Structure-Activity Relationships for Receptor Binding of Environmental Estrogens: Analogies with β-Cyclodextrin Inclusion Complexes

Takahiro SUZUKI* and Stuart SHAPIRO**

Abstract

A group contribution model for calculating relative binding affinities (RBAs) of structurally diverse ligands to rat uterine estrogen receptor-α (ER-α) was developed. Least squares regression yielded an equation correlating RBA with the first-order molecular connectivity index $^1\chi$ (a measure of molecular bulk) and weighted structural fragments (measures of molecular forces between the receptor and ligand). This model showed a very good index of determination ($r^2 = 0.848$) between experimental and calculated $\log_{10}$ RBAs for a training set comprised of 128 compounds. Similarities between ER-α : (xeno) estrogen binding and formation of β-cyclodextrin : (xeno)estrogen inclusion complexes were explored by comparing experimental free energies for (xeno)estrogens bound to ER-α with those for the same ligands complexed with natural and 2-hydroxpropyl-β-cyclodextrin. Indices of determination in the range 0.6-0.7 were obtained for the ER-α and β-cyclodextrin systems. Development of a group contribution method for *in silico* screening of potential xenoestrogens represents a practical alternative to tedious, costly receptor-ligand binding assays and animal experimentation.

Keywords: environmental estrogen, group-contribution method, molecular connectivity, QSAR, xenoextrogen

1. Introduction

Cyclodextrins (CDs) are cyclic polysaccharides usually comprised of six (α-CD), seven (β-CD), or eight (γ-CD) α-**d**-glucosyl units connected by 1,4'-*O*-glycosidic bonds. CDs assume an open conical shape, with their hydroxyl moieties directed towards the outside, leaving a relatively non-polar interior which can accommodate a variety of organic and inorganic "guest" molecules (Connors, 1997). The natural CDs

* Natural Science Laboratory, Toyo University, 5-28-20 Hakusan, Bunkyo-ku, Tokyo 112-8606, Japan
** Basilea Pharmaceutica International AG, Grenzacherstrasse 487, Postfach, CH-4005 Basel, Switzerland

have been modified by adding alkyl or hydroxyalkyl groups to the glucosyl hydroxyls. The binding properties of natural and modified CDs have led to their being used in a wide range of practical applications, including solubilization/stabilization of pharmaceuticals and stationary phases for chromatographic separation (Szejtli, 1988; Szejtli, 1998; Hedges, 1998). CDs also represent interesting models for interactions of ligands with enzyme and receptor binding sites (Szejtli, 1988).

Most steroids and many other chemicals that interact with endocrine systems (Brunstrom et al., 2003; Colborn et al., 1993; Hutchinson and Pickford, 2002; Singleton and Khan, 2003) fall within the size range of guest molecules for β-CD cavities. Considerable effort has been invested to understand the toxic activities of xenoestrogens ("endocrine disruptors"), defined by the United States Environmental Protection Agency as substances which interfere with the synthesis, secretion, transport, binding, action, or elimination of hormones responsible for the maintenance of homeostasis, reproduction, development, and/or behavior (Kavlock, 1996). To date the European Union has classified 553 synthetic substances as potential xenoestrogens (European Commission, 2001). *In vitro* bioassays for endocrine disruptive activity using MCF-7 human breast adenocarcinoma cells and other estrogen-sensitive cell lines are very costly, time-consuming, and labor-intensive procedures; therefore, quantitative structure-activity relationship (QSAR) models, using traditional Hansch and molecular connectivity methods as well as sophisticated 3D techniques, have been applied to analyzing ligand binding to estrogen receptors (ERs) (Blair et al., 2000; Bradbury et al., 1996; Gantchev et al., 1994; Gao et al., 1999a,b; Gao, 2001; Hutchinson and Pickford, 2002; Kramer and Giesy, 1999; Roncaglioni et al., 2004; Sadler et al., 1998; Shi et al., 2001; Suzuki et al., 2001; Tong et al., 1997a,b; Tong et al., 1998; Waller et al., 1995, 1996; Waller, 2004; Zheng and Tropsha, 2000).

Steroids (Wallimann et al., 1997) and other compounds (Brunstrom et al., 2003; Colborn et al., 1993; Hutchinson and Pickford, 2002; Singleton and Khan, 2003) that interact with endocrine systems also form inclusion complexes with natural or modified β-CDs in aqueous milieux. The molecular shape and size of the guest molecule, desolvation of the guest molecule (release of "high energy" water from the "host" CD cavity), relief of conformational strain in the uncomplexed CD, van der Waals interactions, and hydrogen bonds between moieties of the guest and host are believed to contribute in varying degrees to the formation and stability of CD inclusion complexes (Connors, 1997). Schematic representations of the inclusion complex formed between β-CD and 17β-estradiol, and of the binding of 17β-estradiol to estrogen receptor-α (ER-α) (Brzozowski et al., 1997) are shown in Figure 1. Though there are differences in cavity size and shape between estrogen receptors and β-cyclodextrins, similar steric and energetic interactions likely drive both complexation of 17β-estradiol by β-CD and binding of 17β-estradiol by ER-α. Therefore, CD binding of xenoestrogens may func-

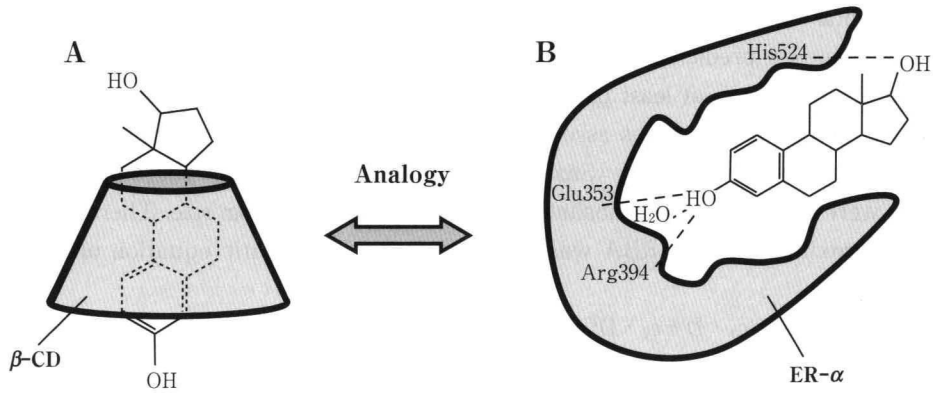


Fig. 1. β-CD:17β-estradiol complex (A); ER-α:17β-estradiol complex (B).

tion as surrogates for ER-α binding of xenoestrogens.

Diverse computational tools, including molecular modeling, neural networks, and 2D- and 3D- QSAR/QSPR, have been used to elucidate the factors governing host: guest interactions in CD inclusion complexes and to predict their thermodynamic stabilities (Katritzky et al., 2004; Klein et al., 2000; Lipkowitz, 1998; Liu and Guo, 1999; Matsui et al., 1985; Park and Nah, 1994; Suzuki et al., 2000). Nonlinear group contribution models (GCMs) for calculating binding constants or free energies of complexation of guest molecules with natural α- and β-CDs have been reported (Suzuki, 2001). Here we present a GCM, derived from experimental binding data for a structurally diverse set of molecules, for calculating relative binding affinities (RBAs) of ligands to ER-α, as well as a correlation between free energies of complexation of molecules with native and modified β-CDs and RBAs for xenoestrogens with ER-α.

2. Methods

Relative binding affinities

Binding affinities for rat uterine cytosolic estrogen receptor ER-α were determined by competitive binding assay with [^{3}H] 17β-estradiol (E_2) (Blair et al., 2000; Shi et al., 2001; Waller et al., 1996). The relative binding affinity (RBA) for a compound "X" is a dimensionless number defined as 100 times the molar ratio E_2/X required to decrease E_2 binding to ER-α by 50%. On the RBA scale, 17β-estradiol is assigned a value of 100, with lower affinity analogues having lower values and higher affinity analogues having higher values.

QSAR derivation

A training set of 128 molecules, ranging from 4-methylphenol (molecular weight 108.14) to ICI-164384 (molecular weight 525.82), is presented in Table 1, along with

experimental $\log_{10}$ RBA values. The structures for 128 molecules are shown in Figure 2.

Derivation of a predictive model for (xeno)estrogen RBAs follows from the assumption that $\log_{10}$ RBA is at least partially describable in terms of ligand fragment contributions, dimensionless values associated with single atoms or multiatom assemblies. Moreover, in an earlier study (Suzuki et al., 2000) natural β-CD:guest binding constants were found to have a nonlinear dependence on molecular size. Therefore, as a first approximation $\log_{10}$ RBA was assumed to be a quadratic equation of general form:

$$\log_{10} \mathrm{RBA} = c_1 \cdot D + c_2 \cdot D^2 + \sum(n_i \cdot G_i) + c_0 \quad (1)$$

where

D = descriptor representing molecular bulk or size;

c_1, c_2 = regression coefficients;

G_i = value assigned to ligand structural fragment i;

n_i = number of times molecular fragment i occurs in the ligand; and

c_0 = regression constant.

As possible descriptors of molecular bulk or size, zeroth-order ($^0\chi$) and first-order ($^1\chi$) molecular connectivity indices (Kier and Hall, 1986), as well as molecular weights (Mw), were calculated for each molecule in the data set (Table 1). Regression coefficients were determined by the method of least squares, using Excel Multivariate Analysis v3.0 (Esumi Co., Ltd., Tokyo, Japan) on a microcomputer running the Windows XP operating system.

Derivation of fragment constants

The quality of a GCM (i.e., the precision of its predictions and their reliability) depends upon the definition of molecular fragments and their assigned weightings. After considering the available number of data points, facile identification of groups contributing molecular structure, and implicit inclusion of the constitutive factors, a set of fragments similar to that used earlier by Lydersen (1955) for estimating critical properties was employed. Fragment constants G_i for 26 chemical groups are listed in Table 2. $I_{OH\text{-}OH}$ is an indicator variable, equal to either 0 or 1 depending upon the presumptive absence or presence, respectively, of cooperative hydrogen bonding involving oxygen atoms between the ligand and ER-α. Three examples of the decomposition of molecules into defined fragments are shown in Figure 3.

Determination of binding constants for β-CD:guest inclusion complexes

The thermodynamic stability of a 1:1 CD:guest inclusion complex can be expressed as the dissociation constant of that complex, $K^{diss}_{(1:1)}$. Values of $K^{diss}_{(1:1)}$ for inclusion complexes between natural β-CD or 2-hydroxypropyl-β-CD (2-HP-β-CD) (Yoshida et al., 1988) and 11 estrogens or xenoestrogens (kepone, 4-nonylphenol, bisphenol A, methoxychlor, estrone, 17β-estradiol, 17β-estriol, ethynylestradiol, di-*n*-butyl phthalate, butyl benzyl phthalate, endosulfan) were determined by the solubility method

Table 1. Experimental and calculated $\log_{10}$ RBA values for the rat uterine estrogen receptor (ER-α) for 128 molecules.

no.	compound	formula	$\log_{10}$ RBA		$I^*_{OH\text{-}OH}$	ref.**
			obs'd	calc'd		
1	4-methylphenol	C_7H_8O	−4.50	−3.76	0	3
2	4-chloro-3-methylphenol	C_7H_7ClO	−3.38	−3.34	0	3
3	2-chloro-4-methylphenol	C_7H_7ClO	−3.66	−3.34	0	3
4	4-chloro-2-methylphenol	C_7H_7ClO	−3.67	−3.34	0	3
5	3-ethylphenol	$C_8H_{10}O$	−3.87	−3.67	0	3
6	4-ethylphenol	$C_8H_{10}O$	−4.17	−3.67	0	3
7	methyl-4-hydroxybenzoate	$C_8H_8O_3$	−3.44	−3.47	0	2
8	ethyl-4-hydroxybenzoate	$C_9H_{10}O_3$	−3.22	−3.32	0	2
9	4-*sec*-butylphenol	$C_{10}H_{14}O$	−3.37	−2.99	0	3
10	2-*sec*-butylphenol	$C_{10}H_{14}O$	−3.54	−2.99	0	3
11	4-*tert*-butylphenol	$C_{10}H_{14}O$	−3.61	−2.81	0	3
12	kepone	$C_{10}Cl_{10}O$	−1.89	−2.20	0	3
13	propyl paraben	$C_{10}H_{12}O_3$	−3.22	−3.13	0	3
14	4-*tert*-amylphenol	$C_{11}H_{16}O$	−3.26	−2.58	0	3
15	butyl-4-hydroxybenzoate	$C_{11}H_{14}O_3$	−3.07	−2.94	0	3
16	2,4'-dichlorobiphenyl	$C_{12}H_8Cl_2$	−3.61	−2.64	0	3
17	4-phenylphenol	$C_{12}H_{10}O$	−3.04	−3.20	0	3
18	3-phenylphenol	$C_{12}H_{10}O$	−3.44	−3.20	0	3
19	2,3,4,5-tetrachloro-4'-biphenylol	$C_{12}H_6Cl_4O$	−0.64	−1.17	0	3
20	2',5'-dichloro-4-biphenylol	$C_{12}H_8Cl_2O$	−1.44	−2.19	0	3
21	4-chloro-4'-biphenylol	$C_{12}H_9ClO$	−2.18	−2.69	0	3
22	2-chloro-4-biphenylol	$C_{12}H_9ClO$	−2.77	−2.64	0	3
23	4,4'-sulfonyldiphenol	$C_{12}H_{10}O_4S$	−3.07	−3.69	0	3
24	4-heptyloxyphenol	$C_{13}H_{20}O_2$	−2.88	−3.17	0	3
25	4-benzyloxyphenol	$C_{13}H_{12}O_2$	−3.44	−3.46	0	3
26	2,2'-methylenebis (4-chlorophenol)	$C_{13}H_{10}Cl_2O_2$	−2.45	−2.46	0	3
27	bis (4-hydroxyphenyl) methane	$C_{13}H_{12}O_2$	−3.02	−2.48	0	3
28	4,4'-dihydroxybenzophenone	$C_{13}H_{10}O_3$	−2.46	−2.81	0	3
29	2,4-dihydroxybenzophenone	$C_{13}H_{10}O_3$	−2.61	−2.81	0	3
30	4-phenethylphenol	$C_{14}H_{14}O$	−2.69	−2.83	0	3
31	4-tert-octylphenol	$C_{14}H_{22}O$	−1.82	−1.59	0	3
32	4-octylphenol	$C_{14}H_{22}O$	−2.34	−2.54	0	2
33	4,4'-dihydroxystilbene	$C_{14}H_{12}O_2$	−0.55	−0.69	1	3
34	4,4'-ethylene diphenol	$C_{14}H_{12}O_2$	−1.44	−0.69	1	3
35	2,2',4,4'-tetrahydroxybenzil	$C_{14}H_{10}O_6$	−0.68	−0.25	1	3
36	*o,p'*-DDT	$C_{14}H_9Cl_5$	−2.85	−1.53	0	3
37	4,4'-(dichlorovinylidene) diphenol	$C_{14}H_{10}Cl_2O_2$	0.42	−0.10	0	3
38	bisdemethylmethoxychlor	$C_{14}H_{11}Cl_3O_2$	−0.60	−1.48	0	3
39	benzyl 4-hydroxybenzoate	$C_{14}H_{10}O_3$	−2.54	−2.69	0	3
40	heptyl 4-hydroxybenzoate	$C_{14}H_{20}O_3$	−2.09	−2.38	0	3
41	coumestrol	$C_{15}H_8O_5$	−0.05	−0.25	0	3
42	chalcone	$C_{15}H_{12}O$	−2.82	−3.42	0	3
43	4-nonylphenol	$C_{15}H_{24}O$	−1.53	−2.36	0	3
44	4'-hydroxychalcone	$C_{15}H_{12}O_2$	−2.43	−2.91	0	3
45	4-hydroxychalcone	$C_{15}H_{12}O_2$	−2.55	−2.85	0	3

Table 1 (continued)

no.	compound	formula	$\log_{10}$ RBA		$I^*_{OH\text{-}OH}$	ref.**
			obs'd	calc'd		
46	bisphenol A	$C_{15}H_{16}O_2$	−2.11	−1.59	0	3
47	4-[2,2-dichloro-1-(4-methoxy-phenyl) ethenyl]phenol	$C_{15}H_{12}Cl_2O_2$	−0.64	−0.37	0	3
48	monodemethylmethoxychlor	$C_{15}H_{13}Cl_3O_2$	−0.89	−1.77	0	3
49	6-hydroxyflavone	$C_{15}H_{10}O_3$	−3.41	−3.49	0	3
50	4'-hydroxyflavanone	$C_{15}H_{12}O_3$	−2.65	−3.20	0	3
51	3'-hydroxyflavanone	$C_{15}H_{12}O_3$	−2.78	−3.23	0	3
52	6-hydroxyflavanone	$C_{15}H_{12}O_3$	−3.05	−3.22	0	3
53	7-hydroxyflavanone	$C_{15}H_{12}O_3$	−3.73	−3.13	0	3
54	equol	$C_{15}H_{14}O_3$	−0.82	−0.29	1	3
55	2-ethylhexyl 4-hydroxybenzoate	$C_{15}H_{22}O_3$	−1.74	−1.90	0	3
56	6,4'-dihydroxyflavone	$C_{15}H_{10}O_4$	−0.82	−1.21	1	3
57	4,2',4'-trihydroxychalcone	$C_{15}H_{12}O_4$	−1.26	−1.81	0	3
58	daidzein	$C_{15}H_{10}O_4$	−1.65	−1.17	1	3
59	7,3',4'-trihydroxyisoflavone	$C_{15}H_{10}O_5$	−0.35	−0.66	1	3
60	3,6,4'-trihydroxyflavone	$C_{15}H_{10}O_5$	−0.35	−0.66	1	3
61	genistein	$C_{15}H_{10}O_5$	−0.36	−0.65	1	3
62	apigenin	$C_{15}H_{10}O_5$	−1.55	−2.22	0	3
63	baicalein	$C_{15}H_{10}O_5$	−3.35	−2.21	0	3
64	naringenin	$C_{15}H_{12}O_5$	−2.13	−2.18	0	3
65	phloretin	$C_{15}H_{14}O_5$	−1.16	−1.33	0	3
66	kaempferol	$C_{15}H_{10}O_6$	−1.61	−2.86	0	3
67	fisetin	$C_{15}H_{10}O_6$	−2.35	−2.85	0	3
68	morin	$C_{15}H_{10}O_7$	−3.35	−2.33	0	3
69	myricetin	$C_{15}H_{10}O_8$	−2.75	−1.85	0	3
70	formononetin	$C_{16}H_{12}O_4$	−2.98	−3.04	1	3
71	1,3-diphenyltetramethyldisiloxane	$C_{16}H_{22}OSi_2$	−3.16	−3.16	0	3
72	dimethylstilbestrol	$C_{16}H_{16}O_2$	1.16	1.80	1	3
73	bisphenol B	$C_{16}H_{18}O_2$	−1.07	−1.43	0	3
74	methoxychlor	$C_{16}H_{15}Cl_3O_2$	−3.20	−2.07	0	1
75	biochanin A	$C_{16}H_{12}O_5$	−2.37	−2.52	0	3
76	prumetin	$C_{16}H_{12}O_5$	−2.74	−2.60	0	3
77	dimethyl allenolic acid	$C_{17}H_{20}O_3$	−0.02	−1.08	0	3
78	diphenolic acid	$C_{17}H_{18}O_4$	−3.13	−3.32	0	3
79	3-deoxyestrone	$C_{18}H_{22}O$	−2.20	−0.72	0	0
80	estra-1,3, 5(10)-trien-3-ol	$C_{18}H_{24}O$	1.26	0.70	0	0
81	3-deoxyestradiol	$C_{18}H_{24}O$	−0.30	−1.04	0	3
82	17-deoxyestradiol	$C_{18}H_{24}O$	1.14	0.70	0	0
83	4-dodecylphenol	$C_{18}H_{30}O$	−1.73	−1.80	0	2
84	dienestrol	$C_{18}H_{18}O_2$	1.57	2.23	1	3
85	diethylstilbestrol	$C_{18}H_{20}O_2$	2.60	2.17	1	3
86	estron	$C_{18}H_{22}O_2$	0.86	−0.17	0	2
87	3-hydroxy-estra-1,3,5 (10)-trien-16-one	$C_{18}H_{22}O_2$	−0.29	−0.18	0	2
88	*meso*-hexestrol	$C_{18}H_{22}O_2$	2.48	0.61	1	3
89	**dl**-hexestrol	$C_{18}H_{22}O_2$	0.56	0.61	1	2
90	17α-estradiol	$C_{18}H_{24}O_2$	0.49	1.05	1	2

Table 1 (continued)

no.	compound	formula	$\log_{10}$ RBA		$I^*_{OH\text{-}OH}$	ref.**
			obs'd	calc'd		
91	17β-estradiol	$C_{18}H_{24}O_2$	2.00	1.05	1	2
92	17β-estriol	$C_{18}H_{24}O_3$	0.99	−0.15	1	2
93	6α-hydroxy-estradiol	$C_{18}H_{24}O_3$	−0.15	−0.15	1	2
94	4-hydroxy-estradiol	$C_{18}H_{24}O_3$	1.82	1.59	1	3
95	2-hydroxy-estradiol	$C_{18}H_{24}O_3$	1.47	1.64	1	3
96	3,3'-dihydroxyl hexestrol	$C_{18}H_{22}O_4$	1.19	1.72	1	3
97	nordihydroguariaretic acid	$C_{18}H_{22}O_4$	−1.51	0.14	0	3
98	4-ethyl-7-OH-3-(*p*-methoxyphenyl)-dihydro-1-benzopyran-2-one	$C_{18}H_{18}O_4$	−0.05	−0.28	0	3
99	α-zearalenol	$C_{18}H_{24}O_5$	1.63	0.32	1	3
100	β-zearalenol	$C_{18}H_{24}O_5$	−0.69	0.25	1	3
101	zearalanone	$C_{18}H_{24}O_5$	0.32	−0.33	0	3
102	α-zearalanol	$C_{18}H_{26}O_5$	1.48	0.90	1	3
103	β-zearalanol	$C_{18}H_{26}O_5$	−0.19	0.84	1	3
104	hexestrol monomethyl ether	$C_{19}H_{24}O_2$	0.97	0.63	1	3
105	3-methylestriol	$C_{19}H_{26}O_3$	−1.65	−1.53	0	3
106	*p*-(α,β-diethyl-*p*-methylphenethyl)-mesophenol	$C_{19}H_{24}O$	0.60	−0.91	0	3
107	mestilbol	$C_{19}H_{22}O_2$	1.31	0.34	0	3
108	5α-androstane-3α,17β-diol	$C_{19}H_{30}O_2$	−0.92	0.01	1	3
109	doisynoestrol	$C_{19}H_{22}O_3$	−2.74	−2.42	0	3
110	aurin	$C_{19}H_{14}O_3$	−1.50	−1.15	0	3
111	phenol red	$C_{19}H_{14}O_5S$	−3.25	−2.63	0	3
112	2,6-dimethyl hexestrol	$C_{20}H_{26}O_2$	1.11	0.43	0	3
113	triphenylethylene	$C_{20}H_{16}$	−2.78	−1.43	0	3
114	ethynylestradiol	$C_{20}H_{24}O_2$	2.28	2.16	1	3
115	diethylstilbestrol dimethyl ether	$C_{20}H_{24}O_2$	−1.25	0.04	0	3
116	norethynodrel	$C_{20}H_{26}O_2$	−0.65	0.35	0	3
117	16β-hydroxy-16-methyl-17β−estradiol 3-methyl ether	$C_{20}H_{28}O_3$	−1.48	−1.82	0	3
118	phenolphthalein	$C_{20}H_{14}O_4$	−1.87	−0.33	0	3
119	phenolphthalin	$C_{20}H_{16}O_4$	−3.67	−2.73	0	3
120	mestranol	$C_{21}H_{26}O_2$	0.35	0.32	0	3
121	moxestrol	$C_{21}H_{26}O_3$	1.14	0.29	0	3
122	tamoxifen	$C_{26}H_{29}NO$	0.21	0.15	0	3
123	toremifene	$C_{26}H_{28}ClNO$	0.14	0.16	0	3
124	clomiphene	$C_{26}H_{28}ClNO$	−0.14	0.16	0	3
125	4-hydroxytamoxifen	$C_{26}H_{29}NO_2$	2.24	0.69	0	3
126	droloxifene	$C_{26}H_{29}NO_2$	1.18	0.69	0	3
127	nafoxidine	$C_{29}H_{31}NO_2$	−0.14	−0.14	0	3
128	ICI 164384	$C_{34}H_{55}NO_3$	1.16	2.96	1	3

* See Table 2.

** 1) Waller et al., 1996; 2) Blair et al., 2000; 3) Shi et al., 2001.

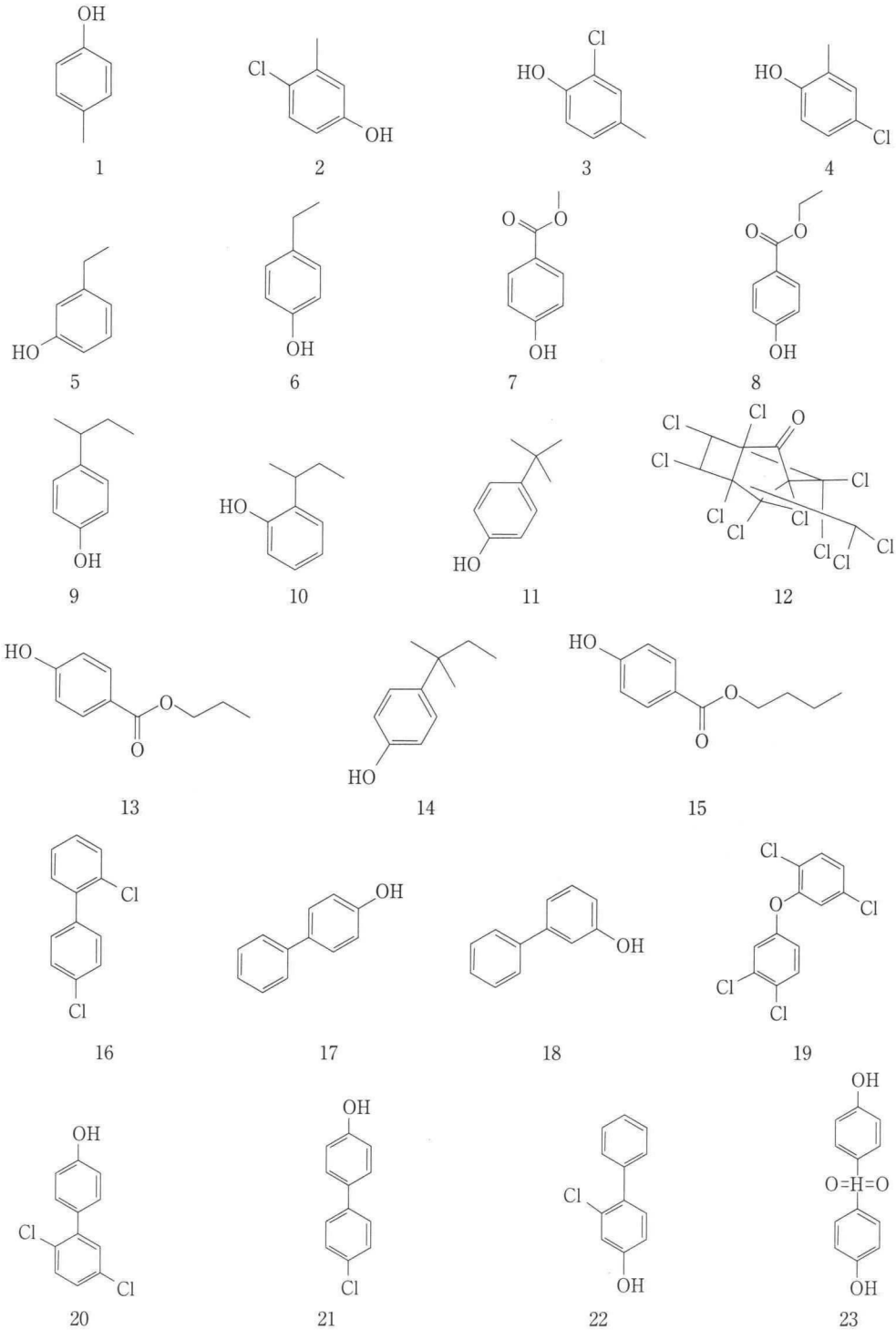
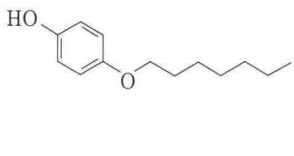
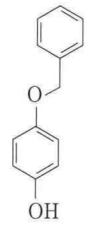


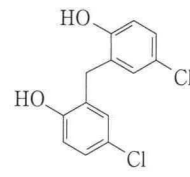
Fig. 2. Chemical structures of the compounds in Table 1.



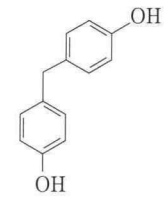
24



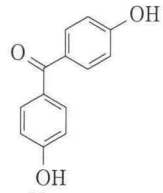
25



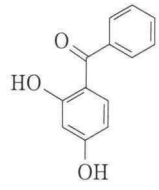
26



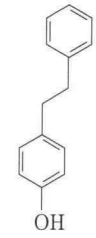
27



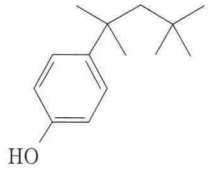
28



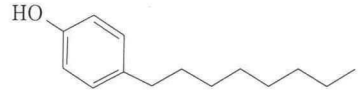
29



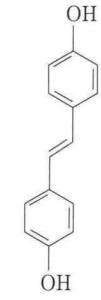
30



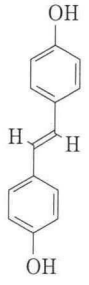
31



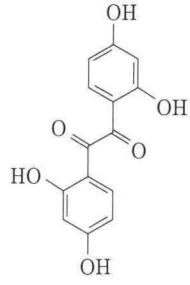
32



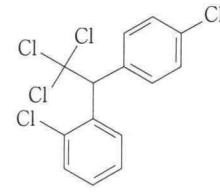
33



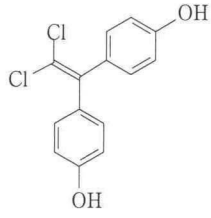
34



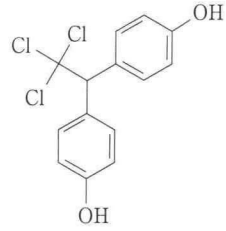
35



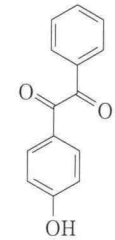
36



37

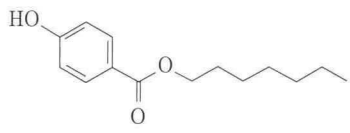


38

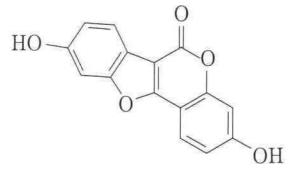


39

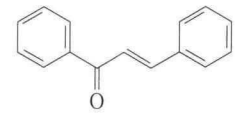
Fig. 2 (continued)



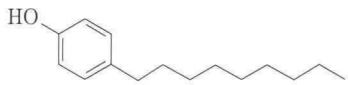
40



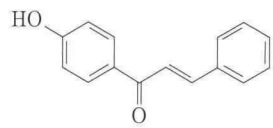
41



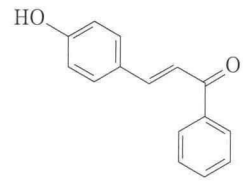
42



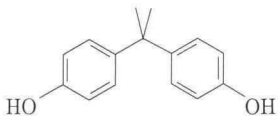
43



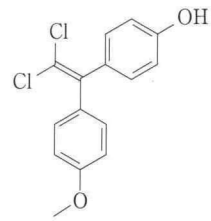
44



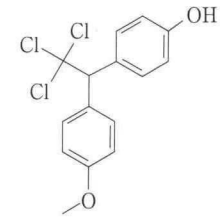
45



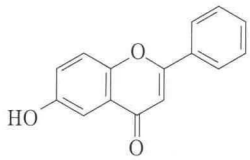
46



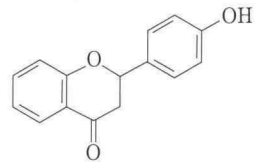
47



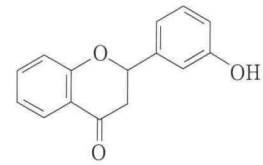
48



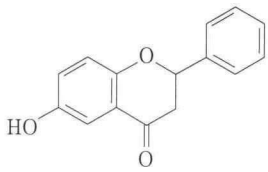
49



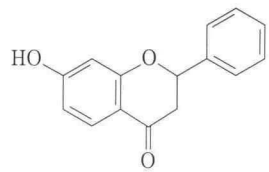
50



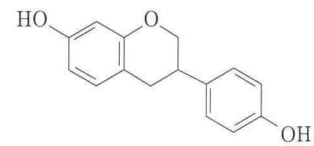
51



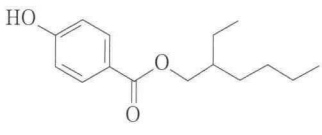
52



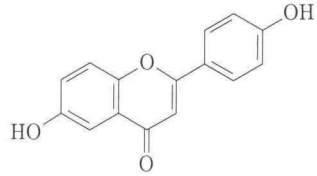
53



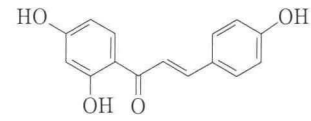
54



55

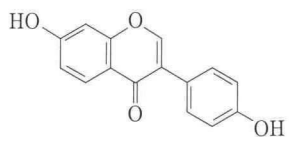


56

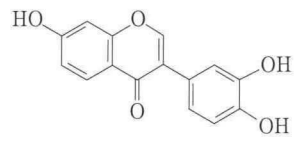


57

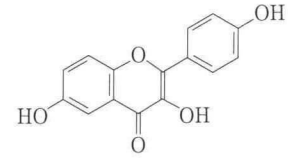
Fig. 2 (continued)



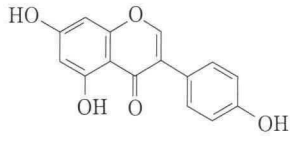
58



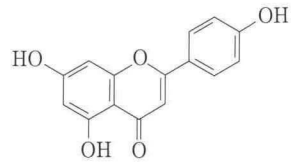
59



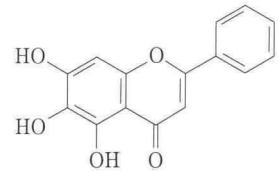
60



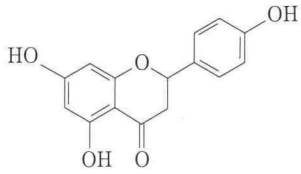
61



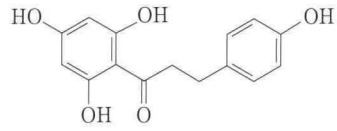
62



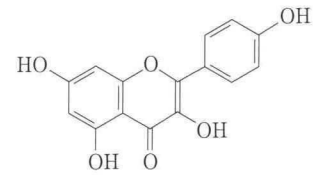
63



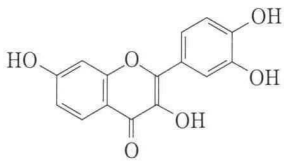
64



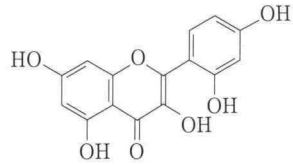
65



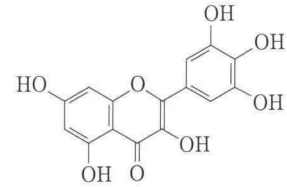
66



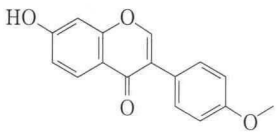
67



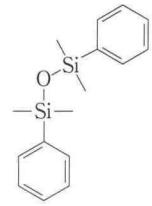
68



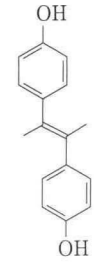
69



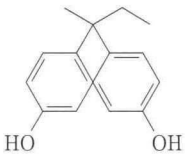
70



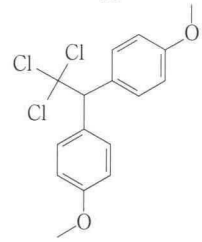
71



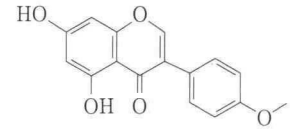
72



73

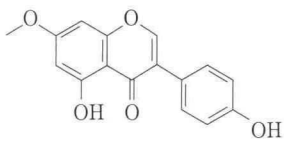


74

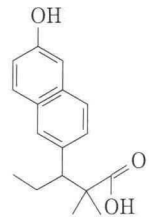


75

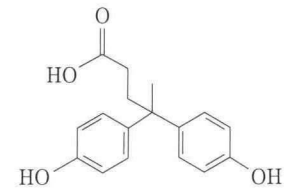
Fig. 2 (continued)



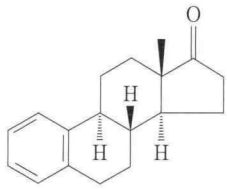
76



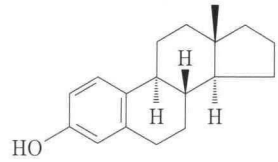
77



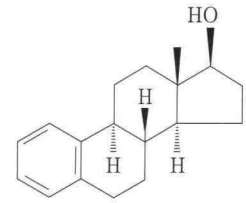
78



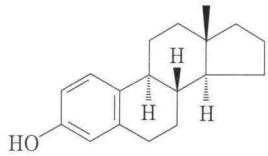
79



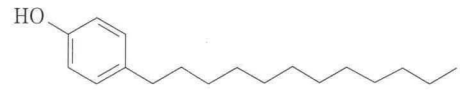
80



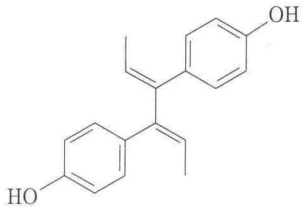
81



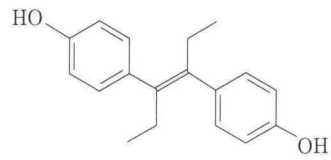
82



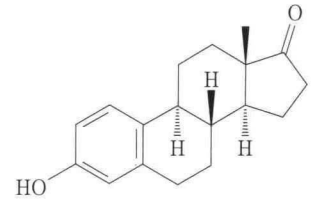
83



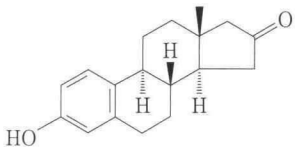
84



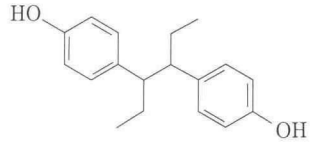
85



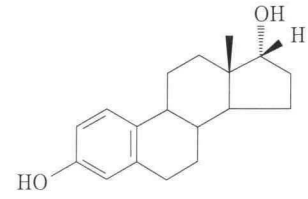
86



87

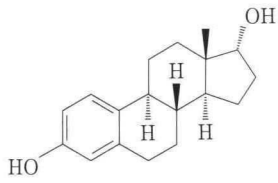


88, 89

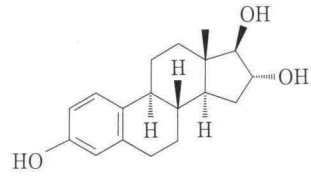


90

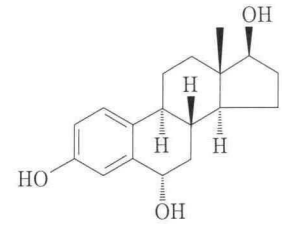
Fig. 2 (continued)



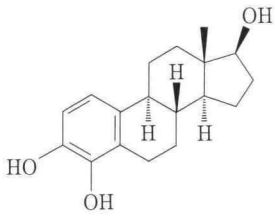
91



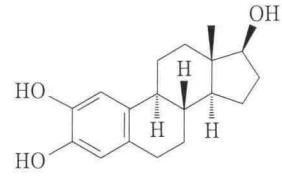
92



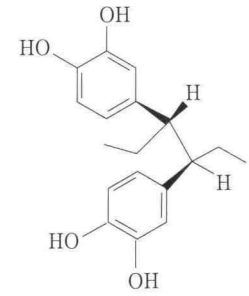
93



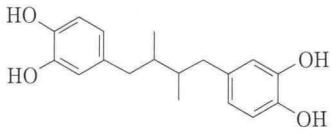
94



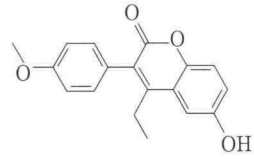
95



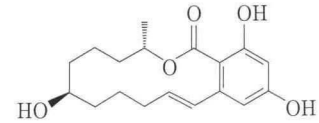
96



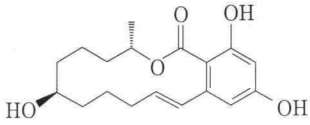
97



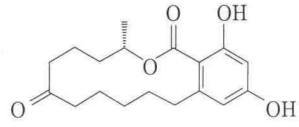
98



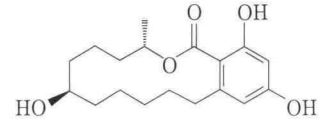
99



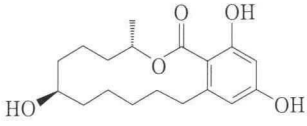
100



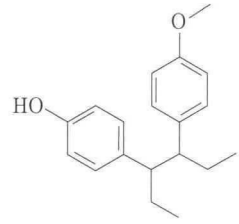
101



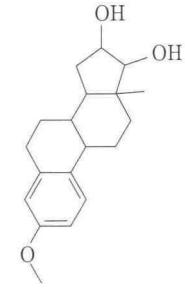
102



103

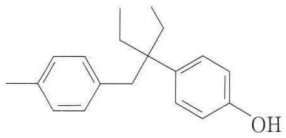


104

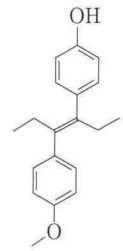


105

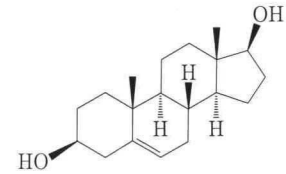
Fig. 2 (continued)



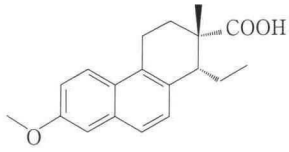
106



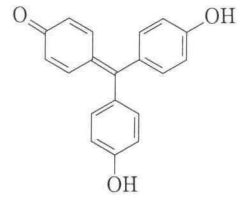
107



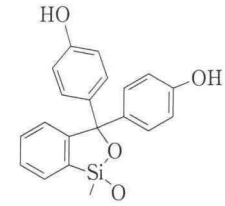
108



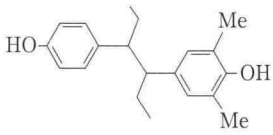
109



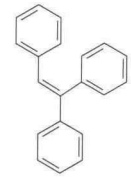
110



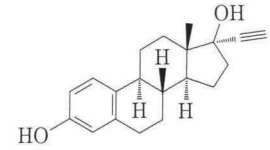
111



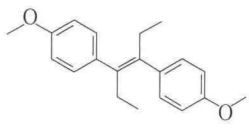
112



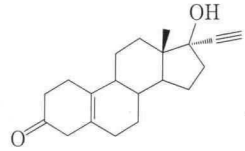
113



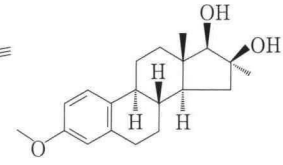
114



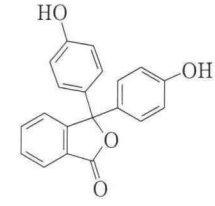
115



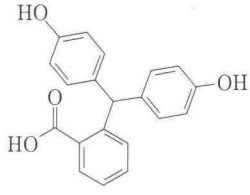
116



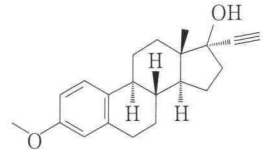
117



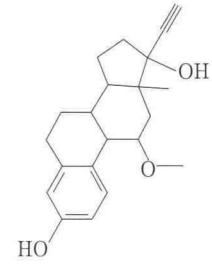
118



119

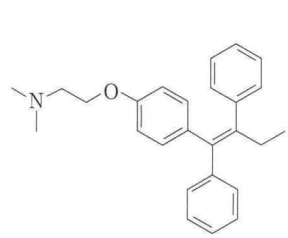


120

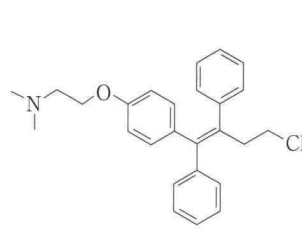


121

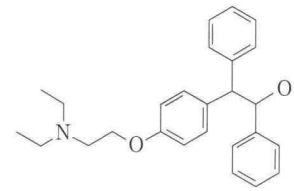
Fig. 2 (continued)



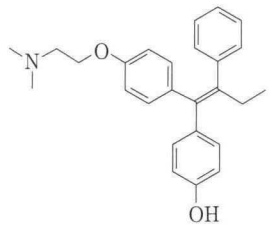
122



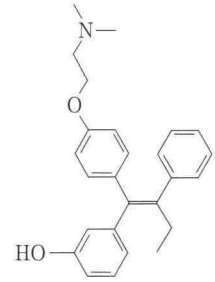
123



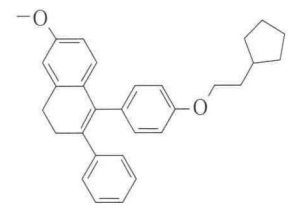
124



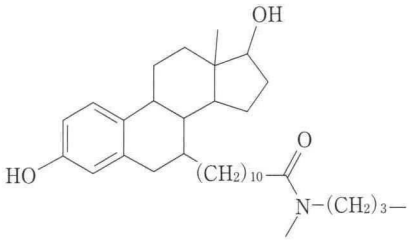
125



126



127



128

Fig. 2 (continued)

(Higuchi et al., 1965). An excess of guest molecule was added to solutions containing increasing concentrations of β-CD or 2-HP-β-CD; after equilibration by shaking at 25 ℃ for 7-10 days, an aliquot was centrifuged the supernate filtered, and the filtrate analyzed spectrophotometrically at 220 nm (25℃) after dilution with water. $K^{diss}_{(1:1)}$ was calculated from the initial slope of phase solubility diagrams and aqueous solubility according to eq (2) (Uekama, 1983).

$$K^{diss}_{(1:1)} = \text{slope}/[\text{intercept} \cdot (1 - \text{slope})] \quad (2)$$

Table 2. Structural fragments and their contribution to log RBA for ER-α.

fragment or parameter	no. of compds	freq of occurrence	contribution
constant			−4.89
$^1\chi$	128	128	0.0415
$^1\chi^2$	128	128	-6.1×10^{-5}
Nonring Increments			
$-CH_3$	77	133	−0.09
$-CH_2-$	43	122	0.17
>CH-	15	21	0.73
>C<	12	13	1.18
=CH-	8	15	0.19
=C<	14	26	1.51
-C≡CH	4	4	0.87
Ring Increments			
$-CH_2-$	33	172	0.23
>CH-	32	93	0.42
>C<	24	38	0.63
=CH-	126	819	−0.05
=C<	127	522	0.71
Oxygen Increments			
-OH (alcohol)	23	28	−1.41
-OH (phenol)	108	181	−0.23
-O- (nonring)	23	25	−0.46
-O- (ring)	23	23	−0.94
>C=O (nonring)	9	10	−0.18
>C=O (ring)	27	27	−0.66
-COOH (acid)	4	4	−2.21
-COO- (ester)	15	15	0.29
OH---OH*	26	26	1.54
Nitrogen Increments			
>N- (nonring)	5	5	−0.42
>N- (ring)	1	1	−1.69
Other Increments			
-Cl	18	45	−0.27
$-SO_2-$	2	2	−1.07
>Si<	1	2	0.59

*The incidence of this parameter is shown in Table 1.

3. Results and Discussion

RBAs for the ER-α:xenoestrogen system

Many experimental and theoretical descriptors are related to molecular size or bulk, but unlike parameters such as molecular volume, molecular surface area, or molar refractivity, molecular weight and topological indices can be calculated easily with high precision without any conditional assumptions. Therefore, regression analyses were performed using the experimental RBAs listed in Table 1, assuming linear and quadratic

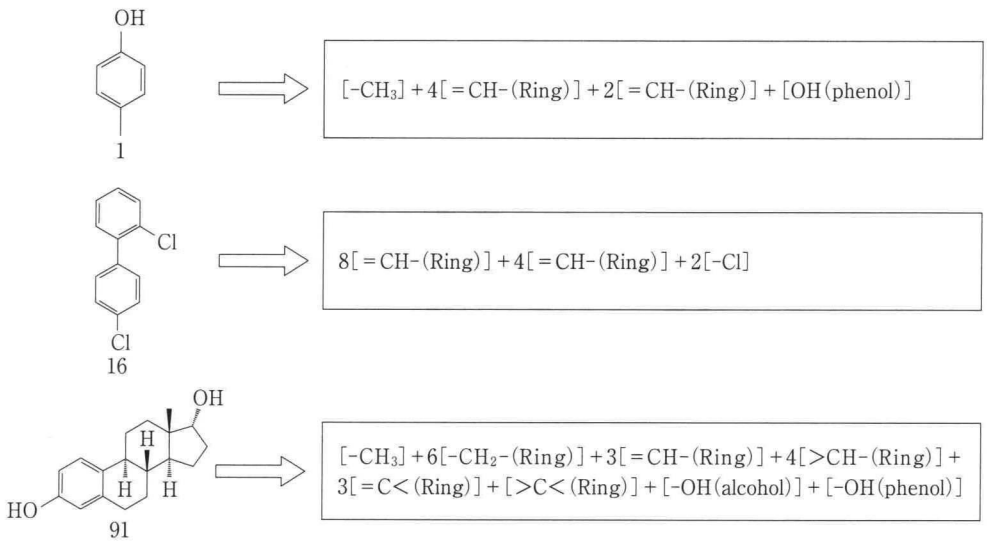


Fig. 3. Examples of decomposition of molecules to molecular fragments.

terms for molecular weight, ${}^0\chi$, or ${}^1\chi$ and summation of fragment constants $\sum(n_i \cdot G_i)$. For this data set the best GCM was obtained with the first-order Kier-Hall molecular connectivity index ${}^1\chi$:

$$\log_{10} \text{RBA} = 0.0415\ {}^1\chi - 6.1 \times 10^{-5}({}^1\chi)^2 + \sum(n_i \cdot G_i) - 4.89 \quad (3)$$
$$(n = 128,\ r^2 = 0.848,\ q^2 = 0.654,\ \text{SD} = 0.345,\ \text{absolute ME} = 0.538,\ F = 19.8)$$

where
n = sample size,
r^2 = index of determination,
q^2 = cross-validated leave-one-out index of determination,
SD = standard deviation,
absolute ME = absolute mean error, and
F = Fisher distribution statistic.

The statistical quality of eq (3) was slightly reduced if the ${}^1\chi$ and $({}^1\chi)^2$ terms were omitted. Good models also were obtained using ${}^0\chi$ or Mw instead of ${}^1\chi$.

A plot of experimental $\log_{10}$ RBA values *vs.* calculated $\log_{10}$ RBA values obtained from eq (3) for the complete data set (n = 128) is shown in Figure 4. A reasonable fit was found for molecules of diverse chemical type comprising the data set. Except one parameter, *viz.* two oxygen atoms having cooperative hydrogen bonding with ER-α, such as occurs in molecules like 17β-estradiol and diethylstilbestrol ($I_{OH\text{-}OH}$), all struc-

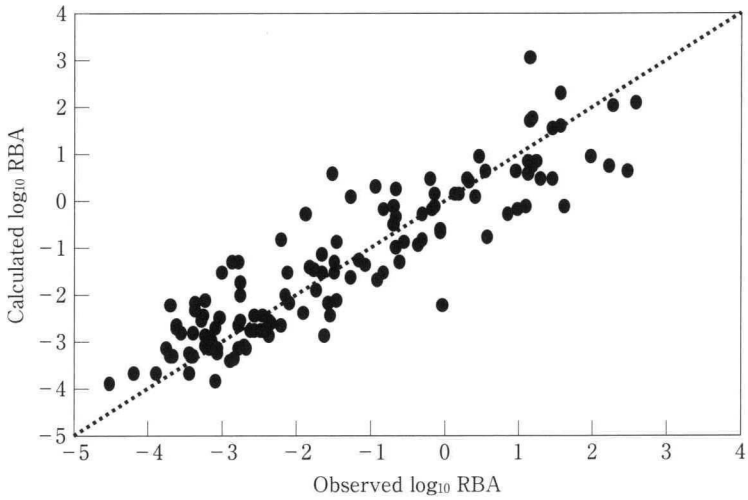


Fig. 4. Correlation between calculated (eq(3)) and experimental $\log_{10}$ RBAs for ER-α with 128 (xeno)estrogens.

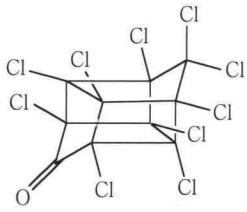
tural fragments correspond to those described in a previous study on free energies of inclusion complexation of organic molecules (Suzuki, 2001).

Group contributions to RBA

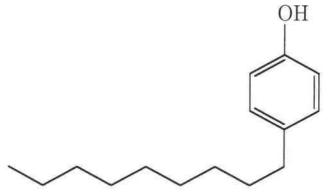
The contributions of $^1\chi$, $(^1\chi)^2$, and ligand fragments to the RBA for ER-α is reflected in the values of the regression coefficients for the topological indices and weightings of individual fragments (Table 2). Fragment constants for most carbon moieties (except $-CH_3$ and =CH-(ring), whose contributions are very close to 0) were>0, *i.e.* they enhanced ligand binding to ER-α, indicating a positive impact of van der Waals forces on the stability of receptor-ligand complexes. In contrast, with two exceptions (-COO-(ester), >Si<) fragments containing heteroatoms impacted negatively on the stability of receptor-ligand complexes. The high fragment constant for OH---OH, + 1.54 is attributable to the importance of hydrogen bonding for ligand binding to ER-α. However, the fragment constant for phenolic hydroxyl is only −0.23, implying that phenols per se modestly destabilize the receptor-ligand complex; whereas the fragment constant of −1.41 for non-aromatic hydroxyl groups indicates that this moiety has a pronouonced destabilizing effect on receptor-ligand complexes. A negative effect of aliphatic -OH groups on the stability of β-CD:guest inclusion complexes has been described previously (Suzuki, 2001).

Correlation of binding stabilities of receptor-ligand complexes and β-CD:guest inclusion complexes

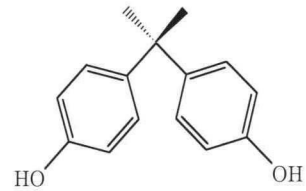
Gibbs' free energies of complexation for β-CD:guest complexes with 11 structurally diverse (xeno)estrogens (Figure 5), and Gibbs' free energies for these same compounds bound to ER-α were calculated from



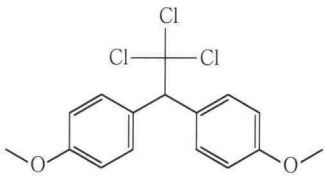
12. kepone



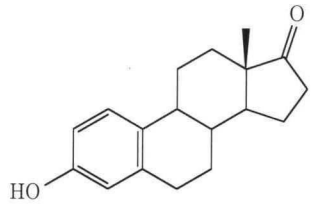
42.4-nonylphenol



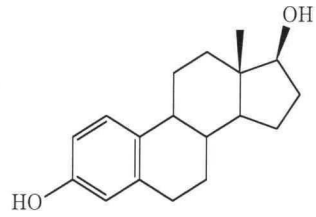
45. bisphenol A



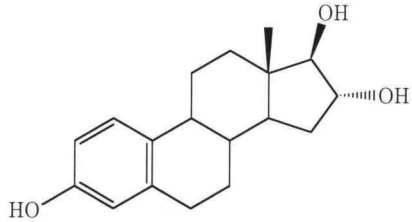
74. methoxychlor



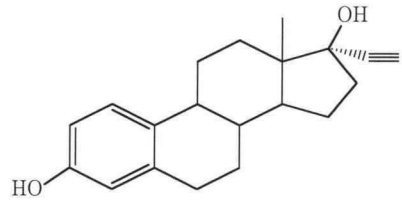
86. estrone



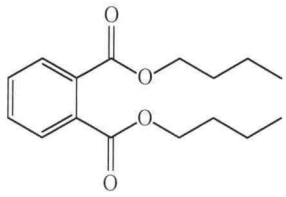
91. 17β-estradiol



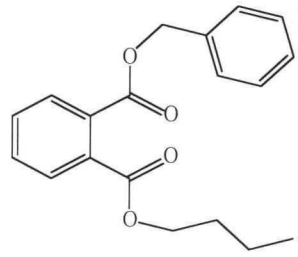
92. 17β-estriol



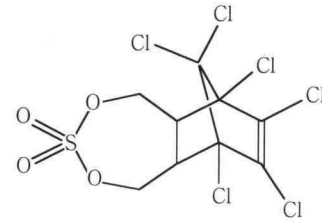
114. ethynyl estradiol



129. di-*n*-butyl phthalate



130. butyl benzyl phthalate (BBP)



131. α-benzoepin

Fig. 5. Xenoestrogens used to construct a correlation between their ΔG of binding to ER-α and ΔG of complexation to β-CDs.

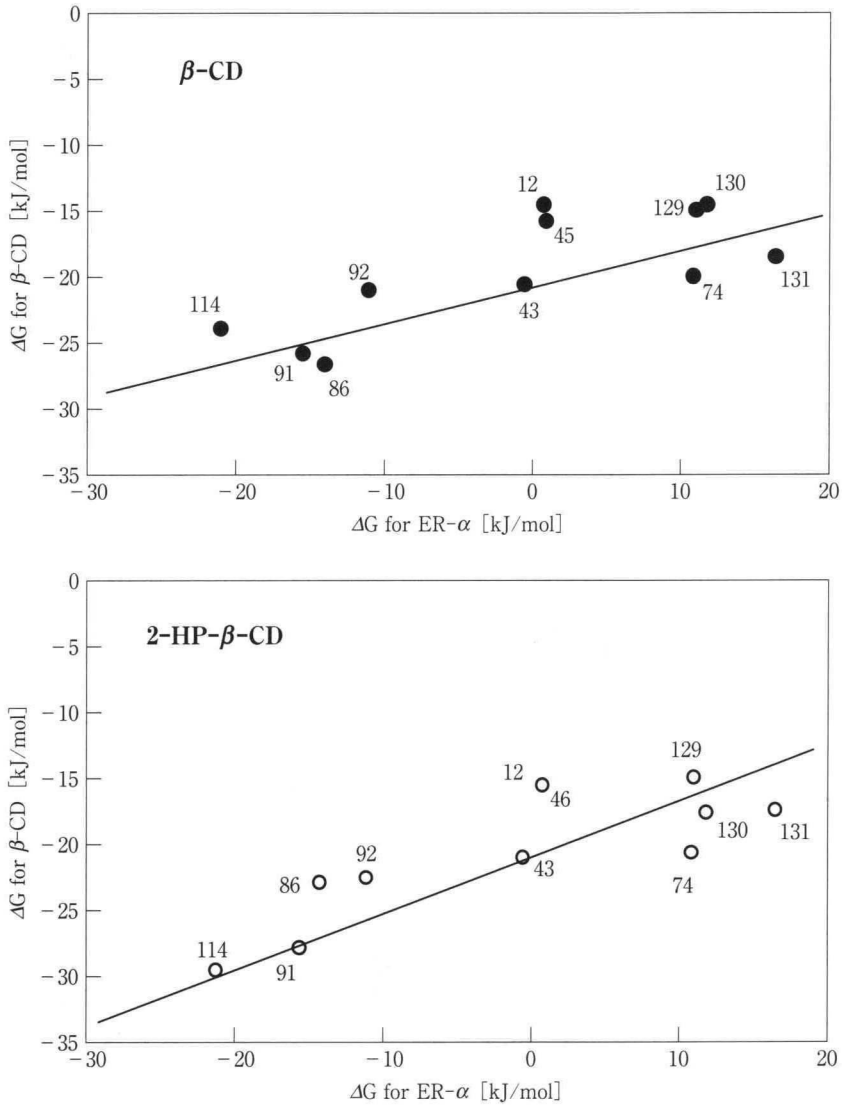


Fig. 6. Plots of experimental ΔG of binding of (xeno)extrogens to ER-α *vs.* experimental ΔG of complexation of (xeno)extrogens to β-CDs.

$$\Delta G = -2.303\ RT\ \log_{10} K, \tag{4}$$

where $K = K^{diss}{}_{(1:1)}$ for the β-CD:guest complex and K = equilibrium dissociation constant (Uekama et al., 1983) for the receptor-ligand complex. Figure 6 compares experimental free energies of complexation for (xeno)estrogens with β-CDs *vs.* experimental free energies of binding of these (xeno)estrogens with ER-α. A slightly better correlation was obtained for the free energies of 2-HP-β-CD:(xeno)estrogen *vs.* ER-α:

(xeno) estrogen ($r^2=0.661$) than for the free energies of natural β-CD:(xeno)estrogen *vs.* ER-α:(xeno)estrogen ($r^2=0.581$).

$$\Delta G \text{ (natural } \beta\text{-CD)} = 0.265\, \Delta G[\text{ER-}\alpha\text{:(xeno)estrogen}] - 19.46 \quad (5)$$
$$(n=11,\ r^2=0.581,\ SD=3.03,\ F=12.6)$$

$$\Delta G \text{ (2-HP-}\beta\text{-CD)} = 0.315\, \Delta G[\text{ER-}\alpha\text{:(xeno)estrogen}] - 20.31 \quad (6)$$
$$(n=11,\ r^2=0.661,\ SD=3.04,\ F=17.5)$$

The slight discrepancy in statistical quality between eqs (5) and (6) may, to some extent, reflect more precise $K^{diss}_{(1:1)}$ values for the 2-HP-β-CD complexes, since 2-HP-β-CD is much more soluble in aqueous milieux than native β-CD (Loftsson and Duchêne, 2007). Nonetheless, reasonably linear relationships exist for (xeno)estrogen binding to β-CDs *vs.* (xeno)estrogen binding to ER-α, suggesting that β-CD:(xeno)estrogen systems, and particularly the 2-HP-β-CD:(xeno)estrogen system, can function as alternatives to receptor binding studies and animal experimentation for projecting potential xenoestrogenic activity expressed as relative binding affinity to rat uterine estrogen receptor-α.

4. Conclusions

A GMC for predicting the estrogenic activity of chemicals was developed. Results obtained using a training set comprised of 128 compounds demonstrated that prediction of RBAs of molecules to rat uterine ER-α can be achieved with reasonable accuracy using the first-order molecular connectivity index $^1\chi$ in combination with molecular fragments. To the best knowledge of the authors, this work represents the first application of GCMs or additive schemes to the quantification of ligand binding to ERs. This method should prove especially useful for the prioritization or preliminary screening of chemicals for their potential to cause endocrine disruption.

Acknowledgement

The author acknowledges Mr. Teppei Ito for his help collecting the data for this study.

References

Blair, R.M.; Fang, H.; Branham, W. S.; Hass, B. S.; Dial, S. L.; Moland, C. L.; Tong, W.; Shi, L.; Perkins, R.; Sheehan, D. M. (2000) The estrogen receptor relative binding affinities of 188 natural and xenochemicals : structural diversity of ligands. *Toxicol. Sci.* **54**, 138-153.

Bradbury, S. P.; Mekenyan, O. G.; Ankley, G. T. (1996) Quantitative structure-activity relationships for polychlorinated hydroxybiphenyl estrogen receptor binding affinity : An

assessment of conformer flexibility ? *Environ. Toxicol. Chem.*, **15**, 1945-1954.

Brunstrom, B.; Axelsson, J.; Halldin, K. (2003) Effects of endocrine modulators on sex differentiation in birds. *Ecotoxicology* **12**, 287-295.

Brzozowski, A. M.; Pike, A. C. W.; Dauter, Z.; Hubbard, R. E.; Bonn, T.; Engström, O.; Öhman, L.; Geene, G.L.; Gustafsson, J.-Å.; Carlquist, M. (1997) Molecular basis of agonism and antagonism in the oestrogen receptor. *Nature* **389**, 753-758.

Colborn, T.; Saal, F.S.V.; Soto, A.M. (1993) Developmental effects of endocrine disrupting chemicals in wildlife and humans. *Environ. Health Perspect.* **101**, 378-384.

Connors, K. A. (1997) The stability of cyclodextrin complexes in solution. *Chem. Rev.* **97**, 1325-1357.

European Commission (2001) Communication from the Commission to the Council and the European Parliament on the Implementation of the Community Strategy for Endocrine Disruptors - a Range of Substances Suspected of Interfering with the Hormone Systems of Humans and Wildlife, COM 262 final, Brussels, 14 June 2001. http://europa.eu.int/comm/environment/docum/01262_en.htm#bkh.

Gantchev, T., T. G.; Ali, H.; van Lier, J.E. (1994) Quantitative structure-activity relationships/comparative molecular field analysis (QSAR/CoMFA) for receptor-binding properties of halogenated estradiol derivatives. *J. Med. Chem.* **37**, 4164-4176.

Gao, H.; Williams, C.; Labute, P.; Bajorath, J. (1999a) Binary quantitative structure-activity relationship (QSAR) analysis of estrogen receptor ligands. *J. Chem. Inf. Comput. Sci.* **39**, 164-168.

Gao, H.; Katzenellenbogen, J. A.; Garg, R.; Hansch, C. (1999b) Comparative QSAR analysis of estrogen receptor ligands. *Chem. Rev.* **99**, 723-744.

Gao, H. (2001) Application of BCUT metrics and genetic algorithm in binary QSAR analysis. *J. Chem. Inf. Comput. Sci.* **41**, 402-407.

Hedges, A.R. (1998) Industrial applications of cyclodextrins. *Chem. Rev.* **98**, 2035-2044.

Higuchi, D.; Connors, K.A., in Reilley, C.N. (1965) *Advances in Analytical Chemistry and Instrumentation*, vol. 4, Interscience, New York, p.132.

Hutchinson, T.H.; Pickford, D.B. (2002) Ecological risk assessment and testing for endocrine disruption in the aquatic environment. *Toxicology* **181**, 383-387.

Katritzky, A.R.; Fara, D. C.; Yang, H.; Karelson, M.; Suzuki, T.; Solov'ev, V. P.; Varnek, A. (2004) Quantitative structure-property relationship modeling of β-cyclodextrin complexation free energies. *J. Chem. Inf. Comput. Sci.* **44**, 529-541.

Kavlock, R.J. (1996) Research needs for risk assessment of health and environmental effects of endocrine disruptors : a review of the U.S. EPA-sponsored workshop. *Environ. Health Perspect.* **104**, 715-740.

Kier, L. B.; Hall, L. H. (1986) *Molecular Connectivity in Structure-Activity Analysis*; Research Studies Press : Letchworth, Herts., U.K.

Klein, C. T.; Polheim, D.; Viernstein, H.; Wolschann, P. (2000) A method for predicting the free energies of complexation between β-cyclodextrin and guest molecules. *J. Inclusion Phenom. Macrocyclic Chem.* **36**, 409-423.

Kramer, V. J.; Giesy, J. P. (1999) Specific binding of hydroxylated polychlorinated biphenyl metabolites and other substances to bovine calf uterine estrogen receptor : structure-binding relationships. *Sci. Total Environ.* **233**, 141-161.

Lipkowitz, K. B. (1998) Applications of computational chemistry to the study of cyclodextrins. *Chem. Rev.* **98**, 1829-1873.

Liu, L.; Guo, Q.-X. (1999) Wavelet neural network and its application to the inclusion of β-cyclodextrin with benzene derivatives. *J. Chem. Inf. Comput. Sci.* **39**, 133-138.

Loftsson, T.; Duchêne, D. (2007) Cyclodextrins and their pharmaceutical applications. *Int. J. Pharm.* **329**, 1-11.

Lydersen, A. L. (1955) Estimation of critical properties of organic compounds by group contributions. *Univ. Wisconsin. Coll. Eng. Expt. Stn. Rep.*, no. 3.

Matsui, Y.; Nishioka, T.; Fujita, T. (1985) Quantitative structure-reactivity analysis of the inclusion mechanism by cyclodextrins. *Top. Curr. Chem.* **128**, 61-89.

Park, J. H.; Nah, T.H. (1994) Binding forces contributing to the complexation of organic molecules with β-cyclodextrin in aqueous solution. *J. Chem. Soc., Parkin Trans.* 2, 1359-1362.

Roncaglioni, A.; Novič, M.; Vračko, M.; Benfenati, E. (2004) Classification of potential endocrine disruptors on the basis of molecular structure using a nonlinear modeling method. *J. Chem. Inf. Comput. Sci.* **44**, 300-309.

Sadler, B.R.; Cho, S. J.; Ishaq, K.S.; Chae, K.; Korach, K.S. (1998) Three-dimensional quantitative structure-activity relationship study of nonsteroidal estrogen receptor ligands using the comparative molecular field analysis/cross-validated r^2-guided region selection approach. *J. Med. Chem.* **41**, 2261-2267.

Shi, L. M.; Fang, H.; Tong, W.; Wu, J.; Perkins, R.; Blair, R. M.; Branham, W. S.; Dial, S. L.; Moland, C. L.; Sheehan, D. M. (2001) QSAR models using a large diverse set of estrogens. *J. Chem. Inf. Comput. Sci.* **41**, 186-195.

Singleton, D.W.; Khan, S.A. (2003) Xenoestrogen exposure and mechanisms of endocrine disruption. *Front. Biosci.* **8**, 110-118.

Suzuki, T.; Ishida, M.; Fabian, W.M.F. (2000) Classical QSAR and comparative molecular field analyses of the host-guest interaction of organic molecules with cyclodextrins. *J. Comput.-Aided Mol. Des.* **14**, 669-678.

Suzuki, T.; Ide, K.; Ishida, M.; Shapiro, S. (2001) Classification of environmental estrogens by physicochemical properties using principal component analysis and hierarchical cluster analysis. *J. Chem. Inf. Comput. Sci.* **41**, 718-726.

Suzuki, T. (2001) A nonlinear group contribution method for predicting the free energies of inclusion complexation of organic molecules with α- and β-cyclodextrins. *J. Chem. Inf. Comput. Sci.* **41**, 1266-1273.

Szejtli, J. (1988) *Cyclodextrin Technology*, Kluwer Academic Publishers: Dordrecht.

Szejtli, J. (1998) Introduction and general overview of cyclodextrin chemistry. *Chem. Rev.* **98**, 1743-1753.

Tong, W.; Perkins, R.; Xing, L.; Welsh, W.J.; Sheehan, D.M. (1997a) QSAR models for binding of estrogenic compounds to estrogen receptor α and β subtypes. *Endocrinology* **138**, 4022-4025.

Tong, W.; Perkins, R.; Strelitz, R.; Collantes, E.R.; Keenan, S.; Welsh, W.J.; Branham, W.S.; Sheehan, D.M. (1997b) Quantitative structure-activity relationships (QSARs) for estrogenic binding to the estrogen receptor : Predictions across species. *Environ. Health Perspect.* **105**, 1116-1124.

Tong, W.; Lowis, D.R.; Perkins, R.; Chen, Y.; Welsh, W.J.; Goddette, D.W.; Heritage, T.W.;

Sheehan, D.M. (1998). Evaluation of quantitative structure-activity relationship methods for large-scale prediction of chemicals binding to the estrogen receptor. *J. Chem. Inf. Comput. Sci.* **38**, 669-677.

Uekama, K.; Fujinaga, T.; Hirayama, F.; Otagiri, M.; Yamasaki, M; Seo, H.; Hashimoto, T.; Tsuruoka, M. (1983) Improvement of the oral bioavailability of digitalis glycosides by cyclodextrin complexation. *J. Pharm. Sci.* **72**, 1338-1341.

Waller, C.L.; Minor, D.L; McKinney, J.D. (1995) Using three-dimensional quantitative structure-activity relationships to examine estrogen receptor binding affinities of polychlorinated hydroxybiphenyls. *Environ. Health Perspect.* **103**, 702-707.

Waller, C.L.; Oprea, T.I.; Chae, K.; Park, H.-K.; Korach, K.S.; Laws, S.C.; Wiese, T.E.; Kelce, W.R.; Gray, L.E., Jr. (1996) Ligand-based identification of environmental estrogens. *Chem. Res. Toxicol.* **9**, 1240-1248.

Waller, C.L. (2004) A comparative QSAR study using CoMFA, HQSAR, and FRED/SKEYS paradigms for estrogen receptor binding affinities of structurally diverse compounds. *J. Chem. Inf. Comput. Sci.* **44**, 758-765.

Wallimann, P.; Marti, T.; Fürer, A.; Diederich, F. (1997) Steroids in molecular recognition. *Chem. Rev.* **97**, 1567-1608.

Yoshida, A.; Yamamoto, M.; Hirayama, F.; Uekama, K. (1988) Improvement of chemical instability in aqueous solution by complexation with β-cyclodextrin Derivatives. *Chem. Pharm. Bull.* **36**, 4075-4080.

Zheng, W.; Tropsha, A. (2000) A novel variable selection QSAR approach based on the k-nearest neighbor principle. *J. Chem. Inf. Comput. Sci.* **40**, 185-194.

要　旨

種々のエストロゲン様物質の β-シクロデキストリン複合体形成能とのアナロジーに基づく構造活性相関

鈴木孝弘，Stuart Shapiro

化学物質の内分泌かく乱作用を調べるためには、培養細胞や動物を使った生物学的検定法が有効とされている。しかし、膨大な数に増え続けている化学物質を、一つ一つ生物学的手法で調べることは時間とコスト面から実質的に困難である。本研究では、種々のエストロゲン様化学物質128種について、エストロゲン受容体への結合能と化学構造との相関を調べ、化学物質のエストロゲン様作用を分子のサイズを考慮した原子団寄与法により予測する手法を構築した。さらに、この化学物質:エストロゲン受容体結合系と化学物質:β-シクロデキストリン複合体形成系とのアナロジーに着目して、これら2つの系の関係を検討したところ、r^2 の値で0.6—0.7と統計的に有意な相関を見出した。これより、これら2つの系には、化学物質が相互作用する空間的な形の違いはあるが、その結合に関するドライビングフォースには、きわめて類似の性質があることを確認した。