Pergamon

Bioorganic & Medicinal Chemistry 6 (1998) 903–910

BIOORGANIC & MEDICINAL CHEMISTRY

# Quantitative Structure–Activity Studies of Octopaminergic Agonists and Antagonists against Nervous System of *Locusta migratoria*

Akinori Hirashima,[a,*] Canping Pan,[b] Kenji Shinkai,[a] Jun Tomita,[a] Eiichi Kuwano, Eiji Taniguchi[a] and Morifusa Eto[c]

[a]*Department of Agricultural Chemistry, Kyushu University, Fukuoka 812-8581, Japan*
[b]*Department of Applied Chemistry, Beijing Agricultural University, Beijing 100094, People's Republic of China*
[c]*Miyakonojo National College of Technology, 473-1 Yoshio-cho, Miyakonojo 885, Japan*

Received 12 November 1997; accepted 16 January 1998

**Abstract**—The quantitative structure–activity relationship (QSAR) of octopaminergic agonists and antagonists against the thoracic nerve cord of the migratory locust, *Locusta migratoria* L., was analyzed using physicochemical parameters and regression analysis. The hydrophobic effect, dipole moment, and shape index were important in terms of $K_i$: the more hydrophobic, the greater dipole moment, and the smaller shape index of the molecules, the greater the activity. A receptor surface model (RSM) was generated using some subset of the most active structures. Three-dimensional energetics descriptors were calculated from RSM/ligand interaction and these three-dimensional descriptors were used in QSAR analysis. This data set was studied further using molecular shape analysis. 

## Introduction

Octopamine (OA), which has been found to present in high concentrations in various insect tissues, is the monohydroxillic analogue of the vertebrate hormone noradrenaline. It has been well established that OA acts as neurohormone, neurotransmitter, and neuromodulator in insects.[1] Three different receptor classes OAR1, OAR2A, and OAR2B had been distinguished from non-neuronal tissues.[2] In the nervous system of locust, a particular receptor class was characterized and established as a new class OAR3 by pharmacological investigations of the [$^3$H]OA binding site using various agonists and antagonists.[3–7] Structure–activity studies of various types of OA-agonists and antagonists were also reported using the nervous tissue of the migratory locust, *Locusta migratoria* L. However, information on the structural requirements of these OA-agonists and antagonists for high OA-receptor ligands is still limited. Hence, this paper deals with the quantitative structure–activity relationship (QSAR) of OA-agonists and antagonists against the nervous tissue of the migratory locust followed by molecular modeling of these superimposed OA ligands to elucidate mechanisms of OA ligand–receptor interaction.

Key words: *Locusta migratoria*; quantitative structure–activity relationship (QSAR); receptor surface model (RSM); molecular shape analysis.
*Corresponding author. Tel: +81-92-642-2856; Fax: +81-92-642-2864/642-2804; E-mail. ahirasim@agr.kyushu-u.ac.jp

0968-0896/98/$19.00 
*PII:* S0968-0896(98)00024-8

## Results

In order to quantitatively understand the dependence of biological activities on physicoparameters of OA agonists and antagonists, regression analysis was applied to representative compounds listed in Figure 1 and Table 1, leading to eq (1).

$$\begin{aligned} \mathrm{p}K_i = {} & 6.723(\pm 0.279) + 0.620(\pm 0.071)\log P \\ & + 0.226(\pm 0.019)DM - 0.449(\pm 0.049)SIO \end{aligned} \quad (1)$$

where $n = 49$, $r = 0.914$, $s = 0.446$, and $F = 75.604$. According to eq (1), the positive log $P$, $DP$, and the

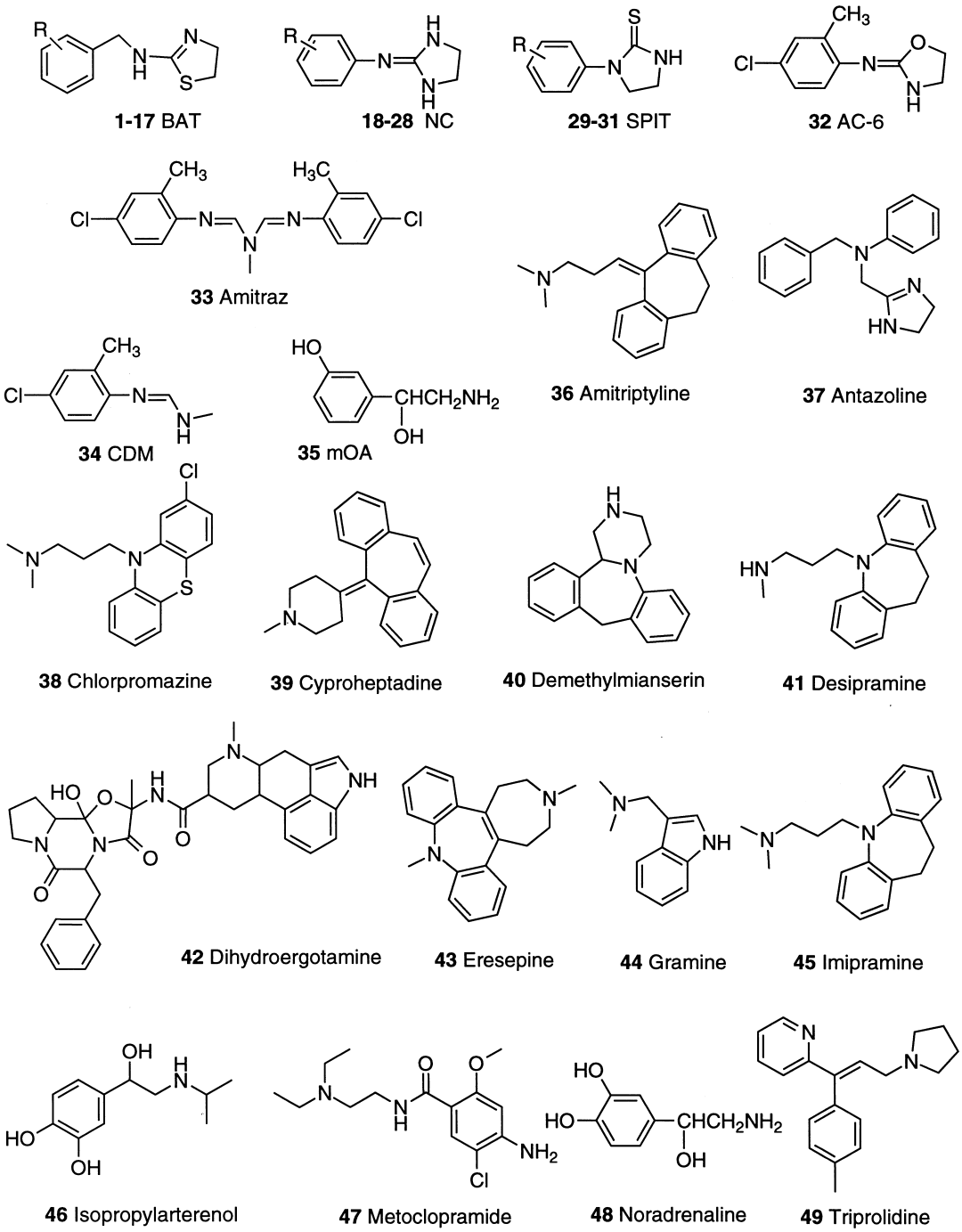


**Figure 1.** Structures of OA agonists and antagonists used for regression analysis.

negative *SIO* terms mean that the more hydrophobic, the greater dipole moment, and the smaller shape index of the molecules, the greater the activity. Values of p$K_i$, the logs of the reciprocals of $K_i$, were used as OA-agonist and antagonist activity index. $K_i$ is the concentration of OA agonist and antagonist necessary for half-maximal inhibition of OA binding. Following physicoparameters were used: dipole moment (*DP*), the magnitude of the molecule's dipole; log *P*, the octanol–water partition coefficient calculated using the atom typing scheme of

**Table 1.** Regression analysis of structure-OA agonist and antagonist activity

| Compd | | $K_i$ (nM) | p$K_i$ | | | | | |
|---|---|---|---|---|---|---|---|---|
| no. | R | | Obsvd | Calcd[a] | Dev. | log $P$[b] | $DM$[c] | $SIO$[d] |
| **1**[e] | BAT H | 280 ± 59[g] | 6.55 | 6.363 | 0.187 | 2.676 | 1.047 | 5.024 |
| **2**[e] | BAT 2-Cl | 440 ± 189[g] | 6.36 | 6.448 | −0.088 | 3.194 | 0.324 | 5.186 |
| **3**[e] | BAT 2-F | 447 ± 125[g] | 6.35 | 6.556 | −0.206 | 2.815 | 2.419 | 5.186 |
| **4**[e] | BAT 2-$CH_3$ | 650 ± 360[g] | 6.19 | 6.528 | −0.338 | 3.143 | 0.816 | 5.186 |
| **5**[e] | BAT 2-$CF_3$ | 290 ± 203[g] | 6.54 | 6.818 | −0.278 | 3.559 | 2.097 | 5.760 |
| **6**[e] | BAT 3-Cl | 95 ± 66[g] | 7.02 | 6.540 | 0.480 | 3.194 | 0.731 | 5.186 |
| **7**[e] | BAT 3-F | 250 ± 75[g] | 6.60 | 6.615 | −0.015 | 2.815 | 2.098 | 5.186 |
| **8**[e] | BAT 3-$CH_3$ | 34 ± 26[g] | 7.47 | 6.619 | 0.851 | 3.143 | 1.218 | 5.186 |
| **9**[e] | BAT 3-$CF_3$ | 380 ± 266[g] | 6.42 | 7.136 | −0.716 | 3.559 | 3.505 | 5.760 |
| **10**[e] | BAT 3-$NO_2$ | 185 ± 33[g] | 6.73 | 6.917 | −0.187 | 2.629 | 5.706 | 6.074 |
| **11**[e] | BAT 4-Cl | 89 ± 6[g] | 7.05 | 6.575 | 0.475 | 3.194 | 0.885 | 5.186 |
| **12**[e] | BAT 4-F | 460 ± 124[g] | 6.34 | 6.409 | −0.069 | 2.815 | 1.190 | 5.186 |
| **13**[e] | BAT 2,3-$Cl_2$ | 42 ± 19[g] | 7.38 | 7.070 | 0.310 | 3.712 | 2.011 | 5.365 |
| **14**[e] | BAT 2-Cl,4-F | 184 ± 90[g] | 6.74 | 6.690 | 0.050 | 3.333 | 1.368 | 5.365 |
| **15**[e] | BAT 2,5-$Cl_2$ | 132 ± 83[g] | 6.88 | 6.867 | 0.013 | 3.712 | 1.114 | 5.365 |
| **16**[e] | BAT 3-Cl,4-F | 109 ± 59[g] | 6.96 | 6.791 | 0.169 | 3.333 | 1.816 | 5.365 |
| **17** | BAT 3,4-$F_2$ | 445 ± 196[g] | 6.35 | 6.556 | −0.206 | 2.955 | 1.810 | 5.365 |
| **18** | NC H | 23 ± 5[h] | 7.63 | 7.960 | −0.330 | 1.808 | 9.037 | 4.297 |
| **19** | NC 4-Br | 15 ± 3[h] | 7.83 | 8.516 | −0.686 | 2.599 | 9.696 | 4.481 |
| **20** | NC 2,4-$Cl_2$ | 0.81 ± 0.18[h] | 9.09 | 8.460 | 0.630 | 2.844 | 9.172 | 4.680 |
| **21** | NC 2,4-$(CH_3)_2$ | 1.02 ± 0.42[h] | 8.99 | 8.272 | 0.718 | 2.742 | 8.619 | 4.680 |
| **22** | NC 2,6-$Cl_2$ | 47 ± 18[h] | 7.32 | 7.154 | 0.166 | 2.844 | 3.401 | 4.680 |
| **23** | NC 2,6-$(CH_3)_2$ | 20 ± 7[h] | 7.70 | 8.390 | −0.690 | 2.742 | 9.142 | 4.680 |
| **24** | NC 2,4,5-$Cl_3$ | 2.27 ± 0.89[h] | 8.64 | 8.326 | 0.314 | 3.362 | 7.574 | 4.888 |
| **25** | NC 2,4,6-$Cl_3$ | 19 ± 3[h] | 7.73 | 8.512 | −0.782 | 3.362 | 8.397 | 4.888 |
| **26** | NC 2,6-$Cl_2$,4-$NH_2$ | 58 ± 16[h] | 7.24 | 7.015 | 0.225 | 2.060 | 5.343 | 4.888 |
| **27** | NC 2,4,6-$(CH_3)_3$ | 4.38 ± 1.30[h] | 8.36 | 8.511 | −0.151 | 3.209 | 8.812 | 4.888 |
| **28** | NC 2,6-$(CH_2CH_3)_2$,4-$N_3$ | 1.05 ± 0.47[h] | 8.98 | 8.958 | 0.022 | 4.380 | 13.144 | 7.695 |
| **29**[f] | SPIT 4-Cl | 280 ± 134[g] | 6.55 | 6.744 | −0.194 | 2.842 | 3.462 | 4.481 |
| **30**[f] | SPIT 2-$CH_3$,4-Cl | 20 ± 9[g] | 7.70 | 8.390 | −0.690 | 3.309 | 3.603 | 4.680 |
| **31**[f] | SPIT 2,6-$(CH_2CH_3)_2$ | 170 ± 51[g] | 6.77 | 6.500 | 0.270 | 4.051 | 3.792 | 6.074 |
| **32** | AC-6 | 0.95 ± 0.24[h] | 9.02 | 8.829 | 0.191 | 2.966 | 10.469 | 4.680 |
| **33** | Amitraz | 22 ± 5[h] | 7.67 | 7.139 | 0.531 | 5.809 | 3.266 | 8.741 |
| **34** | CDM | 137 ± 70[h] | 6.91 | 7.277 | −0.367 | 3.494 | 2.847 | 5.024 |
| **35** | mOA | 5050 ± 1860[h] | 5.30 | 5.624 | −0.324 | 0.507 | 1.951 | 4.133 |
| **36** | Amitriptyline | 1570 ± 330[i] | 5.80 | 6.452 | −0.652 | 4.519 | 1.326 | 7.513 |
| **37** | Antazoline | 118 ± 62[i] | 6.93 | 6.408 | 0.522 | 4.195 | 3.029 | 8.022 |
| **38** | Chlorpromazine | 766 ± 54[i] | 6.12 | 6.184 | −0.064 | 3.820 | 1.142 | 7.051 |
| **39** | Cyproheptadine | 844 ± 356[i] | 6.07 | 6.176 | −0.106 | 3.565 | 0.971 | 6.630 |
| **40** | Demethylmianserin | 55 ± 3[i] | 7.26 | 6.637 | 0.623 | 3.632 | 0.408 | 5.413 |
| **41** | Desipramine | 3210 ± 610[i] | 5.49 | 6.040 | −0.550 | 3.645 | 1.517 | 7.320 |
| **42** | Dihydroergotamine | 272 ± 30[i] | 6.57 | 6.062 | 0.508 | 2.828 | 11.768 | 11.313 |
| **43** | Eresepine | 474 ± 166[i] | 6.32 | 5.931 | 0.389 | 3.208 | 9.566 | 6.481 |
| **44** | Gramine | 1840 ± 1012[i] | 5.74 | 5.857 | −0.117 | 1.242 | 0.747 | 4.022 |
| **45** | Imipramine | 1450 ± 333[i] | 5.84 | 6.124 | −0.284 | 4.006 | 1.281 | 7.513 |
| **46** | Isopropylarterenol | 18600 ± 4840[h] | 4.77 | 5.117 | −0.347 | 1.442 | 1.949 | 6.554 |
| **47** | Metoclopramide | 52601 ± 6500[i] | 4.28 | 4.441 | −0.161 | 1.273 | 4.487 | 9.107 |
| **48** | Noradrenaline | 475 ± 42[h] | 6.32 | 5.931 | 0.389 | 0.222 | 4.770 | 4.297 |
| **49** | Triprolidine | 2900 ± 145[i] | 5.54 | 6.239 | −0.699 | 4.629 | 1.093 | 8.022 |

[a]Calculated by eq (1).
[b]Log $P$ was calculated using the atom typing scheme of Ghose et al.[8]
[c]$DM$ is the magnitude of the molecule's dipole.
[d]$SIO$ is a topological index quantifying the shape.
[e]Obtained by the hydrochloric acid-catalyzed cyclization of the corresponding thioureas.
[f]Synthesized by the cyclization of monoethanolamine hydrogen sulfate with the corresponding arylisothiocyanates in the presence of sodium hydroxide.
[g]Personal communication (T. Roeder, Hamburg University).
[h,i]Cited from Refs 6 and 7, respectively.

Ghose et al.[8]; shape index order 2 (*SIO*), a topological index quantifying the shape. The figures in parentheses are 95% confidence intervals and all terms are justified more than 99% by a *t*-test; *n*, the number of data points; *r*, the correlation coefficient; *s*, the standard deviation; *F*, the value of the *F*-test, The use of other parameters instead of or the addition of other parameters to eq (1) did not improve the correlation: e.g. molar reflectivity (MR) was used as the steric parameter, which was calculated using the atom typing scheme of Ghose et al.,[8] the heat of formation (the energy released or used when a molecule is formed from elements in their standard states) determined after optimizing the molecular geometry first using augmented MM2 then using the Molecular Orbital Package (MOPAC) with PM3 parameters.[9] HOMO energy (the energy required to remove an electron from the highest occupied molecular orbital), the dielectric energy (a portion of the total energy of a molecule embedded in a dielectric, the stabilizing portion that results from screening the charges in the molecule by a dielectric, calculated at an optimized geometry in water),[10] electron affinity (the change in the total energy of a molecule when an electron is added), conformation minimum energy (energy calculated for an optimized conformation), log *P* estimated by the log octanol–water partition coefficient program (KOWWIN),[11] solvent accessible surface (SAS) area when the solvent is water, shape index order 1 and 3 (a topological index quantifying the shape), LUMO energy (the energy gained when an electron is added to the lowest uncoupled orbital), total energy (the work required to separate the electron and nuclei infinitely far apart), and steric energy (the sum of the molecular mechanics potential energies calculated for the bonds, bond angles, dihedral angles, nonbonded atoms, and so forth.

The descriptors generated for receptor surface model (RSM), the activities of the compounds, and the predictions using their first model are shown in Table 2. A RSM represents essential information about the hypothetical receptor site as a three-dimensional surface with associated properties mapped onto the surface model. The location and shape of the surface represent information about the steric nature of the receptor site: the associated properties represent other information of interest, such as hydrophobicity, partial charge, electrostatic potential, and hydrogen-bonding propensity. The isosurface procedure produces a surface that entirely encloses the molecules over which it is generated. The surface has no holes and is known as a closed model. RSMs are best constructed from a set of the most active analogues that are chosen to cover the variety of steric and electrostatic variations likely to appear in the test data. The approach we took was to automatically build a set of different RSMs from different combinations of the most active analogues, and then use a variable-selection technique such as genetic function approximation (GFA) to discover the RSM whose descriptors yield the best QSARs of the full training set. GFA allows the discovery and use of nonlinear descriptors by using spline-based terms. A RSM was generated (Figure 2) using some subset of the most active structures (**20**, **21**, **28**, and **32**). The rationale underlying this model is that the most active structures tend to explore the best spatial and electronic interactions with receptor, while the least active do not and tend to have unfavorable steric or electronic interactions. The best model generated using the descriptors from the closed RSM is given in eq (2).

$$\begin{aligned} \mathrm{p}K_i = {} & 5.313 + 0.00170\ IntraEnergy - 0.364\ Rotlbonds \\ & + 0.569\ H_{\text{bond donor}} + 0.803 \log\ P \\ & - 0.0217\ InterVDWEnergy - 0.00766\ PMI - X \end{aligned} \tag{2}$$

where $n = 49$, $r^2 = 0.776$, $F = 24.302$, $CV - r^2 = 0.695$, and $Bsr^2 = 0.777$: *r*-squared ($r^2$), the square of the correlation coefficient; cross-validated $r^2$ ($CV - r^2$), a squared, correlation coefficient generated during a validation procedure using the equation; bootstrap $r^2$ ($Bsr^2$), the average squared correlation coefficient calculated during the validation procedure. *r*-Squared is used to describe the goodness of fit of the data in the study table to the QSAR model. A cross-validated $r^2$ is usually smaller than the overall $r^2$ for a QSAR equation. It is used as a diagnostic tool to evaluate the predictive power of an equation generated using the multiple linear regression or partial least squares method. A bootstrap $r^2$ is computed from the subset of variables used one-at-a-time for the validation procedure. It can be used more than one time in computing the $r^2$ statistic. Once a reasonable RSM has been defined, a series of structures can be evaluated against the model. The results of the minimization procedure can be used as descriptors either to refine the model or to predict activity. Three-dimensional energetics descriptors were calculated from RSM/ligand interaction. These three-dimensional descriptors were used in QSAR analysis, leading to eq (3):

$$\mathrm{p}K_i = 6.862 + (\text{ELE terms}) + (\text{VDW terms}) \tag{3}$$

where $n = 49$, $r = 0.873$, $F = 24.302$, and $CV - r^2 = 0.673$: ELE; energy of electrostatic term; VDW, energy of van der Waals term. The ability to minimize a structure against a model allows one to flexibility fit a structure into the model. The minimization can also be used as a shape- (and not atom-) based alignment technique. Applying the fitting process to a set of molecules will force all the molecules to adopt a shape that is consistent with the model. This data set

**Table 2.** Receptor surface descriptors and prediction using top model

| No. | *IntraEnergy*[a] | *Rotlbonds*[b] | $H_{\text{bond donor}}$[c] | Log *P*[d] | *InterVDWEnergy*[e] | *PMI* − *X*[f] | p$K_i$ | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | Obsvd | Calcd[g] | Dev. |
| **1** | 28.51 | 3 | 1 | 2.53 | −1.13 | 45.45 | 6.55 | 6.55 | 0 |
| **2** | 401.52 | 3 | 1 | 3.05 | 16.92 | 79.27 | 6.36 | 6.95 | −0.59 |
| **3** | 59.76 | 3 | 1 | 2.67 | 2.30 | 60.34 | 6.35 | 6.53 | −0.18 |
| **4** | 88.53 | 3 | 1 | 3.00 | 3.60 | 61.35 | 6.19 | 6.80 | −0.61 |
| **5** | 369.05 | 3 | 1 | 3.42 | 13.38 | 108.77 | 6.54 | 7.04 | −0.50 |
| **6** | 35.15 | 3 | 1 | 3.05 | −1.06 | 51.25 | 7.02 | 6.93 | 0.09 |
| **7** | 32.21 | 3 | 1 | 2.67 | −0.99 | 48.19 | 6.60 | 6.64 | −0.04 |
| **8** | 36.08 | 3 | 1 | 3.00 | −1.05 | 48.72 | 7.47 | 6.91 | 0.56 |
| **9** | 91.65 | 3 | 1 | 3.42 | 5.44 | 74.09 | 6.42 | 7.00 | −0.58 |
| **10** | 113.93 | 4 | 1 | 2.49 | 4.99 | 59.35 | 6.73 | 6.06 | 0.67 |
| **11** | 105.95 | 3 | 1 | 3.05 | 5.81 | 57.77 | 7.05 | 6.85 | 0.20 |
| **12** | 38.73 | 3 | 1 | 2.67 | 1.21 | 52.31 | 6.34 | 6.58 | −0.24 |
| **13** | 92.25 | 3 | 1 | 3.57 | 7.84 | 82.54 | 7.38 | 7.01 | 0.37 |
| **14** | 220.37 | 3 | 1 | 3.19 | 4.08 | 99.79 | 6.74 | 6.87 | −0.13 |
| **15** | 365.91 | 3 | 1 | 3.57 | 10.05 | 161.41 | 6.88 | 6.82 | 0.06 |
| **16** | 76.01 | 3 | 1 | 3.19 | 10.20 | 66.00 | 6.96 | 6.76 | 0.20 |
| **17** | 57.24 | 3 | 1 | 2.81 | 1.91 | 59.77 | 6.35 | 6.65 | −0.30 |
| **18** | 16.75 | 1 | 2 | 1.78 | −3.75 | 29.90 | 7.63 | 7.40 | 0.23 |
| **19** | 12.61 | 1 | 2 | 2.57 | −4.57 | 34.91 | 7.83 | 8.00 | −0.17 |
| **20** | 12.41 | 1 | 2 | 2.81 | −4.92 | 86.22 | 9.09 | 7.82 | 1.27 |
| **21** | 14.34 | 1 | 2 | 2.71 | −4.85 | 53.05 | 8.99 | 7.99 | 1.00 |
| **22** | 55.92 | 1 | 2 | 2.81 | −1.22 | 115.59 | 7.32 | 7.58 | −0.26 |
| **23** | 36.52 | 1 | 2 | 2.71 | 1.52 | 66.13 | 7.70 | 7.79 | −0.09 |
| **24** | 44.40 | 1 | 2 | 3.33 | −0.97 | 136.43 | 8.64 | 7.82 | 0.82 |
| **25** | 64.76 | 1 | 2 | 3.33 | −1.12 | 125.53 | 7.73 | 7.94 | −0.21 |
| **26** | 71.68 | 2 | 4 | 2.03 | −1.26 | 118.24 | 7.24 | 7.74 | −0.50 |
| **27** | 57.47 | 1 | 2 | 3.18 | 1.51 | 69.58 | 8.36 | 8.17 | 0.19 |
| **28** | 1331.64 | 4 | 2 | 4.20 | 16.50 | 154.28 | 8.98 | 9.09 | −0.11 |
| **29** | 47.49 | 1 | 1 | 2.96 | 2.13 | 39.85 | 6.55 | 7.62 | −1.07 |
| **30** | 171.38 | 1 | 1 | 3.42 | 10.34 | 57.54 | 7.70 | 7.89 | −0.19 |
| **31** | 176.21 | 3 | 1 | 4.16 | 17.28 | 141.92 | 6.77 | 6.97 | −0.20 |
| **32** | 10.53 | 1 | 1 | 3.08 | −4.76 | 48.19 | 9.02 | 7.74 | 1.28 |
| **33** | 728.29 | 4 | 0 | 5.32 | 37.72 | 102.40 | 7.67 | 7.76 | −0.09 |
| **34** | 13.08 | 2 | 0 | 2.79 | −2.99 | 40.53 | 6.91 | 6.60 | 0.31 |
| **35** | 16.46 | 5 | 4 | 0.51 | 0.01 | 40.39 | 5.30 | 5.90 | −0.60 |
| **36** | 1589.96 | 3 | 0 | 4.52 | 152.01 | 189.99 | 5.80 | 5.79 | 0.01 |
| **37** | 1829.74 | 5 | 1 | 4.05 | 83.17 | 186.71 | 6.93 | 7.18 | −0.25 |
| **38** | 1513.33 | 4 | 0 | 3.82 | 74.44 | 256.52 | 6.12 | 5.91 | 0.21 |
| **39** | 707.11 | 0 | 1 | 3.56 | 52.78 | 242.04 | 6.07 | 6.94 | −0.87 |
| **40** | 236.98 | 0 | 1 | 3.63 | 24.73 | 183.67 | 7.26 | 7.26 | 0 |
| **41** | 416.35 | 4 | 1 | 3.64 | 32.38 | 217.57 | 5.49 | 5.69 | −0.20 |
| **42** | 5540.21 | 6 | 3 | 3.24 | 319.38 | 463.65 | 6.57 | 6.35 | 0.22 |
| **43** | 1374.13 | 0 | 0 | 3.21 | 80.15 | 252.90 | 6.32 | 6.54 | −0.22 |
| **44** | 46.40 | 2 | 1 | 1.65 | −1.05 | 48.40 | 5.74 | 6.21 | −0.47 |
| **45** | 485.29 | 4 | 0 | 4.01 | 41.63 | 226.37 | 5.84 | 5.26 | 0.58 |
| **46** | 67.40 | 8 | 4 | 1.44 | 3.80 | 62.18 | 4.77 | 5.39 | −0.62 |
| **47** | 684.66 | 9 | 3 | 1.27 | 44.22 | 164.43 | 4.28 | 3.71 | 0.57 |
| **48** | 30.69 | 6 | 5 | 0.22 | 0.47 | 41.15 | 6.32 | 5.88 | 0.44 |
| **49** | 496.99 | 4 | 0 | 4.63 | 39.62 | 270.34 | 5.54 | 5.49 | 0.05 |

[a]The internal energy of a molecule as it sits in and constrained by the receptor model.
[b]Number of rotatable bonds.
[c]Number of hydrogen bond donors.
[d]The octanol–water partition coefficient calculated using the atom typing scheme of Ghose et al.[8]
[e]The van der Waals interaction energy of the molecule with the receptor.
[f]Principal moment of inertia.
[g]Calculated by eq (2).

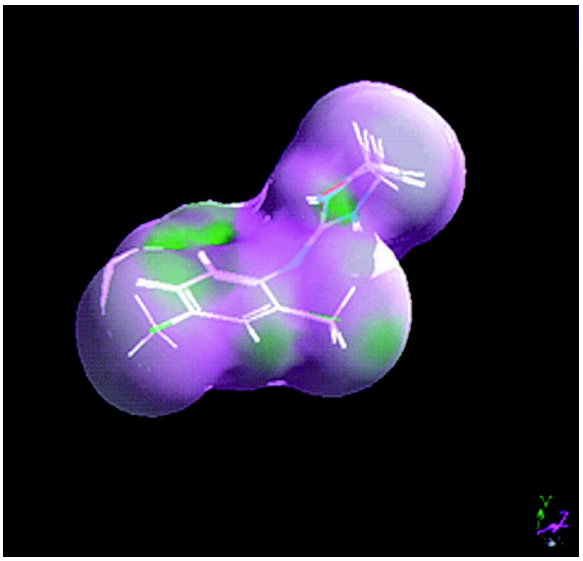

**Figure 2.** A closed RSM generated from four most active compounds (**20**, **21**, **28**, and **32**) in OA binding data set.

was studied further using molecular shape analysis (MSA),[12,13] leading to eq (4).

$$pK_i = 7.755 + 0.482H_{\text{bond donor}} + 0.788 \log P + 1.315\, F_o - 0.006PMI - X - 0.387\, Rotlbonds + 0.357H_{\text{bond acceptor}} - 3.523\, Density \quad (4)$$

where $n = 49$, $r = 0.878$, and $F = 19.638$: $H_{\text{bond donor}}$, the number of hydrogen bond donors; log of partition coefficient (log $P$), the octanol/water partition coefficient (in this atom-based approach, each atom of the molecule is assigned to a particular class, with additive contributions to the total value of log $P$); common overlap volume ratio ($F_o$), the common overlap steric volume descriptor divided by the volume of the individual molecule; principal moment of inertia (*PMI*), the moments of inertia are computed for a series of straight fines through the center of mass; number of rotatable bonds (*Rotlbonds*), the number of bonds in the current molecule having rotations that are considered to be meaningful for molecular mechanics is counted and all terminal H atoms are ignored (for example, methyl groups are not rotatable); $H_{\text{bond acceptor}}$, number of hydrogen bond acceptors; *Density*, a 3D spatial descriptor that is defined as the ratio of molecular weight to molecular volume (units of g/ml). The density reflects the types of atoms and how tightly they are packed in a molecule. Density can be related to transport and melt behavior. The moments of inertial are given by $I = \Sigma m_i d_i^2$. Distances are established along each line proportional to the reciprocal of the square root of $I$ on either side of the center of mass. The locus of these distances forms an ellipsoidal surface. The principal moments are associated with the principal axes of the ellipsoid. If all three moments are equal, the molecule is considered to be a symmetrical top. If no moments are equal, the molecule is considered to be an unsymmetrical top.

NC derivative substituted with 2,4-$Cl_2$ (**20**) had the highest potency, followed by AC-6 **32**, 2,4-$(CH_3)_2$ **21**, and 2,6-$(CH_3CH_2)_2$,4-$N_3$ **28** in binding to OAR3 of locust nervous system (Table 1). Introduction of a 3-Cl, 3-$CH_3$, 4-Cl, 2,3-$Cl_2$, 2,5-$Cl_2$, or 3-Cl,4-F to the phenyl of unsubstituted BAT **1** increased the potency, leading to **6**, **8**, **11**, **13**, **15**, or **16**, respectively. Similarly, the introduction of a 4-Br, 2,4-$Cl_2$, 2,4-$(CH_3)_2$, 2,4,5-$Cl_3$, 2,4,6-$(CH_3)_3$, or 2,6-$(CH_2CH_3)_2$,4-$N_3$ to the phenyl of unsubstituted NC **18** increased the potency, leading to **15**, **20**, **21**, **24**, **27**, or **28**, respectively. The above data suggest that phenyl ring substitution requirements for BAT and NC derivatives active as octopaminergic agonists differ substantially from each other.

## Discussion

QSAR modeling is an area of research pioneered by Hansch and Fujita. QSAR attempts to model the activity of a series of compounds using measured or computed properties of the compounds as shown in eq (1). More recently, QSAR has been extended by including the analysis three-dimensional information about the series by three-dimensional shape descriptors, as illustrated by the MSA approach in eq (2).

In drug discovery, it is common to have measured activity data for a set of compounds acting upon a particular protein but not to have knowledge of the three-dimensional structure of the protein active site. In the absence of such three-dimensional information, one can attempt to build a hypothetical model of the receptor site that can provide insight about receptor site characteristics. Such a model is known as a RSM, which provides compact and quantitative descriptors which capture three-dimensional information about a putative receptor site. These descriptors were used for predictive QSAR models. This approach is effective for the analysis of data sets where activity information is available but the structure of the receptor site is unknown. RSM attempts to postulate and represent the essential features of a receptor site itself, rather than the common features of the molecules that bind to it. RSMs differ from pharmacophore models in that the former tries to capture essential information about the receptor, while the latter captures information about the commonality of compounds that bind. Pharmacophore models tend

to be geometrically underconstrained (while topologically overconstrained); this steric underconstraint leads to false positives, that is, compounds that are deemed active by the model but which are inactive when tested. RSMs, on the other hand, tend to be geometrically overconstrained (and topologically neutral) since, in the absence of steric variation in a region, they assume the tightest steric surface which fits all training compounds.

OA is not likely to penetrate either the cuticle or the central nervous system of insects effectively, since it is fully ionized at physiological pH. Derivatization of the polar groups would be one possible solution to this problem in trying to develop potential pest-control agents. The above QSAR, RSM, and MSA studies show that agonists or antagonists with certain substitutions can be potential ligands to OA receptors. They may help to point the way towards developing extremely potent and relatively specific OA agonists and antagonists, leading to potential post-control agents. In order to optimize the activities of these compounds as OA-agonists and antagonists, more detailed experiments are in progress.

## Experimental Procedure

### Compounds and their bioassay activities

Forty-nine molecules that employ a single [$^3$H]OA binding criteria were selected from recent publications by Roeder.[3–7] The respective $IC_{50}$ values of the antagonists had been corrected according to the Cheng–Prussof correction [$K_i = IC_{50}/(1 + K/K_d)$] in order to obtain the $K_i$ values. Their chemical structures and experimental activities are listed in Figure 1 and Table 1.

### QSAR calculations

QSAR was calculated with a Macintosh personal computer, an NEC PC-9801VM personal computer system using QSAR 2.05, a program designed for Hansch-Fujita's QSAR analysis (Tanabe Pharmac. Co., Ltd., Osaka, Japan), or Molecular Simulations Incorporated's Cerius$^2$ QSAR environment (Burlington, MA, USA).

### Molecular modeling

Molecules were built using the Tektronix CAChe WorkSystem™ package of programs (Oxford Molecular Group, OR, USA) on an Apple PowerMac 8100/112 (100 M RAM, OS J-7.5.1). Molecular geometries in cartesian coordinates of (*R*)-OA was obtained from available X-ray crystallographic data (Cambridge data base system) and used as one of the starting points for subsequent calculations. Starting from the characteristic average value angles of each structure, it was energy-minimized using Molecular Mechanics (MM2) and MOPAC (PM3) program (version 94) included in the CAChe system (version 3.8) by connecting to a network server IBM R86000. The geometries of the energy-minimized molecules were checked using the CAChe 3D visualization tools. Energy-minimized molecules were generated as described as follows: firstly, a thorough conformational search had been carried out within the molecular mechanics technique (MM2 force field) using initially the steepest descents minimization methods followed by the conjugate gradient until the gradient was below 0.001 Kcal/mol in no more than 5000 steps. Then the unique minima was fully optimized with MOPAC (PM3 force field). Thus a non-linear least squares (NLLSQ) gradient minimization route by Bartel's method was applied for gradient norm, while with Broyden–Fletcher–Goldfarb–Shanno (BFGS) method, the nearest minimum-energy geometry was located for optimization. And a full Mulliken population was checked to analyze the final restricted Hartree–Fock (RHF) wavefunction. The PRECISE keyword was also used in order to increase the geometric and electronic convergency criteria. In some cases, the GEO-OK keyword was used to prevent BFGS routine from unexpected termination. Other settings were used as default.

### Receptor surface modeling[14,15]

All experiments were conducted on a Silicon Graphics Indigo/Extream, running under the IRIX 5.3 operating system. RSMs were applied against previously described data sets and their functionality is available as part of Molecular Simulations Incorporated's Cerius$^2$ modeling environment (Burlington, MA). Once the desired RSM has been constructed, all the structures in the training and test sets were evaluated against the model. The evaluation consists of computing several energetic descriptors that are based upon the interactions between ligand and model. For this work, six descriptors were generated. Receptor descriptors: *IntraEnergy*, the internal energy of a molecule as it sits in and constrained by the receptor model; *InterVDWEnergy*, the van der Waals interaction energy of the molecule with the receptor. Structural descriptors: *Rotlbands*, number of rotatable bonds; $H_{\text{bond donor}}$, number of hydrogen bond donors. Spatial descriptor: *PMI*, principal moment of inertia. Thermodynamic descriptor: log *P*, the octanol–water partition coefficient calculated using the atom typing scheme of Ghose et al.[8]

## Acknowledgements

The authors thank Professor Emeritus Toshio Fujita of Kyoto University and Tanabe Pharmacology Co. Japan for donating the QSAR program. They are grateful to

Professor Takaaki Sonoda at the Institute of Advanced Material Study, Kyushu University, Japan for allowing us to use the CAChe Groupserver IBM RS6000. This work was supported in part by a Grant-in-Aid for scientific research from the Ministry of Education, Science, and Culture of Japan.

## References

1. Evans, P. D. *Comp. Mol. Neurobiol.* **1993**, 287.
2. Evans, P. D. *J. Physiol.* **1981**, *318*, 99.
3. Roeder, T.; Nathanson, J. A. *Neurochem. Res.* **1993**, *18*, 921.
4. Roeder, T.; Gewecke, M. *Biochem. Pharm.* **1990**, *39*, 1793.
5. Roeder, T. *Life Science* **1991**, *50*, 21.
6. Roeder, T. *Br. J. Pharmac.* **1995**, *114*, 210.
7. Roeder, T. *Eur. J. Pharmac.* **1990**, *191*, 221.
8. Ghose, A. K.; Pritchett, A.; Crippen, G. M. *J. Comp. Chem.* **1988**, *9*, 80.
9. Stewart, J. J. P. *J. Comp. Chem.* **1989**, *10*, 221.
10. Klamt, K.; Schüürmann, G. *J. Chem. Soc. Perkin Trans.* **1993**, *2*, 799.
11. Meylan, W. M.; Howard, P. H. *J. Pharm. Sci.* **1995**, *84*, 83.
12. Burke, B. L.; Hopfinger, A. J. *J. Med. Chem.* **1990**, *33*, 274.
13. Burke, B. L.; Hopfinger, A. J. In *Molecular Similarity*; Johnson, M. A.; Maggiora, G. M., Eds; John Wiley and Sons: New York. **1990**, pp 11–73.
14. Hahn, M. *J. Med. Chem.* **1995**, *38*, 2080.
15. Hahn, K.; Rogers, D. *J. Med. Chem.* **1995**, *38*, 2091.