:1, 100 mL) at −50 °C under Ar atmosphere. After stirring for 24 h, the reaction was neutralized by an excess amount of Na_2CO_3 solid at −50 °C. Then the dark-green suspension was warmed to room temperature and vigorously stirred under O_2 atmosphere for 30 min. The light-gray suspension was extracted with EtOAc, and the extracts were washed with water and brine. The organic layer was dried over $MgSO_4$, filtered, and concentrated in vacuo to give diamine **10** (10.8 g). $HgCl_2$ (4.37 mg, 16.1 mmol) was added to a solution of crude diamine **10** (10.3 g), bis(Boc)-2-methyl-2-thiopseudourea (**16**) (4.68 g, 16.1 mmol), and Et_3N (7.48 mL, 53.7 mmol) in DMF (40 mL) at room temperature under Ar atmosphere. After stirring for 1 h, the reaction mixture was diluted with EtOAc (150 mL) and filtered through a pad of Celite. The filtrates were washed with H_2O and brine. The organic layer was dried over $MgSO_4$, filtered, and concentrated in vacuo. The residue was purified by column chromatography on silica gel (hexane/EtOAc = 15:1 to 6:1) to give guanidine **11** (12.1 g, 82%, two steps). Spectral data for guanidine **11**: $[\alpha]_D^{17} = +0.6$ ($c = 2.6$ in $CHCl_3$); ^{1}H NMR (500 MHz, $CDCl_3$): $\delta = 10.31$ (s, 1H), 7.33 (m, 5H), 7.21 (d, $J = 8.6$ Hz, 2H), 6.83 (d, $J = 8.6$ Hz, 2H), 5.69 (br, 1H), 5.44 (d, $J = 9.7$ Hz, 1H), 5.09 (d, $J = 12.0$ Hz, 1H), 5.04 (d, $J = 12.0$ Hz, 1H), 4.47 (d, $J = 2.9$ Hz, 1H), 4.42 (d, $J = 12.0$ Hz, 1H), 4.39 (d, $J = 9.7$ Hz, 1H), 4.33 (d, $J = 12.0$ Hz, 1H), 3.92 (d, $J = 10.9$ Hz, 1H), 3.89 (dd, $J = 6.3$, 6.9 Hz, 1H), 3.77 (s, 3H), 3.59 (dt, $J = 8.0$, 10.9 Hz, 1H), 3.50 (m, 2H), 3.31 (dd, $J = 6.9$, 9.7 Hz, 1H), 2.35 (m, 1H), 1.96 (dd, $J = 6.9$, 13.8 Hz, 1H), 1.49 (s, 9H), 1.43 (s, 9H), 0.97 ppm (m, 21H); ^{13}C NMR (125 MHz, $CDCl_3$): $\delta = 161.3$, 159.1, 157.6, 156.5, 150.2, 136.4, 130.2, 129.6, 128.4, 128.0, 127.9, 113.6, 82.3, 79.5, 73.0, 72.5, 69.8, 66.8, 55.2, 53.0, 46.8, 32.1, 28.1, 28.0, 17.8, 12.0 ppm; HRMS (ESI): *m/z* calcd for $C_{43}H_{68}N_4O_{10}SiNa$: 851.4569 [$M$+Na$^+$]; found: 851.4602.

Cyclic guanidine 12: Chloromethanesulfonyl chloride (5.72 mL, 47.6 mmol) was added to a solution of guanidine **11** (9.90 g, 11.9 mmol) and *i*Pr$_2$NEt (12.3 mL, 71.4 mmol) in CH_2Cl_2 (60 mL) at 0 °C under an N_2 atmosphere. After stirring for 12 h, the reaction was quenched with saturated aqueous $NaHCO_3$, and the organic layer was extracted with CH_2Cl_2. The combined organic layer was dried over $MgSO_4$, filtered, and concentrated in vacuo. The residue was purified by column chromatography on silica gel (hexane/EtOAc = 10:1 to 8:1) to give cyclic guanidine **12** (8.35 mg, 87%). Spectral data for cyclic guanidine **12**: $[\alpha]_D^{18} = +70.9$ ($c = 2.0$ in $CHCl_3$); ^{1}H NMR (400 MHz, $CDCl_3$): $\delta = 7.33$ (m, 5H), 7.20 (d, $J = 8.24$ Hz, 2H), 6.82 (d, $J = 8.7$ Hz, 2H), 5.09 (d, $J = 12.2$ Hz, 1H), 5.05 (br, 1H), 4.99 (d, $J = 12.2$ Hz, 1H), 4.45 (d, $J = 10.3$ Hz, 1H), 4.42 (m, 1H), 4.29 (d, $J = 10.3$ Hz, 1H), 4.25 (m, 1H), 3.99 (m, 1H), 3.80 (m,1H), 3.78 (s = 3H), 3.71 (dd, $J = 3.7$, 10.1 Hz, 1H), 3.58 (m, 1H), 3.38 (dd, $J = 7.3$, 9.6 Hz, 1H), 3.18 (d, $J = 11.0$ Hz, 1H), 1.76 (m, 2H), 1.49 (s, 9H), 1.48 (s, 9H), 0.98 ppm (m, 21H); ^{13}C NMR (100 MHz, $CDCl_3$): $\delta = 159.6$, 159.0, 155.6, 151.6, 135.9, 130.4, 129.6, 128.6, 128.3, 128.2, 113.4, 82.6, 78.1, 74.1, 72.9, 71.4, 68.6, 67.2, 59.4, 55.2, 54.5, 46.0, 32.6, 28.5, 28.1, 17.9, 12.0 ppm; HRMS (ESI): *m/z* calcd for $C_{43}H_{67}N_4O_9Si$: 811.4677 [M+H$^+$]; found: 811.4678.

Bis(guanidine) 13: TBAF (3.23 g, 12.4 mmol) was added to a solution of cyclic guanidine **12** (8.35 g, 10.3 mmol) in a solution of THF (50 mL) at 0 °C. After stirring for 1 h, the reaction mixture was quenched with saturated aqueous NH_4Cl and extracted with EtOAc. The organic layer was dried over $MgSO_4$, filtered, and concentrated in vacuo. The residue was purified by column chromatography on silica gel (hexane/EtOAc; 3:1 to EtOAc) to give alcohol (6.27 g, 93%). Pd(OH)$_2$/C (5%, 33.3 mg) was added to a solution of alcohol (665 mg 1.02 mmol) in methanol (15 mL), and the mixture was vigorously stirred under an H_2 atmosphere (balloon) at room temperature for 3 h. The reaction mixture was filtered through a pad of Celite, and the filtrates were concentrated in vacuo to give amine (538 mg). $HgCl_2$ (331 mg, 1.22 mmol), Et_3N (425 μL, 3.05 mmol), and 1,3-bis(Cbz)-2-methyl-2-thiopseudourea (**17**) (437 mg, 1.22 mmol) were added to a solution of amine (538 mg) in DMF (10 mL) at room temperature for 2 h. The reaction mixture was diluted with EtOAc, and the resulting solution was filtered through a pad of Celite. The filtrates were washed with H_2O and brine twice, and the organic layer was dried over $MgSO_4$, filtered, and concentrated in vacuo. The residue was purified by column chromatography on Chromatorex NH (hexane/EtOAc = 3:2 to 1:2) to give bis(guanidine) **13** (718.2 mg, 85%, two steps). Spectral data for bis(guanidine) **13**: $[\alpha]_D^{21} = +122.8$ ($c = 0.46$ in MeOH); ^{1}H NMR (400 MHz, $CDCl_3$): $\delta = 11.59$ (s, 1H), 8.97 (d, $J = 6.41$ Hz, 1H), 7.35 (m, 10H), 7.22 (d, $J = 8.42$ Hz, 2H), 6.82 (d, $J = 8.42$ Hz, 2H), 5.48 (brs, 1H), 5.16 (dd, $J = 11.91$, 17.40 Hz, 2H), 5.13 (dd, $J = 12.36$, 45.34 Hz, 2H), 4.62 (m, 1H), 4.51 (t, $J = 5.50$, 1H), 4.39 (dd, $J = 11.45$, 27.93 Hz, 2H), 3.90 (dd, $J = 7.33$, 19.00 Hz, 1H), 3.77 (dd, $J = 4.12$, 9.62 Hz, 1H), 3.73 (s, 3H), 3.68 (m, 1H), 3.47 (dd, $J = 8.24$, 9.62 Hz, 1H), 3.17 (dd, $J = 4.12$, 7.79 Hz, 1H), 3.09 (m, 1H), 2.01 (m, 1H), 1.67 (m, 1H), 1.53 (s, 9H), 1.49 ppm (s, 9H); ^{13}C NMR (100 MHz, $CDCl_3$): $\delta = 162.73$, 159.61, 155.96, 153.58, 136.42, 134.40, 130.17, 129.56, 129.09, 128.83, 128.67, 128.18, 127.87, 113.93, 83.86, 78.85, 74.93, 73.33, 69.34, 68.79, 68.69, 67.43, 55.36, 54.95, 45.17, 30.13, 28.53, 28.06 ppm; HRMS (ESI): *m/z* calcd for $C_{43}H_{55}N_6O_{11}$: 831.3928 [M+H$^+$]; found: 831.3896.

Mono-Cbz-guanidine 14: NaOMe (9.0 mg, 0.167 mmol) was added to a solution of bis(guanidine) **13** (139 mg, 0.167 mmol) in THF/MeOH (1:1, 4 mL) at 0 °C under an N_2 atmosphere, and the mixture was stirred for 5 h. Saturated aqueous NH_4Cl was added to the reaction mixture, and organic layer was extracted with EtOAc. The extracts were washed with brine, dried over $MgSO_4$, filtered, and concentrated in vacuo. The residue was purified by column chromatography on Chromatorex NH (hexane/

EtOAc=3:1 to 1:2) to give mono-Cbz-guanidine **14** (75.7 mg, 65%). Spectral data for mono-Cbz-guanidine **14**: $[\alpha]_D^{14}=+202.3$ ($c=1.0$ in MeOH); ^{1}H NMR (500 MHz, CD_3OD): $\delta=7.17$ (m, 5H), 7.08 (d, $J=8.70$ Hz, 2H), 6.67 (d, $J=8.70$ Hz, 2H), 4.98 (dd, $J=12.82$, 39.38 Hz, 2H), 4.32 (d, $J=10.53$ Hz, 1H), 4.19 (m, 2H), 4.10 (d, $J=3.21$ Hz, 1H), 3.67 (m, 3H), 3.63 (s, 3H), 3.41 (m, 1H), 3.20 (m, 2H), 1.75 (m, 1H), 1.61 (m, 1H), 1.38 (m, 18H); ^{13}C NMR (125 MHz, CD_3OD): $\delta=162.07$, 160.11, 159.89, 152.55, 138.36, 129.96, 128.86, 128.17, 128.03, 113.94, 83.39, 78.93, 73.32, 73.02, 71.32, 67.93, 60.84, 55.11, 49.34, 46.76, 32.09, 28.42, 27.87 ppm; HRMS (ESI): m/z calcd for $C_{35}H_{49}N_6O_9$: 697.3561 $[M+H^+]$; found: 697.3557.

Fused-type bicyclic guanidine 15: IBX (117 mg, 0.416 mmol) was added to a solution of mono-Cbz-guanidine **14** (48 mg, 0.0694 mmol) in DMSO (1 mL), and the resulting suspension was stirred for 10 min at room temperature. Then the reaction mixture was warmed to 70°C and stirred for 6 h. The reaction was quenched with 10% aqueous $Na_2S_2O_3$ and saturated aqueous $NaHCO_3$, and the solution was diluted with EtOAc. The organic layer was dried over $MgSO_4$, filtered, and concentrated in vacuo. The residue was purified by column chromatography on Chromatorex NH (hexane/EtOAc; 1:2 to pure EtOAc) to give fused-type bicyclic guanidine **15** (11.8 mg, 24%). Spectral data for fused-type bicyclic guanidine **15**: $[\alpha]_D^{15}=+16.9$ ($c=0.47$ in MeOH); ^{1}H NMR (500 MHz, CD_3OD): $\delta=$ 7.20 (m, 7H), 6.79 (d, $J=8.59$ Hz, 2H), 4.94 (d, $J=1.72$ Hz, 2H), 4.61 (m, 1H), 4.51 (d, $J=10.88$ Hz, 1H), 4.39 (d, $J=10.88$ Hz, 1H), 4.00 (d, $J=$ 1.72 Hz, 1H), 3.90 (dd, $J=7.45$, 9.74 Hz, 1H), 3.72 (m, 1H), 3.68 (s, 3H), 3.52 (m, 1H), 3.44 (m, 1H), 2.27 (m, 1H), 2.08 (m, 1H), 1.34 ppm (m, 18H); ^{13}C NMR (125 MHz, CD_3OD): $\delta=160.95$, 160.37, 153.10, 150.41, 138.66, 131.02, 130.75, 129.38, 128.86, 114.79, 86.42, 85.32, 82.96, 80.05, 74.15, 69.82, 67.48, 59.99, 55.69, 45.08, 33.10, 28.77, 28.27 ppm; HRMS (ESI): m/z calcd for $C_{35}H_{47}N_6O_{10}$: 711.3353 $[M+H^+]$; found: 711.3400.

(−)-FD-dcSTX (4b): $Pd(OH)_2/C$ (2.5 mg) was added to a solution of **15** (4.8 mg, 0.00676 mmol) in methanol (500 μL), and the suspension was vigorously stirred under an H_2 atmosphere (balloon) at room temperature. After 30 min, the reaction mixture was filtered through a pad of Celite. The filtrates were concentrated in vacuo to give bis-Boc-bis-aminal (3.1 mg). HCl (3N, 500 μL) was added to the resulting bis-Boc-bis-aminal at 0°C. After stirring for 1 h at room temperature, the reaction mixture was concentrated in vacuo to give FD-dcSTX (**4b**) as a 2HCl salt (1.7 mg, 78%). Spectral data for FD-dcSTX (**4b**): $[\alpha]_D^{15}=-63.9$ ($c=0.11$ in MeOH); ^{1}H NMR (400 MHz, D_2O): $\delta=4.06$ (s, 1H), 3.82 (m, 3H), 3.70 (t, $J=10.07$ Hz, 1H), 3.48 (dd, $J=10.07$, 17.40 Hz, 2H), 2.50 (m, 1H), 2.34 ppm (m, 1H); ^{13}C NMR (100 MHz, 4% CD_3CO_2D in D_2O): $\delta=152.91$, 152.16, 85.53, 81.49, 62.52, 55.24, 45.78, 43.00, 31.44 ppm; HRMS (ESI): m/z calcd for $C_9H_{17}N_6O_3$: 257.1362 $[M+H^+]$; found: 257.1319.

Alcohol 19: Ac_2O (2 mL) and pyridine (2 mL) were added to bis-guanidine **13** (230 mg, 0.277 mmol) at room temperature. The reaction mixture was warmed to 50°C and stirred for a further 3 h, then cooled to room temperature, quenched with toluene, and concentrated in vacuo. The resulting residue was purified by column chromatography on Chromatorex NH (hexane/EtOAc; 6:1 to 3:1) to give acetate (245 mg, 96%). DDQ (377 mg, 1.66 mmol) was added to a solution of acetate (144 mg, 0.166 mmol) in CH_2Cl_2/H_2O (2:1, 4.5 mL) at room temperature. The reaction mixture was stirred for 5 h, saturated aqueous $NaHCO_3$ was added, and the resulting solution was extracted with EtOAc three times. The combined organic layer was dried over $MgSO_4$, filtered, and concentrated in vacuo. The residue was purified by column chromatography on Chromatorex NH (hexane/EtOAc; 6:1 to 1:1) to give alcohol **19** (117 mg, 94%). Spectral data for alcohol **19**: $[\alpha]_D^{23}=+171.7$ ($c=0.3$ in MeOH); ^{1}H NMR (400 MHz, $CDCl_3$): $\delta=11.76$ (s, 1H), 8.65 (d, $J=8.70$ Hz, 1H), 7.36 (m, 10H), 5.23 (dd, $J=11.91$, 20.15 Hz, 2H), 5.11 (dd, $J=12.82$, 16.94 Hz, 2H), 4.48 (m, 1H), 4.08 (m, 1H), 4.00 (d, $J=11.91$ Hz, 1H), 3.90 (t, $J=8.70$, 1H), 3.65 (m, 3H), 3.52 (t, $J=9.63$ Hz, 1H), 2.05 (m, 5H), 1.49 ppm (m, 18H); ^{13}C NMR (100 MHz, $CDCl_3$): $\delta=170.83$, 163.55, 156.44, 153.82, 136.63, 134.44, 128.82, 128.59, 128.52, 127.87, 83.51, 80.11, 75.41, 68.63, 67.33, 64.91, 51.36, 46.46, 28.77, 28.38, 28.20, 21.16 ppm; HRMS (ESI): m/z calcd for $C_{37}H_{49}N_6O_{11}$: 753.3459 $[M+H^+]$; found: 753.3496.

Carbamate 20: Trichloroacetyl isocyanate (19 μL, 0.16 mmol) was added to a solution of alcohol **19** (60 mg, 0.080 mmol) in CH_2Cl_2 (1 mL) at 0°C under an N_2 atmosphere. The reaction was stirred for 1 h. Then K_2CO_3 and methanol were added at 0°C and warmed to room temperature. The reaction mixture was stirred for 3 h, quenched with H_2O, and extracted with EtOAc twice. The combined organic layer was dried over $MgSO_4$, filtered, and concentrated in vacuo. The residue was purified by preparative TLC (EtOAc, four times) to give carbamate **20** (16.6 mg, 34%). Spectral data for carbamate **20**: $[\alpha]_D^{22}=+52.3$ ($c=0.32$ in MeOH); ^{1}H NMR (400 MHz, CD_3OD): $\delta=7.30$ (m, 5H), 5.04 (dd, $J=3.66$, 12.82 Hz, 2H), 4.25 (m, 2H), 3.58 (m, 4H), 3.28 (m, 2H), 1.84 (m, 1H), 1.69 (m, 1H), 1.41 ppm (m, 18H); ^{13}C NMR (100 MHz, CD_3OD): $\delta=$ 162.6, 160.2, 159.1, 154.6, 152.8, 138.8, 129.4, 128.6, 84.2, 79.7, 73.3, 68.3, 64.8, 59.8, 55.6, 47.2, 45.5, 32.8, 28.8, 28.7 ppm; HRMS (ESI): m/z calcd for $C_{28}H_{42}N_7O_9$: 620.3044 $[M+H^+]$; found: 620.3074.

Fused-type bicyclic guanidine 21: IBX (62 mg, 0.223 mmol) was added to a solution of **20** (23 mg, 0.0371 mmol) in DMSO (1 mL), and the resulting suspension was stirred for 10 min at room temperature. Then, the reaction mixture was warmed to 70°C and stirred for 6 h. The reaction was quenched with 10% aqueous $Na_2S_2O_3$ and saturated aqueous $NaHCO_3$, and the solution was diluted with EtOAc. The organic layer was dried over $MgSO_4$, filtered, and concentrated in vacuo. The residue was purified by column chromatography on Chromatorex NH (hexane/EtOAc; 1:8 to EtOAc/MeOH; 10:1) to give **21** (4.6 mg, 20%). Spectral data for fused-type bicyclic guanidine **21**: $[\alpha]_D^{23}=-7.4$ ($c=0.1$ in MeOH); ^{1}H NMR (400 MHz, CD_3OD): $\delta=7.34$ (m, 5H), 5.02 (d, $J=1.37$ Hz, 2H), 4.53 (dd, $J=5.5$, 11.9 Hz, 2H), 4.28 (dd, $J=8.24$, 11.91 Hz, 2H), 4.02 (d, $J=3.96$ Hz, 1H), 3.63 (m, 3H), 2.37 (m, 1H), 2.19 (m, 1H), 1.48 (m, 9H), 1.44 ppm (m, 9H); ^{13}C NMR (100 MHz, CD_3OD): $\delta=164.5$, 160.6, 159.3, 153.0, 149.5, 138.6, 130.0, 129.5, 86.9, 86.2, 83.9, 80.8, 68.1, 65.3, 59.8, 45.8, 33.5, 29.3, 28.9 ppm; HRMS (ESI): m/z calcd for $C_{28}H_{40}N_7O_{10}$: 634.2836 $[M+H^+]$; found: 634.2884.

(−)-FD-STX (4a): $Pd(OH)_2/C$ (3.0 mg) was added to a solution of **21** (7.0 mg, 0.011 mmol) in methanol (600 μL) under an N_2 atmosphere, and the suspension was vigorously stirred under an H_2 atmosphere (balloon) at room temperature. After 7 h, the reaction mixture was filtered through a pad of Celite. The filtrates were concentrated in vacuo to give bis-Boc-bis-aminal (5.3 mg). TFA (500 μL) was added to the resulting bis-Boc-bis-aminal (5.3 mg) in CH_2Cl_2 (1 mL) at room temperature. Methanol was added to the reaction mixture, and it was concentrated in vacuo to give FD-STX (**4a**) as a 2TFA salt (4.8 mg, 82%). Spectral data for FD-STX (**4a**): $[\alpha]_D^{23}=-34.9$ ($c=0.32$ in MeOH); ^{1}H NMR (400 MHz, D_2O): $\delta=4.39$ (dd, $J=8.70$, 10.99 Hz, 1H), 4.17 (dd, $J=4.58$, 10.99 Hz, 1H), 4.05 (s, 1H), 3.95 (dd, $J=4.58$, 8.70 Hz, 1H), 3.68 (t, $J=10.7$ Hz, 1H), 3.45 (m, 1H), 2.47 (m, 1H), 2.33 ppm (dd, $J=7.79$, 13.28 Hz, 1H); ^{13}C NMR (100 MHz, 4% CD_3CO_2D in D_2O): $\delta=159.1$, 153.0, 152.1, 85.6, 81.7, 65.5, 52.5, 45.3, 43.2, 31.4 ppm; HRMS (ESI): m/z calcd for $C_{10}H_{18}N_7O_4$: 300.1420 $[M+H^+]$; found: 300.1390.

(−)-FD-doSTX (4c): TFA (0.5 mL) was added to a solution of (−)-**3** (12.1 mg, 0.0199 mmol) in CH_2Cl_2 (1 mL) at 0°C. After being stirred for 1 h at room temperature, the reaction mixture was concentrated under reduced pressure to give bis-Cbz-bis-aminal. Pd/C (catalytic amount) was added to a solution of bis-Cbz-bis-aminal in methanol (1 mL), and the reaction mixture was stirred at room temperature under an atmosphere of hydrogen (balloon). After 4 h, the reaction mixture was filtered through a pad of Celite, and the filtrates were acidified with 10% HCl and concentrated in vacuo to give FD-doSTX (**4c**) (5.7 mg, 92%, two steps). Spectral data for **4c**: $[\alpha]_D^{20}=-62.7$ ($c=0.9$ in MeOH); ^{1}H NMR (400 MHz, D_2O): $\delta=3.88$ (d, $J=1.0$ Hz, 1H), 3.75 (dd, $J=1.0$, 7.1 Hz, 1H), 3.63 (dd, $J=9.6$, 9.6 Hz, 1H), 3.42 (ddd, $J=7.6$, 9.6, 10.1 Hz, 1H), 2.43 (ddd, $J=9.6$, 10.1, 13.6 Hz, 1H), 2.27 (dd, $J=7.6$, 13.6 Hz, 1H), 1.38 ppm (d, $J=7.1$ Hz, 3H); ^{13}C NMR (100 MHz, CD_3OD): $\delta=$ 152.4, 151.4, 85.2, 81.8, 49.4, 48.4, 42.6, 30.9, 19.3 ppm; HRMS (ESI): m/z calcd for $C_9H_{17}N_6O_2$: 241.1413 $[M+H]$; found: 241.1434.

Cell culture: HEK293 was grown in Dulbecco's modified Eagle's medium (DMEM; Life Technologies, Grand Island, NY) supplemented with 10% fetal bovine serum (30 units; BioWhittaker, Walkersville, MD) and penicillin-streptomycin (30 μg mL^{-1}; Gibco, Grand Island, NY), re-

spectively, at 37 °C under 5% CO_2 in air. Neuro-2a was grown in RPMI1640 (Sigma, St Louis, MO) supplemented with 10% fetal bovine serum (50 units) and penicillin-streptomycin (50 $\mu g\,mL^{-1}$), respectively.

Inhibitory activity to cytotoxicity of ouabain and veratridine: Two days' worth of cultured cells (Neuro-2a) were harvested with a trypsin/EDTA solution (0.5%/0.2%) (Sigma) in phosphate buffer saline (PBS), and cells were suspended at a density of 5×10^5 cells mL^{-1} in growth medium. This suspension (100 μL) was plated into each well of 96-well plates and cultured at 37 °C for 24 h. The test medium (100 μL), which contained FD-STXs at various concentrations with 1:3 serial dilutions, and ouabain (200 μM) (Sigma) and veratridine (100 μM) (Sigma) were poured into each well. After incubation at 37 °C for 24 h, the number of viable cells was determined using Cell Counting Kit-8 (Dojindo, Kumamoto, Japan). Cell Counting Kit-8 solution (5 μL) was added to each well, and the mixture was incubated at 37 °C for 3 h, and then absorbance at 450/655 nm was measured using a microplate reader (Model 550, Bio-Rad, Hercules, CA). Three wells were used for one concentration of FD-STXs, and the same experiments were performed three times for each compound. The cell viability of the well that contained the test medium in the presence of only ouabain and veratridine (in the absence of FD-STXs) was taken as 0%, and that of the well that contained the medium in the absence of all toxins was taken as 100%. The IC_{50} were determined from the Hill plot.

Transient transfection: One day before transfection, HEK293 cells were plated in a 60 mm dish (10^6 cells $dish^{-1}$). Complexes for transfection were prepared as follows: DNA (1 μg) was diluted in Buffer EC (150 μL), the DNA condensation buffer for the Effectene (Hilden, Germany). Next, enhancer (8 μL) from the kit was added to the solution of DNA. After 2 min, the above medium (150 μL) and Effectene Transfection Reagent kit (25 μL; Hilden, Germany) were mixed and incubated for 5 min. Then, the medium from 50%-confluent cells was removed and replaced with fresh medium (3 mL). The mixed complexes with fresh medium (1 mL) were added to the dish, which was swirled by rocking the plate. Incubation was continued at 37 °C under humidified 5% CO_2 in air for at least 4 h. The medium was removed and replaced with fresh medium (4 mL). The transfected cells were used for electrophysiological experiments within 24–72 h after transfection.

Electrophysiological recording: Macroscopic sodium ion currents from the transfected cells were measured using a whole-cell patch-clamp method using an Axopatch 200B amplifier (Axon Instruments; Foster City, CA, USA); more than 80% of the series resistance was compensated to minimize voltage errors. The bath solution contained: NaCl (70 mM), *N*-methyl-D-glucamine (67 mM), $CaCl_2$ (1 mM), $MgCl_2$ (1.5 mM), glucose (10 mM), and HEPES (5 mM) (pH 7.4). The pipette solution contained: CsF (70 mM), CsCl (60 mM), NaF (12 mM), EGTA (5 mM), and HEPES (5 mM) (pH 7.4). Whole-cell patch pipettes with a resistance of 1.5–2 MΩ were used to achieve optimum voltage control. Recordings were started 10 min after establishing a whole-cell recording configuration. Whole-cell membrane currents were digitized at a sampling rate of 50–100 kHz using a 12-bit analogue-to-digital converter (DigiData 1321A interface; Axon Instruments), controlled by pClamp software (version 8; Axon Instruments). All experiments were conducted at room temperature (22–24 °C). Data are presented as mean ± SD along with the number of observations (n), unless otherwise stated.

Acknowledgements

Financial support for this work was provided in part by a grant-in-aid for scientific research from JSPS (no. 203130). O.I. is grateful for financial support from JSPS in the form of a Pre-doctoral Fellowship for Young Scientists.

[1] W. A. Catterall, A. L. Goldin, S. G. Waxman, *Pharmacol. Rev.* **2005**, *57*, 397–409.

[2] Y. Qu, R. Curtis, D. Lawson, K. Gilbride, P. Ge, P. S. DiStefano, I. Silos-Santiago, W. A. Catterall, T. Scheuer, *Mol. Cell. Neurosci.* **2001**, *18*, 570–580.

[3] a) J. J. Clare, S. N. Tate, M. Nobbs, M. A. Romanos, *Drug Discovery Today* **2000**, *5*, 506–520; b) M. Noda, T. Ikeda, T. Kayano, H. Suzuki, H. Takeshima, M. Kurasaki, H. Takahashi, S. Numa, *Nature* **1986**, *320*, 188–192; c) L. S. Meadows, Y. H. Chen, A. J. Powell, J. J. Clare, D. S. Ragsdale, *Neuroscience* **2002**, *114*, 745–753; d) J. S. Trimmer, S. S. Cooperman, S. A. Tomiko, J. Zhou, S. M. Crean, M. B. Boyle, R. G. Kallen, Z. Sheng, R. L. Barchi, F. J. Sigworth, *Neuron* **1989**, *3*, 33–49; e) M. Chahine, P. B. Bennett, A. L. George, Jr., R. Horn, *Pflugers Arch. Eur. J. Physiol.* **1994**, *427*, 136–142; f) P. S. Dietrich, J. G. McGivern, S. G. Delgado, B. D. Koch, R. M. Eglen, J. C. Hunter, L. Sangameswaran, *J. Neurochem.* **1995**, *64–65*, 2262–2272; g) M. R. Smith, R. D. Smith, N. W. Plummer, M. H. Meisler, A. L. Goldin, *J. Neurosci.* **1998**, *18*, 6093–6102; h) N. Klugbauer, L. Lacinova, V. Flockerzi, F. Hofmann, *EMBO. J.* **1995**, *14*, 1084–1090; i) T. R. Cummins, J. R. Howe, S. G. Waxman, *J. Neurosci.* **1998**, *18*, 9607–9619; j) L. Sangameswaran, L. M. Fish, B. D. Koch, D. K. Rabert, S. G. Delgado, M. Ilnicka, L. B. Jakeman, S. Novakovic, K. Wong, P. Sze, E. Tzoumaka, G. R. Stewart, R. C. Herman, H. Chan, R. M. Eglen, J. C. Hunter, *J. Biol. Chem.* **1997**, *272*, 14805–14809.

[4] a) J. Satin, J. W. Kyle, M. Chen, P. Bell, L. L. Cribbs, H. A. Fozzard, R. B. Rogart, *Science* **1992**, *256*, 1202–1205; b) L. Sangameswaran, S. G. Delgado, L. M. Fish, B. D. Koch, L. B. Jakeman, G. R. Stewart, P. Sze, J. C. Hunter, R. M. Eglen, R. C. Herman, *J. Biol. Chem.* **1996**, *271*, 5953–5956; c) A. N. Akopian, L. Sivilotti, J. N. Wood, *Nature* **1996**, *379*, 257–262; d) S. D. Dib-Hajj, L. Tyrrell, J. A. Black, S. G. Waxman, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 8963–8968; e) S. Dib-Hajj, J. A. Black, T. R. Cummins, S. G. Waxman, *Trends Neurosci.* **2002**, *25*, 253–259.

[5] E. Watanabe, A. Fujikawa, H. Matsunaga, Y. Yasoshima, N. Sako, T. Yamamoto, C. Saegusa, M. Noda, *J. Neurosci.* **2000**, *20*, 7743–7751.

[6] a) B. Fontaine, T. S. Khurana, E. P. Hoffman, G. A. Bruns, J. L. Haines, J. A. Trofatter, M. P. Hanson, J. Rich, H. McFarlane, D. M. Yasek, D. Romano, J. F. Gusella, R. H. Brown, *Science* **1990**, *250*, 1000–1002; b) K. Jurkat-Rott, F. Lehmann-Horn, *Neurotherapeutics* **2007**, *4*, 216–224.

[7] M. T. Keating, M. C. Sanguinetti, *Cell* **2001**, *104*, 569–580.

[8] a) C. Rosker, B. Lohberger, D. Hofer, B. Steinecker, S. Quasthoff, W. Schreibmayer, *Am. J. Physiol.* **2007**, *293*, C783–C789; b) B. S. Williams, J. P. Felix, B. T. Priest, R. M. Brochu, K. Dai, S. B. Hoyt, C. London, Y. S. Tang, J. L. Duffy, W. H. Parsons, G. J. Kaczorowski, M. L. Garcia, *Biochemistry* **2007**, *46*, 14693–14703; c) I. Drizin, R. J. Gregg, M. J. Scanio, L. Shi, M. F. Gross, R. N. Atkinson, J. B. Thomas, M. S. Johnson, W. A. Carroll, B. E. Marron, M. L. Chapman, D. Liu, M. J. Krambis, C. C. Shieh, X. Zhang, G. Hernandez, D. M. Gauvin, J. P. Mikusa, C. Z. Zhu, S. Joshi, P. Honore, K. C. Marsh, R. Roeloffs, S. Werness, D. S. Krafte, M. F. Jarvis, C. R. Faltynek, M. E. Kort, *Bioorg. Med. Chem.* **2008**, *16*, 6379–6386; d) M. E. Kort, I. Drizin, R. J. Gregg, M. J. Scanio, L. Shi, M. F. Gross, R. N. Atkinson, M. S. Johnson, G. J. Pacofsky, J. B. Thomas, W. A. Carroll, M. J. Krambis, D. Liu, C. C. Shieh, X. Zhang, G. Hernandez, J. P. Mikusa, C. Zhong, S. Joshi, P. Honore, R. Roeloffs, K. C. Marsh, B. P. Murray, J. Liu, S. Werness, C. R. Faltynek, D. S. Krafte, M. F. Jarvis, M. L. Chapman, B. E. Marron, *J. Med. Chem.* **2008**, *51*, 407–416; e) C. London, S. B. Hoyt, W. H. Parsons, B. S. Williams, V. A. Warren, R. Tschirret-Guth, M. M. Smith, B. T. Priest, E. McGowan, W. J. Martin, K. A. Lyons, X. Li, B. V. Karanam, N. Jochnowitz, M. L. Garcia, J. P. Felix, B. Dean, C. Abbadie, G. J. Kaczorowski, J. L. Duffy, *Bioorg. Med. Chem. Lett.* **2008**, *18*, 1696–1701; f) S. England, M. J. de Groot, *Br. J. Pharmacol.* **2009**, *158*, 1413–1425.

[9] a) M. F. Jarvis, P. Honore, C. C. Shieh, M. Chapman, S. Joshi, X. Zhang, M. Kort, W. Carroll, B. Marron, R. Atkinson, J. Thomas, D. Liu, M. Krambis, Y. Liu, S. McGaraughty, K. Chu, R. Roeloffs, C. Zhong, J. P. Mikusa, G. Hernandez, D. Gauvin, C. Wade, C. Zhu, M. Pai, M. Scanio, L. Shi, I. Drizin, R. Gregg, M. Matulenko, A.

Hakeem, M. Gross, M. Johnson, K. Marsh, P. K. Wagoner, J. P. Sullivan, C. R. Faltynek, D. S. Krafte, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 8520–8525; b) B. T. Priest, G. J. Kaczorowski, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 8205–8206.

[10] a) E. Moczydlowski, S. Hall, S. S. Garber, G. S. Strichartz, C. J. Miller, *J. Gen. Physiol.* **1984**, *84*, 687–704; b) C. Y. Kao, A. Nishiyama, *J. Physiol.* **1965**, *180*, 50–66.

[11] a) K. A. Schug, W. Lindner, *Chem. Rev.* **2005**, *105*, 67–114; b) T. Narahashi, H. G. Haas, E. F. Therrien, *Science* **1967**, *157*, 1441–1442.

[12] a) G. Choudhary, L. Shang, X. Li, S. C. Dudley, Jr., *Biophys. J.* **2002**, *83*, 912–919; b) D. B. Tikhonov, B. S. Zhorov, *Biophys. J.* **2005**, *88*, 184–197; c) we prepared the picture in Scheme 1 based on the schematic view of docking studies by Dudley and co-workers in reference [12a].

[13] L. E. Llewellyn, *Nat. Prod. Rep.* **2006**, *23*, 200–222.

[14] a) B. M. Andresen, J. Du Bois, *J. Am. Chem. Soc.* **2009**, *131*, 12524–12525; b) C. Robillot, D. Kineavy, J. Burnell, L. E. Llewellyn, *Toxicon* **2009**, *53*, 460–465.

[15] a) O. Iwamoto, H. Koshino, D. Hashizume, K. Nagasawa, *Angew. Chem.* **2007**, *119*, 8779–8782; *Angew. Chem. Int. Ed.* **2007**, *46*, 8625–8628; b) O. Iwamoto, R. Shinohara. K. Nagasawa, *Chem. Asian J.* **2009**, *4*, 277–285; c) O. Iwamoto, K. Nagasawa, *Org. Lett.* **2010**, *12*, 2150–2153.

[16] The structure of (+)-STX (**1a**) and compound **3**, which is a synthetic intermediate for FD-doSTX (**4c**), were superimposed by MOE. Please see Figure S1 in the Supporting Information.

[17] M. Bashir-Uddin Surfraz, M. Akhtar, R. K. Allemann, *Tetrahedron Lett.* **2004**, *45*, 1223–1226.

[18] K. S. Kim, L. Qian, *Tetrahedron Lett.* **1993**, *34*, 7677–7680.

[19] a) I. J. McAlpine, R. W. Armstrong, *Tetrahedron Lett.* **2000**, *41*, 1849–1853; b) K. Nagasawa, H. Koshino, T. Nakata, *Tetrahedron Lett.* **2001**, *42*, 4155–4158.

[20] a) K. Kogure, M. L. Tamplin, U. Simidu, R. R. Colwell, *Toxicon* **1988**, *26*, 191–197; b) R. L. Manger, L. S. Leja, S. Y. Lee, J. M. Hungerford, M. M. Wekell, *Anal. Biochem.* **1993**, *214*, 190–194; c) R. L. Manger, L. S. Leja, S. Y. Lee, J. M. Hungerford, Y. Hokama, R. W. Dickey, H. R. Granade, R. Lewis, T. Yasumoto. M. M. Wekell, *J. AOAC Int.* **1995**, *78*, 521–527; d) T. Yasumoto, M. Fukui, K. Sasaki, K. Sugiyama, *J. AOAC Int.* **1995**, *78*, 574–582; e) M. Yotsu-Yamashita, D. Urabe, M. Asai, T. Nishikawa, M. Isobe, *Toxicon* **2003**, *42*, 557–560.

[21] See the Supporting Information for calculation of IC_{50} values.

[22] The voltage that provided a peak current (maximum active potential) for $rNa_V1.4$ and $rNa_V1.5$ was −10 mV.

[23] The holding potentials of $Na_V1.4$ and $Na_V1.5$ are −100 and −120 mV, respectively.

[24] Under the same conditions, (+)-dc-STX (**1b**) showed inhibitory activities against $Na_V1.4$ and $Na_V1.5$ with IC_{50} values of (29±1.8) nM and (1.2±0.1) μM, respectively. Compound (+)-dc-STX (**1b**) interacted reversibly with both channels.

[25] The IC_{50} values were determined according to the following equation; $IC_{50}=[\text{inhibitor } \mathbf{4}](I_{\text{inhibitor}4}/I_{\max})/(1-(I_{\text{inhibitor}4}/I_{\max}))$, in which $I_{\max}$ is the peak current under control conditions and $I_{\text{inhibitor}4}$ is the equilibrium peak current in the presence of an inhibitor **4**: J. L. Penzotti, H. A. Fozzard, G. M. Lipkind, S. C. Dudley, Jr., *Biophys. J.* **1998**, *75*, 2647–2657.

Received: April 7, 2011
Revised: July 20, 2011
Published online: September 16, 2011

JOURNAL OF CHEMICAL INFORMATION AND MODELING

ARTICLE

pubs.acs.org/jcim

Correlation Analyses on Binding Affinity of Sialic Acid Analogues and Anti-Influenza Drugs with Human Neuraminidase Using ab Initio MO Calculations on Their Complex Structures — LERE-QSAR Analysis (IV)

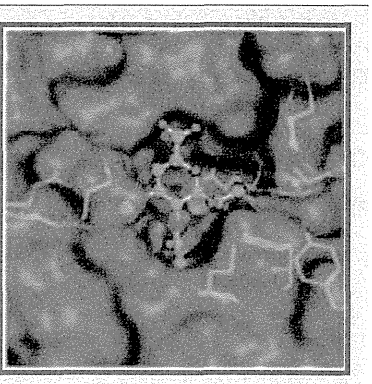
Seiji Hitaoka,[†] Hiroshi Matoba,[†] Masataka Harada,[†] Tatsusada Yoshida,[†] Daisuke Tsuji,[†] Takatsugu Hirokawa,[‡] Kohji Itoh,[†] and Hiroshi Chuman*[,†]

[†]Institute of Health Biosciences, The University of Tokushima Graduate School, 1-78 Shomachi, Tokushima 770-8505, Japan

[‡]Computational Biology Research Center (CBRC), National Institute of Advanced Industrial Science and Technology (AIST), 2-42 Aomi, Koto-ku, Tokyo 135-0064, Japan

Supporting Information

ABSTRACT: We carried out full ab initio fragment molecular orbital (FMO) calculations for complexes comprising human neuraminidase-2 (hNEU2) and sialic acid analogues including anti-influenza drugs zanamivir (Relenza) and oseltamivir (Tamiflu) in order to examine the variation in the observed inhibitory activity toward hNEU2 at the atomic and electronic levels. We recently proposed the LERE (linear expression by representative energy terms)-QSAR (quantitative structure—activity relationship) procedure. LERE-QSAR analysis quantitatively revealed that the complex formation is driven by hydrogen-bonding and electrostatic interaction of hNEU2 with sialic acid analogues. The most potent inhibitory activity, that of zanamivir, is attributable to the strong electrostatic interaction of a positively charged guanidino group in zanamivir with negatively charged amino acid residues in hNEU2. After we confirmed that the variation in the observed inhibitory activity among sialic acid analogues is excellently reproducible with the LERE-QSAR equation, we examined the reason for the remarkable difference between the inhibitory potencies of oseltamivir as to hNEU2 and influenza A virus neuraminidase-1 (N1-NA). Several amino acid residues in close contact with a positively charged amino group in oseltamivir are different between hNEU2 and N1-NA. FMO-IFIE (interfragment interaction energy) analysis showed that the difference in amino acid residues causes a remarkably large difference between the overall interaction energies of oseltamivir with hNEU2 and N1-NA. The current results will be useful for the development of new anti-influenza drugs with high selectivity and without the risk of adverse side effects.



1. INTRODUCTION

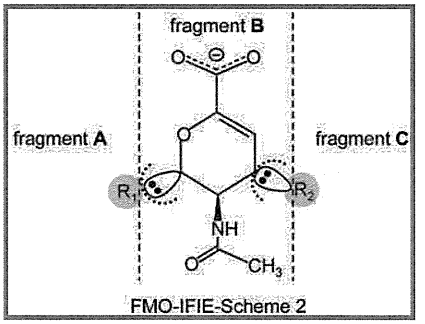
Neuraminidases (NAs, also called sialidases) are ubiquitous exoglycohydrolases that hydrolyze the terminal sialic acids of glycoproteins, glycopeptides, gangliosides, oligosaccharides, and polysaccharides and have been proved to play important roles in various biological processes through regulation of cellular sialic acid contents. NAs are widely distributed in nature, from viruses, and microorganisms such as bacteria, protozoa, and fungi to higher animals and humans.[1] Structural and functional studies on NAs have been widely reported so far. Most of them were on influenza virus NAs,[2−4] a smaller number on human neuraminidases (hNEUs) having also been reported.[5−8] The NA inhibitor first developed, 2-deoxy-2,3-didehydro-*N*-acetyl-neuraminic acid (Neu5Ac2en, DANA),[9] was designed as a putative transition state analogue of sialic acid in the late 1960s. Although the inhibition by DANA is not selective for influenza virus NAs,[10,11] DANA shows moderate inhibitory activity toward a variety of NAs, suggesting that amino acid residues in the active sites of NAs are highly conserved among a number of species including influenza virus and human NAs.

hNEUs are classified into four types,[12] the lysosomal neuraminidase (hNEU1),[13] cytosolic neuraminidase (hNEU2),[14] plasma membrane-associated neuraminidase (hNEU3),[15] and lysosomal/mitochondrial membrane-associated neuraminidase (hNEU4),[16] according to several characteristics, such as their subcellular distributions, enzymatic properties, and substrate specificities. They are involved in various cellular functions such as cell proliferation, differentiation, and apoptosis:[17] hNEU1 is essential for the lysosomal catabolism of sialylated glycoconjugates,[18−20] hNEU2 is able to recognize not only sialosyl linkages but also saccharide units close to sialic acids and the supramolecular organization of such as gangliosides, hNEU3 is expected to be involved in cell surface functions by modulating gangliosides, and hNEU4 is active against a broad range of sialylated glycoconjugates and has been implicated in the catabolism of glycolipids, but the details are not known at present. The crystallographic structure of hNEU2[21] is the only available one among those of the four hNEUs.

As mentioned previously, DANA inhibits influenza virus NAs moderately but is not used as an anti-influenza drug because of its practically low potency. Through structure-based approaches

Received: May 27, 2011
Published: August 28, 2011

Table 1. Chemical Structure and Inhibitory Potency of Compounds 1−5 against hNEU2



compound no.	name	R_1 (fragment A)	R_2 (fragment C)	hNEU2 log K_i[a]
1	IEM[b]	$OCH_2CH(CH_3)_2$	OH	−3.06
2	HEM[c]	$OCH_2CH_2CH_2(OH^1)$	OH	−3.13
3	DEM[d]	$OCH_2CH(OH^1)CH_2(OH^2)$	OH	−2.85
4	DANA[e]	$CH(OH^1)CH(OH^2)CH_2(OH^3)$	OH	−3.85
5	zanamivir	$CH(OH^1)CH(OH^2)CH_2(OH^3)$	$NHC(=NH_2^+)NH_2$	−4.77

[a] Taken from ref 33. [b] Isobutyl ether DANA mimetic. [c] 3-Hydroxypropyl ether DANA mimetic. [d] 2,3-Dihydroxypropyl ether DANA mimetic. [e] 2-Deoxy-2,3-didehydro-*N*-acetyl-neuraminic acid.

and the structure−activity relationships (SAR) of analogues of DANA, four anti-influenza drugs have been developed so far: zanamivir (Relenza),[22] oseltamivir (Tamiflu),[23] peramivir (Rapiacta),[24] and laninamivir (Inavir).[25] These drugs have the potential to affect the activities of endogenous hNEUs, which can recognize similar substrates. Indeed, zanamivir and oseltamivir have been reported to inihibit hNEUs at the micromolar level or more, while they exhibit inhibitory activity in the nanomolar range against influenza virus NAs.[26] Over a thousand adverse events have been reported so far to have occurred with the use of oseltamivir, including a small number of rare but severe neuropsychiatric adverse events, especially in children.[27] Therefore, it is of great importance to elucidate the differences in sensitivity to these anti-influenza drugs among human and influenza virus NAs from the aspect of the inhibitory interaction mechanism at the atomic and electronic levels.

We previously reported correlation analyses of the interaction energies of influenza A virus neuraminidase-1 (N1-NA) with a series of eight sialic acid analogues including oseltamivir involving ab initio FMO (fragment molecular orbital)-IFIE (interfragment interaction energy) calculations at the Møller−Plesset (MP2) level.[28] Our previous study revealed that the dispersion and/or hydrophobic interaction energies of the alkoxy side chain in each ligand with several amino acid residues in N1-NA, which surrounds the former, govern the overall free-energy change in the binding interaction. Recently, we proposed a novel quantitative structure−activity relationship (QSAR) procedure called linear expression by representative energy terms (LERE) involving molecular calculations such as ab initio FMO and generalized Born/surface area (GB/SA) ones.[29,30] The proposed LERE-QSAR procedure was demonstrated to provide detailed quantitative information as to the ligand−protein interaction of carbonic anhydrase with a series of substituted benzenesulfonamides at the atomic and electronic levels, which cannot be obtained with any other procedures.

There are a small number of compounds of which the inhibitory potency against hNEU2 have been measured and reported.[26,31−34]

In this work, we constructed complexes of hNEU2 with six sialic acid analogues exhibiting relatively potent inhibitory activity, zanamivir, oseltamivir, DANA, and three analogues of DANA, by using molecular mechanics calculations and then carried out ab initio FMO and continuum solvation model calculations for each complex structure. Subsequently, we performed LERE-QSAR analysis and quantitatively examined the binding mechanism at the atomic and electronic levels. Finally, we discuss the observed difference between the sensitivities of oseltamivir to hNEU2 and N1-NA in terms of structure and interaction energy.

2. METHODS

2.1. Compound Set. Table 1 shows the chemical structures of four sialic acid analogues [IEM (compound **1**), HEM, (**2**), DEM (**3**), and DANA (**4**)] and an anti-influenza drug, zanamivir (**5**), along with their inhibitory constants K_i against hNEU2.[33] Compounds **1**−**4** have a hydroxyl group as a common substituent (R_2, fragment **C**) and variable substituents (R_1, fragment **A**), hydroxyl groups existing in compounds **2**−**4**. Compounds **4** and **5** have an identical substituent R_1, but compound **5** has a guanidino group ($NHC(=NH_2^+)NH_2$) in fragment **C**, instead of the hydroxyl one of compounds **1**−**4**. Thus, a total of five compounds with the common skeletal structure of 3-acetamido-3,4-dihydro-2*H*-pyran-6-carboxylic acid (fragment **B**) were subjected to analyses. Another anti-influenza drug, oseltamivir (compound **6**), was added to the above compound set in Section 3.2.

2.2. Modeling of Complex Structures. The X-ray structures of complexes of hNEU2 with compounds **1** (PDB code: 2F11, complex **1**), **2** (2F12, complex **2**), **3** (2F13, complex **3**), **4** (DANA, 1VCU, complex **4**), and **5** (zanamivir, 2F0Z, complex **5**) are all available. The structure of hNEU2 varies slightly among the crystallographic structures of complexes **1**−**5**. The initial geometries of complexes **1**−**3** and **5** were taken from the crystallographic structure of complex **4** having a minimum number of missing residues, by replacing DANA in complex **4** with each of compounds **1**−**3** and **5**. Loops **A** (Ile103−Thr122) and **B**

(Val265−Ser275) in hNEU2 are located in the vicinity of fragment **A** in compound **4**. Loops **A** and **B** in the crystallographic structure of complex **5** occupy very close positions to those in complex **4**, respectively. However, loops **A** and **B** in the crystallographic structure of complex **2** take on significantly different positions from those in complexes **4** and **5**. Although several amino acid residues in loop **A** are not found (missing residues) in the crystallographic structures of complexes **1** and **3**, the atomic positions of the corresponding amino acid residues other than the missing ones are nearly the same among the crystallographic structures of complexes **1**, **2**, and **3**. Therefore, the initial geometries of complexes **1**−**3** were reconstructed by replacing loops **A** and **B** in the crystallographic structure of complex **4** with those in the crystallographic one of complex **2**.

In our previous work,[28] we showed that two crystal water molecules inside the active site of N1-NA play an important role in the binding interaction: ligands having guanidino and amino groups in fragment **C** bind with only one water molecule (W1) and with two molecules (W1 and W2), respectively. The X-ray structures of nine complexes of hNEU2 with ligands (PDB codes: 2F11, 2F12, 2F13, 2F25 (dimer), 2F27 (dimer), and 1VCU (dimer)) were carefully examined. The root-mean-square deviation values for the atomic positions of W1 and W2 are both within 0.44 Å, indicating that their positions were nearly fixed. W1 and W2 for complexes **1**−**4** and W1 for complex **5** were placed at the respective average positions determined by the above procedure. Five and four sodium ions (Na^+) were added to neutralize the total −5 and −4 charge of the complexes of hNEU2 with negatively charged ligands (compounds **1**−**4**) and with neutral ones (compounds **5** and **6**), respectively. Each complex was then solvated in a truncated octahedral box of TIP3P waters extending 12 Å from hNEU2.

Three-step energy minimizations were performed as follows: (i) hydrogen atoms, (ii) water molecules and counterions, and (iii) the entire system were sequentially relaxed for 2000, 6000, and 6000 steps, respectively. During these minimizations, the atomic positions of Glu111, Tyr179, Tyr181, Glu218, and Gln270, which could be involved in electrostatic and/or hydrogen-bonding interactions with fragment **A** in compounds **1**−**5**, were constrained with a harmonic potential of 10 kcal/mol/Å^2. The finally obtained energy minimized complex structures were used in the following calculations. All molecular mechanics calculations were carried out using the AMBER10 package[35] with parameters of parm99[36] and general AMBER force field (GAFF)[37] for hNEU2 and each compound, respectively. The X-ray structures of compounds **1**−**5** bound with hNEU2 were subjected to partial geometry optimization (bond lengths and angles), fixing dihedral angles to the corresponding X-ray ones with HF/6-31G* calculations (Gaussian03 program[38]). Similarly, that of compound **6** bound with N1-NA was geometry optimized. Then, partial atomic charges in each compound were determined according to the restrained electrostatic potential (RESP) fitting procedure[39−41] with HF/6-31G* calculations. The atomic charges in compounds **1**−**6** are consistent with those of atoms in amino acid residues, because the latter charges were determined with the same procedure in the Amber (parm99) force field generally used for proteins.[36,37] Atom names, atom types, and partial atomic charges of compounds **1**−**6** are listed in Table S1 (Supporting Information). The summation of partial atomic charges q within each fragment [Σq_i (**X**) (summation over atoms in fragment **X**), **X** = **A**, **B**, and **C**] for zanamivir (compound **5**) and oseltamivir (compound **6**) is similar with that reported by Udommaneethanakit et al.[41]

2.3. LERE Formalism. We recently proposed LERE analysis involving molecular calculations such as the ab initio FMO and GB/SA ones and reported LERE-QSAR analyses of a series of substituted benzenesulfonamides with bovine carbonic anhydrase.[29,30]

A basic assumption when dissecting free-energy terms is that they are all additive[28,42,43]

$$\Delta G_{obs} = \Delta G_{bind} + \Delta G_{sol} + \Delta G_{dis} + \Delta G_{others} \quad (I)$$

ΔG_{obs} on the left-hand side of eq I is the overall free-energy change obtained from the observed inhibitory constant K_i: $\Delta G_{obs} = 2.303\,RT \log K_i$ (T = 310 K). ΔG_{bind}, ΔG_{sol}, and ΔG_{dis} in eq I are the intrinsic interaction energy between an inhibitor and protein, the solvation free-energy change associated with complex formation, and the dissociation free-energy change of an inhibitor, respectively. ΔG_{others}, which represents the sum of free-energy terms such as the deformation energies of a protein and ligand other than ΔG_{bind}, ΔG_{sol}, and ΔG_{dis}, is assumed to be linear with the sum of representative free-energy terms ΔG_{bind}, ΔG_{sol}, and ΔG_{dis} in eq I, and ΔG_{others} is expected to act as a penalty term ($\beta < 0$), i.e., LERE approximation

$$\Delta G_{others} = \beta(\Delta G_{bind} + \Delta G_{sol} + \Delta G_{dis}) + const \text{ (LERE approximation)} \quad (II)$$

During complex formation between compounds **1**−**5** and hNEU2, ΔG_{dis} does not need to be considered in the current case, because a carboxylate group common to compounds **1**−**5** and a guanidino one in compound **5** take on nearly complete ionized forms [COO^- ($pK_a \sim 2.4$) and $NHC(=NH_2^+)NH_2$ ($pK_a \sim 13$), respectively] before and after complex formation at experimental pH = 5.6.[33,44]

ΔG_{obs} very probably exhibits a ubiquitous energetic relation on complex formation between a protein and a series of ligands having a similar structure, the entropy−enthalpy compensation rule[45−48]

$$T\Delta S_{obs} = \alpha\Delta H_{obs} + const \text{ (entropy−enthalpy compensation)} \quad (IIa)$$

The value of α ($\alpha > 0$) on complex formation between a sialic acid analogue and hNEU2 is supposed to be within the range of 0.70 ± 0.20, judging from the results of isothermal titration calorimetry (ITC) experiments on complex formation of various sugars and sugar mimetics with their binding proteins.[49−52] Frederick et al.[53] showed that a change in the protein conformational entropy (ΔS_{bind}, which includes changes in the vibrational, rotational, and translational entropies) is linearly correlated with a change in overall entropy on the formation of complexes of calmodulin with its binding peptides, using nuclear magnetic resonance (NMR) relaxation methods. They proposed the possibility that ΔS_{bind} makes a major contribution to the overall binding entropy change. In the case where the complex formation is mostly driven by the electrostatic interaction energy between a charged ligand and protein, the intrinsic binding enthalpy term ΔH_{bind} and the polar component of solvation free-energy change ΔG_{sol}^{pol}, calculated with continuum solvation models such as the generalized Born (GB) and Poisson−Boltzmann (PB) ones, are probably major contributors to the overall free-energy changes. Avbelj et al.[54] recently showed that the observed solvation (hydration) free-energy changes of the polar parts in several dipeptides consist mostly of the enthalpy term arising

from the electrostatic interaction, the energy of which can be well reproduced with PB type calculation. Directly from the original Born equation: $\Delta G_{sol}^{pol} = -(q^2/2r)(1 - 1/\varepsilon)$, the enthalpy change is given by the expression: $\Delta H_{sol} = \Delta G_{sol}^{pol} - T(\partial G_{sol}^{pol}/\partial T)_P = -(q^2/2r)[1 - 1/\varepsilon - (T/\varepsilon^2)(\partial\varepsilon/\partial T)_P]$, where r, q, ε, and T are the cavity radius formed by the ion in a particular solvent, ion charge, dielectric constant, and temperature, respectively.[55,56] The reported value of $[(T/\varepsilon)(\partial\varepsilon/\partial T)_P]$ for water (ε = 78.4) at 298 K is −1.357,[55] giving 0.983 for the ratio between ΔG_{sol}^{pol} and ΔH_{sol} ($\Delta G_{sol}^{pol}/\Delta H_{sol}$ = 0.983). The ratio close to unity suggests that ΔH_{sol} is generally replaceable by ΔG_{sol}^{pol} with satisfactory accuracy. We showed the possibility that the energy term (E_{sol}) calculated with the ab initio MO-continuum solvation model (COSMO) corresponds to the effective enthalpic energy term associated with the partition equilibrium between the aqueous and organic solvent phases rather than the overall solvation free-energy term, when we analyzed more than 200 log $P_{sol/w}$ values.[57]

ΔH_{bind} can be replaced by ΔE_{bind}, because of the lack of significant volume and pressure changes in solution. ΔE_{bind} consists of ΔE_{bind}^{HF} and E^{disp}, which are the intrinsic binding energy at the Hartree−Fock (HF) level and the dispersion (van der Waals) energy, respectively. ΔG_{sol} consists of ΔG_{sol}^{pol} and the nonpolar free-energy change ΔG_{sol}^{nonpol}. ΔE_{bind}^{HF} and ΔG_{sol}^{pol} are considered to be "electrostatic energy terms"[58] involved in the electrostatic interactions among a ligand, protein, and solvent medium. E^{disp} and ΔG_{sol}^{nonpol} are considered to be "local energy terms", which are effective only at a short distance. We reported that ΔG_{sol}^{pol} shows an excellent anticorrelation with ΔE_{bind}^{HF} during complex formation between carbonic anhydrase and a series of substituted benzenesulfonamides.[29,30] We assumed that the sum of the electrostatic energy terms represents the effective overall enthalpic change, i.e., the effective binding enthalpy change in a solvent, $\Delta H_{bind}^{sol} = \Delta E_{bind}^{HF} + \Delta G_{sol}^{pol}$, on the condition that contributions from the local energy terms to ΔH_{bind}^{sol} are significantly smaller than those from the electrostatic ones. This condition is equivalent to that the statistical variances of E^{disp} and ΔG_{sol}^{nonpol} are negligibly smaller than those of ΔE_{bind}^{HF} and ΔG_{sol}^{pol} in a correlation equation. The entropy−enthalpy compensation is now expressed as

$$T\Delta S_{bind}^{sol} = \alpha(\Delta E_{bind}^{HF} + \Delta G_{sol}^{pol}) + const \quad \text{(III)}$$

Assuming eq III, the effective overall free-energy change ΔG_{bind}^{sol} is expressed as $(1 - \alpha)(\Delta E_{bind}^{HF} + \Delta G_{sol}^{pol}) + const$. Although the protein conformational entropic term ($T\Delta S_{bind}$) is often estimated using normal mode and molecular dynamics calculations, the calculated one is known to exhibit significant large uncertainties.[59−61] Alternatively, eqs I, II, and III yield eq IV without estimation of $T\Delta S_{bind}$[62,63] directly

$$\Delta G_{obs} = (1 + \beta)(1 - \alpha)(\Delta E_{bind}^{HF} + \Delta G_{sol}^{pol}) + const \quad \text{(IV)}$$

Equation IV appears to take a similar form to the solvated interaction energy (SIE) model proposed by Naïm et al.[64] Although they introduced several adjusting parameters such as the effective interior dielectric constant in order to adjust the calculated overall free-energy change to the observed one, eq IV has only two scaling factors, α and β, which effectively and naturally consider the entropy−enthalpy compensation and penalty effects, respectively. It is obvious but noteworthy that the α and β values depend on a set of a series of ligands and their target proteins. Also, it should be noted that the above procedure is probably applicable to a congeneric series of ligands having the same skeleton but probably not to a set containing structurally diverse ones.

2.4. Calculation of Representative Energy Terms. ΔE_{bind}^{HF} (= E^{HF}(complex) − [E^{HF}(protein) + E^{HF}(ligand)]) and E^{disp} (= $\Delta E_{bind}^{MP2} - \Delta E_{bind}^{HF}$) were calculated using ab initio fragment molecular orbital (FMO) calculations with one residue (including W1 and W2) per FMO-fragment partition (FMO-IFIE-Scheme 1)[65−68] at the HF and MP2/6-31G levels, respectively, using the ABINIT-MP program.[69,70] Our previous studies involving FMO calculations[28−30,71,72] demonstrated that interfragment interaction energy (IFIE) analysis can provide valuable information independent from that obtained with the LERE-QSAR one. In the present work, the structures of compounds **1**−**5** were further divided into three FMO-fragments **A**, **B**, and **C** (FMO-IFIE-Scheme 2, shown in Table 1),[28,73−75] in order to clarify the contribution of each fragment to the overall interaction energy. Amino acid residues, which are closely located in fragments **A**, **B**, and **C**, are denoted by pockets **A**, **B**, and **C**, respectively. IFIE^{HF} (fragment **X**, Y) denotes the IFIE value between **X** (fragment **X** in a ligand) and Y (amino acid residue Y in a protein) at the HF level, and ΣIFIE^{HF} (fragment **X**, pocket **Z**) represents summation of IFIE^{HF} (fragment **X**, Y) over all the amino acid residues Y in pocket **Z**. As expected,[28,29,71,72] ΔE_{bind}^{HF} is almost completely linear with ΣΣIFIE^{HF} (fragment **X**, pocket **Z**) (summation over **X**, **Z** = **A**, **B**, and **C**) for compounds **1**−**5** (r = 0.999). This fact guarantees that the total intrinsic binding energy can be decomposed into local fragment−pocket interaction energies. The solvation free-energy terms ΔG_{sol}^{pol} and ΔG_{sol}^{nonpol} were calculated with the molecular dynamics (MD)-PB/SA (Poisson−Boltzmann/surface area) and MD-GB/SA (generalized Born/surface area) procedures[76,77] included in the AMBER10 package. These terms were obtained as ensemble averages of snapshot structures in the MD trajectory, $<G_{sol}^{pol\ (PB)}>$, $<G_{sol}^{pol\ (GB)}>$, and $<G_{sol}^{nonpol}>$. It is necessary to average G_{sol}, because G_{sol}^{pol} is supposed to be sensitive to subtle conformational changes occurring in domains other than the binding one.[62,63,78,79] Short-time MD with 200 ps of a production run at 300 K was performed to refine each complex structure (structure annealing), and snapshot structures were collected for the last 150 ps trajectory at 3.0 ps intervals. We confirmed that $<\Delta G_{sol}^{pol\ (PB)}>$ and $<\Delta G_{sol}^{pol\ (GB)}>$ were not significantly changed during a longer simulation time, by performing 1 and 10 ns simulations for (1) the complex of hNEU2 with compound **4** and (2) that of N1-NA with compound **6**, respectively. The absolute differences of $<\Delta G_{sol}^{pol}>$ between the 200 ps and longer time simulations are less than 1.4 and 3.3 kcal/mol for (1) and (2), respectively.

3. RESULTS AND DISCUSSION

3.1. LERE-QSAR Analysis of Complex Formation of hNEU2 with Sialic Acid Analogues and Zanamivir. Table 2 lists the observed and calculated overall free-energy changes along with the electrostatic interaction energy terms ΔE_{bind}^{HF} and $<\Delta G_{sol}^{pol}>$, and the local interaction energy terms E^{disp} and $<\Delta G_{sol}^{nonpol}>$, of compounds **1**−**5** listed in Table 1. Figure 1 shows the variations in these energy terms. The variances of ΔE_{bind}^{HF} and $<\Delta G_{sol}^{pol\ (PB)}>$, which mainly arise from the electrostatic interaction, are considerably large (3989 and 1967 kcal^2/mol^2, respectively). On the other hand, those of E^{disp} and $<\Delta G_{sol}^{nonpol}>$ are much smaller (5.06 and 0.01 kcal^2/mol^2, respectively). Thus, the variation in overall free-energy change is governed by ΔE_{bind}^{HF} and $<\Delta G_{sol}^{pol\ (PB)}>$, satisfying the condition for eq IV.

Table 2. Overall Free-Energy Change ΔG and Representative Energy Terms[a] (hNEU2)

compound no.	name	ΔG_{obs}[b]	ΔG_{calc}[c]	ΔG_{calc}[d]	ΔE_{bind}^{HF}	$<\Delta G_{sol}^{pol\ (PB)}>$[e]	$<\Delta G_{sol}^{pol\ (GB)}>$[e]	E^{disp}	$<\Delta G_{sol}^{nonpol}>$[e,f]
1	IEM	−4.33	−4.03	−3.95	−78.27	79.45 (6.79)	67.62 (7.06)	−38.88	−5.00 (0.05)
2	HEM	−4.44	−4.44	−4.44	−85.77	78.82 (4.76)	67.35 (5.31)	−37.32	−4.99 (0.05)
3	DEM	−4.05	−4.30	−4.41	−98.06	93.87 (7.66)	80.17 (8.50)	−40.34	−5.08 (0.06)
4	DANA	−5.47	−5.65	−5.86	−154.28	123.44 (9.27)	113.31 (8.93)	−36.81	−4.87 (0.11)
variance[g]					889.36	327.25	351.25	1.92	0.01
5	zanamivir	−6.77	−6.62	−6.38	−247.22	197.14 (7.71)	197.98 (6.05)	−43.03	−5.22 (0.05)
variance[g]					3989.09	1967.09	2429.14	5.06	0.01

[a] In kcal/mol. [b] $\Delta G_{obs} = 2.303\ RT \log K_i$ (T = 310 K). [c] Calculated from eq 1. [d] Calculated from eq 2. [e] Average value and standard deviation (in parentheses). [f] $\Delta G_{sol}^{nonpol} = \gamma\ \Delta ASA$, where ASA is the water accessible surface area and γ is taken to be 0.0072 kcal/mol/Å^2.[80] [g] In kcal^2/mol^2.

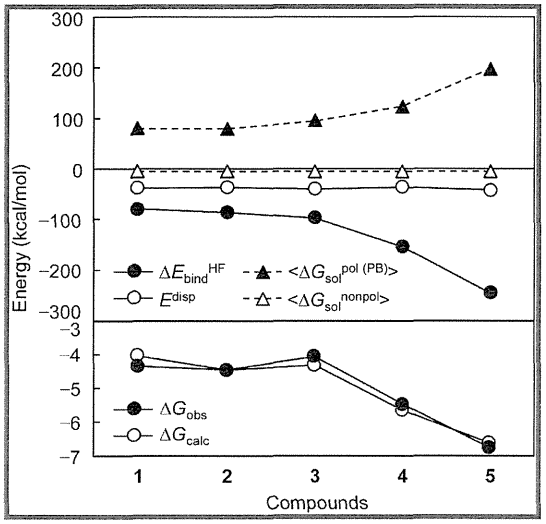


Figure 1. Variations of ΔE_{bind}^{HF}, $<\Delta G_{sol}^{pol\ (PB)}>$, E^{disp}, $<\Delta G_{sol}^{nonpol}>$, ΔG_{obs}, and ΔG_{calc} values (hNEU2). ΔG_{calc} is calculated by eq 1.

The correlation coefficient between ΔE_{bind}^{HF} and $<\Delta G_{sol}^{pol\ (PB)}>$ is −0.996 for compounds **1**−**5** (−0.988 for compounds **1**−**4**). This means that these two energy terms are dependent on and compensate for each other. The variance of ΔE_{bind}^{HF} for compounds **1**−**5** is ~200% larger than that of $<\Delta G_{sol}^{pol\ (PB)}>$ (~270% for compounds **1**−**4**), and the difference between $|\Delta E_{bind}^{HF}|$ and $<\Delta G_{sol}^{pol}>$ is distinctly large for compounds **4** (DANA) and **5** (zanamivir). That is, the stabilization due to ΔE_{bind}^{HF} overwhelms the destabilization due to $<\Delta G_{sol}^{pol}>$. The sum of ΔE_{bind}^{HF} and $<\Delta G_{sol}^{pol}>$ is excellently correlated with ΔG_{obs}, as shown in eqs 1 and 2

$$\Delta G_{obs} = (1 + \beta)(1 - \alpha)(\Delta E_{bind}^{HF} + <\Delta G_{sol}^{pol(PB)}>) - 4.09$$
$$n = 5, r = 0.979, s = 0.265, F = 68.3, (1 + \beta)(1 - \alpha) = 0.0505 \quad (1)$$

$$\Delta G_{obs} = (1 + \beta)(1 - \alpha)(\Delta E_{bind}^{HF} + <\Delta G_{sol}^{pol(GB)}>) - 3.28$$
$$n = 5, r = 0.940, s = 0.439, F = 22.9, (1 + \beta)(1 - \alpha) = 0.0630 \quad (2)$$

Although the statistical quality of eq 1 using PB is slightly better than that of eq 2 using GB, it is probably difficult and mostly pointless to discuss the difference, because each of the two computing procedures, PB and GB, involves several approximations and parameters. The correlation coefficient between $<\Delta G_{sol}^{pol\ (PB)}>$ and $<\Delta G_{sol}^{pol\ (GB)}>$ is 0.999, clearly showing that these two $<\Delta G_{sol}^{pol}>$ values are equivalent in terms of statistical correlation, at least. When compound **5**, which shows the largest $|\Delta E_{bind}^{HF}|$ and $|<\Delta G_{sol}^{pol}>|$, is removed (n = 4), the statistical qualities of eqs 1 and 2 are poorer but still statistically significant (r = 0.941 and 0.914, respectively). The absolute values of intercept in eqs 1 and 2 are small, 4.09 and 3.28, respectively. The β values in eqs 1 and 2 are similar, −0.83 and −0.79, respectively, when α is assumed to be 0.70.[49−52] These results indicate that the penalty energy term ΔG_{others} in eq II is nearly proportional to $(\Delta E_{bind}^{HF} + <\Delta G_{sol}^{pol}>)$ and makes a negative contribution to the observed overall free-energy change ΔG_{obs}, as we expected. As noted above, there is an excellent anticorrelation between ΔE_{bind}^{HF} and $<\Delta G_{sol}^{pol}>$, and the variation of the former overwhelms that of the latter in the current case.

We then attempted to determine the reason why compound **5** has the smallest ΔE_{bind}^{HF} value among compounds **1**−**5** in terms of structure and interaction. As shown in Table 2, the large variance in ΔE_{bind}^{HF} (889 kcal^2/mol^2) among compounds **1**−**4** arises from differences in fragment **A**, because these compounds are different only in fragment **A**. We carefully examined ΔE_{bind}^{HF} arising from interaction of fragment **A** with pocket **A** in hNEU2. Figure 2 shows the IFIE^{HF} values of fragment **A** with amino acid residues in pocket **A**. Pocket **A** in hNEU2 consists of Glu111, Tyr179, Tyr181, Leu217, Glu218, and Gln270. The side chain of Glu218 is located in close to fragment **A** (R_1) but does not form a stable hydrogen-bond with hydroxyl groups in fragment **A** of compounds **2**−**4**. The weak stabilization of IFIE^{HF} (fragment **A**, Glu218) commonly found in compounds **1**−**4** is probably due to the charge (electron) transfer (CT) from negatively charged Glu218 to neutral fragment **A** in compounds **1**−**4**. In addition to the effect of CT, weak hydrogen-bonding/electrostatic interactions between hydroxyl groups on fragment **A** in compounds **2**−**4** and oxygen atoms in the side chain of Glu218 are possibly involved. The IFIE^{HF} (fragment **A**, pocket **A**) values for compounds **1** and **2** indicate that fragment **A** in these compounds does not interact with pocket **A** effectively. The negative IFIE^{HF} (fragment **A**/compound **3**, Tyr179) value is due to a weak hydrogen-bonding interaction between the side chain of Tyr179 and the OH^1 group in fragment **A** (distance (H^{1}O---HO−Tyr179) = 2.04 Å). The OH^1 group attached to fragment **A** in compound **4** shows remarkably high attractive interaction energy with Glu111, because the OH^1 group and oxygen atoms in the side chain of Glu111 take on atomic positions optimum for effective

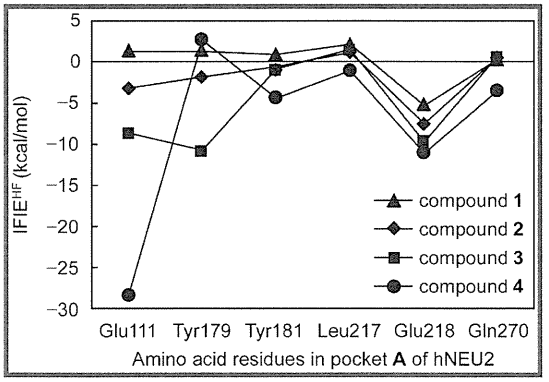


Figure 2. IFIE^{HF} values of fragment **A** in compounds **1**−**4** with amino acid residues in pocket **A** of hNEU2.

hydrogen-bonding and electrostatic interactions [distance (H^1---O−Glu111) = 1.54 Å, angle (O−H^1---O−Glu111) = 176 deg, and angle (H^1---O−C_δ−Glu111) = 116 deg].

The variance analysis of the IFIE^{HF} value between residue X and fragment **A** in a ligand among compounds **1**−**4** (denoted as var.[IFIE^{HF} (fragment **A**, X)]) shows that var.[IFIE^{HF} (fragment **A**, Glu111)], var.[IFIE^{HF} (fragment **A**, Tyr179)], and var.[IFIE^{HF} (fragment **A**, Arg237)] are 128.7, 27.1, and 24.8 kcal^2/mol^2, respectively, and that var.[IFIE^{HF} (fragment **A**, X)] values other than the above three ones are less than 5.5 kcal^2/mol^2. This result clearly suggests that these three residues govern the overall variation in the electrostatic interaction energy among compounds **1**−**4**. Glu111 is a key residue which interacts with the OH^1 group (fragment **A**) in compound **4** through the electrostatic and/or hydrogen-bonding interactions. In addition, Arg237 (pocket **B**) in close proximity to fragment **A** significantly contributes to the stabilization of the electrostatic interaction energy with compound **4**, but not with compounds **1**−**3**. In fact, IFIE^{HF} (fragment **A**, Arg237) of compound **4** is 10−13 kcal/mol smaller than those of compounds **1**−**3**. This is because the OH^2 group in compound **4** only takes favorable position for the electrostatic interaction with the positively charged side chain of Arg237.

Fragment **A** in compound **4** undergoes the largest stabilization through the hydrogen-bonding and electrostatic interactions. Consequently, compound **4** exhibits the most potent inhibitory activity among compounds **1**−**4**.

The above results quantitatively confirm that the hydrogen-bonding interaction energy of fragment **A** in compounds **1**−**4** with residues in pocket **A** mostly determines the observed overall free-energy change and suggest that the position of the hydroxyl group in fragment **A** is crucial for the optimum hydrogen-bond formation with residues in pocket **A**. Although the interaction energy of fragment **A** with amino acid resides in pocket **A** is not the most important contributor to the stabilization of the total binding energy, it should be noted that the structural factors that stabilize the total binding energy are different from those that govern the variation in the total binding energy among compounds **1**−**4**.

Next, we examined the most potent compound **5** (zanamivir), which has the same skeletal structure (fragment **B**) as compounds **1**−**4** but has a positively charged guanidino group instead of a neutral hydroxyl group (fragment **C**). Notably, compounds **4** and **5** have the same fragment **A** as well as fragment **B** structure. Differences in the interaction energy between compounds **4** and **5** are unquestionably attributable to those in fragment **C**.

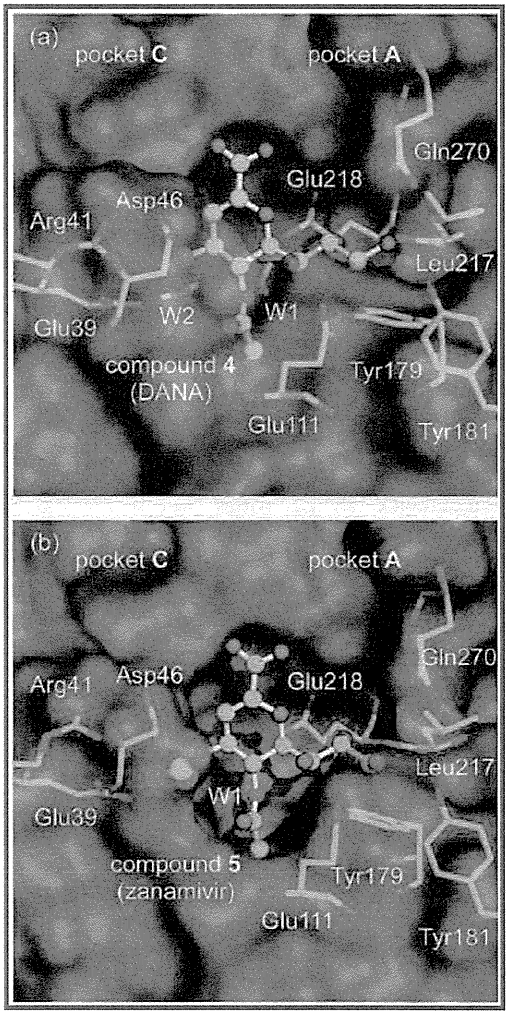


Figure 3. (a) Compound **4** (DANA) and (b) compound **5** (zanamivir) accommodated in the active site of hNEU2.

Figure 3 (a) and (b) shows the structures of compounds **4** and **5** bound at the active site of hNEU2, respectively. The guanidino group (fragment **C**) in compound **5** is involved in electrostatic interactions with Glu39 and Asp46 in hNEU2, and the corresponding hydroxyl group in compound **4** forms a hydrogen-bond with Arg41.

As can be seen in Figure 4, the IFIE^{HF} (fragment **C**/compound **5**, Glu39) and IFIE^{HF} (fragment **C**/compound **5**, Asp46) values represent large negative energies, and the IFIE^{HF} (fragment **C**/compound **5**, Arg41) value is the positive energy due to the electrostatic repulsive interaction between the positively charged guanidino group and the side chain of Arg41. The relatively small negative IFIE^{HF} (fragment **C**/compound **4**, Arg41) value represents the hydrogen-bonding interaction, which causes considerably weaker stabilization than electrostatic interactions of the guanidino group in compound **5** with Glu39 and Asp46. In total, $\Delta E_{bind}{}^{HF}$ of compound **5** is 93 kcal/mol more stable than that of compound **4**.

3.2. Comparison of Inhibitory Activities of Anti-Influenza Drugs against hNEU2 and N1-NA. Although anti-influenza drugs zanamivir (compound **5**) and oseltamivir (compound **6**) exhibit excellent inhibitory activity toward various types of influenza virus NAs including N1-NA, their activity toward

hNEU2 is much lower. It is notable that compound **6** has a skeletal structure (fragment **B**) different from that in compounds **1**−**5**. Table 3 shows the ΔG_{obs} values of compounds **5** and **6** with hNEU2 and N1-NA. The large difference in ΔG_{obs} between compound **5** with hNEU2 and N1-NA as well as that between compound **6** with hNEU2 and N1-NA can be attributed to those between the binding interaction modes of hNEU2 and N1-NA with the two compounds.

The whole protein structures of hNEU2 and N1-NA take on a close topology: six bladed β-propeller folds commonly conserved among neuraminidases are arranged at topologically almost equivalent positions. Although the overall amino acid sequence homology between hNEU2 and N1-NA is considerably low (identity ~16%, similarity ~25%), several residues in their active sites seem to be conserved: an "arginine triad" (Arg21, 237, and 304 in hNEU2; Arg118, 292, and 371 in N1-NA) and tyrosine (Tyr334 in hNEU2; Tyr406 in N1-NA), which are essential for expression of their catalytic activities. We carefully compared the atomic positions of amino acid residues located in the active sites of the two PDB crystal structures, 1VCU (hNEU2)[21] and 2HU4 (N1-NA).[82] Each of the above corresponding pairs of residues takes on nearly the same position in hNEU2 and N1-NA (within 0.76 Å (C_α atoms)). Figure 5 (a) and (b) schematically represents the interactions of compound **6** with hNEU2 and N1-NA, respectively.

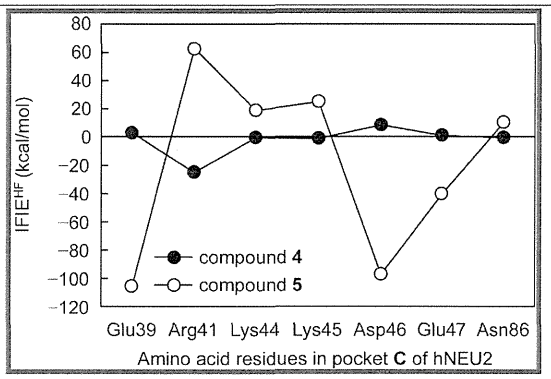


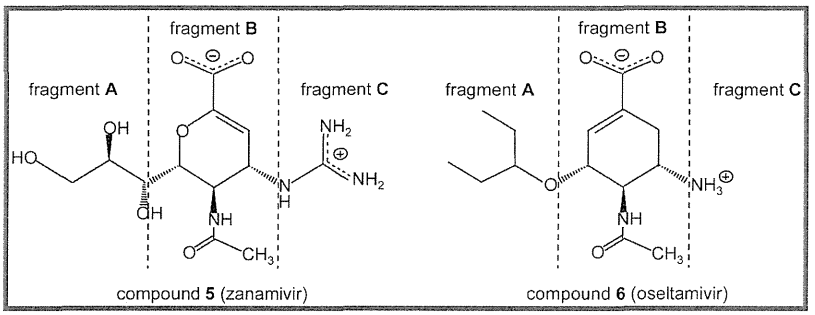
Figure 4. $IFIE^{HF}$ values of fragment **C** in compounds **4** (DANA) and **5** (zanamivir) with amino acid residues in pocket **C** of hNEU2.

In pocket **B**, one Tyr and three Arg residues (arginine triad) common in hNEU2 and N1-NA surround a negatively charged carboxylate group in fragment **B** of compound **6**, exhibiting a very close geometry.

A few amino acid residues in pocket **A** are different between hNEU2 and N1-NA. For example, a negatively charged Glu111 in hNEU2 is replaced by nonpolar Ile222 in N1-NA. Although fragment **A** in compound **5** as well as compound **4** form a hydrogen-bond with residues in pocket **A** of hNEU2, fragment **A** in compound **6** is not able to form a hydrogen-bond because it does not have a hydroxyl group. Both compounds **5** and **6** exhibit potent inhibitory potency as to N1-NA, probably because the hydrogen-bonding interaction between fragment **A** and pocket **A** in N1-NA is not dominant, but the dispersion type of interaction (E^{disp} and $\Delta G_{sol}{}^{nonpol}$) governs the interaction energy between fragment **A** and pocket **A** in N1-NA, as demonstrated in our previous study.[28]

We then examined the differences between the $IFIE^{HF}$ (fragment **C** (NH_3^+)/compound **6**, pocket **C**/hNEU2) and $IFIE^{HF}$ (fragment **C**/compound **6**, pocket **C**/N1-NA) values. Figure 6 shows the $IFIE^{HF}$ (fragment **C**/compound **6**, pocket **C**/hNEU2) and $IFIE^{HF}$ (fragment **C**/compound **6**, pocket **C**/N1-NA) values. The $\Sigma IFIE^{HF}$ (fragment **C**, pocket **C**/hNEU2) and $\Sigma IFIE^{HF}$ (fragment **C**, pocket **C**/N1-NA) values are mostly due to those of fragment **C** with ionized residues: negatively charged Glu39 and Asp46 and positively charged Arg41 in hNEU2 and negatively charged Glu119, Asp151, and Glu227 and positively charged Arg156 in N1-NA. Asp46 in hNEU2 and Asp151 in N1-NA stabilize, and Arg41 in hNEU2 and Arg156 in N1-NA destabilize the interaction energy. The above stabilization and destabilization commonly occur in the two complexes. Ile22, Glu39, Met85, and Asn86 in pocket **C** of hNEU2 are replaced by Glu119, Leu134, Trp178, and Glu227 in N1-NA, respectively. Among the above replacements, that from neutral Ile22 in hNEU2 to negatively charged Glu119 in N1-NA makes the interaction energy of fragment **C** with N1-NA 105 kcal/mol more stable than that with hNEU2. Similarly, the replacement of Asn86 with Glu227 makes the interaction energy 54 kcal/mol

Table 3. Chemical Structure of Zanamivir and Oseltamivir along with Inhibitory Potency against hNEU2 and N1-NA



compound				$\Delta G_{obs}{}^a$	
no.	name	R_1 (fragment **A**)	R_2 (fragment **C**)	hNEU2[b]	N1-NA[c]
5	zanamivir	$CH(OH^1)CH(OH^2)CH_2(OH^3)$	$NHC(=NH_2^+)NH_2$	−6.77	−12.80
6	oseltamivir	$OCH(CH_2CH_3)_2$	NH_3^+	−3.26	−13.26

[a] In kcal/mol. [b] $\Delta G_{obs} = 2.303\ RT \log K_i$ (T = 310 K). K_i value is taken from ref 33. [c] $\Delta G_{obs} = 2.303\ RT \log (IC_{50})$ (T = 310 K). IC_{50} value is taken from ref 81.

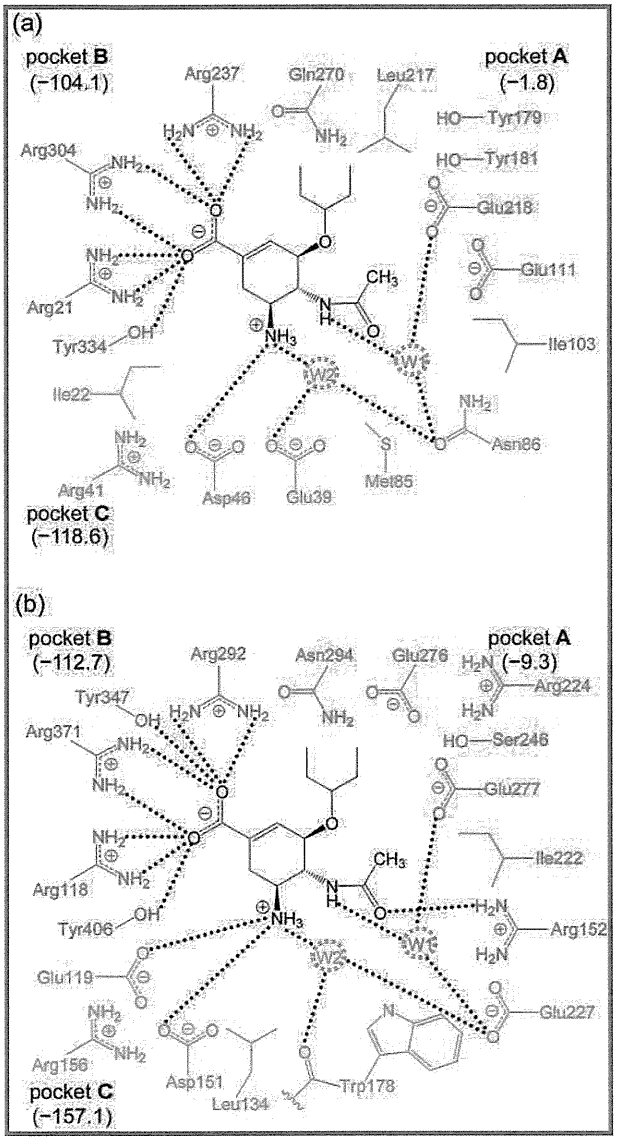


Figure 5. Interaction of compound **6** (oseltamivir) with amino acid residues in the active sites of (a) hNEU2 and (b) N1-NA. $\Sigma IFIE^{HF}$ (fragment **X**, pocket **X**: **X** = **A** (red), **B** (blue), and **C** (green)) values are shown in parentheses (kcal/mol). W1 and W2 represent water molecules in the active sites. A dotted line represents an attractive interaction.

more stable. Conversely, the replacement of negatively charged Glu39 with neutral Leu134 stabilizes the complex of hNEU2 with compound **6** by 70 kcal/mol. The energy differences $\Delta\Sigma IFIE^{HF}$ (= $\Sigma IFIE^{HF}$ (fragment **C**/compound **6**, pocket **C**/hNEU2) − $\Sigma IFIE^{HF}$ (fragment **C**/compound **6**, pocket **C**/N1-NA)) reaches a stabilization value of 38 kcal/mol.

In summary, compound **6** shows remarkably less inhibitory activity toward hNEU2 than that toward N1-NA, because of differences in pockets **A** and **C** between the two proteins.

We finally examined the difference between the inhibitory activities of compounds **5** and **6** with hNEU2. In our previous paper,[28] we showed that $\Sigma IFIE^{HF}$ (fragment **C**/compound **6**, pocket **C**/N1-NA) is not far from $\Sigma IFIE^{HF}$ (fragment **C** in the compound where the amino group in compound **6** is replaced by a guanidino one, pocket **C**/N1-NA) (−157 and −145 kcal/mol, respectively). However, there is a significantly large difference

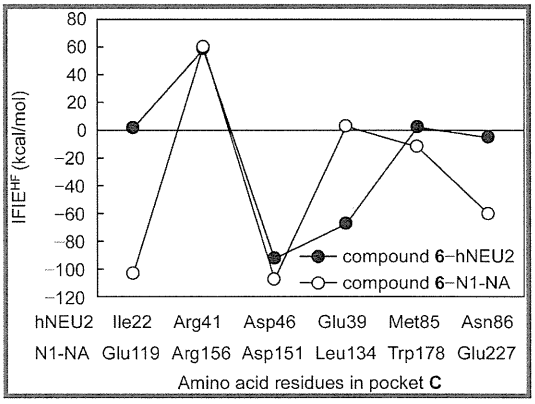


Figure 6. $IFIE^{HF}$ values of fragment **C** in compound **6** (oseltamivir) with amino acid residues in pocket **C** of hNEU2 and N1-NA.

between $\Sigma IFIE^{HF}$ (fragment **C** (NH_3^+)/compound **6**, pocket **C**/hNEU2) and $\Sigma IFIE^{HF}$ (fragment **C** ($NHC(=NH_2^+)NH_2$)/compound **5**, pocket **C**/hNEU2). This difference between $\Sigma IFIE^{HF}$ (fragment **C**/compound **6**, pocket **C**/hNEU2) and $\Sigma IFIE^{HF}$ (fragment **C**/compound **5**, pocket **C**/hNEU2) (−119 and −149, respectively) mostly arises from the interaction energy of fragment **C** with Glu39 [$IFIE^{HF}$ (fragment **C**/compound **6**, Glu39) and $IFIE^{HF}$ (fragment **C**/compound **5**, Glu39) are −68 (see Figure 6) and −106 (Figure 4) kcal/mol, respectively]. This is because the positively charged guanidino group in compound **5** is located closer to negatively charged Glu39 in hNEU2 than the amino group in compound **6**. The nearest interatomic (heavy atoms) distances between Glu39 and fragment **C** are 2.76 and 4.48 Å for compounds **5** and **6**, respectively. The above differences as well as those in the interactions between fragment **A** and pocket **A** make the inhibitory activity of compound **6** toward hNEU2 considerably less potent than that of compound **5**.

4. SUMMARY AND CONCLUSION

In our previous study,[28] we showed that the variation in the inhibitory potency as to N1-NA among eight analogues (including oseltamivir) having alkoxy side chains (OC_nH_{2n+1}, n = 0, 3, 4, and 5) at R_1 (fragment **A**) is governed by the dispersion (E^{disp}) and/or hydrophobic interaction ($\Delta G_{sol}{}^{nonpol}$) energies. In the present study, we showed that the observed overall free-energy changes of DANA, three analogues of DANA, and an anti-influenza drug zanamivir, all of which have a common skeletal structure (fragment **B**), are linearly correlated with the calculated one using FMO and MD-PB(GB)/SA calculations (LERE-QSAR). Unlike the case reported in the previous study, the statistical variances of the local energy terms E^{disp} and $<\Delta G_{sol}{}^{nonpol}>$ are both 3−6 orders of magnitude smaller than those of the electrostatic energy terms $\Delta E_{bind}{}^{HF}$ and $<\Delta G_{sol}{}^{pol}>$. Because of penalty and entropy−enthalpy compensation effects, the coefficient of ($\Delta E_{bind}{}^{HF} + <\Delta G_{sol}{}^{pol}>$) in the correlation equations eqs 1 and 2 is considerably smaller than unity (~0.05). There is an excellent anticorrelation between the electrostatic energy terms $\Delta E_{bind}{}^{HF}$ and $<\Delta G_{sol}{}^{pol}>$ (r = ~ −0.99). Comparison of the variances of the two electrostatic energy terms showed that the contribution (stabilization) from the intrinsic binding energy term $\Delta E_{bind}{}^{HF}$ to the overall free-energy change overwhelms that (destabilization) from the polar solvation free-energy term $<\Delta G_{sol}{}^{pol}>$. $\Delta E_{bind}{}^{HF}$ can be expressed as the sum of the FMO-IFIE terms for an inhibitor and hNEU2. FMO-IFIE analysis revealed that the

variation in the overall free-energy changes among four silalic acid analogues is determined by hydrogen-bonding interactions between fragment **A** in a sialic acid analogue and amino acid residues in pocket **A** in hNEU2. DANA and zanamivir, both of which have the same fragment **A** structure, show excellent inhibitory activities, because of the effective hydrogen-bond formation of a hydroxyl group in fragment **A** with Glu111 in hNEU2. The hydroxyl group in fragment **C** of DANA is replaced by a positively charged guanidino group in zanamivir. This replacement causes great stabilization of the interaction energy through the electrostatic interactions of the guanidino group with negatively charged Glu39 and Asp46, resulting in that zanamivir exhibits more potent activity than DANA.

Second, we determined the reason why an anti-influenza drug, oseltamivir, shows very weak inhibitory activity toward hNEU2, although it exhibits potent activity toward N1-NA. The observed free-energy change for complex formation of hNEU2 with oseltamivir is 10 kcal/mol larger (less stable) than that in the case of N1-NA. There are significant differences in amino acid residues in the active site between the two proteins: notably Ile22, Glu39, and Asn86 in hNEU2 are Glu119, Leu134, and Glu227 in N1-NA, respectively. These three residues surround the hydroxyl and guanidino groups in fragment **C** of DANA and zanamivir, respectively. IFIE analysis showed that these replacements cause a large difference in the free-energy change between the oseltamivir−hNEU2 and oseltamivir−N1-NA complexes. In addition to the difference in the interaction between fragment **C** and pocket **C**, fragment **A** in oseltamivir, unlike that in zanamivir, is not able to form hydrogen bonds with pocket **A**. These differences reasonably account for the inefficiency of oseltamivir as to hNEU2.

The present study could provide clues for understanding the neuropsychiatric side effects of oseltamivir, which are similar to the known symptoms of endogenous hNEUs related disorders. Li et al.[83] proposed that the hNEU2 variation caused by the Arg41Gln mutation could be associated with the adverse side effects of oseltamivir. They demonstrated that replacement of positively charged Arg41 with neutral Gln41 increases the sensitivity of hNEU2 to oseltamivir and reduces the catalytic activity of hNEU2. This observation is strongly supported by the present results, which suggest that the unfavorable electrostatic repulsive interaction between Arg41 and the positively charged amino group in oseltamivir can be abolished with the mutation from Arg41 to Gln41. In fact, $IFIE^{HF}$ (fragment **C**, Gln41) is 67 kcal/mol more stable than $IFIE^{HF}$ (fragment **C**, Arg41) (−9 and 58 kcal/mol, respectively).

The current and previous results[28] obtained with the LERE-QSAR analysis successfully reveal the quantitative contribution of each fragment in neuraminidases (hNEU2 and N1-NA) and sialic acid analogues including anti-influenza drugs to the overall interaction free-energy change. Consequently, the results bring a new and consistent understanding of the binding interaction mechanism, which cannot be obtained with any other procedures.

The present work will provide useful information for drug development and for the curing of influenza and other diseases in which neuraminidases are involved.

■ ASSOCIATED CONTENT

Supporting Information. Atom names, atom types, and partial atomic charges of compounds **1**−**6**. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Phone: +81-88-633-7257. Fax: +81-88-633-9508. E-mail: hchuman@ph.tokushima-u.ac.jp.

■ ACKNOWLEDGMENT

This work was supported by the Ministry of Agriculture, Forestry and Fisheries (Agri-Health Translational Research Project) and Grants-in-Aid for Scientific Research (No. 20590036) from the Ministry of Education, Culture, Sports, Science and Technology.

■ REFERENCES

(1) Achyuthan, K. E.; Achyuthan, A. M. Comparative enzymology, biochemistry and pathophysiology of human *exo*-α-sialidases (neuraminidases). *Comp. Biochem. Physiol., Part B: Biochem. Mol. Biol.* **2001**, *129*, 29–64.

(2) De Clercq, E. Antiviral agents active against influenza A viruses. *Nat. Rev. Drug Discovery* **2006**, *5*, 1015–1025.

(3) von Itzstein, M. The war against influenza: discovery and development of sialidase inhibitors. *Nat. Rev. Drug Discovery* **2007**, *6*, 967–974.

(4) Liu, Y.; Zhang, J.; Xu, W. Recent progress in rational drug design of neuraminidase inhibitors. *Curr. Med. Chem.* **2007**, *14*, 2872–2891.

(5) Monti, E.; Preti, A.; Venerando, B.; Borsani, G. Recent development in mammalian sialidase molecular biology. *Neurochem. Res.* **2002**, *27*, 649–663.

(6) Miyagi, T.; Wada, T.; Yamaguchi, K.; Hata, K. Sialidase and malignancy: a minireview. *Glycoconjugate J.* **2004**, *20*, 189–198.

(7) Buschiazzo, A.; Alzari, P. M. Structural insights into sialic acid enzymology. *Curr. Opin. Chem. Biol.* **2008**, *12*, 565–572.

(8) Monti, E.; Bonten, E; d'Azzo, A; Bresciani, R; Venerando, B; Borsani, G; Schauer, R; Tettamanti, G. Sialidases in vertebrates: a family of enzymes tailored for several cell functions. *Adv. Carbohydr. Chem. Biochem.* **2010**, *64*, 403–479.

(9) Meindl, P.; Tuppy, H. 2-Deoxy-2,3-dehydrosialic acids. II. Competitive inhibition of *Vibrio cholerae* neuraminidase by 2-deoxy-2,3-dehydro-*N*-acetylneuraminic acids. *Hoppe-Seyler's Z. Physiol. Chem.* **1969**, *350*, 1088–1092.

(10) Meindl, P.; Bodo, G.; Palese, P.; Schulman, J.; Tuppy, H. Inhibition of neuraminidase activity by derivatives of 2-deoxy-2,3-dehydro-*N*-acetylneuraminic acid. *Virology* **1974**, *58*, 457–463.

(11) Holzer, C. T.; von Itzstein, M.; Jin, B.; Pegg, M. S.; Stewart, W. P.; Wu, W-.Y. Inhibition of sialidases from viral, bacterial and mammalian sources by analogues of 2-deoxy-2,3-dihydro-*N*-acetylneuraminic acid modified at the C-4 position. *Glycoconjugate J.* **1993**, *10*, 40–44.

(12) Miyagi, T.; Kato, K.; Ueno, S.; Wada, T. Aberrant expression of sialidase in cancer. *Trends Glycosci. Glycotechnol.* **2004**, *16*, 371–381.

(13) Pshezhetsky, A. V.; Richard, C.; Michaud, L.; Igdoura, S.; Wang, S.; Elsliger, M.-A.; Qu, J.; Leclerc, D.; Gravel, R.; Dallaire, L.; Potier, M. Cloning, expression and chromosomal mapping of human lysosomal sialidase and characterization of mutations in sialidosis. *Nat. Genet.* **1997**, *15*, 316–320.

(14) Monti, E.; Preti, A.; Rossi, E.; Ballabio, A.; Borsani, G. Cloning and characterization of NEU2, a human gene homologous to rodent soluble sialidases. *Genomics* **1999**, *57*, 137–143.

(15) Wada, T.; Yoshikawa, Y.; Tokuyama, S.; Kuwabara, M.; Akita, H.; Miyagi, T. Cloning, expression, and chromosomal mapping of a human ganglioside sialidase. *Biochem. Biophys. Res. Commun.* **1999**, *261*, 21–27.

(16) Monti, E.; Bassi, M. T.; Bresciani, R.; Civini, S.; Croci, G. L.; Papini, N.; Riboni, M.; Zanchetti, G.; Ballabio, A.; Preti, A.; Tettamanti, G.; Venerando, B.; Borsani, G. Molecular cloning and characterization of NEU4, the fourth member of the human sialidase gene family. *Genomics* **2004**, *83*, 445–453.

(17) Miyagi, T. Aberrant expression of sialidase and cancer progression. *Proc. Jpn. Acad., Ser. B* **2008**, *84*, 407–418.

(18) Pshezhetsky, A. V.; Elsliger, M.-A.; Vinogradova, M. V.; Potier, M. Human lysosomal β-galactosidase-cathepsin A complex: definition of the β-galactosidase–binding interface on cathepsin A. *Biochemistry* **1995**, *34*, 2431–2440.

(19) Bonten, E. J.; Campos, Y.; Zaitsev, V.; Nourse, A.; Waddell, B.; Lewis, W.; Taylor, G.; d'Azzo, A. Heterodimerization of the sialidase NEU1 with the chaperone protective protein/cathepsin A prevents its premature oligomerization. *J. Biol. Chem.* **2009**, *284*, 28430–28441.

(20) d'Azzo, A.; Hoogeveen, A.; Reuser, A. J.; Robinson, D.; Galjaard, H. Molecular defect in combined β-galactosidase and neuraminidase deficiency in man. *Proc. Natl. Acad. Sci. U.S.A.* **1982**, *79*, 4535–4539.

(21) Chavas, L. M. G.; Tringali, C.; Fusi, P.; Venerando, B.; Tettamanti, G.; Kato, R.; Monti, E.; Wakatsuki, S. Crystal structure of the human cytosolic sialidase Neu2. Evidence for the dynamic nature of substrate recognition. *J. Biol. Chem.* **2005**, *7*, 469–475.

(22) von Itzstein, M.; Wu, W.-Y.; Kok, G. B.; Pegg, M. S.; Dyason, J. C.; Jin, B.; Phan, T. V.; Smythe, M. L.; White, H. F.; Oliver, S. W.; Colman, P. M.; Varghese, J. N.; Ryan, D. M.; Woods, J. M.; Bethell, R. C.; Hotham, V. J.; Cameron, J. M.; Penn, C. R. Rational design of potent sialidase-based inhibitors of influenza virus replication. *Nature* **1993**, *363*, 418–423.

(23) Kim, C. U.; Lew, W.; Williams, M. A.; Liu, H.; Zhang, L.; Swaminathan, S.; Bischofberger, N.; Chen, M. S.; Mendel, D. B.; Tai, C. Y.; Laver, W. G.; Stevens, R. C. Influenza neuraminidase inhibitors possessing a novel hydrophobic interaction in the enzyme active site: design, synthesis, and structural analysis of carbocyclic sialic acid analogues with potent anti-influenza activity. *J. Am. Chem. Soc.* **1997**, *119*, 681–690.

(24) Babu, Y. S.; Chand, P.; Bantia, S.; Kotian, P.; Dehghani, A.; El-Kattan, Y.; Lin, T.-H.; Hutchison, T. L.; Elliott, A. J.; Parker, C. D.; Ananth, S. L.; Horn, L. L.; Laver, G. W.; Montgomery, J. A. BCX-1812 (RWJ-270201): discovery of a novel, highly potent, orally active, and selective influenza neuraminidase inhibitor through structure-based drug design. *J. Med. Chem.* **2000**, *43*, 3482–3486.

(25) Honda, T.; Masuda, T.; Yoshida, S.; Arai, M.; Kobayashi, Y.; Yamashita, M. Synthesis and anti-influenza virus activity of 4-guanidino-7-substituted Neu5Ac2en derivatives. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 1921–1924.

(26) Hata, K.; Koseki, K.; Yamaguchi, K.; Moriya, S.; Suzuki, Y.; Yingsakmongkon, S.; Hirai, G.; Sodeoka, M.; von Itzstein, M.; Miyagi, T. Limited inhibitory effects of oseltamivir and zanamivir on human sialidases. *Antimicrob. Agents Chemother.* **2008**, *52*, 3484–3491.

(27) Edwards, E. T.; Truffa, M. M. One-year post pediatric exclusivity postmarketing adverse events review drug: oseltamivir phosphate, 2005. U.S. Food and Drug Administration Centre for Drug Evaluation and Research. http://www.fda.gov/ohrms/dockets/AC/05/briefing/2005-4180b_06_01_Tamiflu%20AE_reviewed.pdf (accessed May 1, 2011).

(28) Hitaoka, S.; Harada, M.; Yoshida, T.; Chuman, H. Correlation analyses on binding affinity of sialic acid analogues with influenza virus neuraminidase-1 using ab initio MO calculations on their complex structures. *J. Chem. Inf. Model.* **2010**, *50*, 1796–1805.

(29) Yoshida, T.; Munei, Y.; Hitaoka, S.; Chuman, H. Correlation analyses on binding affinity of substituted benzenesulfonamides with carbonic anhydrase using ab initio MO calculations on their complex structures. *J. Chem. Inf. Model.* **2010**, *50*, 850–860.

(30) Munei, Y.; Shimamoto, K.; Harada, M.; Yoshida, T.; Chuman, H. Correlation analyses on binding affinity of substituted benzenesulfonamides with carbonic anhydrase using ab initio MO calculations on their complex structures (II). *Bioorg. Med. Chem. Lett.* **2011**, *21*, 141–144.

(31) Magesh, S.; Moriya, S.; Suzuki, T.; Miyagi, T.; Ishida, H.; Kiso, M. Design, synthesis, and biological evaluation of human sialidase inhibitors. Part 1: selective inhibitors of lysosomal sialidase (NEU1). *Bioorg. Med. Chem. Lett.* **2008**, *18*, 532–537.

(32) Magesh, S.; Savita, V.; Moriya, S.; Suzuki, T.; Miyagi, T.; Ishida, H.; Kiso, M. Human sialidase inhibitors: design, synthesis, and biological evaluation of 4-acetamido-5-acylamido-2-fluoro benzoic acids. *Bioorg. Med. Chem.* **2009**, *17*, 4595–4603.

(33) Chavas, L. M. G.; Kato, R.; Suzuki, N.; von Itzstein, M.; Mann, M. C.; Thomson, R. J.; Dyason, J. C.; McKimm-Breschkin, J.; Fusi, P.; Tringali, C.; Venerando, B.; Tettamanti, G.; Monti, E.; Wakatsuki, S. Complexity in influenza virus targeted drug design: interaction with human sialidases. *J. Med. Chem.* **2010**, *53*, 2998–3002.

(34) Li, Y.; Cao, H.; Yu, H.; Chen, Y.; Lau, K.; Qu, J.; Thon, V.; Sugiarto, G.; Chen, X. Identifying selective inhibitors against the human cytosolic sialidase NEU2 by substrate specificity studies. *Mol. Biosyst.* **2011**, *7*, 1060–1072.

(35) Case, D. A.; Cheatham, T. E., III.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M., Jr.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. The Amber biomolecular simulation programs. *J. Comput. Chem.* **2005**, *26*, 1668–1688.

(36) Wang, J.; Cieplak, P.; Kollman, P. A. How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J. Comput. Chem.* **2000**, *21*, 1049–1074.

(37) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and testing of a general Amber force field. *J. Comput. Chem.* **2004**, *25*, 1157–1174.

(38) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. Gaussian 03, Revision C.02; Gaussian, Inc.: Wallingford, CT, 2004.

(39) Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *J. Phys. Chem.* **1993**, *97*, 10269–10280.

(40) Singh, U. C.; Kollman, P. A. An approach to computing electrostatic charges for molecules. *J. Comput. Chem.* **1984**, *5*, 129–145.

(41) Udommaneethanakit, T.; Rungrotmongkol, T.; Bren, U.; Frecer, V.; Stanislav, M. Dynamic behavior of avian influenza A virus neuraminidase subtype H5N1 in complex with oseltamivir, zanamivir, peramivir, and their phosphonate analogues. *J. Chem. Inf. Model.* **2009**, *49*, 2323–2332.

(42) Bren, U.; Martínek, V.; Florián, J. Decomposition of the solvation free energies of deoxyribonucleoside triphosphates using the free energy perturbation method. *J. Phys. Chem. B* **2006**, *110*, 12782–12788.

(43) Bren, M.; Florián, J.; Mavri, J.; Bren, U. Do all pieces make a whole? Thiele cumulants and the free energy decomposition. *Theor. Chem. Acc.* **2007**, *117*, 535–540.

(44) Tringali, C.; Papini, N.; Fusi, P.; Croci, G.; Borsani, G.; Preti, A.; Tortora, P.; Tettamanti, G.; Venerando, B.; Monti, E. Properties of recombinant human cytosolic sialidase HsNEU2. The enzyme hydrolyzes monomerically dispersed GM1 ganglioside molecules. *J. Biol. Chem.* **2004**, *279*, 3169–3179.

(45) Dunitz, J. D. Win some, lose some: enthalpy–entropy compensation in weak intermolecular interactions. *Chem. Biol.* **1995**, *2*, 709–712.

(46) Gallicchio, E.; Kubo, M. M.; Levy, R. M. Entropy−enthalpy compensation in solvation and ligand binding revisited. *J. Am. Chem. Soc.* **1998**, *120*, 4526–4527.
(47) Sharp, K. Entropy−enthalpy compensation: fact or artifact? *Protein Sci.* **2001**, *10*, 661–667.
(48) Freire, E. A thermodynamic approach to the affinity optimization of drug candidates. *Chem. Biol. Drug Des.* **2009**, *74*, 468–472.
(49) Schwarz, F. P.; Puri, K. D.; Bhat, R. G.; Surolia, A. Thermodynamics of monosaccharide binding to concanavalin A, pea (Pisum sativum) lectin, and lentil (Lens culinaris) lectin. *J. Biol. Chem.* **1993**, *268*, 7668–7677.
(50) Gupta, D.; Cho, M.; Cummings, R. D.; Brewer, C. F. Thermodynamics of carbohydrate binding to galectin-1 from Chinese hamster ovary cells and two mutants. A comparison with four galactose-specific plant lectins. *Biochemistry* **1996**, *35*, 15236–15243.
(51) Swaminathan, C. P.; Surolia, N.; Surolia, A. Role of water in the specific binding of mannose and mannooligosaccharides to concanavalin A. *J. Am. Chem. Soc.* **1998**, *120*, 5153–5159.
(52) Gloster, T. M.; Meloncelli, P.; Stick, R. V.; Zechel, D.; Vasella, A.; Davies, G. J. Glycosidase inhibition: an assessment of the binding of 18 putative transition-state mimics. *J. Am. Chem. Soc.* **2007**, *129*, 2345–2354.
(53) Frederick, K. K.; Marlow, M. S.; Valentine, K. G.; Wand, A. J. Conformational entropy in molecular recognition by proteins. *Nature* **2007**, *448*, 325–329.
(54) Avbelj, F.; Baldwin, R. L. Origin of the change in solvation enthalpy of the peptide group when neighboring peptide groups are added. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 3137–3141.
(55) Noyes, R. M. Thermodynamics of ion hydration as a measure of effective dielectric properties of water. *J. Am. Chem. Soc.* **1962**, *84*, 513–522.
(56) Rashin, A. A.; Honig, B. Reevaluation of the Born model of ion hydration. *J. Phys. Chem.* **1985**, *89*, 5588–5593.
(57) Chuman, H.; Mori, A.; Tanaka, H.; Yamagami, C.; Fujita, T. Analyses of the partition coefficient, log P, using ab initio MO parameter and accessible surface area of solute molecules. *J. Pharm. Sci.* **2004**, *93*, 2681–2697.
(58) Kukić, P.; Nielsen, J. E. Electrostatics in proteins and protein-ligand complexes. *Future Med. Chem.* **2010**, *2*, 647–666.
(59) Case, D. A. Normal mode analysis of protein dynamics. *Curr. Opin. Struct. Biol.* **1994**, *4*, 285–290.
(60) Hayward, S.; Kitao, A.; Go, N. Harmonicity and anharmonicity in protein dynamics: a normal mode analysis and principal component analysis. *Proteins* **1995**, *23*, 177–186.
(61) Grünberg, R.; Nilges, M.; Leckner, J. Flexibility and conformational entropy in protein−protein binding. *Structure* **2006**, *14*, 683–693.
(62) Hou, T.; Wang, J.; Li, Y.; Wang, W. Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations. *J. Chem. Inf. Model.* **2011**, *51*, 69–82.
(63) Hou, T.; Wang, J.; Li, Y.; Wang, W. Assessing the performance of the molecular mechanics/Poisson Boltzmann surface area and molecular mechanics/generalized Born surface area methods. II. The accuracy of ranking poses generated from docking. *J. Comput. Chem.* **2011**, *32*, 866–877.
(64) Naïm, M.; Bhat, S.; Rankin, K. N.; Dennis, S.; Chowdhury, S. F.; Siddiqi, I.; Drabik, P.; Sulea, T.; Bayly, C. I.; Jakalian, A.; Purisima, E. O. Solvated interaction energy (SIE) for scoring protein−ligand binding affinities. 1. Exploring the parameter space. *J. Chem. Inf. Model.* **2007**, *47*, 122–133.
(65) Kitaura, K.; Ikeo, E.; Asada, T.; Nakano, T.; Uebayasi, M. Fragment molecular orbital method: an approximate computational method for large molecules. *Chem. Phys. Lett.* **1999**, *313*, 701–706.
(66) Fedorov, D. G.; Kitaura, K. Extending the power of quantum chemistry to large systems with the fragment molecular orbital method. *J. Phys. Chem. B* **2007**, *111*, 6904–6914.
(67) Mochizuki, Y.; Yamashita, K.; Fukuzawa, K.; Takematsu, K.; Watanabe, H.; Taguchi, N.; Okiyama, Y.; Tsuboi, M.; Nakano, T.; Tanaka, S. Large-scale FMO-MP3 calculations on the surface proteins of influenza virus, hemagglutinin (HA) and neuraminidase (NA). *Chem. Phys. Lett.* **2010**, *493*, 346–352.
(68) Mazanetz, M. P.; Ichihara, O.; Law, R. J.; Whittaker, M. Prediction of cyclin-dependent kinase 2 inhibitor potency using the fragment molecular orbital method. *J. Cheminf.* **2011**, *3*, 2.
(69) Nakano, T.; Mochizuki, Y.; Kato, A.; Fukuzawa, K.; Ishikawa, T.; Amari, S.; Kurisaki, I.; Tanaka, S. Developments of FMO methodology and graphical user interface in ABINIT-MP. In *The Fragment Molecular Orbital Method: Practical Applications to large Molecular Systems*; Fedorov, D. G., Kitaura, K., Eds.; CRC Press: Boca Raton, FL, 2009; pp 37−59.
(70) Mochizuki, Y.; Nakano, T.; Koikegami, S.; Tanimori, S.; Abe, Y.; Nagashima, U.; Kitaura, K. A parallelized integral-direct second-order Møller−Plesset perturbation theory method with a fragment molecular orbital scheme. *Theor. Chem. Acc.* **2004**, *112*, 442–452.
(71) Yoshida, T.; Yamagishi, K.; Chuman, H. QSAR study of cyclic urea type HIV-1 PR inhibitors using ab initio MO calculation of their complex structures with HIV-1 PR. *QSAR Comb. Sci.* **2008**, *27*, 694–703.
(72) Yoshida, T.; Fujita, T.; Chuman, H. Novel quantitative structureactivity studies of HIV-1 protease inhibitors of the cyclic urea type using descriptors derived from molecular dynamics and molecular orbital calculations. *Curr. Comput.-Aided Drug Des.* **2009**, *5*, 38–55.
(73) Iwata, T.; Fukuzawa, K.; Nakajima, K.; Aida-Hyugaji, S.; Mochizuki, Y.; Watanabe, H.; Tanaka, S. Theoretical analysis of binding specificity of influenza viral hemagglutinin to avian and human receptors based on the fragment molecular orbital method. *Comput. Biol. Chem.* **2008**, *32*, 198–211.
(74) Ishikawa, T.; Ishikura, T.; Kuwata, K. Theoretical study of the prion protein based on the fragment molecular orbital method. *J. Comput. Chem.* **2009**, *30*, 2594–2601.
(75) Sawada, T.; Fedorov, D. G.; Kitaura, K. Structural and interaction analysis of helical heparin oligosaccharides with the fragment molecular orbital method. *Int. J. Quantum Chem.* **2009**, *109*, 2033–2045.
(76) Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; Donini, O.; Cieplak, P.; Srinivasan, J.; Case, D. A.; Cheatham, T. E., III. Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Acc. Chem. Res.* **2000**, *33*, 889–897.
(77) Massova, I.; Kollman, P. A. Combined molecular mechanical and continuum solvent approach (MM-PBSA/GBSA) to predict ligand binding. *Perspect. Drug Discovery Des.* **2000**, *18*, 113–135.
(78) Kuhn, B.; Gerber, P.; Schulz-Gasch, T.; Stahl, M. Validation and use of the MM-PBSA approach for drug discovery. *J. Med. Chem.* **2005**, *48*, 4040–4048.
(79) Li, Y.; Liu, Z.; Wang, R. Test MM-PB/SA on true conformational ensembles of protein−ligand complexes. *J. Chem. Inf. Model.* **2010**, *50*, 1682–1692.
(80) Sitkoff, D.; Sharp, K. A.; Honig, B. Accurate calculation of hydration free energies using macroscopic solvent models. *J. Phys. Chem.* **1994**, *98*, 1978–1988.
(81) Gubareva, L. V.; Webster, R. G.; Hayden, F. G. Comparison of the activities of zanamivir, oseltamivir, and RWJ-270201 against clinical isolates of influenza virus and neuraminidase inhibitor-resistant variants. *Antimicrob. Agents Chemother.* **2001**, *45*, 3403–3408.
(82) Russell, R. J.; Haire, L. F.; Stevens, D. J.; Collins, P. J.; Lin, Y. P.; Blackburn, G. M.; Hay, A. J.; Gamblin, S. J.; Skehel, J. J. The structure of H5N1 avian influenza neuraminidase suggests new opportunities for drug design. *Nature* **2006**, *443*, 45–49.
(83) Li, C.-Y.; Yu, Q.; Ye, Z.-Q.; Sun, Y.; He, Q.; Li, X.-M.; Zhang, W.; Luo, J.; Gu, X.; Zheng, X.; Wei, L. A nonsynonymous SNP in human cytosolic sialidase in a small Asian population results in reduced enzyme activity: potential link with severe adverse reactions to oseltamivir. *Cell Res.* **2007**, *17*, 357–362.

Molecular Systems Biology 7; Article number 472; doi:10.1038/msb.2011.5
Citation: *Molecular Systems Biology* 7:472

www.molecularsystemsbiology.com

molecular systems biology

Analysis of multiple compound–protein interactions reveals novel bioactive molecules

Hiroaki Yabuuchi[1,5], Satoshi Niijima[1,5], Hiromu Takematsu[2], Tomomi Ida[1], Takatsugu Hirokawa[3], Takafumi Hara[4], Teppei Ogawa[1], Yohsuke Minowa[1], Gozoh Tsujimoto[4] and Yasushi Okuno[1,*]

[1] Department of Systems Biosciences for Drug Discovery, Graduate School of Pharmaceutical Sciences, Kyoto University, Kyoto, Japan, [2] Laboratory of Membrane Biochemistry and Biophysics, Graduate School of Biostudies, Kyoto University, Kyoto, Japan, [3] Computational Biology Research Center, National Institute of Advanced Industrial Science and Technology, Tokyo, Japan and [4] Department of Genomic Drug Discovery Science, Graduate School of Pharmaceutical Sciences, Kyoto University, Kyoto, Japan
[5] These authors contributed equally to this work
* Corresponding author. Department of Systems Biosciences for Drug Discovery, Graduate School of Pharmaceutical Sciences, Kyoto University, 46-29 Yoshida-Shimo-Adachi-cho, Sakyo-ku, Kyoto 606-8501, Japan. Tel.: +81 75 753 4559; Fax: +81 75 753 4559; E-mail: okuno@pharm.kyoto-u.ac.jp

Received 21.7.10; accepted 20.1.11

The discovery of novel bioactive molecules advances our systems-level understanding of biological processes and is crucial for innovation in drug development. For this purpose, the emerging field of chemical genomics is currently focused on accumulating large assay data sets describing compound–protein interactions (CPIs). Although new target proteins for known drugs have recently been identified through mining of CPI databases, using these resources to identify novel ligands remains unexplored. Herein, we demonstrate that machine learning of multiple CPIs can not only assess drug polypharmacology but can also efficiently identify novel bioactive scaffold-hopping compounds. Through a machine-learning technique that uses multiple CPIs, we have successfully identified novel lead compounds for two pharmaceutically important protein families, G-protein-coupled receptors and protein kinases. These novel compounds were not identified by existing computational ligand-screening methods in comparative studies. The results of this study indicate that data derived from chemical genomics can be highly useful for exploring chemical space, and this systems biology perspective could accelerate drug discovery processes.
Molecular Systems Biology **7:** 472; published online 1 March 2011; doi:10.1038/msb.2011.5
Subject Categories: bioinformatics; computational methods
Keywords: chemical genomics; data mining; drug discovery; ligand screening; systems chemical biology

Introduction

Experimental perturbations of biological systems, such as genetic mutation and chemical exposure, have been used as powerful approaches to deepen our systems-level understanding of biological processes and to discover unprecedented biological phenomena (Lehár *et al*, 2008). In particular, perturbations by chemical probes provide broader applications not only for analysis of complex systems but also for intentional manipulations of these systems, e.g., a medicine is a small molecule designed for the purpose of clinical therapy that can actively manipulate biological systems from disordered to well-ordered states. Unfortunately, the number of well-characterized chemical probes is highly limited, which has bottlenecked their wide range of application.

The set of all possible small organic molecules, referred to as chemical space, has been estimated to consist of more than 10^{60} compounds (Dobson, 2004). Chemical space is as vast as the diversity of biological systems, and the vastness of the two domains creates difficulty in comprehensive understanding of the interface between chemical space and biological systems (Lipinski and Hopkins, 2004; Renner *et al*, 2009). Recently, chemical genomics has emerged as a promising area of research applicable to exploration of novel bioactive molecules, and researchers are currently striving toward the identification of all possible ligands for all target protein families (Wang *et al*, 2009). Large-scale data sets of compound–protein interactions (CPIs) are being collected, and chemical genomics studies have shown that patterns of protein–ligand interactions are too diverse to be understood as simple one-to-one events. For example, multiple structurally different compounds have been shown to bind the same protein or express similar biological activities (Eckert and Bajorath, 2007; Young *et al*, 2008). In other cases, one drug has been shown to affect multiple targets from different protein families (MacDonald *et al*, 2006; Paolini *et al*, 2006). This phenomenon, termed polypharmacology, is thought to be one critical cause of adverse drug effects (Hopkins, 2008). There-

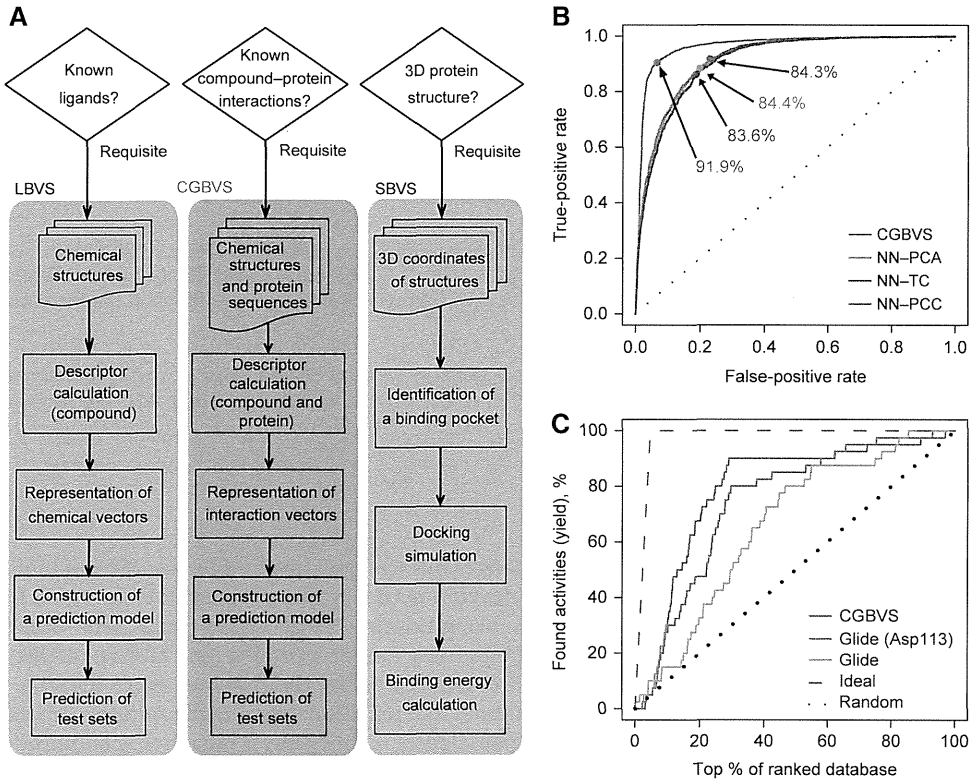


Figure 1 Overview and performance of CGBVS. (**A**) Comparison of the strategies used for CGBVS, LBVS, and SBVS. CGBVS used multiple CPI data, represented the CPI in vector form, and used SVM for CPI pattern learning. (**B**) ROC curves obtained by fivefold cross-validation using compound–GPCR interactions for the CGBVS (red) and LBVS methods. The best accuracy rate for each method is also shown. NN–PCA (light blue), nearest neighbor (NN) method with the Pearson correlation coefficient in the constructed space using principal component analysis (PCA); NN–TC (navy), NN method with the Tanimoto coefficient (TC); and NN–PCC (purple), NN method with the Pearson correlation coefficient (PCC) in the original space. (**C**) Enrichment curves obtained by the CGBVS and SBVS methods. Glide was used both without constraints and with constrained hydrogen bonding between the compounds and Asp113, a residue known to be crucial for ligand binding to ADRB2 (Strader *et al*, 1987). Information regarding interactions between ADRB2 and its ligands was not used in test set for machine learning for CGBVS.

fore, an integrative understanding of multiple interactions among chemical and biological components beyond a one-compound/one-target simplification could open up new opportunities in drug development, but the need to develop appropriate data mining methods for characterizing and visualizing the full complexity of interactions between chemical space and biological systems is urgent (Oprea *et al*, 2007). Recently, mining of multiple CPI data sets has been used to identify new protein targets for known drugs and thereby predict unreported polypharmacology (Keiser *et al*, 2009). However, this approach only identifies additional targets for known drugs. No existing screening approach has so far succeeded in identifying novel bioactive compounds using multiple interactions among compounds and target proteins, and the potential application of analysis of multiple CPIs to identify novel bioactive molecules remains unknown.

High-throughput screening (HTS) and computational screening have greatly aided in the identification of early lead compounds for drug discovery. However, the large numbers of assays required for HTS to identify drugs that target multiple proteins render this process very costly and time-consuming. Therefore, interest in using *in silico* strategies for screening has been increasing. The most common computational approaches, ligand-based virtual screening (LBVS) and structure-based virtual screening (SBVS; Oprea and Matter, 2004; Muegge and Oloff, 2006; McInnes, 2007; Figure 1A), have been used for practical drug development. Unfortunately, these methods have important limitations. LBVS aims to identify molecules that are very similar to known active molecules and generally has difficulty identifying compounds with novel structural scaffolds that differ from reference molecules. Attempts to scaffold-hop using LBVS are prone to identification of increased numbers of false positives (Eckert and Bajorath, 2007). Therefore, the primary objective of virtual screening, reduction of the number of candidate compounds to be assayed, remains unachievable using this method. The other popular strategy, SBVS, is constrained by the number of three-dimensional crystallographic structures available and, more importantly, by the difficulty of accurately simulating molecular docking processes for targets, including membrane-spanning G-protein-coupled receptors (GPCRs). To circumvent these limitations, we have shown that a new computational screening strategy, chemical genomics-based virtual screening (CGBVS), has the potential to identify novel, scaffold-hopping compounds and assess their polypharmacology by using a machine-learning method to recognize conserved molecular patterns in comprehensive CPI data sets.

Results

Theoretical framework for CGBVS

The CGBVS strategy is made up of five steps: CPI data collection, descriptor calculation, representation of interaction vectors, predictive model construction using training data sets, and predictions from test data (Figure 1A and Supplementary Figure S1). Importantly, step 1, the construction of a data set of chemical structures and protein sequences for known CPIs, does not require the three-dimensional protein structures needed for SBVS. We chose GPCRs, important pharmaceutical targets (Hopkins and Groom, 2002), as our first target proteins for virtual ligand screening. In total, 5207 CPIs (including 317 GPCRs and 866 ligands) retrieved from the GLIDA database (Okuno *et al*, 2006) were used as experimental data (Supplementary Table S1). In step 2, compound structures and protein sequences were converted into numerical descriptors using 929-dimensional and 400-dimensional feature vectors, respectively. A wide variety of chemical descriptors was used to describe the substructures, as well as the physicochemical and molecular properties of the small molecules. Descriptors for protein sequences were created using a string kernel (see Materials and methods section and Supplementary information for details). These descriptors were used to construct chemical or biological spaces, in which decreasing distance between vectors corresponded to increasing similarity of compound structures or protein sequences. In step 3, we represented multiple CPI patterns by concatenating these chemical and protein descriptors (in $929+400$ dimensions). Using these interaction vectors, we could quantify the similarity of molecular interactions for compound–protein pairs, despite the fact that the ligand and protein similarity maps differed substantially (Keiser *et al*, 2007). In step 4, concatenated vectors for CPI pairs (positive samples) and non-interacting pairs (negative samples) were input into a support vector machine (SVM; Vapnik, 1995), an established machine-learning technique widely applied to pattern-recognition problems (Schölkopf *et al*, 2004; Shawe-Taylor and Cristianini, 2004). Using training sets, an SVM classifier was generated as a hyperplane dividing positive and negative samples into two distinct classes representing interaction and non-interaction. By mapping the samples into high-dimensional feature space using a nonlinear kernel function, samples that were linearly inseparable in the original input space could be linearly separated in the feature space. As a nonlinear SVM can extract patterns from data sets with nonlinear characteristics, non-intuitive interaction rules can be obtained from multiple CPI patterns, creating the potential to identify novel CPIs. In the final step, the SVM classifier constructed using training sets was applied to test data. Along with providing simple yes/no outputs, the calculated prediction scores also ranked all test compound–GPCR pairs in the order of binding probability.

Computational evaluation of CGBVS

To evaluate the predictive value of CGBVS, we compared its performance with that of LBVS methodologies using respective data sets of 5207 interacting and non-interacting pairs. The performance of each method was tested by repeating fivefold cross-validations 20 times. CGBVS performed with a considerably higher accuracy ($91.9 \pm 0.3\%$) than LBVS ($84.4 \pm 0.3\%$, at best). We also recorded the number of true-positive interactions as a function of false positives and plotted receiver operating characteristic (ROC) curves (Hanley and McNeil, 1982), as shown in Figure 1B and Supplementary Table S2. ROC analysis revealed that CGBVS performed better than all LBVS methods in terms of the score ranking of CPI pairs.

Recently, the crystal structure of the β2-adrenergic receptor (ADRB2) has been determined (Cherezov *et al*, 2007; Rasmussen *et al*, 2007). Therefore, we were able to compare CGBVS and SBVS in a retrospective virtual screening based on the human ADRB2 using a representative docking program, Glide (Friesner *et al*, 2004). For CGBVS, we constructed a predictive model based on 5167 CPI pairs, excluding 40 known ADRB2-related CPIs to avoid any bias in favor of CGBVS during the machine-learning step. Using both methods, we predicted scores for the same 866 known GPCR ligands, including the 40 known ADRB2 ligands as positive controls. We asked whether the scores for these 40 known positive compounds were higher than scores for other compounds. Figure 1C and Supplementary Table S3 show that CGBVS provided higher enrichment factors (EFs) and hit rates than did SBVS. These results suggest that CGBVS is more successful than conventional approaches for prediction of CPIs.

Polypharmacological interactions for ADRB2

We also evaluated the ability of the CGBVS method to predict the polypharmacology of ADRB2 by attempting to identify novel ADRB2 ligands from the above ligand data set. As an established, well-studied pharmaceutical target (Waldeck, 2002), we expected that novel ADRB2 ligands would be difficult to find, and that searching for novel scaffolds for such a well-known target would be a stringent assessment of the predictive ability of CGBVS. After training an SVM classifier using all 5207 CPIs, we ranked the prediction scores for the interactions of 826 reported GPCR ligands (excluding the 40 known ADRB2 ligands) with ADRB2, and then analyzed the 50 highest-ranked compounds in greater detail. To complement the less-than-comprehensive binding data available in the original GLIDA database, a literature search was performed using SciFinder (Wagner, 2006) and PubMed (http://www.ncbi.nlm.nih.gov/pubmed/). This search identified 15 of the top 50 compounds as known ADRB2 ligands. Importantly, these compounds were identified as ligands of other GPCRs, not of ADRB2, in training sets used in the machine-learning step. ADRB2 ligands already reported in the literature were excluded from analysis, but the remainder were tested in *in vitro* binding assays. From the remaining 35 ligands, 21 were commercially available. Of these 21, 11 were not previously reported, but were discovered to bind to ADRB2 (Figure 2A and Supplementary Table S5). These compounds included ligands for the acetylcholine, serotonin, dopamine, and neuropeptide Y receptors (Figure 2E), indicating the presence of potential polypharmacological interactions with ADRB2.

To substantiate the novelty of the ligands identified by CGBVS, we compared the predictive scores estimated by the three virtual screening approaches (CGBVS, LBVS, and SBVS).