Comparative QSAR Analysis of Estrogen Receptor Ligands

Hua Gao,[†] John A. Katzenellenbogen,[‡] Rajni Garg,[†] and Corwin Hansch*[,†]

Departments of Chemistry, Pomona College, Claremont, California 91711, and University of Illinois, 600 South Mathews Avenue, Urbana, Illinois 61801

Received April 27, 1998 (Revised Manuscript Received November 18, 1998)

Contents

* To whom correspondence may be addressed.
[†] Pomona College.
[‡] University of Illinois.

I. Introduction

In recent years, the need for agents for the treatment of hormone-responsive cancers and the mechanistic toxicological studies of environmental pollutants have sparked a great deal of scientific interest and research into the chemical and biological interactions between estrogen receptors and their ligands. Breast cancer is the most frequently diagnosed cancer in women and a major cause of cancer.[1] Since estrogens are known to play a role in the development and growth of many breast cancers, a logical approach for the treatment of estrogen-sensitive breast cancer is the use of antiestrogens that block the interaction of estrogens with their specific receptor. Several classes of antiestrogens have been developed for the treatment of breast cancer, but unfortunately, although the initial response of breast cancer to hormone therapy can be substantial, resistance to estrogen antagonist therapy often develops.[2,3] Tamoxifen (Nolvadex), the antiestrogen most frequently used in breast cancer hormone therapy has mixed agonist−antagonist properties; other more recently developed antiestrogens, such as ICI 182,780, are pure antiestrogens and may prove to be more effective in breast cancer treatment.[4] Antiestrogens are also being studied as agents to prevent breast cancer in women at high risk.[5]

Estrogens are also widely prescribed in menopausal women as hormone replacement therapy to maintain bone mineral density and preserve cardiovascular health. The use of agents, such as Raloxifene (Evista), that have a favorable balance of agonist activities in certain tissues (bone, liver, vasculature) and antagonist activities in other tissues (uterus and breast)[6,7] is favored for such uses because of their reduced risk in promoting tumor development and growth in these tissues, and there are many active programs of research in academic and industrial laboratories aimed at the development of such tissue-selective estrogens.

The finding that compounds from the environment, of both synthetic and natural origin, can interfere with sexual development and reproductive function has led to intense investigations of these endocrine-disruptive substances.[8] Many (though not all[9,10]) endocrine disrupters are estrogens and exert agonist and antagonist effects through the estrogen receptor: those that are synthetic include pesticides, food antioxidants, and metabolites of nonionic surfactants;[11] naturally occurring ones are plant secondary metabolites and mold metabolites.[8] Among the latter

10.1021/cr980018g CCC: $35.00
Published on Web 02/19/1999

Hua Gao received his Ph.D. in Pharmaceutical Sciences at the University of Southern California. He joined Professor Corwin Hansch in 1995 as a postdoctoral research associate and worked with BioByte Corporation as a scientist. Currently, he is working with MDS Panlabs as a scientist in Computational Chemistry and Informatics. His research interests include QSAR, computer-assisted drug design (CADD), and natural products.

John Katzenellenbogen did his undergraduate and graduate studies in chemistry at Harvard University with E. J. Corey, receiving his Ph.D. degree in 1969 for work on the development of methods for stereospecific olefin synthesis and their application to natural product synthesis. That same year he joined the faculty of the Department of Chemistry at the University of Illinois, where he is now the Swanlund Professor of Chemistry. Dr. Katzenellenbogen's research interests have focused on various aspects of the structure, function, and use of steroid receptors: He prepared affinity labeling agents for the estrogen receptor and used them to study receptor structure, function, and dynamics; he developed an extensive series of steroid receptor-based agents labeled with fluorine-18 and technetium-99m for imaging receptor-positive tumors of the breast and prostate by radioimaging methods; and he has designed fluorescent probes for steroid receptors that enable receptor dynamics to be followed in individual cells. John Katzenellenbogen has trained more than 70 doctoral and postdoctoral students and published more than 300 articles. He is a fellow of the American Academy of Arts and Sciences. He received the 1995 Aebersold Award of the Society of Nuclear Medicine and a Cope Scholar Award from the American Chemical Society in 1999.

category are common natural secondary plant metabolites—flavanoids and lignins—that are present at high levels in common foods, especially those derived from soy and whole grains; these compounds are considered to be beneficial to human health.

The estrogen receptor (ER), the target of these agents, is a ligand-modulated transcription factor that regulates the activity of certain genes.[12] A member of the nuclear hormone receptor gene superfamily, ER has a multidomain structure, with two conserved domains that are responsible for DNA binding on one hand, and ligand binding, dimerization, and transcriptional activation on the other.[12] The binding of ligands to the hormone-binding domain of ER stabilizes the interaction of the receptor with target sequences in the regulatory region of these genes. This binding may be either directly to specific DNA enhancer sequences or, in some cases, to AP1 enhancers through the AP1 transcription factors Fos and Jun. The activation or repression of these genes by the ligand receptor complex is then mediated by the recruitment by ER of a variety of

Rajni Garg received her M.Sc. in Chemistry (1984) from Meerut University and M.Phil. (1988) degree from Delhi University, India. Her M.Phil. dissertation work was on peptide synthesis. She was a faculty member in the Chemistry Department of Birla Institute of Technology and Science, Pilani, India, from 1991 to 1996, where she taught organic and physical chemistry. She received her Ph.D. degree in 1996 under the supervision of Professor S. P. Gupta. Her doctoral work was on QSAR Studies on Anti-HIV Agents. In February 1997, she joined Professor Corwin Hansch as a postdoctoral researcher, and she is currently involved in building a C-QSAR databank. Her research interests include QSAR and computer-assisted drug design.

Corwin Hansch received his undergraduate education at the University of Illinois and his Ph.D. in Organic Chemistry from New York University in 1944. After working with the DuPont Company, first on the Manhattan Project and then in Wilmington, DE, he joined the Pomona College faculty in 1946. He has remained at Pomona except for two sabbaticals: one at the Federal Institute of Technology in Zurich with Professor Prelog and the other at the University of Munich with Professor Huisgen. The Pomona group published the first paper on the QSAR approach relating chemical structure with biological activity in 1962. Since then, QSAR has received widespread attention. In *Current Contents* (1981/41) he was named as one of the 300 most cited scientists out of over one million publishing in all fields of science for the period 1965–1978. In 1986, Hansch was cited in *Current Contents* as being one of the 250 most cited primary authors for 1984. He is an honorary fellow of the Royal Society of Chemistry and recently received the ACS Award for Computers in Chemical and Pharmaceutical Research for 1999.

coregulatory proteins that interact with components of the basal transcriptional complex and have enzymatic activity that alters the architecture of chromatin.[13]

There are two estrogen receptors, designated estrogen receptor α and β (ERα and ERβ).[14−16] These receptor subtypes are related in both structure and function but have somewhat different tissue distributions.[17−19] Ligands of remarkable structural diversity (these include steroidal analogues, nonsteroidal compounds such as stilbenes, triarylethylenes, phenylindoles, phenylindenes, and coumarins[20−25]) bind to the estrogen receptor, in many cases with good affinity, a fact that has remained a curiosity for many years.

Recently, the X-ray crystal structure of the ligand-binding domain of ERα, complexed with estradiol and the nonsteroidal selective estrogen antagonist Raloxifene, was reported.[26] Like other members of the nuclear hormone receptor superfamily, the ligand-binding domain of ER has the canonical antiparallel α-helical triple sandwich topology, with the ligand being completely engulfed within the lower portion of the domain.[26−30] The polar ends of the ligand estradiol are involved in hydrogen bonds with the only polar residues within this binding pocket; E353 and R394 are the hydrogen-bonding partners for the A-ring phenol, and H524 for the 17β-hydroxyl group. While the A- and D-ring ends of the ligand are held rather tightly by the receptor, the B- and C-ring regions are surrounded by considerable empty volume, especially below the B-ring and above the C-ring, a situation that was predicted on the basis of a limited analysis of ligand structure−affinity relationships.[31] These voids surrounding the ligand, which are almost as large as the ligand itself, are lined with nonpolar residues and are thought to permit the sort of molecular compliance required to accommodate ligands of different structure.[26]

The immediate interest of some of us in estrogens comes from the study of the toxic action of phenols to rapidly growing cells, although activity probably is not mediated by the estrogen receptor. In a first analysis of data from Oglesby et al.[26a] at the EPA, we found that the malformation of rat embryos in vitro by simple phenols is correlated with the Hammett parameter σ^+.[32,33] We decided to follow up this finding by studying the action of a set of simple 3- and 4-substituted phenols on rapid growing cancer cells of L1210 leukemia. From this analysis, it was clear that σ^+ was the most important parameter and that log *P* (octanol/water partition coefficient) played a small but significant role. We then decided to examine some estrogens in this assay: 4-octylphenol, 4-nonylphenol, bisphenol A, diethylstilbestrol, and estradiol. The activity of these hormones fit the same equation as the simple phenols (see eq 27). It was necessary to estimate σ^+ for the last three of the above compounds. To circumvent this problem, however, we used AM1 calculations to establish the LUMO−HOMO gap for the simple and complex phenols, and we found this to give a slightly better correlation.[33a] We have recently found that three of the four components of the drug Premarin are also well fit, esterone was too insoluble to test (unpublished results).

With this background in the quantitative structure activity relationship (QSAR) analysis of estrogens, we became interested in undertaking a wider investigation of the interaction of estrogens with the estrogen receptor. In building our C-QSAR database (which at present contains 12 500 examples),[34] we have derived a number of QSAR equations based on the interactions between the estrogen receptor and its ligands. Although in these experiments, different receptor preparations (e.g., mouse, rat, lamb, and calf uterine cytosol) and different ligands were used, there seems to be a common feature among these QSAR equations, that is, the lack of a positive hydrophobic parameter, except at the 11β-position of estradiol derivatives. This is surprising, considering the fact that the ligand-binding domain in the estrogen receptor is very hydrophobic[26,31] and that the ligand binding pocket is lined with residues that are, except for a few, all hydrophobic.[26] Thus, our results show that the interactions between the receptor and the substituents of the ligands can, in many cases, be quantitatively explained simply by steric and electronic parameters. However, hydrophobicity does appear to play a role for 11β-substituted estradiols, as well as in the in vivo activities of estrogens.

II. Methods

The binding affinity of estrogen receptor ligands has been collected from the literature (see individual data set for detailed references). These binding affinity determinations, which have often been done by different competitive binding affinity assay methods, using different receptor preparations (see above), were all placed on a common "relative binding affinity" (RBA) scale. Values on this scale were calculated as a percent from the ratio of IC_{50} values of test compounds to that of estradiol to displace 50% of [^{3}H] estradiol from estrogen receptor preparations (generally uterine cytosol fractions, which are largely estrogen receptor-α). Thus, on the RBA scale, estradiol by definition has a value of 100, with lower affinity ligands having lower values and higher affinity ligands, higher values. It probably does not make a great difference what species and target tissue is used as the source of the estrogen receptor for these binding studies, because there is little evidence for species difference in structure−affinity relationships, and, in most of the target tissues used, the estrogen receptor-α subtype will predominate.[16,19] In each case, however, we note the species and tissue source of the receptor. Competitive binding affinity assays, however, are sometimes run at different temperatures, 0 or 25 °C (and sometimes higher), and for different incubation times; the lower temperatures are used to maintain receptor stability, but the higher temperatures allow more rapid equilibration of the binding of both the competing ligand and the labeled tracer. RBA values for the same compound can be quite different at different temperatures, which is thought to be largely the result of incomplete equilibration of the assay at the lower temperature, a problem that can be especially severe with high-affinity ligands for which dissociation as well as association rates can be rather slow.[35] In each correlation, however, we have specified the assay tem-

perature, and we treat data from assays at different temperatures separately.

All of the physicochemical parameters were automatically loaded from our C-QSAR database, and the QSAR regression analyses were executed with the C-QSAR program. The utility of the QSAR program in correlation analysis has been discussed.[34,36] Included in the program are all of the commonly used substituent parameters.[37]

The parameters used in this report have been discussed in detail along with their applications.[41] Here we provide a brief definition. E_S is the classic Taft parameter derived from the rate of hydrolysis of aliphatic esters. It is normally most useful for intramolecular steric effects, but with relatively small substituents, it sometimes accounts for intermolecular interactions. MR is calculated as follows:

$$\left(\frac{n^2-1}{n^2+2}\right)\left(\frac{\mathrm{MW}}{d}\right)$$

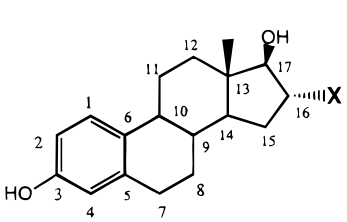
where n is the refractive index, MW is the molecular weight, and d is the density of a substance. Since there is rather little variation in n, MR is largely a measure of volume with a small correction for polarizability. We have scaled our values by 0.1. MR can be used for a substituent or for the whole molecule. MgVol is the molar volume calculated by the method of McGowan. B1, B5, and L are the sterimol parameters for substituents. B1 is a measure of the width of the first atom of a substituent, B5 is an attempt to define the overall volume, and L is for substituent length. ClogP is the calculated[80] octanol/water partition coefficient. The electronic parameters σ, σ^+, and σ^- apply to substituent effects on aromatic systems, and σ^* applies to aliphatic systems. RBA stands for relative binding affinity. The numbers in parentheses in the QSAR equations are for the 95% confidence intervals. All of these parameters and their applications have been discussed.[41]

III. Results

1. Estradiol Derivatives

A number of estradiol derivatives have been synthesized and their relative binding affinities with estrogen receptors have been evaluated over the years. We have derived the following QSAR from these studies.

A. Relative Binding Affinities (RBA) of 16α-Substituted Estradiols with Rat Uterine Estrogen Receptor at 0 °C (Table 1)[38]



$$\log \mathrm{RBA} = -0.48(\pm 0.10)\mathrm{MR} + 2.08(\pm 0.28)$$
$$n = 22 \quad r^2 = 0.84 \quad s = 0.432 \quad F_{1,20} = 102 \qquad (1)$$

The compounds in this series have a variety of substituent groups at the 16α-position; some of them were prepared in radiohalogen labeled form, for in vitro ligand binding assays[39,40] and for in vivo imaging of estrogen receptor-positive breast tumors. The best prediction of their binding affinity is made with the parameter MR, for molar refraction. MR is essentially a measure of the volume of the substituent X, with a small correction for the polarizability of this group.[41] We found poor correlation with the hydrophobic parameter π, although there is some collinearity between π and MR ($r^2 = 0.74$), so some component of hydrophobicity is embodied in the MR term. From the negative term in eq 1, it is evident that the estrogen receptor has a limited tolerance to steric effects at the 16α-position. Since the relationship with MR is linear, this implies that the receptor has some flexibility at this site. In this regard, it is of note that, in the two X-ray structures of the estrogen receptor ligand binding domain, the two bound ligands, the steroidal ligand estradiol and the nonsteroidal ligand Raloxifene, have rather different structures in the D-ring portion.[26] The receptor accommodates both ligands well by minor repositioning of residues surrounding the D-ring region of the ligand.

All of the terms in all of the equations except one term in eq 20 and one in eq 24, pass the F test at the 0.95 level of significance. In the case of the two examples where one term fails, the overall F test passes the 0.95 level.

Table 1. Relative Binding Affinities of 16α-Substituted Estradiol Derivatives[38]

	log RBA			
substituents	obsd	calcd from eq 1	Δ	MR
Br	2.09	1.65	0.44	0.89
H	2.00	2.03	−0.03	0.10
Cl	2.00	1.79	0.21	0.60
CH_2Br	1.97	1.44	0.53	1.34
F	1.91	2.04	−0.13	0.09
I	1.90	1.40	0.49	1.39
CH_2I	1.88	1.18	0.70	1.86
CH_2N_3	1.81	1.49	0.32	1.22
CN	1.80	1.78	0.02	0.63
CH_2Cl	1.74	1.58	0.16	1.05
$CH_2C{\equiv}CH$	1.64	1.46	0.18	1.29
$CH_2CH{=}CH_2$	1.58	1.38	0.20	1.45
OH	1.28	1.95	−0.67	0.29
$CH_2CH(Me)F$	0.85	1.40	−0.56	1.41
$CH_2CH{=}CHCH_2OC_6H_5$[a]	0.85	−0.09	0.94	4.49
C_3H_7	0.70	1.40	−0.66	1.41
$CH_2CH(Me)F$	0.70	1.36	−0.59	1.50
$CH_2C_6H_5$	0.70	0.63	0.07	3.00
$CH_2C{\equiv}CCH_2NH{-}NAP$[a]	0.58	−0.81	1.39	5.97
CH_2OH[a]	0.38	1.74	−1.36	0.72
$CH_2CH_2CH_2CN$[a]	−0.05	1.14	−1.19	1.94
$CH_2CH_2CH_2CH_2NHCOMe$	−0.10	0.54	−0.64	3.19
$CH_2CH{=}CHCH_2NH{-}NBD$	−0.52	−0.72	0.20	5.80
$CH_2CH_2CH_2CH_2NH{-}NBD$[b]	−0.70	−0.86	0.16	6.07
$CH_2CH{=}CHCH_2NH{-}NAP$[b]	−0.70	−0.74	0.04	5.83
$CH_2C{\equiv}CCH_2NH{-}NBD$	−1.00	−0.68	−0.32	5.70

[a] Data points not used in deriving equation. [b] NBD and NAP:

Table 2. Relative Binding Affinities of 11β-Substituted Estradiol Derivatives at 0 °C[42]

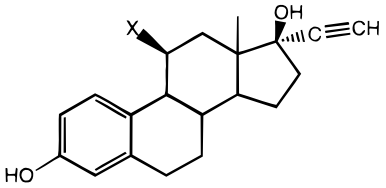
	log RBA			
substituents	obsd	calcd from eq 2	Δ	MR
H	2.05	2.05	0.00	0.103
n-C_3H_7	1.98	1.95	0.03	1.496
$CH_2CH{=}CH_2$	1.97	1.95	0.02	1.449
$CH{=}CH_2$	1.96	1.98	−0.02	1.099
$CHMe_2$	1.92	1.95	−0.03	1.496
$CH_2C_6H_5$	1.89	1.84	0.05	3.001
2-thienyl	1.88	1.88	−0.00	2.404
C_6H_4-4-OMe	1.79	1.83	−0.04	3.174
C_6H_4-2-OMe[a]	1.18	1.83	−0.65	3.174

[a] Data points not used in deriving equation.

Table 3. Relative Binding Affinities of 11β-Substituted Estradiol Derivatives at 25 °C[42]

	log RBA				
substituents	obsd	calcd from eq 3	Δ	π	MR
n-C_3H_7	2.86	2.90	−0.17	0.00	1.496
$CHMe_2$	2.85	2.88	−0.04	1.55	1.496
$CH_2CH{=}CH_2$	2.74	2.46	0.29	1.53	1.449
$CH{=}CH_2$	2.62	2.49	0.13	0.82	1.099
H	2.39	2.56	−0.17	1.10	0.103
2-thienyl	1.96	2.10	−0.14	1.82	2.404
C_6H_4-4-OMe	1.82	1.59	0.23	1.26	3.174
$CH_2C_6H_5$	1.72	1.96	−0.23	1.61	3.001
C_6H_4-2-OMe	0.93	0.98	−0.05	2.01	3.174

B. Relative Binding Affinities of 11β-Substituted Estradiols with Mouse Uterine Estrogen Receptor (Tables 2 and 3)[42]



a. binding at 0 °C (2 h)

$$\log \text{RBA} = -0.07(\pm 0.03)\text{MR} + 2.06(\pm 0.06) \quad (2)$$

$$n = 8 \quad r^2 = 0.86 \quad s = 0.032 \quad F_{1,6} = 38$$

b. binding at 25 °C (5 h)

$$\log \text{RBA} = 1.08(\pm 0.50)\pi - 0.96(\pm 0.28)\text{MR} + 2.66(\pm 0.42) \quad (3)$$

$$n = 9 \quad r^2 = 0.92 \quad s = 0.207$$
$$F_{1,7} = 9.22(\text{MR}) \quad F_{1,6} = 8.42(\pi)$$

Compounds in this series include some of the highest affinity ligands for the estrogen receptor and ligands that have been developed as estrogen antagonists,[43] high potency pharmaceuticals,[44] or tumor imaging agents.[45] QSAR 2 and 3 are derived from the binding affinity of 11β-substituted estradiol derivatives determined at 0 and 25 °C, respectively (see Methods). They indicate that the receptor has certain tolerance for size and preference for hydrophobic character with ligand substituents at this site.

The large, positive coefficient for the substituent parameter π in QSAR 3 (from the data at 25 °C) indicates that hydrophobic substituents are favored, whereas the large, negative coefficient of MR indicates, as was the case in QSAR 1, that sterically bulky substituents lower the binding affinity. With these parameters, when the substituents become larger, the negative steric interaction becomes dominant and affinity drops. Although there is some collinearity between MR and π ($r^2 = 0.65$), eq 3 is much better than an equation that is parabolic in either MR or π. The coefficient with π is what we have come to expect when complete desolvation of the substituent occurs.[46]

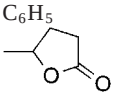
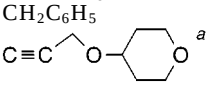
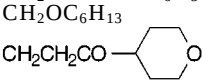
The X-ray crystal structure of the receptor shows that there is some very hydrophobic space above the B-ring of the ligand at position 11,[26] presumably sufficient to accommodate groups of moderate size without contacting the binding surface of the receptor; larger substituents would require some movement of ligand or receptor for a complex to form. Thus, the conclusions from QSAR at the 11β site are similar to those made from an earlier analysis of structure−affinity relationships.[31] Others have noted the presence of a possible hydrophobic pocket at the 11β-position.[86,87]

As was noted in Methods, the measured receptor binding affinity can be a function of both equilibrium time and temperature.[35] It is interesting that QSAR 2, obtained from assays at 0 °C, indicates that affinity measurements at this temperature are rather insensitive to substituent size at the 11β-position and that no compounds have affinities substantially greater than that of estradiol. This may be symptomatic of incomplete equilibration of the ligands in the assay at this temperature, where the relative fraction of unlabeled competitor and labeled tracer that are bound to the receptor reflects the rather similar rates of association of all of these ligands, the distribution of receptor-bound ligands being, in essence, a kinetic one. By contrast, when binding assays were run at 25 °C, the relative binding affinities of some of the compounds turned out to be higher than that of estradiol. One would imagine that, at the higher temperature, the ligands would reach a complete thermodynamic equilibrium, such that the very low dissociation rates of the high-affinity ligand would lead to the occupation of a greater fraction of the binding sites than was the case when true equilibrium had not been reached. Assays performed under these latter conditions (25 °C) are, therefore, more sensitive to ligand structure.

C. Relative Binding Affinities of 17α-Substituted Estradiols with Rat Uterine Estrogen Receptor at 0 °C (Table 4)[47,48]

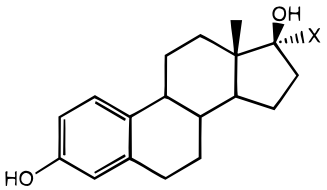
Salman et al.[47,48] published two sets of relative binding affinities for 17α-substituted estradiol derivatives; compounds in this series include the orally active 17α ethynyl estrogens as well as other analogues developed to evaluate the potential for fluo-

Table 4. Relative Binding Affinities of 17α-Substituted Estradiol Derivatives[47,48]

substituents	log RBA obsd	log RBA calcd from eq 4	Δ	MgVol	*I*
C≡CH	2.02	1.76	0.25	2.40	0
H	2.00	2.13	−0.13	2.20	0
Me	1.86	1.87	0.00	2.34	0
$CH_2C{\equiv}CH$	1.69	1.50	0.19	2.54	0
C≡CMe	1.51	1.50	0.00	2.54	0
C≡CI	1.49	1.29	0.20	2.65	0
$CH_2C{\equiv}CI$	1.45	1.03	0.42	2.79	0
$CH_2CH{=}CH_2$	1.26	1.42	−0.17	2.58	0
C_6H_5	1.08	1.00	0.08	2.81	0
	0.92	0.33	0.59	2.73	1
	0.90	0.89	0.01	2.43	1
$C{\equiv}CC{\equiv}CC_6H_5$	0.85	0.28	0.58	3.20	0
$C{\equiv}CH_2OH$	0.81	0.58	0.23	2.59	1
$C{\equiv}C_6H_5$	0.76	0.64	0.13	3.00	0
C_3H_7	0.69	1.35	−0.65	2.62	0
C_4H_9	0.67	1.08	−0.41	2.76	0
$C{\equiv}CC_4H_8I$	0.67	0.24	0.43	3.22	0
$CH_2C_6H_5$	0.63	0.74	−0.11	2.95	0
 [a]	0.60	−0.63	1.23	3.25	1
$C{\equiv}CC_6H_{11}$	0.51	0.40	0.11	3.13	0
$C{\equiv}CC{\equiv}CC_5H_{11}$	0.40	0.10	0.30	3.30	0
	0.32	0.62	−0.30	2.57	1
CH_2CH_2CHOH	0.15	0.42	−0.28	2.68	1
$CH_2C{\equiv}CC_5H_{11}$	0.04	0.20	−0.16	3.24	0
$CH_2C{\equiv}CC_6H_{13}$	0.04	0.14	−0.10	3.27	0
$CH_2C{\equiv}CC{\equiv}CC_4H_9$	0.04	0.10	−0.06	3.30	0
$C{\equiv}CC_6H_{13}$	−0.05	0.20	−0.24	3.24	0
$C{\equiv}CC_6H_{12}I$	−0.10	−0.28	0.18	3.50	0
 [a]	−0.28	0.62	−0.90	2.57	1
$CH_2CH{=}CHC_5H_{11}$	−0.30	0.12	−0.42	3.28	0
$CH_2C{\equiv}CC{\equiv}CC_6H_5$	−0.40	0.02	−0.41	3.34	0
$CH_2OC_6H_{13}$	−0.70	−0.62	−0.08	3.24	1
	−0.96	−0.79	−0.17	3.34	1

[a] Data points not used in deriving equation.

rescence assays of the estrogen receptor. We combined these two sets of data and derived eq 4.



$$\log \text{RBA} = -0.18(\pm 0.28)I - 1.85(\pm 0.2)\text{MgVol} + 6.20(\pm 0.94) \quad (4)$$

$$n = 31 \quad r^2 = 0.85 \quad s = 0.310$$
$$F_{1,29} = 55.8(\text{MgVol}) \quad F_{1,28} = 36.5(I)$$

In this equation, *I* is an indicator variable for any substituent groups containing oxygen atoms (i.e., *I* = 1 for the presence of oxygen and 0 for other substituents). From the QSAR equation, we can see that the estrogen receptor has a limitation for the size of the substituents at the 17α-position. There is a negative MgVol (calculated molecular volume by McGowan's method[49]) term in this equation as well.

Table 5. Relative Binding Affinities of 17α-XCH=CH-estradiol Derivatives at 0 °C[50]

substituents	log RBA obsd	log RBA calcd from eq 5	Δ	*I*	B5
I(*Z*)	2.31	2.16	0.14	1	2.15
Br(*Z*)	2.29	2.19	0.10	1	1.95
Cl(*Z*)	2.10	2.21	−0.11	1	1.80
$SC_6H_5(Z)$[a]	2.07	1.63	0.44	1	6.42
Cl(*E*)	2.01	1.97	0.04	0	1.80
estradiol	2.00	2.07	−0.07	0	1.00
Br(*E*)	1.89	1.95	−0.06	0	1.95
I(*E*)	1.89	1.92	−0.04	0	2.15
$SeC_6H_5(E)$	1.51	1.39	0.12	0	6.42
$SeC_6H_5(Z)$	1.49	1.63	−0.14	1	6.42
$SC_6H_5(E)$	1.39	1.39	0.00	0	6.42

[a] Data points not used in deriving equation.

Table 6. Relative Binding Affinities of 17α-XCH=CH-estradiol Derivatives at 25 °C[50]

substituents	log RBA obsd	log RBA calcd from eq 6	Δ	*I*	B5
I(*Z*)	2.89	2.59	0.30	1	2.15
Br(*Z*)	2.82	2.61	0.21	1	1.95
Cl(*Z*)	2.30	2.63	−0.33	1	1.80
$SC_6H_5(Z)$	2.03	2.21	−0.18	1	6.42
estradiol	2.00	1.97	0.03	0	1.00
Cl(*E*)	1.90	1.89	0.01	0	1.80
I(*E*)	1.79	1.86	−0.07	0	2.15
Br(*E*)	1.75	1.88	−0.13	0	1.95
$SC_6H_5(E)$	1.71	1.47	0.24	0	6.42
$SeC_6H_5(E)$	1.40	1.47	−0.07	0	6.42
$SeC_6H_5(Z)$[a]	1.39	2.21	−0.82	1	6.42

[a] Data points not used in deriving equation.

There is high collinearity between MR, log *P*, and MgVol so that one cannot say that one particular parameter is the proper one. We have used MgVol simply because it yields a slightly better correlation. We believe that a negative steric effect is operative for this set of congeners. However, the fact that correlation with *I* is large and negative, which indicates that any compound with a polar oxygen group has a markedly lower binding affinity, suggests that this region still does not favor substituents that are very polar. Although the portion of the estrogen receptor that accommodates the D-ring region of ligands appears to be more polar than the regions more interior in the ligand binding pocket,[26] there are probably insufficient possibilities for hydrogen bonding and other productive interactions for stabilizing polar functions to compensate for the desolvation energy that must be expended for these substituents to leave water and enter the ligand-binding pocket of the receptor.

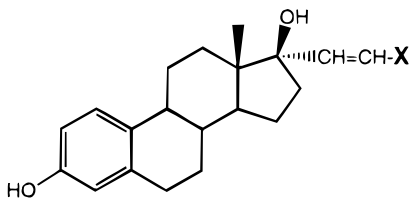
D. Relative Binding Affinities of 17α-XCH=CH-estradiols with Rat Uterine Estrogen Receptor (Tables 5 and 6)[50]

Compounds in this series include radiohalogen-labeled estrogens that are being developed for imaging estrogen receptor-positive breast tumors.[39,40] In eqs 5 and 6, which were developed from the data at 0 and 25 °C, respectively, *I* is an indicator parameter for geometric isomerism (*I* = 1 for *Z*-isomers and 0 for *E*-isomers). B5 is a sterimol steric parameter;[36,37]

Table 7. Relative Binding Affinities of 11β-, 16α-, and 17α-Substituted Estradiols[51]

substituents			log RBA					
11β	16α	17α	obsd	calcd from eq 7	Δ	π,11	MR,11	*I*
Et	H	H	2.60	2.21	0.39	1.02	1.03	0
H	H	C≡CH	2.05	1.81	0.24	0.00	0.10	0
H	H	H	2.00	1.81	0.19	0.00	0.10	0
H	H	CH=CHMe(CIS)	2.00	1.81	0.19	0.00	0.10	0
Et	H	C≡CH	1.94	2.21	−0.26	1.02	1.03	0
Et	OH	Me	1.93	1.50	0.43	1.02	1.03	1
Et	H	Me	1.92	2.21	−0.29	1.02	1.03	0
Et	OH	C≡CH	1.90	1.50	0.40	1.02	1.03	1
Et	H	C≡CMe	1.82	2.21	−0.39	1.02	1.03	0
H	H	Me	1.76	1.81	−0.05	0.00	0.10	0
H	H	C≡CMe	1.67	1.81	−0.14	0.00	0.10	0
H	H	C_6H_5	1.48	1.81	−0.33	0.00	0.10	0
Et	OH	H	1.45	1.50	−0.05	1.02	1.03	1
OMe	H	CH=CHMe	1.42	1.04	0.38	−0.02	0.79	0
H	OH	H	1.32	1.10	0.23	0.00	0.10	1
H	OH	C≡CH	1.30	1.10	0.21	0.00	0.10	1
Et	OH	C≡CMe	1.26	1.50	−0.24	1.02	1.03	1
OMe	H	C≡CMe	1.26	1.04	0.22	−0.02	0.79	0
OMe	H	C_6H_5	1.18	1.04	0.14	−0.02	0.79	0
H	OH	Me	1.15	1.10	0.05	−0.02	0.10	1
OMe	H	C≡CH	1.15	1.04	0.11	0.00	0.79	0
H	OH	C_6H_5	0.90	1.10	−0.19	0.00	0.10	1
OMe	H	H	0.90	1.04	−0.14	−0.02	0.79	0
OMe	H	Me	0.78	1.04	−0.26	−0.02	0.79	0
H	OH	C≡CMe	0.70	1.10	−0.40	0.00	0.10	1
OMe	OH	C_6H_5	0.70	0.33	0.37	−0.02	0.79	1
OMe	OH	Me	0.60	0.33	0.27	−0.02	0.79	1
OMe	OH	H	0.00	0.33	−0.33	−0.02	0.79	1
OMe	OH	C≡CH	0.00	0.33	−0.33	−0.02	0.79	1
OMe	OH	C≡CMe	−0.10	0.33	−0.43	−0.02	0.79	1

it measures the bulkiness, more or less, in terms of the width of the substituent. Both of the QSAR equations have small but negative coefficients with B5, indicating a small negative steric effect. Again, there is no evidence for a role for substituent hydrophobicity at this site.



a. relative binding affinities at 0 °C

$$\log \text{RBA} = 0.24(\pm 0.17)I - 0.13(\pm 0.04)\text{B5} + 2.19(\pm 0.17) \quad (5)$$

$$n = 10 \quad r^2 = 0.91 \quad s = 0.111$$
$$F_{1,8} = 26.2(\text{B5}) \quad F_{1,7} = 11.0(I)$$

b. relative binding affinities at 25 °C

$$\log \text{RBA} = 0.73(\pm 0.34)I - 0.09(\pm 0.08)\text{B5} + 2.06(\pm 0.34) \quad (6)$$

$$n = 10 \quad r^2 = 0.83 \quad s = 0.225$$
$$F_{1,8} = 14.9(I) \quad F_{1,7} = 7.38(\text{B5})$$

It is interesting that the receptor has a stereospecificity preference for vinyl substituents at the 17α-position: All of the *Z*-isomers have higher affinity than the *E*-isomers. With the increase in temperature, the receptor becomes more sensitive to the stereospecificity (coefficient with *I* increases from 0.24 to 0.73 in eq 5 vs 6); this may also reflect the "affinity leveling effect" in binding assays that are conducted (at 0 °C) without sufficient equilibration times, which was noted above for the 11β-substituted estrogens, in QSAR 2 vs 3. Equations 5 and 6 are not very satisfying since almost all of the information is correlated by the *I* term. It does, however, bring out in numerical form the overriding importance of cis−trans isomerism.

E. Relative Binding Affinities of 11β-, 16α-, and 17α-Substituted Estradiols with Lamb Uterine Estrogen Receptor (Table 7)[51]

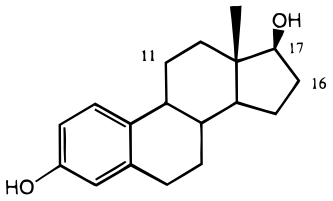
A set of relative binding affinity data at 0 °C for various 11β-, 16α-, and 17α-substituted estradiols has been published by Napolitano et al.[51] We analyzed the data (Table 7) and obtained QSAR eq 7. In this equation, *I* is an indicator variable for compounds containing a 16α-OH that takes the value of 1 for the presence of the OH group (i.e., the estriol series).

It is gratifying that for this series of multiply substituted compounds we see at the 11β-position the same hydrophobic preference (positive correlation with π) and steric interference effects (negative correlation with MR) as we saw for the 11β singly substituted analogues in eq 3. As expected, the polar 16α-OH group reduces binding affinity (negative *I* term). There is no evidence for significant substituent effects at position 17α, although the size range of 17α

Table 8. Relative Binding Affinities of 7α-Undecylestradiol Derivatives[52]

	log RBA			
substituents	obsd	calcd from eq 8	Δ	*L*
OH	0.30	0.22	0.08	2.74
F	0.18	0.29	−0.11	2.65
Cl	−0.16	−0.35	0.20	3.52
Br	−0.70	−0.58	−0.12	3.82
I	−1.00	−0.88	−0.12	4.23
OC_6H_5	−1.00	−1.09	0.09	4.51

substituents in this series is much smaller than that in Tables 5 and 6.

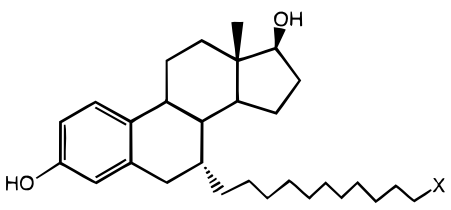


$$\log \text{RBA} = 1.38(\pm 0.32)\pi,11 - 1.08(\pm 0.39)\text{MR},11 - 0.71(\pm 0.23)I + 1.92(\pm 0.24) \quad (7)$$

$$n = 30 \quad r^2 = 0.82 \quad s = 0.300$$
$$F_{1,28} = 10.6(I) \quad F_{1,27} = 18.5(\pi,11)$$
$$F_{1,26} = 34.7(\text{MR},11)$$

F. Relative Binding Affinities of 7α-Undecylestradiol Derivatives with Calf Uterine Estrogen Receptor (Table 8)[52]

DaSilva and van Lier[52] studied the relative binding affinities of 7α-undecylestradiol derivatives at 4 °C (Table 8). From their data we derived eq 8, which shows a correlation with *L*, the sterimol parameter for substituent length. The negative coefficient with *L* indicates a detrimental effect of the length of substituent groups on receptor−ligand binding. It is surprising that the substituent X that is 11 carbons removed from the core of the estrogen still appears to contact the receptor. This suggests that the ligand is bound deeply within the receptor, so that even lengthy substituents are still within the protein. This view is consistent with the crystal structure of the estrogen receptor ligand binding domain.[26] It is of particular interest that the large phenoxyl group is fit as well as the smaller halogens using *L* calculated by Verloop et al.[36,37] The angular attachment of the OC_6H_5 group of the 7α-undecylestradiol derivative means a shorter effective *L* value.

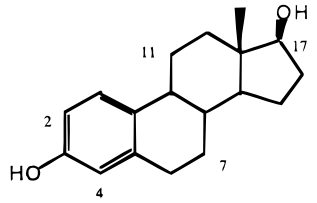


$$\log \text{RBA} = -0.74(\pm 0.25)L + 2.25(\pm 0.92) \quad (8)$$
$$n = 6 \quad R^2 = 0.94 \quad s = 0.155 \quad F_{1,4} = 66.5$$

G. Relative Binding Affinities of Multisubstituted Estradiols with Receptors of Mouse Mammary Epithelial Cells (Table 9)[53]

Gantchev et al.[53] studied the relative binding affinities of estradiol derivatives with multiple substitutions at 2-, 4-, 7α-, 11β-, and 17α-positions with estrogen receptor in the cytosol of mouse mammary epithelial cells at 0 °C. They did a 3-D analysis with CoMFA, which revealed the importance of electrostatic and steric fields. From their data we obtained eq 9.

It seems that groups at positions ortho to the 3-OH group (at C-2 and C-4) have strong detrimental effects on the receptor−ligand binding, as indicated by the negative coefficient with MR,(2&4). This is not surprising, considering that the A-ring of estradiol is tightly constrained within the ligand binding pocket of the estrogen receptor.[26] With this series, similar hydrophobic (π) and steric (MR) effects of substituents at the 11β-position were obtained, although the coefficients were much lower than the ones contained in eqs 3 and 7.



$$\log \text{RBA} = 0.55(\pm 0.34)\pi,11 - 0.49(\pm 0.19)\text{MR},11 - 1.78(\pm 0.23)\text{MR},(2\&4) + 2.09(\pm 0.13) \quad (9)$$

$$n = 48 \quad r^2 = 0.85 \quad s = 0.231$$
$$F_{1,46} = 145(\text{MR, 2\&4}) \quad F_{1,45} = 296(\text{MR},11)$$
$$F_{1,44} = 10.8(\pi,11)$$

In summary, from the results of eqs 1−9, we can see that the estrogen receptor has a certain limited tolerance to steric effects of the substituents. In all of the equations, the only place that we see positive hydrophobic interactions is at the 11β-position. Why is there a lack of hydrophobic interactions at other sites between a *hydrophobic binding domain in the receptor site*[26,31] and these very hydrophobic ligands? The answer might be that most of the hydrophobic area of the ligand binding pocket is, in fact, already well covered by the steroid portion of the ligand, leaving little hydrophobic space for the substituents except at the 11β-position. The X-ray structure of the estrogen receptor ligand binding domain and estradiol speaks otherwise, however.[26] There is, in fact, nearly 200 Å^3 of empty volume surrounding the ligand, and much of it is above the B- and C-rings near 11β, but there is other space below the B-ring near 7α. However, for substituents at the 7α-position (eq 9), we find no term for hydrophobicity; only H and Me are involved in this correlation, but we find no consistent difference between the two substituents.

Table 9. Relative Binding Affinities of Multisubstituted Estradiol Derivatives[53]

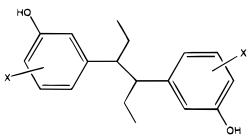
substituents					log RBA					
2	4	7α	11β	17α	obsd	calcd from eq 9	Δ	π,11	MR,11	MR,(2&4)
H	F	H	H	C≡CH	2.08	1.69	0.38	0.00	0.10	0.20
H	H	H	H	H	2.00	1.67	0.33	0.00	0.10	0.20
H	H	H	H	CH≡CH	2.00	1.67	0.33	0.00	0.10	0.20
H	H	H	Et	H	1.89	1.78	0.12	1.02	1.03	0.20
H	F	H	H	$CH=CH_2$	1.87	1.69	0.18	0.00	0.10	0.20
H	H	Me	H	C≡CH	1.86	1.67	0.20	0.00	0.10	0.20
H	F	H	H	CH=CHI(*Z*)	1.83	1.69	0.13	0.00	0.10	0.20
F	H	H	H	C≡CH	1.81	1.75	0.06	0.00	0.10	0.20
H	H	H	Me	H	1.81	1.69	0.12	0.56	0.57	0.20
H	F	Me	H	C≡CH	1.80	1.69	0.10	0.00	0.10	0.20
H	F	H	H	CH=CHI(*E*)	1.75	1.69	0.06	0.00	0.10	0.20
F	H	H	H	CH=CHI(*Z*)	1.73	1.69	0.04	0.00	0.10	0.20
H	H	H	H	C≡CCl	1.73	1.67	0.06	0.00	0.10	0.20
H	H	H	H	CH=CHI(*Z*)	1.70	1.67	0.02	0.00	0.10	0.20
H	H	H	H	CH=C(I)Cl	1.69	1.67	0.00	0.00	0.10	0.20
H	H	Me	H	CH=CHI(*Z*)	1.65	1.67	−0.02	0.00	0.10	0.20
H	H	Me	H	CH=CHI(*E*)	1.64	1.67	−0.04	0.00	0.10	0.20
H	H	H	Et	CH=CHI(*E*)	1.63	1.78	−0.14	1.02	0.10	0.20
H	H	H	H	CH=CHI(*E*)	1.60	1.67	−0.07	0.00	0.10	0.20
H	F	H	OMe	CH=CHI(*Z*)	1.58	1.34	0.24	−0.02	0.79	0.20
H	H	Me	H	C≡CCl	1.58	1.67	−0.09	0.00	0.10	0.20
H	F	Me	H	CH=CHI(*Z*)	1.57	1.69	−0.13	0.00	0.10	0.20
H	H	Me	H	CH=C(I)Cl	1.57	1.67	−0.11	0.00	0.10	0.20
H	F	Me	H	CH=CHI(*E*)	1.56	1.69	−0.13	0.00	0.10	0.20
H	H	H	OMe	CH=C(I)Cl	1.54	1.32	0.22	−0.02	0.79	0.20
F	H	Me	H	C≡CH	1.52	1.69	−0.17	0.00	0.10	0.20
H	H	H	OEt	C≡CCl	1.52	1.32	0.20	0.38	1.25	0.20
H	H	H	OMe	CH=CHI(*Z*)	1.51	1.32	0.18	−0.02	0.79	0.20
H	H	H	OEt	CH=C(I)Cl	1.50	1.32	0.18	0.38	1.25	0.20
H	H	H	OMe	C≡CCl	1.47	1.32	0.15	−0.02	0.79	0.20
F	H	Me	H	CH=CHI(*Z*)	1.46	1.69	−0.24	0.00	0.10	0.20
F	H	H	OMe	CH=CHI(*Z*)	1.45	1.34	0.10	−0.02	0.79	0.20
H	H	H	OMe	CH=CHI(*E*)	1.44	1.34	0.10	−0.02	0.79	0.20
H	F	H	OMe	CH=CHI(*E*)	1.44	1.32	0.12	−0.02	0.79	0.20
H	H	H	OEt	C≡CH	1.32	1.32	0.00	0.38	1.25	0.20
F	H	H	H	$CH=CH_2$	1.25	1.69	−0.45	0.00	0.10	0.20
H	H	H	OEt	CH=CHI(*E*)	1.23	1.32	−0.08	0.38	1.25	0.20
H	F	H	OMe	C≡CH	1.22	1.34	−0.12	−0.02	0.79	0.20
F	H	H	OMe	C≡CH	1.20	1.34	−0.14	−0.02	0.79	0.20
H	H	H	OMe	C≡CH	1.19	1.32	−0.14	−0.02	0.79	0.20
F	H	H	OMe	CH=CHI(*E*)	1.13	1.34	−0.21	−0.02	0.79	0.20
F	H	Me	H	CH=CHI(*E*)	1.08	1.69	−0.60	0.00	0.10	0.20
H	H	H	OEt	H	1.00	1.31	−0.32	0.38	1.25	0.20
H	H	H	OMe	H	0.85	1.32	−0.48	−0.02	0.79	0.20
F	H	H	H	CH=CHI(*E*)[a]	0.79	1.69	−0.89	0.00	0.10	0.20
H	Br	H	H	H	0.70	0.28	0.42	0.00	0.10	0.99
Br	H	H	H	H	0.08	0.28	−0.20	0.00	0.10	0.99
I	H	H	H	H	−0.40	−0.63	0.23	0.00	0.10	1.50
H	I	H	H	H	−1.00	−0.63	−0.37	0.00	0.10	1.50

[a] Data point not used in deriving equation.

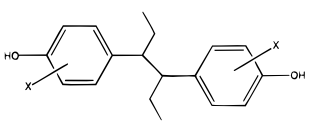
2. Nonsteroidal Compounds

A. Metahexestrol and Hexestrol Derivatives

Hartmann et al.[54,55] synthesized and measured the relative binding affinity of metahexestrol and hexestrol derivatives with estrogen receptors of calf uterine cytosol at 4 °C (Tables 10 and 11). We have derived eqs 10 and 11 for metahexestrol and hexestrol derivatives, respectively:



metahexestrol derivatives



hexestrol derivatives

relative binding affinity of metahexestrol derivatives

$$\log \text{RBA} = -1.25(\pm 0.20)L,4 - 2.85(\pm 0.50)\text{B5-5} - 0.48(\pm 0.35)\sigma^- + 6.23(\pm 0.94) \quad (10)$$

$$n = 14 \quad r^2 = 0.97 \quad s = 0.226 \quad F_{1,12} = 6.6(L,4)$$
$$F_{1,11} = 108(\text{B5-5}) \quad F_{1,10} = 9.4(\sigma^-)$$

relative binding affinity of hexestrol derivatives

$$\log \text{RBA} = -0.62(\pm 0.34)L,2 - 1.25(\pm 0.64)\text{B1-3} - 0.58(\pm 0.18)\text{B5-3} - 0.73(\pm 0.44)\sigma^- + 4.78(\pm 1.53) \quad (11)$$

$$n = 20 \quad r^2 = 0.89 \quad s = 0.320 \quad F_{1,18} = 19.3(\text{B5-3})$$
$$F_{1,17} = 14.7(\sigma) \quad F_{1,16} = 15.5(L,2)$$

Table 10. Relative Binding Affinity of Metahexestrol Derivatives[54,55]

substituents	log RBA obsd	log RBA calcd from eq 10	Δ	*L*,4	B5-5	σ^-
6-Me	1.17	0.89	0.28	2.06	1.00	−0.17
H	1.00	0.81	0.19	2.06	1.00	0.00
6-F	0.84	0.82	0.02	2.06	1.00	−0.03
6-Cl	0.62	0.72	−0.09	2.06	1.00	0.19
4-F	0.04	0.09	−0.04	2.65	1.00	−0.03
4-Me	−0.14	−0.12	−0.01	2.87	1.00	−0.17
4-NH_2	−0.28	0.21	−0.49	2.78	1.00	−0.63
4-Cl	−1.15	−1.11	−0.05	3.52	1.00	0.19
4-Et	−1.39	−1.66	0.27	4.11	1.00	−0.19
4-Br	−1.52	−1.62	0.09	3.82	1.00	0.25
4-CH_2OMe[a]	−1.52	−2.54	1.02	4.78	1.00	−0.10
4-CH_2OH	−1.52	−1.51	−0.01	3.97	1.00	0.08
5-Cl	−1.52	−1.65	0.12	2.06	1.80	0.37
4-NO_2	−1.70	−1.52	−0.18	3.44	1.00	1.27
5-OH	−2.00	−1.90	−0.11	2.06	1.93	0.12

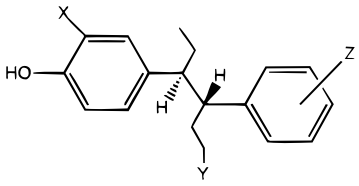
[a] Data point omitted in deriving QSAR.

Table 11. Relative Binding Affinity of Hexestrol Derivatives[54,55]

substituents	log RBA obsd	log RBA calcd from eq 11	Δ	*L*,2	B1-3	B5-3	σ^-
2-OH	1.51	1.14	0.37	2.74	1.00	1.00	0.12
H	1.43	1.65	−0.22	2.06	1.00	1.00	0.00
3-OH	1.30	0.94	0.36	2.06	1.35	1.93	−0.37
3-F	1.20	1.03	0.18	2.06	1.35	1.35	−0.03
2-Me	0.93	1.20	−0.27	2.87	1.00	1.00	−0.07
3-Me	0.91	0.52	0.39	2.06	1.52	2.04	−0.17
2-F	0.81	1.03	−0.22	2.65	1.00	1.00	0.34
3-NH_2	0.75	1.11	−0.36	2.06	1.35	1.97	−0.63
2-C_2H_5	0.57	0.42	0.15	4.11	1.00	1.00	−0.07
2-Cl	0.34	0.47	−0.13	3.52	1.00	1.00	0.37
2-Br	0.25	0.27	−0.01	3.82	1.00	1.00	0.39
3-Cl	0.20	0.04	0.16	2.06	1.80	1.80	0.19
3-C_2H_5	0.08	−0.12	0.20	2.06	1.52	3.17	−0.19
3-CH_2OH	−0.15	−0.05	−0.10	2.06	1.52	2.70	0.08
3-CH_2OMe	−0.26	−0.40	0.14	2.06	1.52	3.40	0.01
3-Br	−0.60	−0.28	−0.33	2.06	1.95	1.95	0.25
3-$CH_2OC_2H_5$	−0.75	−1.01	0.27	2.06	1.52	4.45	0.01
3-I	−0.77	−0.66	−0.11	2.06	2.15	2.15	0.27
3-NO_2	−0.85	−0.99	0.14	2.06	1.70	2.44	1.27
3-CH_2NMe_2	−1.39	−0.80	−0.60	2.06	1.52	4.08	0.01

The only substituents examined in this correlation were ones on the phenolic rings, and for these the receptor−ligand interaction equations contain only steric and electronic parameters. As expected for aromatic ring substituents, steric bulk interferes with binding. However, it seems that electron-donating groups have a weak effect favoring receptor binding. Electron-donating groups will increase the electron density on the phenyl ring, but will also make the phenolic hydroxyl less acidic. Thus, the increased affinity of the derivatives with electron-donating groups could be due to (a) an increased electron-transfer interaction between receptor and ligands or (b) an effect on the electron density on the OH. From the crystal structure of the estrogen receptor ligand binding domain, both effects are reasonable,[26] as the A-ring of estradiol is tightly surrounded by residues and the phenolic hydroxyl group donates one hydrogen bond (which would be weakened by the increased electron density) but accepts two hydrogen bonds.[26] From these two equations, one also can see that substituents that are ortho and meta to the phenolic hydroxy group tend to decrease the receptor binding. Hydrophobicity of the substituents as defined by ClogP or π does not play any significant role in the receptor binding of these two sets of compounds.

Katzenellenbogen's group synthesized a series of 4-substituted deoxyhexestrol derivatives and measured their relative receptor binding affinities with estrogen receptor in lamb uterine cytosol at 0 °C[56−58] (Table 12). From this set, we obtained eq 12, which has negative correlations with the size parameter MR and an electronic term σ. It also contains a small negative hydrophobic term (ClogP). This equation also indicates that the estrogen receptor has limited tolerance to the size of its ligands. The coefficient (−1.19) with σ,z again indicates that high electron density will facilitate receptor binding, an indication of a possible charge-transfer interaction between the receptor and the second phenyl ring.

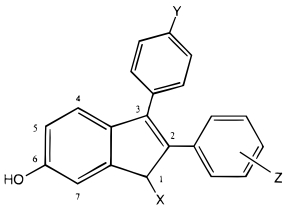


$$\log \text{RBA} = -0.17(\pm 0.14)\text{ClogP} - 0.63(\pm 0.13)\text{MR-}z\text{,sum} - 1.19(\pm 0.25)\sigma,z - 1.39(\pm 1.33)\sigma^-,x + 2.67(\pm 0.76) \quad (12)$$

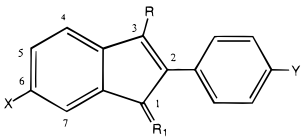
$$n = 33 \quad r^2 = 0.89 \quad s = 0.268$$
$$F_{1,31} = 24.9(\sigma,z) \quad F_{1,30} = 5.50(\text{MR-}z\text{,sum})$$
$$F_{1,29} = 15.9(\text{ClogP}) \quad F_{1,28} = 4.6(\sigma^-,x)$$

B. Indenes and Indenones

Anstead and Katzenellenbogen[59,60] have studied the relative binding affinity of a series of indenes and indenones with rat uterine cytosol estrogen receptors at 0 °C (Tables 13 and 14).



2,3-diarylindenes



2-arylindenes and indenones

relative binding affinity of 2,3-diarylindenes

$$\log \text{RBA} = -1.31(\pm 0.41)\sigma,z + 1.42(\pm 0.58)\text{B1-}z\text{,2} - 0.67(\pm 0.68) \quad (13)$$

$$n = 12 \quad r^2 = 0.90 \quad s = 0.264$$
$$F_{1,10} = 12.2(\sigma,z) \quad F_{1,9} = 30.6(\text{B1-}z\text{,2})$$

relative binding affinity of 2-arylindenes and indenones

$$\log \text{RBA} = 1.94(\pm 0.64)I,x + 1.53(\pm 0.63)I,y - 1.61(\pm 0.63) \quad (14)$$

$$n = 14 \quad r^2 = 0.84 \quad s = 0.514$$
$$F_{1,12} = 8.8(I,x) \quad F_{1,11} = 28.4(I,y)$$

Equation 13, for the 2,3-diarylindenes, is similar to eqs 10 and 11 in that electron-donating groups were found to enhance the receptor binding. Substituents at the ortho position in the pendant C-2 phenyl ring (the Z substituent) facilitate the binding. This

Table 12. Relative Binding Affinity of 4-Substituted Deoxyhexestrol Derivatives with Lamb Uterine Estrogen Receptor[56−58]

substituents			log RBA						
X	Y	Z	obsd	calcd from eq 12	Δ	ClogP	MR,*z*	σ^-,*x*	σ,*z*
H	H	4-OH	2.48	1.99	0.48	5.11	0.39	0.00	−0.37
H	H	2-F,4-OH	2.31	1.92	0.39	5.35	0.39	−0.03	−0.31
H	F	4-OH	2.11	2.12	−0.01	4.38	0.39	0.00	−0.37
H	Cl	4-OH	2.05	2.04	0.01	4.82	0.39	0.00	−0.37
H	H	4-NH_2	1.98	2.27	−0.30	4.55	0.65	0.00	−0.66
H	Br	4-OH	1.85	2.02	−0.17	4.96	0.39	0.00	−0.37
F	H	2-F,4-OH	1.81	1.89	−0.08	5.59	0.38	−0.03	−0.31
H	I	4-OH	1.78	1.95	−0.17	5.35	0.39	0.00	−0.37
H	H	4-CH_2OH	1.60	1.34	0.26	4.74	0.82	0.00	0.00
H	H	4-OCH_2CH_2OH	1.32	1.27	0.05	4.82	1.42	0.00	−0.27
H	H	4-OMe	1.28	1.46	−0.18	5.70	0.89	0.00	−0.27
H	H	2-Br,4-OH	1.20	1.22	−0.02	5.99	0.39	0.25	−0.14
H	H	2-I,4-OH	1.15	1.22	−0.08	6.18	0.39	0.27	−0.19
H	H	4-NO_2[a]	1.04	0.27	0.77	5.52	0.84	0.00	0.78
H	H	4-CH_2Br	0.97	0.48	0.49	6.56	1.44	0.00	0.14
H	H	4-CH_2Cl	0.90	0.72	0.18	6.34	1.15	0.00	0.12
H	H	4-$OCH_2(CHOCH_2)$	0.87	0.96	−0.09	5.26	1.80	0.00	−0.27
H	H	4-$OCH_2CH(OH)CH_2OH$	0.84	0.97	−0.13	4.00	2.13	0.00	−0.27
H	H	4-N_3	0.76	0.81	−0.05	6.22	1.12	0.00	0.08
H	H	4-$OCH_2CH{=}CH_2$	0.73	0.82	−0.09	6.27	1.71	0.00	−0.25
H	H	4-CN	0.72	0.53	0.18	5.21	0.74	0.00	0.66
H	H	4-CH_2OEt	0.69	0.51	0.18	6.08	1.77	0.00	0.01
H	H	4-$OCH_2CH(OH)Me$	0.63	0.87	−0.24	5.13	1.98	0.00	−0.27
H	H	4-COMe	0.61	0.42	0.20	5.22	1.22	0.00	0.50
H	H	4-$OCH_2CH(OH)CH_2Cl$	0.60	0.55	0.05	5.20	2.47	0.00	−0.27
H	H	4-$OCH_2CH(OH)CH_2Br$	0.51	0.35	0.15	5.34	2.75	0.00	−0.27
I	H	2-I,4-OH	0.51	0.23	0.28	7.25	1.68	0.27	−0.19
Br	H	2-Br,4-OH	0.42	0.58	−0.16	6.87	1.17	0.25	−0.14
H	H	4-SO_2N_3	0.26	−0.08	0.34	3.42	1.72	0.00	0.91
H	H	4-COOH	0.26	0.69	−0.44	5.52	0.80	0.00	0.45
H	H	4-$COCH_2Br$	0.04	−0.02	0.06	5.47	1.84	0.00	0.50
H	H	4-$O(CH_2)_2O(CH_2)_2Cl$	−0.05	0.12	−0.16	6.04	2.93	0.00	−0.27
H	H	4-$COCH_2Cl$	−0.36	0.13	−0.49	5.33	1.64	0.00	0.50
H	H	4-SO_2F	−0.48	−0.04	−0.45	5.94	0.97	0.00	0.91

[a] Data point omitted in deriving QSAR.

Table 13. Relative Binding Affinity of 2,3-Diarylindenes[59,60]

substituents			log RBA				
X	Y	Z	obsd	calcd from eq 13	Δ	σ,*z*	B1-*z*,2
H	4-OH	2-CF_3	1.58	1.45	0.13	0.54	1.99
H	4-OH	2-Me	1.48	1.72	−0.24	−0.17	1.52
H	H	4-OH	1.45	1.24	0.21	−0.37	1.00
H	4-OH	4-OH	1.17	1.24	−0.07	−0.37	1.00
Me	H	H	1.08	0.76	0.32	0.00	1.00
H	H	H	0.95	0.76	0.19	0.00	1.00
H	4-OH	H	0.67	0.75	−0.08	0.00	1.00
H	4-OH	4-CN	0.00	−0.11	0.11	0.66	1.00
H	4-OH	4-Br	−0.12	0.45	−0.67	0.23	1.00
H	4-OH	4-NO_2	−0.16	−0.27	0.11	0.78	1.00
H	4-OH	3-NO_2	−0.22	−0.18	−0.04	0.71	1.00
H	H	4-NO_2	−0.34	−0.27	−0.07	0.78	1.00

Table 14. Relative Binding Affinity of 2-Arylindenes and Indenones[59,60]

substituents				log RBA				
R	X	Y	R1	obsd	calcd from eq 14	Δ	*I*,*x*	*I*,*y*
Et	OH	OH	Me,H	1.91	1.86	0.05	1.00	1.00
Et	OH	OH	Et,H	1.89	1.86	0.03	1.00	1.00
C_6H_5	OH	H	O[a]	1.77	0.33	1.44	1.00	0.00
Et	OH	OH	H2	1.20	1.86	−0.66	1.00	1.00
C_6H_5	OH	H	Me,H	1.08	0.33	0.75	1.00	0.00
Et	H	OH	Et,H[a]	0.97	−0.08	1.05	0.00	1.00
C_6H_5	OH	H	H2	0.95	0.33	0.62	1.00	0.00
Et	H	OH	O	0.66	−0.08	0.74	0.00	1.00
Et	H	OH	H2	0.36	−0.08	0.44	0.00	1.00
Et	OH	H	Et,H	0.34	0.33	0.01	1.00	0.00
Et	OH	H	O	0.08	0.33	−0.25	1.00	0.00
Et	OH	H	H2	−0.24	0.33	−0.57	1.00	0.00
C_6H_5	H	OH	O	−0.35	−0.08	−0.27	0.00	1.00
C_6H_5	H	OH	H2	−0.44	−0.08	−0.36	0.00	1.00
C_6H_4-4-OH	H	H	H2	−1.77	−1.61	−0.16	0.00	0.00
C_6H_5	H	H	O	−2.02	−1.61	−0.41	0.00	0.00

[a] Data point omitted in deriving QSAR.

may result from the substituent filling a pocket on the receptor. Alternatively, such substitution results in a sterically enforced twisting of the C-2 phenyl group that would restrict the conformational freedom of this ligand. B1-*z*,2 is the sterimol parameter B1 for the ortho substituents on the 2-phenyl ring.

Equation 14 is not a good QSAR because of the poor substituent selection. *I*,*x* and *I*,*y* are indicator variables for hydroxy groups at the X and Y positions. This equation simply shows that both of the phenolic hydroxy groups are important factors in estrogen receptor binding of the 1-arylindenes and indenones. Considering the confidence limits on these two terms, however, one cannot be sure that there is much difference in their affinity. It is significant that in the crystal structure of the estrogen receptor complexed with the structurally related benzothiophene Raloxifene, both phenolic hydroxyl groups are involved in hydrogen bonds to the receptor.[26]

C. 2-Phenylindole Derivatives

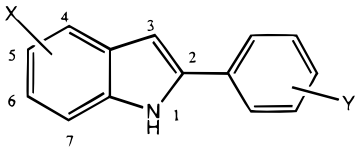
von Angerer et al.[61] reported the relative binding affinities of a series of 2-phenylindoles with calf uterine estrogen receptor at 4 °C (Table 15). Equation

Table 15. Relative Binding Affinities of 2-Phenylindoles[61]

substituents		log RBA				
X	Y	obsd	calcd from eq 15	Δ	*I*,*y*	E_S-*x*,1
6-OH,1-Et-3-Me	4-OH	1.52	0.83	0.69	0	−1.31
5-OH,1-Et-3-Et	4-OH	1.36	0.83	0.53	0	−1.31
6-OH,1,3-Et_2	4-OH	1.32	0.83	0.49	0	−1.31
6-OH,1-C_3H_7,3-Et	4-OH	1.28	1.03	0.25	0	−1.43
5-OH,1-C_3H_7	4-OH	1.26	1.03	0.22	0	−1.43
6-OH,1-Et	4-OH	1.20	0.83	0.38	0	−1.31
5-OH,1-C_3H_7,3-Me	4-OH	1.20	1.03	0.17	0	−1.43
6-OH,1-C_3H_7,3-Me	4-OH	1.11	1.03	0.08	0	−1.43
6-OH,1-$CHMe_2$,3-Me	4-OH	1.11	1.51	−0.39	0	−1.71
6-OH,1,3-Me_2	4-OH	1.00	0.71	0.29	0	−1.24
5-OH,1-Et,3-Me	4-OH	0.98	0.83	0.15	0	−1.31
6-OH,1-C_3H_7	4-OH	0.93	1.03	−0.10	0	−1.43
5-OH,1-C_3H_7,3-Me	3-OH	0.87	0.46	0.41	1	−1.43
6-OH,1-Me,3-Et	4-OH	0.77	0.71	0.06	0	−1.24
5-OH,1-Et	4-OH	0.76	0.83	−0.07	0	−1.31
5-OH,1,3-Me_2	4-OH	0.66	0.71	−0.05	0	−1.24
5-OH,1-C_4H_9,3-Me	4-OH	0.66	1.38	−0.71	0	−1.63
6-OH,1-C_4H_9	4-OH	0.63	1.38	−0.74	0	−1.63
6-OH,1-Me	4-OH	0.58	0.71	−0.13	0	−1.24
5-OH,1-$CHMe_2$,3-Me	4-OH[a]	0.54	1.51	−0.97	0	−1.71
6-OH,1-C_3H_7,3-Me	3-OH	0.54	0.46	0.08	1	−1.43
6-OH,1-Et,3-Me	3-OH	0.48	0.26	0.22	1	−1.31
5-OH,1-C_5H_{11},3-Me	4-OH[a]	0.36	1.39	−1.03	0	−1.64
5-OH,1-Et,3-Me	3-OH	0.34	0.26	0.09	1	−1.31
5-OH,1,3-$(C_3H_7)_2$	4-OH[a]	0.23	1.03	−0.80	0	−1.43
6-OH,1-Et	3-OH	0.23	0.26	−0.03	1	−1.31
5-OH,1-Et	3-OH	0.23	0.26	−0.03	1	−1.31
5-OH,1-Me	4-OH	−0.10	0.71	−0.81	0	−1.24
5-OH,1,3-Me_2	3-OH	−0.22	0.14	−0.36	1	−1.24
6-OH,1,3-Me_2	3-OH	−0.26	0.14	−0.40	1	−1.24
6-OH,3-Et	4-OH	−0.89	−1.41	0.52	0	0.00
6-OH,3-Me	4-OH	−1.22	−1.41	0.18	0	0.00
5-OH,3-Me	4-OH	−1.22	−1.41	0.18	0	0.00
7-OH,1-Et,3-Me	4-OH[a]	−1.70	0.83	−2.53	0	−1.31
6-OH	4-OH	−2.00	−1.41	−0.59	0	0.00
5-OH	4-OH	−2.00	−1.41	−0.59	0	0.00

[a] Data points not used in deriving equation.

15 is derived for this data set. *I*,*y* is an indicator variable (*I* = 1 for 3-OH). There may be a hydrogen-bonding interaction between the receptor and the OH group on the 2-phenyl ring, and the *p*-OH is more geometrically favorable for hydrogen-bonding than *m*-OH. Interestingly, we see a positive steric effect for substituents on the 1-position, i.e., substituents on the nitrogen. E_S is the Taft steric parameter,[41] all values of which (except *H* = 0) are negative; therefore, the *negative* correlation with the E_S term indicates a *preference* for steric bulk at this site. As was the case with the indene derivatives above, substituents at the 1-position may increase binding by filling a pocket or by conformational restriction that engenders a twist of the 2-aryl substituent.



$$\log \text{RBA} = -0.57(\pm 0.35)I,y - 1.71(\pm 0.30)E_s\text{-}x,1 - 1.41(\pm 0.37) \quad (15)$$

$$n = 32 \quad r^2 = 0.83 \quad s = 0.409$$
$$F_{1,30} = 94.8(E_s\text{-}x,1) \quad F_{1,29} = 11.3(I,y)$$

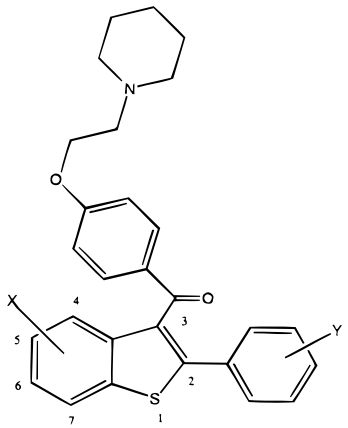
Another set of data for 2-phenylindoles has been reported by Mahboobi et al.[62] (Table 16). The receptor binding affinity was measured in calf uterine cytosol at 0 ∼ 4 °C; eq 16 is derived for this set. *I*, the indicator variable, equals 1 for compounds containing a 4-OCOMe group at the para position of the 2-phenyl ring. This equation contains two negative steric parameters, B1 on the 1-position and MR on the 3-position. It is not clear why, with this set, we see no favorable effect of substituents at these sites as was the case with the indenes (eq 13) and the earlier data set with the indoles (eq 15).

$$\log \text{RBA} = -2.36(\pm 1.12)\text{B1-}x,1 - 1.05(\pm 0.24)\text{MR-}x,3 + 0.46(\pm 0.37)I + 4.66(\pm 1.76) \quad (16)$$

$$n = 23 \quad r^2 = 0.85 \quad s = 0.313 \quad F_{1,21} = 46.5(\text{MR-}x,3)$$
$$F_{1,20} = 10(\text{B-1-}x,1) \quad F_{1,19} = 6.9(I)$$

D. Benzothiophenes

Grese et al.[63] published a large set of data on the relative binding affinities with MCF-7 estrogen receptor at 4 °C and the growth inhibition of MCF-7 cells by benzothiophenes (Tables 17 and 18). We derived eq 17 for the receptor binding and eq 18 for the growth inhibition data.



$$\log \text{RBA} = 1.17(\pm 0.18)I + 0.57(\pm 0.64)\text{MR-}y,2 - 0.41(\pm 0.13)\text{MR-}y,4 - 0.68(\pm 0.47)L\text{-}x,4 + 1.33(\pm 1.02) \quad (17)$$

$$n = 51 \quad r^2 = 0.83 \quad s = 0.278$$
$$F_{1,49} = 73.2(I) \quad F_{1,48} = 38.9(\text{MR-}y,4)$$
$$F_{1,47} = 7.2(L\text{-}x,4) \quad F_{1,46} = 3.7(\text{MR-}y,2)$$

$$\log 1/\text{C} = 1.03(\pm 0.51)\text{ClogP} - 1.23(\pm 0.78)\log(\beta 10^{\text{ClogP}} + 1) + 1.96(\pm 0.34)I - 0.58(\pm 0.27)\text{MR-}y,4 - 0.28(\pm 0.23)\sigma^+,y - 0.06(\pm 3.32) \quad (18)$$

$$n = 46 \quad r^2 = 0.81 \quad s = 0.464$$
$$F_{1,44} = 37.1(I) \quad F_{1,43} = 28.3(\text{MR-}y,4)$$
$$F_{1,42} = 10.9(\text{ClogP}) \quad F_{1,40} = 9.88(\text{bilin(ClogP)})$$
$$F_{1,39} = 4.5(\sigma^+,y)$$

In eq 17, the QSAR for receptor binding, *I* is an indicator variable for compounds containing 6-OH or

Table 16. Relative Binding Affinities of 2-Phenylindoles[62]

substituents		log RBA					
X	Y	obsd	calcd from eq 16	Δ	B1-*x*,1	MR-*x*,3	*I*
6-OCOMe,1-Et,3-Cl	4-OCOMe	1.18	0.91	0.28	1.52	0.603	1
5-OCOMe,1-Et,3-Cl	4-OCOMe	1.02	0.91	0.12	1.52	0.603	1
6-OH,1-Et,3-CHO	4-OH	0.81	0.35	0.46	1.52	0.688	0
6-OH,1-$CH_2C_6H_5$,3-CHO	4-OH	0.74	0.35	0.38	1.52	0.688	0
5-OH,1-Et	4-OH	0.72	0.97	−0.25	1.52	0.103	0
6-OH,1-C_3H_7,3-CHO	4-OH	0.64	0.35	0.28	1.52	0.688	0
6-OCOMe,1-Et,3-CHO	4-OCOMe	0.60	0.82	−0.21	1.52	0.688	1
5-OH,1-$CH_2C_6H_5$,3-CHO	4-OH	0.54	0.35	0.18	1.52	0.688	0
5-OH,1-C_3H_7,3-CHO	4-OH	0.34	0.35	−0.01	1.52	0.688	0
6-OH,1-Et,3-COMe	4-OH	0.28	−0.10	0.37	1.52	1.118	0
6-OH,1-Me,3-CHO	4-OH	0.08	0.35	−0.27	1.52	0.688	0
6-OH,1-Et,3-$CH_2CH(CN)_2$	4-OH	−0.07	−0.41	0.34	1.52	1.419	0
6-OH,1-Et,3-CH_2OH	4-OH	−0.15	0.32	−0.47	1.52	0.719	0
5-OH,1-Et,3-CHO	4-OH	−0.16	0.35	−0.52	1.52	0.688	0
6-OCOMe,1-$SO_2C_6H_5$,3-Cl	4-OCOMe	−0.25	−0.30	0.05	2.03	0.603	1
5-OCOMe,1-$SO_2C_6H_5$,3-Cl	4-OCOMe	−0.35	−0.30	−0.05	2.03	0.603	1
6-OCOMe,1-Et,3-CH_2OCOMe	4-OCOMe	−0.37	−0.19	−0.18	1.52	1.648	1
6-OH,1-Et,3-CH=$CHNO_2$	4-OH	−0.51	−0.65	0.14	1.52	1.642	0
6-OH,1-Et,3-$CH(CH_2NO_2)_2$	4-OH[a]	−0.60	−1.66	1.06	1.52	2.615	0
5-OH,1-Et,3-CH=$CHNO_2$	4-OH	−0.70	−0.65	−0.05	1.52	1.642	0
5-OH,1-Et,3-$CH_2CH(CN)_2$	4-OH	−0.85	−0.41	−0.44	1.52	1.419	0
5-OH,1-Et,3-CH=$CH(CN)_2$	4-OH	−0.96	−0.99	0.03	1.52	1.972	0
6-OH,1-Et,3-CH=$CH(CN)_2$	4-OH	−1.30	−0.99	−0.31	1.52	1.972	0
6-OH,3-CHO	4-OH[a]	−1.52	1.58	−3.11	1.00	0.688	0
5-OH,1-Et,3-$CH(CH_2NO_2)_2$	4-OH	−1.52	−1.66	0.14	1.52	2.615	0
5-OH,1-Et,3-COMe	4-OH[a]	−2.00	−0.10	−1.90	1.52	1.118	0

[a] Data points not used in deriving equation.

Table 17. Relative Binding Affinities of 2-Phenylbenzothiophenes[63]

substituents		log RBA						
X	Y	obsd	calcd from eq 17	Δ	*I*	MR-*y*,2	MR-*y*,4	*L*-*x*,4
6-OH	2-Me,4-OH	1.61	1.31	0.31	1	0.565	0.285	2.06
6-OH	2-Me	1.60	1.38	0.22	1	0.565	0.103	2.06
6-OH	3-F	1.46	1.12	0.34	1	0.103	0.103	2.06
6-OH	3-F,4-OH	1.30	1.04	0.26	1	0.103	0.285	2.06
6-OH	4-F	1.28	1.12	0.16	1	0.103	0.092	2.06
6-OH	3-OH	1.20	1.12	0.08	1	0.103	0.103	2.06
6-OH	2-OMe,4-OH	1.20	1.44	−0.23	1	0.787	0.285	2.06
6-OH	3-Me,4-OH	1.11	1.04	0.07	1	0.103	0.285	2.06
6-OH	3-Cl,4-OH	1.08	1.04	0.04	1	0.103	0.285	2.06
6-OH	3,5-Me_2,4-OH	1.08	1.04	0.04	1	0.103	0.285	2.06
6-OH	4-C≡CH	1.08	0.77	0.31	1	0.103	0.955	2.06
5-OH	4-OH	1.00	1.04	−0.04	1	0.103	0.285	2.06
6-OH	4-CH=CH_2	1.00	0.70	0.29	1	0.103	1.099	2.06
5-F,6-OH	4-OH	0.99	1.04	−0.05	1	0.103	0.285	2.06
6-OH	4-COMe	0.87	0.63	0.17	1	0.103	1.118	2.06
6-OH	4-OMe	0.86	0.84	0.03	1	0.103	0.787	2.06
6-OH	4-Me	0.84	0.93	−0.08	1	0.103	0.565	2.06
5-Me,6-OH	4-OH[a]	0.84	−0.13	0.97	0	0.103	0.285	2.06
6-OH	4-CH_2SEt[a]	0.84	0.18	0.67	1	0.103	2.412	2.06
6-OH	4-COOMe	0.84	0.63	0.21	1	0.103	1.287	2.06
6-OH	H	0.79	1.12	−0.33	1	0.103	0.103	2.06
6-OH	4-COOEt	0.78	0.45	−0.33	1	0.103	1.747	2.06
6-OH	2-OH	0.76	1.22	−0.46	1	0.285	0.103	2.06
4,6-$(OH)_2$	4-OH	0.70	0.58	0.12	1	0.103	0.285	2.74
5,6,7-$(OMe)_3$	4-OMe[a]	0.70	−0.34	1.04	0	0.103	0.787	2.06
6-OH	4-NO_2	0.70	0.86	−0.16	1	0.103	0.736	2.06
6-OH	4-Cl	0.66	0.91	−0.25	1	0.103	0.603	2.06
6-OH	4-SMe	0.60	0.60	0.01	1	0.103	1.380	2.06
6-OH	4-$CONMe_2$	0.60	0.39	0.21	1	0.103	1.880	2.06
6-OH	4-$CONH_2$	0.59	0.76	−0.17	1	0.103	0.981	2.06
6-OH	4-$CHMe_2$	0.48	0.55	−0.07	1	0.103	1.496	2.06
6-C≡CH	4-OH	0.46	−0.13	0.59	0	0.103	0.285	2.06
7-OH	4-OH	0.30	−0.13	0.43	0	0.103	0.285	2.06
6-OH	4-CONHMe	0.20	0.56	−0.36	1	0.103	1.457	2.06
6-OH	4-Et	0.08	0.74	−0.66	1	0.103	1.030	2.06
6-OH	4-COOH[a]	0.08	0.88	−0.80	1	0.103	0.693	2.06
6-OH	4-C_6H_5	0.04	0.12	−0.08	1	0.103	2.536	2.06
5,6-$(OH)_2$	4-OH	0.00	−0.13	0.13	1	0.103	0.285	2.06
6-OH	4-C_4H_9	0.00	0.36	−0.36	1	0.103	1.959	2.06
6-OMe	4-CH_2OH	0.00	−0.31	0.31	0	0.103	0.719	2.06
6-COOMe	4-OH	0.00	−0.13	0.13	0	0.103	0.285	2.06
6-$CONH_2$	4-OH	0.00	−0.13	0.13	0	0.103	0.285	2.06
6-PO_3Et_2	4-OH	0.00	−0.13	0.13	0	0.103	0.285	2.06
6-OH	4-PO_3Et_2	0.00	−0.11	0.11	1	0.103	3.116	2.06
6-OMe	4-OH	−0.10	−0.13	0.03	0	0.103	0.285	2.06
6-OH	4-CF_3[a]	−0.10	0.95	−1.05	1	0.103	0.502	2.06
6-COMe	4-OH	−0.10	−0.13	0.03	0	0.103	0.285	2.06
H	4-OMe	−0.22	−0.34	0.11	0	0.103	0.787	2.06
6-Cl	4-OH	−0.22	−0.13	−0.09	0	0.103	0.285	2.06
5,7-Me_2,6-OH	4-OH	−0.30	−0.13	−0.17	0	0.103	0.285	2.06
6-NMe_2	4-OH	−0.40	−0.13	−0.27	0	0.103	0.285	2.06
H	4-OH	−0.52	−0.13	−0.39	0	0.103	0.285	2.06
H	H	−0.70	−0.06	−0.64	0	0.103	0.103	2.06
6-OMe	4-OMe	−0.70	−0.34	−0.36	0	0.103	0.787	2.06
4-OH	4-OH	−0.70	−0.60	−0.10	0	0.103	0.285	2.74
4,7-Me_2,6-OH	4-OH	−0.70	−0.68	0.02	0	0.103	0.285	2.87

[a] Data points not used in deriving equation.

Table 18. Growth Inhibition of MCF-7 Cells by 2-Phenylbenzothiophenes[63]

substituents		log 1/C						
X	Y	obsd	calcd from eq 18	Δ	ClogP	*I*	MR-*y*,4	σ^+,*y*
6-OH	3-F-4-OH	9.52	8.72	0.81	6.19	1	0.29	−0.58
6-OH	2-Me	9.16	8.82	0.33	6.81	1	0.10	−0.31
6-OH	4-C≡CH	9.09	8.18	0.99	6.88	1	0.96	0.18
6-OH	4-Cl	9.00	8.37	0.63	7.33	1	0.60	0.11
6-OH	3-Me-4-OH	9.00	8.89	0.11	6.46	1	0.29	−0.99
6-OH	2-Me-4-OH	8.70	8.89	−0.20	6.16	1	0.29	−1.23
6-OH	2-OMe-4-OH	8.70	8.66	0.04	5.49	1	0.29	−1.70
6-OH	4-F	8.64	8.76	−0.12	6.76	1	0.09	−0.07
6-OH	3-Cl-4-OH	8.64	8.78	−0.14	6.63	1	0.29	−0.55
6-OH	H	8.60	8.73	−0.13	6.61	1	0.10	0.00
6-OH	3-F[a]	8.60	0.24	−8.36	6.76	1	0.10	0.34
5-F-6-OH	4-OH	8.52	8.71	−0.19	5.93	1	0.29	−0.92
6-OH	3-OH	8.49	8.54	−0.05	5.96	1	0.10	0.12
6-OH	4-Et	8.30	8.19	0.11	7.64	1	1.03	−0.30
6-OH	4-CH=CH_2	8.16	8.16	0.00	7.34	1	1.10	−0.16
6-OH	2-OH	8.00	8.69	−0.69	5.69	1	0.10	−0.92
6-OH	4-C_4H_9	8.00	7.44	0.55	8.70	1	1.96	−0.29
6-C≡CH	4-OH	7.70	6.93	0.77	6.65	0	0.29	−0.92
6-OH	4-$CONMe_2$[a]	7.70	6.81	0.88	5.08	1	1.88	0.36
6-OH	4-$CHMe_2$	7.52	7.84	−0.32	8.04	1	1.50	−0.28
6-COOMe	4-OH	7.52	6.91	0.61	6.41	0	0.29	−0.92
6-OH	4-COMe	7.49	7.89	−0.39	6.06	1	1.12	0.50
H	4-OH	7.46	6.90	0.56	6.38	0	0.29	−0.92
6-OH	4-CONHMe	7.39	7.29	0.11	5.35	1	1.46	0.36
6-OH	4-Me[a]	7.30	0.09	−7.21	7.11	1	0.57	−0.31
6-OH	4-COOMe	7.30	7.91	−0.61	6.59	1	1.29	0.49
6-OH	4-COOEt	7.30	7.63	−0.33	7.12	1	1.75	0.48
6-COMe	4-OH	7.22	6.74	0.48	5.90	0	0.29	−0.92
H	4-OMe	7.00	6.59	0.41	6.95	0	0.79	−0.78
5-OH	4-OH	7.00	0.65	6.35	5.96	1	0.29	−0.92
4,7-Me_2-6-OH	4-OH	7.00	6.93	0.07	6.96	0	0.29	−0.92
6-OH	4-C_6H_5	7.00	7.12	−0.12	8.50	1	2.54	−0.18
6-OH	4-CH_2SEt	7.00	7.44	−0.44	7.85	1	2.41	−0.60
6-OH	3,5-Me_2-4-OH[a]	7.00	0.49	6.51	6.96	1	0.29	−1.06
4-OH	4-OH	6.72	6.62	0.10	5.96	0	0.29	−0.92
6-PO_3Et_2	4-OH	6.70	6.72	−0.02	5.86	0	0.29	−0.92
6-OH	4-$CONH_2$	6.70	7.38	−0.68	5.14	1	0.98	0.36
6-OH	4-PO_3Et_2	6.68	6.73	−0.05	6.08	1	3.12	0.54
6-OMe	4-OH	6.60	6.90	−0.29	6.36	0	0.29	−0.92
H	H	6.52	6.76	−0.24	7.02	0	0.10	0.00
6-OMe	4-OMe	6.52	6.59	−0.07	6.94	0	0.79	−0.78
6-Me	4-OH	6.52	6.93	−0.40	6.88	0	0.29	−0.92
7-OH	4-OH	6.52	6.77	−0.24	5.96	0	0.29	−0.92
6-OH	4-COOH	6.49	6.52	−0.03	4.09	1	0.69	0.42
4,6-$(OH)_2$	4-OH	6.46	0.81	5.64	5.40	1	0.29	−0.92
5,6,7-$(OMe)_3$	4-OMe	6.46	6.51	−0.05	5.74	0	0.79	−0.78
5,6-$(OH)_2$	4-OH	6.39	6.46	−0.07	5.47	1	0.29	−0.92
5,7-Me_2-6-OH	4-OH	6.30	6.92	−0.62	6.96	0	0.29	−0.92
6-OH	4-NO_2[a]	6.30	−0.17	6.47	6.36	1	0.74	0.79
6-OMe	4-CH_2OH	6.22	6.27	−0.05	5.97	0	0.72	−0.04
6-OH	4-OMe[a]	6.00	0.21	5.79	6.54	1	0.79	−0.78
6-Cl	4-OH	6.00	6.91	−0.91	7.11	0	0.29	−0.92
6-OH	4-CF_3[a]	6.00	−0.21	6.21	7.50	1	0.50	0.61
6-$CONH_2$	4-OH	6.00	6.05	−0.05	4.98	0	0.29	−0.92

[a] Data points not used in deriving equation.

5-OH ($I = 1$) without ortho substituents larger than *F*. From this analysis, the OH group at the 5- or 6-position mimics the 3-OH of the estradiol molecule in receptor binding, a fact that is confirmed by the recent crystal structures of these ligands bound to the estrogen receptor.[26] Ortho substituents larger than *F* will have a detrimental effect on receptor−ligand binding as seen in eq 9. Sterically demanding groups are preferred at the ortho position of the 2-phenyl substituent, indicated by a positive correlation with MR-*y*,2. This is reminiscent of the conformational restriction effect that appeared with related substituents in the indenes (eq 13) and indoles (eq 15). Groups on the para position of the 2-phenyl ring, however, lower the receptor binding affinity because of steric effects.

We obtained eq 18 by analysis of the MCF-7 cell growth inhibition data (Table 18). It is similar to eq 17 in that there is a negative steric effect with para substitution on the 2-phenyl ring. It is interesting to note that there is a small negative correlation with the σ^+ value of the Y substituent in this equation that we did not see in the receptor binding data. The more favorable growth promoting potency of compounds with electron-donating groups at this site could be the result of increasing the polarization of the carbonyl group. However, the crystal structure of the estrogen receptor−Raloxifene complex does not show any particular polar interaction with this carbonyl group,[26] and it can be replaced by other functions that link the rings together, such as an ether.[64] There is a bilinear hydrophobic term (ClogP) in this equation, which may arise from a pharmacokinetic factor related to the penetration of these compounds through cell membranes.

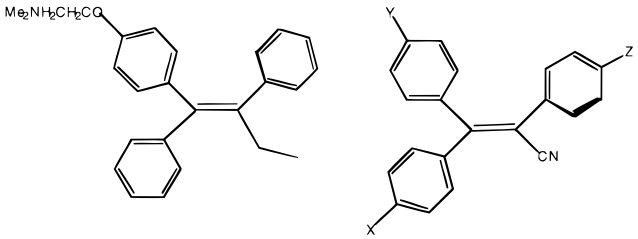
Table 19. Relative Binding Affinity of Triphylacrylonitriles at 0 °C[65]

substituents			log RBA					
X	Y	Z	obsd	calcd from eq 19	Δ	B1-*z*	*L*,*y*	*I*,OH
4-OH	4-OH	H	2.10	1.85	0.25	1.00	8.02	1
4-OH	4-OH	4-Me	1.69	1.73	−0.04	1.52	2.74	1
4-OH	4-OH	4-OH	1.61	1.49	0.12	1.35	2.74	1
4-OH	H	H	1.61	0.88	0.73	1.00	2.06	1
4-OH	4-Me	4-OH	1.56	1.51	0.05	1.35	2.87	1
H	H	4-OH	1.46	1.38	0.08	1.35	2.06	1
4-OH	4-Me	H	1.46	1.02	0.44	1.00	2.87	1
4-OH	4-OH	H	1.45	1.49	−0.04	1.00	2.74	1
4-Me	4-OH	4-OH	1.45	0.99	0.46	1.35	2.74	1
4-OH	4-OH	4-OH	1.43	1.38	0.05	1.35	2.06	1
4-OH	4-OMe	H	1.40	1.19	0.21	1.00	3.98	1
H	4-OH	4-OH	1.28	1.49	−0.21	1.35	2.74	1
4-Me	4-OH	H	0.91	0.99	−0.08	1.00	2.74	1
4-OH	4-$OCHMe_2$	H	0.90	1.33	−0.43	1.00	4.80	1
4-OMe	4-OH	H	0.78	0.99	−0.21	1.00	2.74	1
H	4-OH	H	0.57	0.99	−0.42	1.00	2.74	1
4-$OCH_2CH_2NEt_2$	4-OH	H	0.56	0.99	−0.43	1.00	2.74	1
4-$OCHMe_2$	4-OH	H	0.48	0.99	−0.51	1.00	2.74	1
4-$OCH_2CH_2NEt_2$	4-$OCH_2CH_2NEt_2$	H	−0.38	−0.28	−0.10	1.00	8.02	0
H	H	4-$OCH_2CH_2NEt_2$	−0.77	−0.94	0.17	1.35	2.06	0
4-OMe	4-OMe	H	−0.77	−0.76	−0.01	1.00	3.98	0
4-NMe_2	4-NMe_2	H	−0.92	−1.01	0.09	1.00	3.53	0
H	H	H	−1.40	−1.25	−0.15	1.00	2.06	0
4-$OCHMe_2$	4-$OCHMe_2$	H[a]	−2.00	−0.80	−1.20	1.00	4.80	0

[a] Data point omitted in deriving QSAR.

E. Triphenylacrylonitrile Derivatives (Tables 19 and 20)

This class of compounds is structurally related to Tamoxifen, a known agent for the treatment and prevention of breast cancer. The relative binding affinities of these ligands with calf uterine estrogen receptors are summarized in Tables 19 and 20.[65] Equations 19 and 20 were derived for data obtained at 0 °C for 2 h and 25 °C for 5 h, respectively. *I*,OH is an indicator variable for compounds containing a phenolic OH.



Tamoxifen Triphenylacrylonitriles

relative binding affinity of triphenylacrylonitriles at 0 °C for 2 h

$$\log \text{RBA} = 1.41(\pm 0.85)\text{B1-}z + 0.16(\pm 0.10)L,y + 2.13(\pm 0.36)I,\text{OH} - 3.00(\pm 1.14) \quad (19)$$

$$n = 23 \quad r^2 = 0.91 \quad s = 0.328 \quad F_{1,21} = 90.5(I,\text{OH})$$
$$F_{1,20} = 4.20(L,y) \quad F_{1,19} = 12(\text{B1-}z)$$

relative binding affinity of triphenylacrylonitriles at 25 °C for 5 h

$$\log \text{RBA} = -0.81(\pm 0.45)\text{ClogP} + 0.30(\pm 0.18)L,y + 1.74(\pm 0.76)I,\text{OH} + 1.77(\pm 2.19) \quad (20)$$

$$n = 23 \quad r^2 = 0.86 \quad s = 0.576$$
$$F_{1,21} = 54.8(I,\text{OH}) \quad F_{1,20} = 3.16(\text{ClogP})$$
$$F_{1,19} = 12.8(L,y)$$

$F_{1,20}$(ClogP) is not significant at 0.95 level

Equation 19 reveals that substituents on the Y and Z positions favor receptor binding and that phenolic hydroxy groups are the most important factor for the receptor binding. These equations indicate no positive contribution from substituent hydrophobic interactions. As seen in the cases of estradiol derivatives, different QSAR results were obtained with relative binding affinity data acquired at the different incubation temperatures: eq 20 was derived from the data obtained at 25 °C for an incubation time of 5 h. Surprisingly, a negative hydrophobic (ClogP) term appears, indicating that there is no hydrophobic contribution to the binding with the substituents examined; most of the substituents, however, are relatively polar and would not provide much opportunity for enhanced hydrophobic binding. Neither eq 19 nor 20 contains electronic terms.

F. Increase in the Proliferation of MCF-7 Cells (EC_{50}) with Triphenylacrylonitriles (Table 21)

Triphenylacrylonitriles investigated by Bignon et al.,[65] like Tamoxifen, are partial agonist/antagonists. They can stimulate the proliferation of estrogen receptor positive cells such as the MCF-7 cell line and also partially inhibit estradiol-induced cell proliferation. Equation 21 is derived for the agonistic activity of triphenylacrylonitriles in the stimulation of MCF-7 cell proliferation (Table 21). This equation is very similar to eq 19, with the indicator parameter *I*,OH, for the presence of a hydroxyl group on two of the phenyl groups, as the most important factor. These results also indicate that, in this series of compounds, the agonistic activity (eq 21) is parallel to receptor binding affinity (eqs 19 and 20). However, again there

Table 20. Relative Binding Affinity of Triphenylarylonitriles at 25 °C[65]

substituents			log RBA					
X	Y	Z	obsd	calcd from eq 20	Δ	ClogP	*L*,*y*	*I*,OH
4-OH	4-OH	4-OH	2.22	2.14	0.08	2.74	2.74	1
4-OH	4-$OCH_2CH_2NEt_2$	H	2.03	2.01	0.02	4.89	8.02	1
4-OH	4-OH	4-Me	1.97	1.19	0.78	3.91	2.74	1
4-OH	4-Me	4-OH	1.89	1.23	0.66	3.91	2.87	1
4-OH	H	4-OH	1.87	1.39	0.48	3.41	2.06	1
4-OH	4-OH	H	1.79	1.60	0.19	3.41	2.74	1
4-OH	H	H	1.56	0.85	0.71	4.08	2.06	1
4-OH	4-Me	H	1.45	0.70	0.75	4.58	2.87	1
4-OH	4-OMe	H	1.23	1.50	−0.27	4.00	3.98	1
4-Me	4-OH	4-OH	0.96	1.19	−0.23	3.91	2.74	1
4-OH	4-$OCHMe_2$	H	0.81	1.08	−0.27	4.84	4.80	1
H	4-OH	4-OH	0.78	1.60	−0.82	3.41	2.74	1
4-$OCH_2CH_2NEt_2$	4-OH	H	0.53	0.41	0.12	4.89	2.74	1
H	H	4-OH	0.52	0.85	−0.33	4.08	2.06	1
4-Me	4-OH	H	0.40	0.66	−0.26	4.58	2.74	1
H	4-OH	H	0.34	1.06	−0.72	4.08	2.74	1
4-OMe	4-OH	H[a]	−0.18	1.13	−1.31	4.00	2.74	1
4-$OCH_2CH_2NEt_2$	4-$OCH_2CH_2NEt_2$	H	−0.40	−0.93	0.53	6.38	8.02	0
4-$OCHMe_2$	4-OH	H	−0.44	0.45	−0.89	4.83	2.74	1
H	H	H	−1.05	−1.42	0.37	4.75	2.06	0
4-NMe_2	4-NMe_2	H	−1.40	−0.71	−0.69	5.14	3.53	0
4-OMe	4-OMe	H	−1.40	−1.29	−0.11	4.58	3.98	0
4-$OCHMe_2$	4-$OCHMe_2$	H	−2.00	−1.81	−0.19	6.26	4.80	0
H	H	4-$OCH_2CH_2NEt_2$	−2.00	−2.08	0.08	5.56	2.06	0

[a] Data point omitted in deriving QSAR.

Table 21. Triphenylarylonitrile-Induced Increase of Proliferation of MCF_7 Cells (EC_{50})[65]

substituents			log 1/C					
X	Y	Z	obsd	calcd from eq 21	Δ	*L*,*x*	*L*,*y*	*I*
4-OH	4-$OCH_2CH_2NEt_2$	H	10.70	10.61	0.09	2.74	8.02	1
4-$OCH_2CH_2NEt_2$	4-OH	H	10.52	10.56	−0.04	8.02	2.74	1
4-OH	4-OH	4-OH	10.44	9.79	0.65	2.74	2.74	1
4-OH	4-OH	H	10.23	9.79	0.44	2.74	2.74	1
4-OH	H	4-OH	10.20	9.68	0.52	2.74	2.06	1
H	4-OH	4-OH	10.04	9.69	0.35	2.06	2.74	1
4-OH	4-Me	4-OH	9.98	9.81	0.17	2.74	2.87	1
4-$OCHMe_2$	4-OH	H	9.92	10.09	−0.17	4.80	2.74	1
4-OH	4-OH	4-Me	9.85	9.79	0.06	2.74	2.74	1
4-Me	4-OH	4-OH	9.85	9.81	0.04	2.87	2.74	1
4-OH	4-$OCHMe_2$	H	9.75	10.11	−0.36	2.74	4.80	1
4-OH	H	H	9.72	9.68	0.04	2.74	2.06	1
4-OH	4-OMe	H	9.72	9.98	−0.26	2.74	3.98	1
4-OMe	4-OH	H	9.70	9.97	−0.27	3.98	2.74	1
4-OH	4-Me	H	9.55	9.81	−0.26	2.74	2.87	1
4-$OCH_2CH_2NEt_2$	4-$OCH_2CH_2NEt_2$	H	9.55	9.18	0.37	8.02	8.02	0
H	4-OH	H	9.40	9.69	−0.29	2.06	2.74	1
4-Me	4-OH	H	9.38	9.81	−0.43	2.87	2.74	1
H	H	4-OH	9.26	9.58	−0.32	2.06	2.06	1
4-$OCHMe_2$	H	H	8.30	8.21	0.09	4.80	4.80	0
4-NMe_2	4-NMe_2	H	7.92	7.82	0.10	3.53	3.53	0
H	H	4-$OCH_2CH_2NEt_2$	7.60	7.38	0.22	2.06	2.06	0
H	H	H	7.55	7.38	0.17	2.06	2.06	0
4-OMe	4-OMe	H	6.99	7.96	−0.97	3.98	3.98	0
estradiol			10.70					

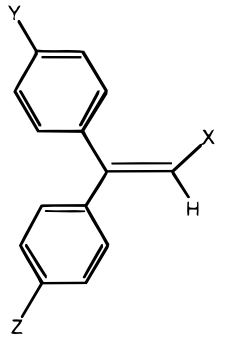
is no evidence for hydrophobic interactions, as was the case in eq 20.

$$\log 1/C = 0.15(\pm 0.12)L,x + 0.16(\pm 0.12)L,y + 2.20(\pm 0.40)I,\mathrm{OH} + 6.75(\pm 0.62) \quad (21)$$

$$n = 24 \quad r^2 = 0.87 \quad s = 0.386 \quad F_{1,22} = 54.8(I,\mathrm{OH})$$
$$F_{1,21} = 13.6(L,y) \quad F_{1,20} = 7(L,x)$$

G. 1,1-Diphenylethylenes (Table 22)

Gilbert et al.[66] reported the relative binding affinity of 1,1-diphenylethylene derivatives with estrogen receptors in lamb uterus cytosol at 25 °C. Equation 22 was derived from their data. *I* is an indicator parameter (*I* = 1 for *Z*-isomer and 0 for *E*-isomer). The phenolic OH is either on the Y position or on Z position. The size of the X group (MR,*x*) has a negative contribution to receptor binding.



1, 1-Diphenylethylenes

$$\log \mathrm{RBA} = 0.63(\pm 0.26)I - 0.64(\pm 0.21)\mathrm{MR},x + 0.36(\pm 0.52) \quad (22)$$

$$n = 13 \quad r^2 = 0.86 \quad s = 0.204$$
$$F_{1,11} = 8.7(\mathrm{MR},x) \quad F_{1,10} = 28.6(I)$$

Table 22. Relative Binding Affinity of 1,1-Diphenylethylenes at 25 °C[66]

substituents			log RBA				
X	Y	Z	obsd	calcd from eq 22	Δ	*I*	MR,*x*
$CHMe_2$	OH	$OCH_2CH_2NEt_2(Z)$	0.00	0.03	−0.03	1	1.50
C_4H_9	OH	$OCH_2CH_2NEt_2(Z)$	−0.13	−0.26	0.14	1	1.96
$CH_2C_6H_5$	OH	$OCH_2CH_2NEt_2(Z)$	−0.59	−0.93	0.34	1	3.00
C_4H_9	$OCH_2CH_2NEt_2$	OH(*E*)	−0.68	−0.89	0.21	0	1.96
C_5H_{11}	OH	$OCH_2CH_2NEt_2(Z)$	−0.68	−0.56	−0.12	1	2.42
C_5H_{11}	OH	$OCH_2CH_2NMe_2(Z)$	−0.70	−0.56	−0.14	1	2.42
$CH_2C_6H_5$	$OCH_2CH_2NEt_2$	OH(*E*)[a]	−0.82	−1.55	0.73	0	3.00
$CHMe_2$	$OCH_2CH_2NEt_2$	OH(*E*)	−0.89	−0.59	−0.29	0	1.50
C_6H_{13}	OH	$OCH_2CH_2NEt_2(Z)$	−0.89	−0.86	−0.03	1	2.89
C_5H_{11}	$OCH_2CH_2NEt_2$	OH(*E*)	−1.00	−1.18	0.18	0	2.42
C_5H_{11}	$C_4H_8N(CH_2)_2O$	OH(*E*)	−1.10	−1.18	0.09	0	2.42
C_5H_{11}	$OCH_2CH_2NMe_2$	OH(*E*)	−1.16	−1.18	0.03	0	2.42
C_8H_{17}	$OCH_2CH_2NEt_2$	OH(*E*)[a]	−1.46	−2.07	0.62	0	3.82
C_8H_{17}	OH	$OCH_2CH_2NEt_2(Z)$	−1.60	−1.45	−0.16	1	3.82
C_6H_{13}	$OCH_2CH_2NEt_2$	OH(*E*)	−1.70	−1.48	−0.22	0	2.89

[a] Data point omitted in deriving QSAR.

Table 23. Antiestrogenicity of 1,1-Diphenylethylenes in MVLN Cells[66]

substituents			log 1/C					
X	Y	Z	obsd	calcd from eq 23	Δ	MR,*x*	MR,*z*	*I*
$CHMe_2$	OH	$OCH_2CH_2NEt_2(Z)$	6.30	6.13	0.17	1.50	3.39	1
C_4H_9	OH	$OCH_2CH_2NEt_2(Z)$[a]	6.30	6.04	0.26	1.96	0.29	1
C_5H_{11}	OH	$OCH_2CH_2NEt_2(Z)$	6.05	5.95	0.10	2.42	3.39	1
$CH_2C_6H_5$	OH	$OCH_2CH_2NEt_2(Z)$	5.89	5.83	0.06	3.00	0.29	1
C_4H_9	$OCH_2CH_2NEt_2$	OH(*E*)	5.82	6.04	−0.21	2.42	3.39	0
C_5H_{11}	OH	$OCH_2CH_2NMe_2(Z)$	5.82	5.76	0.07	1.96	0.29	1
$CH_2C_6H_5$	$OCH_2CH_2NEt_2$	OH(*E*)	5.77	5.85	−0.08	1.50	0.29	0
$CHMe_2$	$OCH_2CH_2NEt_2$	OH(*E*)	5.77	5.55	0.22	3.00	3.39	0
C_5H_{11}	$OCH_2CH_2NMe_2$	OH(*E*)	5.75	5.66	0.08	2.42	0.29	0
C_6H_{13}	OH	$OCH_2CH_2NEt_2(Z)$	5.72	5.85	−0.13	2.89	2.47	1
C_6H_{13}	$OCH_2CH_2NEt_2$	OH(*E*)	5.60	5.57	0.03	2.89	0.29	0
$CHMe_2$	$OCH_2CH_2NPr_2$	$OCH_2CH_2NPr_2(E)$	5.51	5.46	0.05	1.50	3.39	0
C_5H_{11}	$OCH_2CH_2NEt_2$	OH(*E*)	5.50	5.66	−0.17	2.42	0.29	0
C_5H_{11}	$OCH_2CH_2NC_4H_8$	OH(*E*)	5.48	5.66	−0.18	2.42	3.39	0
C_8H_{17}	OH	$OCH_2CH_2NEt_2(Z)$	5.46	5.67	−0.21	3.82	0.29	1
C_8H_{17}	$OCH_2CH_2NEt_2$	OH(*E*)	5.33	5.39	−0.06	3.82	3.39	0
C_5H_{11}	$OCH_2CH_2NEt_2$	$OCH_2CH_2NEt_2(E)$	5.22	5.35	−0.13	2.42	0.29	0
$C_{14}H_{29}$	OH	$OCH_2CH_2NEt_2(Z)$	5.10	5.13	−0.04	6.49	3.39	1
$C_{14}H_{29}$	$OCH_2CH_2NEt_2$	OH(*E*)	5.02	4.85	0.17	6.49	4.23	0

[a] Data point omitted in deriving QSAR.

H. Antiestrogenicity of 1,1-Diphenylethylenes (Table 23)

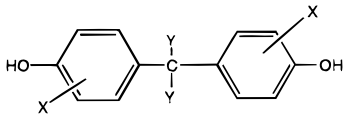
The antiestrogenicity of diphenylethylene derivatives was tested in MVLN cells.[66] The IC_{50} values were expressed as the concentration of test compounds leading to 50% inhibition of the luciferase activity induced by 0.1 nM estradiol. Equation 23 was derived for the antiestrogenicity data. Surprisingly, eq 23 is similar to eq 22. The *Z*-isomer favors the receptor binding, and the size of substituents at X and Z positions decreases the binding. Similar to eq 21, we did not see a positive hydrophobic term as seen in eq 18.

$$\log 1/C = 0.59(\pm 0.25)I - 0.20(\pm 0.06)\mathrm{MR},x - 0.10(\pm 0.08)\mathrm{MR},z + 6.18(\pm 0.22) \quad (23)$$

$$n = 19 \quad r^2 = 0.82 \quad s = 0.164 \quad F_{1,17} = 15.7(\mathrm{MR},x)$$
$$F_{1,16} = 15(I) \quad F_{1,15} = 7.7(\mathrm{MR},z)$$

I. Estrogenic Activity of 4,4′-Dihydroxydiphenylmethanes in Rats

Campbell[67] reported the estrogenic activity of the derivatives of 4,4′-dihydroxydiphenylmethane in rats (Table 24). The estrogenic activity is defined as an arbitrary rate unit (RBR), which is the minimum total dose required to give 100% estrus response in ovariectomized female rats when injected in six doses of the compound as a sesame oil solution over 3 days. In eq 24, RBR is the specific activity expressed as units/mol. *L*-*y*,*b* and B1-*y*,*b* are the steric parameters for the larger of the two Y substituents on the methylene group that links the two phenols.



4,4'-dihydroxydiphenylmethanes

$$\log \mathrm{RBR} = 0.52(\pm 0.20)\mathrm{ClogP} - 0.31(\pm 0.14)L\text{-}y,b - 1.55(\pm 0.73)\mathrm{B1}\text{-}y,b - 1.50(\pm 0.45)\sigma^*,y\text{-sum} + 0.88(\pm 1.20) \quad (24)$$

$$n = 31 \quad r^2 = 0.84 \quad s = 0.351$$
$$F_{1,29} = 35.77(\sigma^*,\mathrm{sum}) \quad F_{1,28} = 7.62(\mathrm{B1}\text{-}y,b)$$
$$F_{1,27} = 29.7(L\text{-}y,b) \quad F_{1,26} = 26.91(\mathrm{ClogP})$$

Equation 24 is similar to eqs 10 and 11 in terms of negative contributions from steric parameters and negative electronic terms. The positive hydrophobic parameter (ClogP), which may be related to the absorption and distribution processes (also in eq 18),

Table 24. Specific Estrogenic Activity of 4,4′-Dihydroxydiphenylmethanes (units/mol) in Rats[67]

substituents		log RBR						
X	Y	obsd	calcd from eq 24	Δ	ClogP	*L*-*y*,*b*	B1-*y*,*b*	σ*,*y*-sum
H	C_3H_7,C_2H_5	0.57	0.02	0.55	5.26	4.920	1.52	−0.21
H	C_3H_7, C_3H_7	0.25	0.32	−0.07	5.79	4.920	1.52	−0.23
H	i-C_4H_9,Me	−0.03	−0.18	0.15	5.13	4.920	1.52	−0.12
H	C_3H_7,Me	−0.11	−0.40	0.29	4.73	4.920	1.52	−0.11
H	Et,Et	−0.11	−0.01	−0.10	4.73	4.110	1.52	−0.20
H	C_4H_9,C_3H_7	−0.17	0.21	−0.38	6.32	6.170	1.52	−0.23
2-Me	(Et)2CH,H	−0.29	−0.38	0.09	6.16	4.720	2.13	−0.22
H	Et,Me	−0.38	−0.44	0.06	4.20	4.110	1.52	−0.10
2-Me	i-C_4H_9,H	−0.41	−0.67	0.26	5.63	4.920	1.52	0.38
H	$CH_2C_6H_5$,H	−0.46	−0.91	0.45	4.69	4.620	1.52	0.28
2-Me	$(C_3H_7)_2$CH,H	−0.52	−0.67	0.15	7.22	6.170	1.90	0.27
2-Me	C_3H_7,H	−0.78	−0.88	0.10	5.23	4.920	1.52	0.38
H	Me,Me	−0.80	−0.47	−0.33	3.67	2.870	1.52	0.00
H	C_4H_9,Me	−0.83	−0.49	−0.34	5.26	6.170	1.52	−0.13
2-Me	i-C_3H_7,H	−0.83	−1.16	0.33	5.10	4.110	1.90	0.30
2-Me	C_6H_{13},H	−0.85	−1.06	0.21	6.82	8.220	1.52	0.36
H	C_6H_5[a],H	−0.87	−3.13	2.26	4.31	6.280	1.71	1.09
3-Me	H,H	−1.00	−0.78	−0.22	3.87	2.060	1.00	0.98
H	C_4H_9,C_4H_9[a]	−1.02	0.53	−1.55	6.85	6.170	1.52	−0.26
2-Me	Me,H	−1.02	−0.95	−0.07	4.17	2.870	1.52	0.49
2-Me	Et,H	−1.15	−0.90	−0.25	4.70	4.110	1.52	0.38
H	$(C_6H_5)_2$CH,H	−1.62	−0.79	−0.83	6.13	5.150	2.01	0.08
H	C_6H_5,Me	−1.86	−2.19	0.33	4.71	6.280	1.71	0.60
3-Me	C_3H_7,C_3H_7	1.20	0.84	0.36	6.78	4.920	1.52	−0.23
3-Me	C_3H_7,Et	1.13	0.54	0.59	6.26	4.920	1.52	−0.21
3-Me	Et,Et	0.55	0.50	0.05	5.73	4.110	1.52	−0.20
3-Me	C_4H_9,C_3H_7	0.49	0.72	−0.23	7.32	6.170	1.52	−0.23
3-Me	C_3H_7, Me	−0.06	0.11	−0.17	5.73	4.920	1.52	−0.11
3-Me	Me,i-C_4H_9	−0.22	0.33	−0.55	6.13	4.920	1.52	−0.12
3-Me	Me,Me[a]	−0.81	0.04	−0.85	4.67	2.870	1.52	0.00
3-Me	Me,Et[a]	−0.83	0.08	−0.91	5.20	4.110	1.52	−0.10
3-Me	Me,C_4H_9[a]	−0.87	0.03	−0.90	6.26	6.170	1.52	−0.13
3-Me	H,i-C_4H_9	−1.00	−1.11	0.11	4.73	4.920	1.52	0.36
3-Me	C_4H_9,C_4H_9[a]	−1.23	1.04	−2.27	7.85	6.170	1.52	−0.26
3-Me	Et,H	−1.50	−1.36	−0.14	3.80	4.110	1.52	0.38
3-Me	Me,C_6H_5	−1.72	−1.68	−0.04	5.71	6.280	1.71	0.60
3-Me	H,C_6H_{13}	−1.89	−1.52	−0.37	5.92	8.220	1.52	0.36

[a] Data points not used in deriving QSAR.

Table 25. Relative Proliferation Rate of MCF-7 Cells by XC_6H_4OH[73]

	log RBR				
substituents	obsd	calcd from eq 25	Δ	ClogP	σ^+
H	0.00	−0.04	0.04	1.48	0.00
4-C_2H_3	0.11	0.20	0.09	2.50	−0.30
4-C_3H_7	0.30	0.49	0.19	3.02	−0.29
4-CH(Me)C_2H_5	0.73	0.65	0.08	3.43	−0.29
4-CMe_3	0.71	0.65	0.06	3.30	−0.26
4-C(Me)$_2C_2H_5$	0.85	0.78	0.07	3.83	−0.29
4-$CH_2CH_2CHMe_2$	0.81	0.80	0.01	3.96	−0.30
4-OC_4H_9	−0.16	−0.26	0.10	3.16	−0.81
4-OC_6H_{13}	0.00	0.08	0.08	4.22	−0.81
4-C_6H_5	0.78	0.79	0.01	3.36	−0.18
4-C_9H_{19}	0.83	0.84	0.01	6.21	−0.31
4-C_6H_4-4-OH[a]	0.47			2.70	unknown

[a] Data point not utilized because of lack of σ^+ value. Activity corrected for presence of 2-OH groups.

is an important finding. The parameter σ* models the field/inductive effect of substituents.[41]

J. MCF-7 Cell Proliferation by Simple Phenols in Cells Free of Natural Estrogens (Table 25)

Equation 21 for cell proliferation by the large triphenylacrylonitriles shows no dependence on hydrophobic interactions. Presumably this is because hydrophobic space is pre-empted by the large parent structures so that substituents have little opportunity to show hydrophobic interactions. In eq 24, the dependence upon log *P* may be associated with drug disposition in the rat and/or receptor binding. However, there are studies on small phenols that provide insight[73,74] on the hydrophobic nature of the ER.

Equation 25 derived from data of Soto et al.[73] comes from cells first treated with charcoal to remove natural estrogens because the much more weakly binding phenols would not be effective in displacing them. They reported relative increases in cell proliferation upon treatment with a standard dose of 10 μM phenol, from which eq 25 was formulated.

$$\log \text{RBR} = 1.54(\pm 0.37)\sigma^+ + 1.11(\pm 0.33)\text{ClogP} - 0.11(\pm 0.04)(\text{ClogP})^2 - 1.39(\pm 0.39) \quad (25)$$

$$n = 11 \quad r^2 = 0.953 \quad s = 0.103 \quad F_{1,9} = 2.53(\sigma^+)$$

$$F_{1,8} = 11.82(\text{ClogP}) \quad F_{1,7} = 39.85((\text{ClogP})^2)$$

$$\text{optimum log } P = 5.1(4.7-5.9)$$

Hydrophobicity plays an important positive role up to ClogP of about 5, after which the role is negative. This is most apparent with nonylphenol. If this point is omitted, another good correlation can be found with

Table 26. Relative Proliferation Rate of MCF-7 Cells by 4-HOC$_6$H$_4$C(X)(Y)C$_6$H$_4$-4′-OH[74]

substituents		log RBR			
		obsd	calcd from eq 26	Δ	ClogP
X	Y				
H	H	−3.00	−3.05	0.05	2.88
Me	Me	−2.00	−2.23	0.23	3.67
Me	H	−3.00	−2.64	0.36	3.27
C_2H_5	H	−2.00	−2.09	0.09	3.80
C_2H_5	Me	−2.00	−1.68	0.32	4.20
C_2H_5	C_2H_5	−1.00	−1.14	0.14	4.73
C_3H_7	C_3H_7	0.00	−0.05	0.05	5.79
CF_3	CF_3[a]	−2.00	−3.44	1.44	2.49
CH_2OH	Me	−4.00	−4.12	0.12	1.84

[a] Data point not used in deriving equation.

ClogP and σ^+, with coefficients of 0.5 and 1.54, respectively, and $r^2 = 0.917$. Nonylphenol is known to be an active estrogen. Hence, it seems likely that if the dose had been varied to attain a standard response, the nonylphenol would have appeared to be more active. Nonspecific cell toxicity is a linear function of log P up to the range of log $P = 5-6$, so that testing lower concentrations of nonylphenol would have mitigated toxicity. Most interesting is the positive σ^+. This means that phenols with negative σ^+ values would be toxic to the cells and, hence, show little proliferation even though they possess the requisite log P. This becomes very interesting in light of eq 27, which correlates the toxicity of phenols to leukemia cells. The σ^+ terms are essentially the same except for opposite signs in eqs 25 and 27. The positive σ^+ term in eq 25 inhibits the cell toxicity that is shown so clearly by 4-OC$_4$H$_9$ and 4-OC$_6$H$_{13}$, which have good log P values but negative σ^+ values. Thus, we believe that σ^+ and (log P)^2 in eq 25 are to be associated with toxicity.

K. MCF-7 Cell Proliferation by 4-HOC$_6$H$_4$C(X)(Y)-C$_6$H$_4$-4′-OH (Table 26)[74]

In this study, RBR represents a relative proliferation rate using cells treated with charcoal to remove the natural estrogens.

$$\log \mathrm{RBR} = 1.03(\pm 0.18)\mathrm{ClogP} - 6.00(\pm 0.72) \qquad (26)$$
$$n = 8 \quad r^2 = 0.97 \quad s = 0.234 \quad F_{1,6} = 191.73$$

In these examples, the substituents X and Y are insulated from the aromatic rings so that no electronic term appears. The strong electron-attracting CF_3 function may be an exception causing its poor fit to the model. Equation 26 is the clearest definition of the hydrophobic character of the ER receptor. Obviously there is an advantage in using cells without a strong binding estrogen in place. Displacing such an estrogen with others is a slow process that may not reach equilibrium in the time allotted for the experiment. The fact that QSAR obtained at two different temperatures yield somewhat different QSAR values (eqs 2, 3; 5, 6; 19, 20) illustrates this point. Equations 25 and 26 suggest that more interesting information on the relationship of binding by smaller estrogens and their role in stimulating cell growth as well as cell toxicity (eq 27) can be obtained with cells free of natural estrogens. Using such a system, one can use all sorts of probes to obtain, via QSAR, a much better definition of the nature of the receptor. All sorts of chemicals obviously unrelated to the natural estrogens have been found to bind and/or cause estrogenic-like effects. From a mechanistic point of view, sets of similar compounds must be studied before one attempts to lump many compounds into a single QSAR.[33] Simply obtaining a statistically valid QSAR does not mean that one can obtain mechanistic insight from it.

IV. Discussion

The estrogen receptor belongs to the steroid/thyroid nuclear hormone receptor superfamily, a group of structurally related, multidomain proteins that act principally as ligand-modulated transcription factors that regulate the expression of specific genes in response to hormonal signals.[68] After the hormonal ligand binds to the receptor, the ligand-binding domain undergoes a conformation change that results in enhanced binding to DNA target sequences and the recruitment of aggregates of coactivator or corepressor proteins that ultimately result in the activation or repression of these genes (for a more complete description, see Introduction).

For a compound to have an estrogenic or antiestrogenic activity through this complex pathway, however, it must first bind to the estrogen receptor protein. There are two estrogen receptors, ERα and ERβ,[14,16] and they have different tissue distributions and somewhat different amino acid sequences in their ligand-binding domains.[19] Thus, these two ER subtypes have somewhat different ligand-binding characteristics and gene-activating activity,[69] although much less is known about ERβ than ERα.

Because estrogens can act through different ER subtypes, and the ligand−ER complex can utilize different genes, a variety of different response elements, and in different cells, varied levels of different coregulatory proteins, it is not surprising that the pharmacology of estrogenic compounds is complex.[70] A compound may be an antagonist on one gene in one tissue, but an agonist on another gene in another tissue. The complexity of these interactions and the differences in the components involved in mediating the action of estrogens in different tissues and cells and at different genes, is used to explain why Tamoxifen, which acts as an antagonist by blocking the growth of some breast cancer cells, can still act like an estrogen in maintaining bone mineral density and in stimulating the uterus, the latter activity contributing to the increased risk of uterine cancer that is found with prolonged Tamoxifen therapy.[71] To complicate matters further, eqs 25 and 27 show that toxicity must also be factored into the picture. Despite these complexities, however, the signal event in the biological action of estrogens is their binding to the estrogen receptor.

From our analysis of the quantitative structure−binding affinity relationships of many types of estrogen receptor ligands with the ERα, we can draw several generalizations:

(a) A phenolic hydroxy group, which mimics the 3-OH on the A-ring of estradiol (and at the corresponding position of nonsteroidal estrogen ligands), appears repeatedly as the most important factor in receptor−ligand interactions.

(b) Substituents that increase the electron density on the phenolic ring also appear to increase binding affinity.

(c) There is, surprisingly, no consistent, positive hydrophobic interaction between ligand substituents and receptor (the same phenomenon has been shown in the CoMFA analysis by Tong et al.[72]). The lack of a consistent hydrophobic contribution from ligand substituents needs to be tempered by the recognition that the ABCD tetracyclic core structure of steroidal estrogens (as well as the corresponding units in nonsteroidal estrogens) is generally very hydrophobic and may contribute to the bulk of ligand binding by a hydrophobic mechanism.

(d) There is a hydrophobic interaction between ERα and substituents on the 11β-position of estradiol. It may be significant that substituents at the 11β-position of estradiol project into the largest unfilled pocket in the ERα ligand-binding domain that surrounds the ligand.[26,31]

(e) In many cases, it is the steric character of substituents, and in other cases their polar character, that results in reduced ligand-binding affinity.

(f) In certain nonsteroidal ligands, substituents that result in conformational restrictions, which in all cases result in enforced deviations from planarity, facilitate receptor−ligand binding.

(g) Positive correlations with potency and hydrophobicity appear frequently in the correlations between ligand structure and biological activity (in cell growth studies, rather than ligand-binding affinity). This is thought to be a reflection of hydrophobic facilitation of membrane penetration.

Finally, one of the reasons that motivated our undertaking this study was that we have discovered that the activity of *estrogenic phenols* such as 4-octylphenol, 4-nonylphenol, diethystilbestrol, estradiol, and estriol as inhibitors of the growth of L1210 leukemia cells is well fit by QSAR eq 27, which is derived from simple phenols.[33] This study has now been extended to the four components of Premarin. Except for estrone, which was too insoluble to test, the others are well fit by eq 27.

$$\log 1/C = -1.58\sigma^+ + 0.211 \log P + 3.10 \quad (27)$$
$$n = 23 \quad r^2 = 0.90 \quad s = 0.19$$
$$F_{1,21} = 30.8(\sigma^+) \quad F_{1,20} = 25.8(\log P)$$

Equation 27 holds only for substituents that have negative σ^+ values. With positive σ^+ values, a different type of toxicity is seen that depends only on log P.[33] The low coefficient with log P in this equation suggests that the receptor for the toxic interaction (which is not likely to be the ER) is not hydrophobic and that the log P term is likely associated with cell penetration. We now plan to test the effect of some of the ligands discussed in this report on growth of leukemia cells.

The problem of outliers is a tough one that has received considerable thought from statisticians.[75] In much of the building of mathematical models, the mechanism cannot be a means for identifying aberrant data points. In chemical−biological interactions, comparative QSAR can be of mechanistic help. We see four major causes for the misfit of datapoints: (1) The shape of the mathematical model may be flawed. (2) The parameters may be poorly designed or simply contain errors in some instances. (3) There may be experimental errors in the data. (4) Side reactions may occur. By side reactions, we mean reactions with receptors other than the principal one responsible for the measured biological response.

The problem of the quality of the parameters is serious, particularly that of log P and steric parameters. Because there are at least a half dozen commercial programs, of varying quality, for calculation of log P, few, indeed, take the time to experimentally determine new values. Of course, in analyzing data from the literature, the molecules are not normally available and, hence, experimental log P values are out of the question. From our experience, we believe that while ClogP may be in error in terms of absolute values, the numbers are surprisingly good in relative terms.[80] Hence, the error would largely be delegated to the intercepts.

Steric parameters are most difficult to define. Two approaches have been tried. We have attempted to use measured or calculated values for the various substituents. The problem with this approach is that not knowing the shape of the receptor site (except in rare examples[46,76]), the only guide is the empirical quality of the QSAR. A more elegant approach is that employed by CoMFA. Here one attempts, by trial and error, to place the members of a data set in a proper conformation from which steric interactions are estimated by exploring the outer surfaces of the set of "congeners". Since the ligands are fixed, the give in the system has to be the receptor wall. However, we know from many examples that steric effects are a linear function of the empirical parameters. That is, as substituent size increases, activity gradually falls. This more or less uniform decline in activity could be due to two effects: the receptor wall could give to some degree, or the positions of the ligands could gradually move, or possibly both mechanisms could operate. In our approach, we do not assume a perfectly uniform mode of binding. Thus, coefficients with steric terms may reflect the complex process of displacement of the ligand and/or the receptor wall.

Of course, we feel that mechanistic-based comparative QSAR will gradually provide more insight. Little such work has been attempted with the CoMFA approach, and as of the present, it does not seem promising.[79] Experimental errors in determining activity are common and not easy to discover. We have also found errors in the calculation of activities.

Probably the most important error is that of side reactions that plague all of organic chemistry. All too often in carrying out a "standard" reaction, one gets yields of less than 50%. The possibilities for side reactions are huge in biological work; even in a simple cell, thousands of reactions are occurring that

may be more or less affected by some of the compounds being tested. It is still surprising that one can get quite useful information studying cells or even mice with a set of 30−40, hopefully, congeners. However, we are very slowly learning how to separate two different types of reactions by different members of closely related chemicals.[33,77]

Some perspective on outliers can be gained from our present database. In the area of pure physical organic chemistry, where one would expect the best possible results, we are far from perfection. Of our 7170 QSAR from physical organic chemistry, 2561 have one or more outliers and 881 have two or more. Most authors of such studies have used fewer than 10 datapoints. Out of 7170 sets, 5400 are based on fewer than 11 data points. In our bio database of 5330 sets, 3543 have more than one outlier and 1170 have more than two. More conservative, better designed sets could do much to produce better QSAR, but synthetic chemists still give little thought to this problem. We believe that approaching the problem with a better understanding of receptor structure and mechanism will be most helpful.[33,46,76,77,81]

Finally, considerable work has been done using the 3-D CoMFA methodology.[82−84,87,88] It is not at all easy to compare this work with that reviewed by us. In the first place, CoMFA, although termed QSAR, does not qualify, as it is generally used, as quantitative SAR. It is qualitative, or at best, semiquantitative. Authors rarely attempt to discuss their results in terms of numbers, but use 3-D pictures. These pictures are not precise enough to be compared with other results, in part because the terms used to formulate a regression-based model are based on principal components. Such terms will have different compositions from data set to data set, so that comparison is only possible via pictures that are not easy to understand. This is not to say that CoMFA cannot provide insight, but it cannot provide the kind we are interested in for mechanistic comparisons.

Of course, hundreds of chemicals, one or a few at a time, have been tested for their ability to bind to the ER or effect some estrogenic-like activity of cells or animals. Many of these do not contain a phenolic OH as do most of the natural female estrogens. As yet, we do not know how many ways chemicals can bind to the ER or how many ways that they can effect cell proliferation. Until some sets of truly congeneric chemicals, but of basically different structures, have been studied with chemicals where the parent structure does not preempt most of the hydrophobic space, we do not think that it is profitable to compare these small studies with QSAR studies. For instance, in the data[74] we used to formulate eq 26, several compounds without phenolic OH groups were studied. They are much less active than the OH-containing compounds, and because they have no structural features in common, we cannot include them in our QSAR.

In conclusion, we believe that the results of this study, when taken with previous analyses,[33] provide new perspective on how estrogens may affect biological processes.

V. Acknowledgment

This research was supported by grants from the National Institutes of Health, 1R01 ES07595-01 (to C.H.) and 1R37 DK15556 (to J.A.K.).

VI. References

(1) American Chemical Society, Facts and Figures; *Cancer Statistics* **1998**. American Chemical Society, Atlanta, GA.
(2) Kimmick, G.; Muss, H. B. *Oncology* **1995**, *9*, 877.
(3) Katzenellenbogen, B. S.; Montano, M. M.; Ekena, K.; Herman, M. E.; McInerney, E. M.; McGuire Lecture, W. L. *Breast Cancer Res. Treat*. **1997**, *44*, 23.
(4) Wakeling, A. E. *Breast Cancer Res. Treat*. **1993**, *25*, 1.
(5) Mckeon, V. A. *J. Obstet. Gynecol. Neonatal Nurs*. **1997**, *26*, 79.
(6) Draper, M. W.; Flowers, D. E.; Huster, W. J.; Neild, J. A.; Harper, K. D.; Arnaud, C. J. *Bone Miner. Res*. **1996**, *11*, 835.
(7) Evans, G.; Bryant, H. U.; Magee, D. E.; Turner, R. T. *Endocrinology* **1996**, *137*, 4139.
(8) Katzenellenbogen, J. A. *Environ. Health Perspect*. **1995**, *103*, 99.
(9) Kelce, W. R.; Lambright, L. R.; Gray, L. E.; Roberts, K. P. *Toxicol. Appl. Pharmacol*. **1997**, *142*, 192.
(10) Kelce, W. R.; Wilson, E. M. *J. Mol. Med*. **1997**, *75*, 198.
(11) Colborn, T.; Dumanoski, D.; Myers, J. P. *Our Stolen Future: Are we threatening our fertility, intelligence, and survival?—A scientific detective story*; Penguin Books USA Inc.: New York, 1996.
(12) Katzenellenbogen, J. A.; Katzenellenbogen, B. S. *Chem. Biol*. **1996**, *3*, 529.
(13) Horwitz, K. B.; Jackson, T. A.; Rain, D. L.; Richer, J. K.; Takimoto, G. S.; Tung, L. *Mol. Endocrinol*. **1996**, *10*, 1167.
(14) Mosselman, S.; Polman, J.; Dijkema, R. *FEBS Lett*. **1996**, *392*, 49.
(15) Katzenellenbogen, B. S.; Korach, K. S. *Endocrinology* **1997**, *138*, 861.
(16) Kuiper, G. G. J. M.; Carlsson, B.; Grandien, K.; Enmark, E.; Haggblad, J.; Nilsson, S.; Gustafsson, J. *Endocrinology* **1997**, *138*, 863.
(17) Couse, J. F.; Lindzey, J.; Grandien, K.; Gustafsson, J. A.; Korach, K. S. *Endocrinology* **1997**, *138*, 4613.
(18) Kuiper, G. G. J. M.; Gustafsson, J. A. *FEBS Lett*. **1997**, *410*, 87.
(19) Kuiper, G. G. J. M.; Carlson, B.; Grandien, K.; Enmark, E.; Haggblad, J.; Nilsson, S.; Gustafsson, J.-A. *Endocrinology* **1997**, *138*, 863.
(20) Wani, M. C.; Rector, D. H.; Christensen, H. D.; Kimmel, G. L.; Cook, C. E. *J. Med. Chem*. **1975**, *18*, 982.
(21) von Angerer, E.; Prekajac, J.; Strohmeier, J. *J. Med. Chem*. **1984**, *27*, 1439.
(22) Mirocha, C. J.; Christensen C. M.; Nelson, G. H. *Cancer Res*. **1968**, *28*, 2319.
(23) Gauthier, S.; Caron, B.; Cloutier, J.; Dory, Y. L.; Favre, A.; Larouche, D.; Mailhot, J.; Ouellet, C.; Schwerdtfeger, A.; Leblanc, G.; Martel, C.; Simard, J.; Merand, Y.; Belanger, A.; Labrie, C.; Labrie, F. *J. Med. Chem*. **1997**, *40*, 2117.
(24) Dodds, E. C.; Golberg, L.; Lawson, W.; Robinson, R. *Nature* **1938**, *142*, 34.
(25) Harris, C. A.; Henttu, P.; Parker, M. G.; Sumpter, J. P. *Environ. Health Perspect*. **1997**, *105*, 802.
(26) Brzozowski, A. M.; Ashley, C.; Pike, A. C. W.; Dauter, Z.; Hubbard, R. E.; Bonn, T.; Engström, O.; Öhman, L.; Greene, G. L.; Gustafsson, J.-A.; Carlquist, M. *Nature* **1997**, *389*, 753. (a) Oglesby, L. A.; Eborn-McCoy, M. T.; Logsdon, T. R.; Copeland, F.; Beyer, P. E.; Kavlock, R. J. *Teratology* **1990**, *41*, 43.
(27) Wagner, R. L.; Apriletti, J. W.; McGrath, M. E.; West, B. L.; Baxter, J. D.; Fletterick, R. J. *Nature* **1995**, *378*, 690.
(28) Renaud, J.-P.; Rochel, N.; Ruff, M.; Vivat, V.; Chambon, P.; Gronemeyer, H.; Moras, D. *Nature* **1995**, *378*, 681.
(29) Bourguet, W.; Ruff, M.; Chambon, P.; Gronemeyer, H.; Moras, D. *Nature* **1995**, *375*, 377.
(30) Wurtz, J.-M.; Bourguet, W.; Renaud, J.-P.; Vivat, V.; Chambon, P.; Moras, D.; Gronemeyer, H. *Nat. Struct. Biol*. **1996**, *3*, 87.
(31) Anstead, G. M.; Carlson, K. E.; Katzenellenbogen, J. A. *Steroids* **1997**, *62*, 268.
(32) Hansch, C.; Gao, H. *Chem. Rev*. **1997**, *97*, 2995.
(33) Selassie, C. D.; De Soyza, T. V.; Rosario, M.; Gao, H.; Hansch, C. *Chem.-Biol. Interact*. **1998**, *113*, 175. (a) Zhang, L.; Gao, H.; Hansch, C.; Selassie, C. D. *J. Chem. Soc., Perkin Trans. 2*, **1998**, 2553.
(34) BioByte Corporation, C-QSAR database; P.O. Box 517, Claremont, CA 91711-0517.
(35) Aranyi, P. *Biochim. Biophys. Acta* **1980**, *628*, 220.
(36) Hansch, C.; Gao, H.; Hoekman, D. In *Comparative QSAR*; Devillers, J., Ed.; Taylor & Francis: Washington, DC, 1998; p 285.

(37) Hansch, C.; Leo, A.; Hoekman, D. *Exploring QSAR. Hydrophobic, Electronic and Steric Constants*; American Chemical Society: Washington, DC, 1995.
(38) Fevig, T. L.; Mao, M. K.; Katzenellenbogen, J. A. *Steroids* **1988**, *51*, 471.
(39) Hochberg, R. B. *Science* **1979**, *205*, 1138.
(40) Katzenellenbogen, J. A. Estrogen and progestin radiopharmaceuticals for imaging breast cancer. In *Estrogens, Progestins, and Their Antagonists*; Pavlik, E. J., Ed.; Birkhauser: Boston, 1996; p 197.
(41) Hansch, C.; Leo, A. *Exploring QSAR: Fundamentals and Applications in Chemistry and Biology*; American Chemical Society: Washington, DC, 1995.
(42) Belanger, A.; Philibert, D.; Teutsch, G. *Steroids* **1981**, *37*, 361.
(43) Nique, F.; Van de Velde, P.; Bremaud, J.; Hardy, M.; Philibert, D.; Teutsch, G. *J. Steroid Biochem. Mol. Biol.* **1994**, *50*, 21.
(44) Sasson, S.; Katzenellenbogen, J. A. *J. Steroid Biochem.* **1989**, *33*, 859.
(45) French, A. N.; Napolitano, E.; VanBrocklin, H. F.; Hanson, R. N.; Welch, M. J.; Katzenellenbogen, J. A. *Nucl. Med. Biol.* **1993**, *20*, 31.
(46) Hansch, C.; Klein, T.; *Acc. Chem. Res.* **1986**, *19*, 392.
(47) Salman, M.; Reddy, B. R.; Delgado, P.; Stotter, P. L.; Fulcher, L. C.; Chamness, G. C. *Steroids* **1991**, *56*, 375.
(48) Salman, M.; Reddy, B. R.; Ray, S.; Stotter, P. L.; Chamness, G. C. *J. Steroid Biochem.* **1986**, *24*, 539.
(49) Abraham, M. H.; McGowan, J. C. *Chromatographia* **1987,** *23*, 243.
(50) Napolitano, E.; Fiaschi, R.; Herman, L. W.; Hanson, R. N. *Steroids* **1996**, *61*, 384.
(51) Napolitano, E.; Fiaschi, R.; Carlson, K. E.; Katzenellenbogen, J. A. *J. Med. Chem.* **1995**, *38*, 429.
(52) DaSilva, J. N.; van Lier, J. E. *J. Med. Chem.* **1990**, *33*, 430.
(53) Gantchev, T. G.; Ali, H.; van Lier, J. E. *J. Med. Chem.* **1994**, *37*, 4164.
(54) Hartmann, R. W.; Heindl, A.; Schonenberger, G. *J. Med. Chem.* **1984**, *27*, 577.
(55) Hartmann, R. W.; Schwarz, W.; Schonenberger, H. *J. Med. Chem.* **1983**, *26*, 1137.
(56) Heiman, D. F.; Senderoff, S. G.; Katzenellenbogen, J. A.; Neeley, R. J. *J. Med. Chem.* **1980**, *23*, 994.
(57) Goswami, R.; Harsy, S. G.; Heiman, D. F.; Katzenellenbogen, J. A. *J. Med. Chem.* **1980**, *23*, 1002.
(58) Katzenellenbogen, J. A.; McGorrin, R. J.; Tatee, T.; Kempton, R. J.; Carlson, K. E.; Kinder, D. H. *J. Med. Chem.* **1981**, *24*, 435.
(59) Anstead, G. M.; Katzenellenbogen, J. A. *J. Med. Chem.* **1988**, *31*, 1754.
(60) Anstead, G. M.; Wilson, S. R.; Katzenellenbogen, J. A. *J. Med. Chem.* **1989**, *32*, 2163.
(61) von Angerer, E.; Prekajac, J.; Strohmeier, J. *J. Med. Chem.* **1984**, *27*, 1439.
(62) Mahboobi, S.; Grothus, G.; von Angerer, E. *Arch. Pharm. (Weinheim, Ger.)* **1994**, *327*, 481.
(63) Grese, T. A.; Cho, S.; Finley, D. R.; Godfrey, A. G.; Jones, C. D.; Lugar, C. W.; Martin, M. J.; Matsumoto, K.; Pennington, L. D.; Winter, M. A.; Adrian, M. D.; Cole, H. W.; Magee, D. L.; Phillips, L. D.; Rowley, E. R.; Short, L. L.; Glasebrook, A. L.; Bryant, H. U. *J. Med. Chem.* **1997**, *40*, 146.
(64) Palkowitz, A. D.; Glasebrook, A. L.; Thrasher, K. J.; Hauser, K. L.; Short, L. L.; Phillips, D. L.; Muehl, B. S.; Sato, M.; Shetler, P. K.; Cullinan, G. J.; Pell, T. R.; Bryant, H. U. *J. Med. Chem.* **1997**, *40*, 1407.
(65) Bignon, E.; Pons, M.; Crastes de Paulet, A.; Dore, J. C.; Gilbert, J.; Abecassis, J.; Miquel, J.-F.; Ojasoo, T.; Raynaud, J.-P. *J. Med. Chem.* **1989**, *32*, 2092.
(66) Gilbert, J.; Fuentes, M.; Ojasoo, T.; Dore, J.-C.; Pons, M. *J. Med. Chem.* **1997**, *40*, 1104.
(67) Campbell, N. R. *Proc. R. Soc. London (B)* **1940**, *129*, 528.
(68) Evans, R. M. *Science* **1988**, *240*, 889−895.
(69) Paech, K.; Webb, P.; Kuiper, G. G. J. M.; Nilsson, S.; Gustafsson, J.; Kushner, P. J.; Scanlan, T. S. *Science* **1997**, *277*, 1508.
(70) Katzenellenbogen, J. A.; O'Malley, B. W.; Katzenellenbogen, B. S. *Mol. Endocrinol.* **1996**, *10*, 119.
(71) Kedar, R. P.; Bourne, T. H.; Powles, T. J.; Collins, W. P.; Ashley, S. E.; Cosgrove, D. O.; Campbell, S. *Lancet* **1994**, *343*, 1318.
(72) Tong, W.; Perkins, R. *Endocrinology* **1997**, *138*, 4022.
(73) Soto, A. M.; Lin, T.-M.; Justicia, H.; Silvia, R. M.; Sonnenschein, C. In *Advances in Modern Environmental Toxicology XXI*; Colborn, T., Clement, C., Eds.; Princeton Scientific Publishing Co.: Princeton, NJ, 1992; Vol. 21, p 295.
(74) Perez, P.; Pulgar, R.; Olea-Serrano, F.; Villalobos, M.; Rivas, A.; Metzler, M.; Pedraza, V.; Olea, N. *Environ. Health Perspect.* **1998**, *106*, 167.
(75) Barnett, V.; Lewis, T. In *Statistic Data*, 3rd ed.; Wiley: New York, 1984.
(76) Selassie, C. D.; Li, R.-L.; Poe, M.; Hansch, C. *J. Med. Chem.* **1991**, *34*, 46.
(77) Smith, C. J.; Hansch, C.; Norton, M. J. *Mutat. Res.* **1997**, *379*, 167.
(78) *3-D QSAR in Drug Design*; Kubinyi, H., Folkers, G., Martin, Y. C., Eds.; Kluwer/Escom: Boston, MA, 1998.
(79) Kim. K. In *Classical and Three-Dimensional QSAR in Agrochemistry*; Hansch, C., Fujita, T., Eds.; ACS Symposium Series 606; American Chemical Society: Washington, DC, 1995; p 302.
(80) Leo, A. *Chem. Rev.* **1993**, *93*, 1281.
(81) Hansch, C. *Acc. Chem. Res.* **1993**, *26*, 147.
(82) Gantchev, T. G.; Ali, H.; van Lier, J. E. *J. Med. Chem.* **1994**, *37*, 4164.
(83) Waller, C.; Opera, T. I.; Chae, K.; Park, H.-K.; Korach, K. S.; Laws, S. C.; Wiese, T. E.; Kelce, W. R.; Gray, L. E., Jr. *Chem. Res. Toxicol.* **1996**, *9*, 1240.
(84) Sadler, B. R.; Cho, S. J.; Ishaq, K. S.; Chae, K.; Korach, K. S. *J. Med. Chem.* **1998**, *41*, 2261.
(85) Waller, C.; Minor, D. L.; McKinney, J. D. *Environ. Health Perspect.* **1995**, *103*, 702.
(86) LeClercq, G.; Heuson, J. C. *Anticancer Res.* **1981**, *1*, 217.
(87) Pouparet, J. H.; Lambert, D. M.; Vamecq, J.; Abul-Hajj, Y. J. *Bioorg. Med. Chem. Lett.* **1995**, *5*, 839.
(88) Tong, W.; Lowis, D. R.; Perkins, R.; Chem, Y.; Welsh, W. J.; Goddette, D. W.; Heritage, T. W.; Sheehan, D. M. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 669.

CR980018G