Structural Analysis of a Group of Phytoestrogens for the Presence of a 2-D Geometric Descriptor Associated with Non-genotoxic Carcinogens and Some Estrogens (44234)

A. R. CUNNINGHAM,* H. S. ROSENKRANZ,*,[1] AND G. KLOPMAN†

Department of Environmental and Occupational Health, University of Pittsburgh, Pittsburgh, Pennsylvania 15261; and Department of Chemistry,† Case Western Reserve University, Cleveland, Ohio 44106*

Abstract. **Analysis of a group of phytoestrogens indicates that the majority of these chemicals are devoid of a 6Å distance descriptor biophore. This 6Å biophore is associated with the carcinogenicity of a set of chemicals in the mouse subset of the Carcinogenic Potency Database assembled by Gold *et al.* The prevalence of non-DNA-reactive carcinogens and chemicals endowed with estrogenic activity included in the group of chemicals possessing 6Å descriptor suggests that it describes a ligand binding site on an estrogen receptor. Evidence is presented that estrogens with and without the 6Å biophore bind to alternate receptors or to similar receptors but with different affinities. Since the 6Å biophore was identified based upon a carcinogenicity database, it is conceivable that binding to the putative receptor that recognizes this biophore is associated with carcinogenicity. Alternatively, estrogens devoid of this 6Å biophore may be noncarcinogenic, suggesting that carcinogenicity and estrogenicity may be separate phenomena.** [P.S.E.B.M. 1998, Vol 217]

Phytoestrogens represent a diverse and naturally occurring source of estrogen exposure in animals and humans. Xenoestrogen exposure has been implicated in the etiology of breast cancer (1) as well as other hormonally related cancers. Alternatively, dietary intake of certain phytoestrogens has been suggested as diminishing the risk of certain cancers by their possible antiestrogenic action (see Ref. 2). The distinction between estrogen and antiestrogen may in fact be obscure. Several of the factors that determine the estrogenicity versus antiestrogenicity of a chemical are species- and tissue-specificity, and dose (reviewed in Ref. 3). Thus a debate exists over the possible risks and benefits from phytoestrogen exposure.

Traditional structure-activity relationship (SAR) analyses related to cancer causation have been based upon the somatic mutation and electrophilic theory of cancer causation (4). Detailed analyses of large cancer bioassay databases by Ashby and colleagues (5, 6) demonstrate the successful utilization of the electrophilic theory in identifying chemical substructures that are implicated in "genotoxic" carcinogenesis.

In fact, the majority of recognized human carcinogens are genotoxic (7, 8, 9). An exception to this generalization is found with hormonal carcinogens. Hormonal carcinogens are thought to act by a receptor-mediated mechanism (10, 11). However, some evidence also exists for the DNA-reactivity of diethylstilbestrol (12, 13) and tamoxifen (14, 15, 16, 17). This apparent multiplicity of action of certain hormonal carcinogens (i.e., estrogenic and genotoxic) must be considered when examining the possible risks (or benefits) of exposure to phytoestrogens.

The suitability of SAR and quantitative SAR (QSAR) involving non-genotoxic carcinogens is a matter of debate. The multiplicity of mechanisms believed to be involved in non-genotoxic carcinogenesis has been seen as making such an endeavor problematic (18, 19, 20, see also Ref. 21).

[1] To whom requests for reprints should be addressed at Department of Environmental and Occupational Health, University of Pittsburgh, Pittsburgh, Pennsylvania 15261.
This investigation was supported by the U.S. Department of Defense (Contract No. DAAA21-93-C-0046) and a Predoctoral Training in Breast Cancer Biology and Therapy Fellowship awarded by the U.S. Army Medical Research and Acquisition Activity.

0037-9727/98/2173-0288$10.50/0

However, in a recent study, we identified a 6Å geometric distance descriptor that appears to be associated with non-genotoxic carcinogens and may, in fact, be indicative of a ligand binding site for an estrogen receptor (22, 23). Additionally, estrogens with and without the 6Å biophore may bind to alternate receptors or with different affinities to the same receptors. Of course the possession of this biophore is only indicative of carcinogenicity and possible estrogenicity. Lipophilicity, steric constraints as well as other physicochemical parameters will influence the binding ability and activities of these chemicals. The results contained herein represent excerpts and extensions of our work (22, 23) to phytoestrogens.

Methods

The computer automated structure evaluation system MULTICASE (MC) (24, 25) was used for these analyses. Basically, the system selects its own descriptors automatically from a learning set of chemicals composed of active and inactive molecules. The descriptors are continuous fragments that are embedded in the molecule. It is assumed that each fragment is not related to activity and will be randomly distributed between inactive and active molecules. Fragments that deviate from a random distribution are considered relevant to activity. The descriptors consist of either activating or inactivating fragments or geometric distance descriptors based on molecular lipophilic centers (see below). MC then selects the most important of these descriptors as a biophore (i.e., the functionality that is associated with the largest number of active molecules and fewest numbers of inactive molecules). These molecules then become a learning subset to identify the chemical and physicochemical properties associated with their activity. This will result in a QSAR equation for this group of molecules. The molecules explained here are removed from the set, and the process is iterated until all molecules are explained. Once the training set has been assimilated, MC can be queried regarding the predicted activity and potency of molecules of unknown activity.

MC incorporates the following rules to identify two-dimensional distance descriptors based upon the presence of lipophilic centers. These two-dimensional distances are calculated from the molecular structure. Heteroatoms and lipophilic carbon atoms are designated as "special" atoms. A carbon atom is designated as a lipophilic center if it is at least four bonds away from a heteroatom and is also the furthest carbon from the heteroatom when its neighbors are considered. After all the "special" atoms are selected, the distances between all possible pairings are calculated. The distribution of these descriptors among active and inactive molecules is analyzed for statistical significance. Various atom groupings are also investigated (i.e., hydrogen bond acceptors and donors as well as halogens).

The learning set used in this investigation was the Carcinogenic Potency Database (CPDB) assembled by Gold *et al.* (26, 27, 28, 29, 30). A mouse database was extracted from this compilation. The rationale for using a species-specific database as opposed to the entire CPDB is that a more coherent set of structural descriptors of carcinogenic activity may be expected in a more strictly defined system.

All dosages for chemicals reported were transformed into gavage equivalents (26). A carcinogenic potency value (TD_{50}) was calculated for each chemical. The calculated TD_{50} is the dose required for 50% of the animals to remain cancer free throughout the course of the experiment (thus accounting for spontaneous cancer) (26, 31). Additionally, the reported TD_{50} value in mg/kg/day was converted into mmol/kg/day. For the purpose of the SAR analyses, TD_{50}

Biophore A:
2D fragment [C-] <--- 6.0 Å ---> [OH-] conjugated and generic

14 out of the known 16 molecules (87%) containing such a biophore are mouse carcinogens with an average activity of 47. (conf.level= 100%)

	Constant	51.8
Modulator 1: OH-CH-	Activating	34.2

The probability that this molecule is a mouse carcinogen is 83.3%

The compound is predicted to be extremely active
The projected mouse carcinogenicity activity 86 CASE units

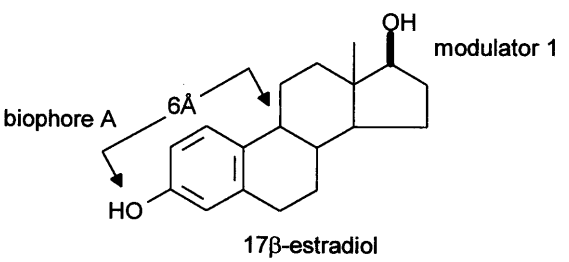

Figure 1. Prediction of the Murine Carcinogenicity of 17β-estradiol. An activity of 86 CASE units indicates a TD_{50} value of 0.001 mmol/kg/day.

Table I. Distribution of the 6Å Distance Descriptor Biophore Among Estrogenic Chemicals

Chemical	Type	6Å
Chalcone	phytoestrogen	-
4-hydroxychalcone	phytoestrogen	-
4,4′-dihydroxychalcone	phytoestrogen	-
4-hydroxy-4′-methoxychalcone	phytoestrogen	-
2′,4,4′-trihydroxychalcone (isoliquiritigenin)	phytoestrogen	-
2′,4,4′6′-tetrahydroxychalcone (naringenin chalcone)	phytoestrogen	-
2′,4,4′,6′-tetrahydroxydihydrochalcone (phloretin)	phytoestrogen	-
flavone	phytoestrogen	-
3-hydroxyflavone	phytoestrogen	-
4′-hydroxyflavone	phytoestrogen	-
6-hydroxyflavone	phytoestrogen	-
7-hydroxyflavone	phytoestrogen	-
3′,6-dihydroxyflavone	phytoestrogen	-
3′,7-dihydroxyflavone	phytoestrogen	-
4′,5-dihydroxyflavone	phytoestrogen	-
4′,6-dihydroxyflavone	phytoestrogen	-
5,7-dihydroxyflavone (chrysin)	phytoestrogen	-
7,8-dihydroxyflavone	phytoestrogen	-
3′,4′,7-trihydroxyflavone	phytoestrogen	-
3,5,7-rihydroxyflavone (galangin)	phytoestrogen	-
4′,5,7-trihydroxyflavone (apigenin)	phytoestrogen	-
4′,7,8-trihydroxyflavone	phytoestrogen	-
3,3′,4′,7-tetrahydroxyflavone (fisetin)	phytoestrogen	-
3′,4′,5,7-tetrahydroxyflavone (luteolin)	phytoestrogen	-
3,4′,5,7-tetrahydroxyflavone (kaempferol)	phytoestrogen	-
3,5,7-trihydroxy-4′-methoxyflavone (kaempferide)	phytoestrogen	-
3,3′,4′,5,7-pentahydroxyflavone (quercetin)	phytoestrogen	-
2′,3,4′,5,7-pentahydroxyflavone (morin)	phytoestrogen	-
3,3′,4′,5,5′,7-hexahydroxyflavone (myricetin)	phytoestrogen	-
flavanone	phytoestrogen	-
4′,7-dihydroxyflavanone	phytoestrogen	-
4′,5,7-trihydroxyflavanone (naringenin)	phytoestrogen	-
3′,5,7-trihydroxy-4′-methoxyflavanone (hesperetin)	phytoestrogen	-
3,3′,4′,5,7-pentahydroxyflavanone (taxifolin)	phytoestrogen	-
3,3′,4′,5,7-flavan pentol flavanone ([±] catechin)	phytoestrogen	-
isoflavone	phytoestrogen	-
4′,7-dihydroxyisoflavone (diadzein)	phytoestrogen	-
7-hydroxy-4′-methoxyisoflavone (formononetin)	phytoestrogen	-
3′,4′,7-trihydroxyisoflavone	phytoestrogen	-
4′,5,7-trihydroxyisoflavone (genistein)	phytoestrogen	-
4′,6,7-trihydroxyisoflavone	phytoestrogen	-
5,7-dihydroxy-4′-methoxyisoflavone (biochanin A)	phytoestrogen	-
coumestrol	phytoestrogen	-
4,4′-dihydroxystilbene	phytoestrogen	+
3,5-dihydroxystilbene	phytoestrogen	-
coumarin	phytoestrogen	-
α-sitosterol	phytoestrogen	-
β-sitosterol	phytoestrogen	-
glucyrrhetinic acid	phytoestrogen	-
zearalenol	phytoestrogen	-
zearalenone	phytoestrogen	-
α-zearalenol	phytoestrogen	-
indenestrol A	phytoestrogen	+
tetrahydrocannabinol	phytoestrogen	-
o,p′-DDE	xenoestrogen	-
chlordecone	xenoestrogen	-
diethylstilbestrol	estrogen	+
3′-hydroxy-*E*-diethylstilbestrol	estrogen	+
4′,4″-diethylstilbestrol quinone	estrogen	-
tamoxifen	antiestrogen	-
3-hydroxytamoxifen	antiestrogen	-
4-hydroxytamoxifen acid	antiestrogen	+
toremifene	antiestrogen	-
4-hydroxy-deamino-hydroxytoremifene	antiestrogen	+
ICI 164,384	antiestrogen	+
ICI 182,780	antiestrogen	+
LY 117018	antiestrogen	-
MER 25	antiestrogen	-
17 β-estradiol	estrogen	+
17α-ethinyl estradiol	estrogen	+
benzestrol	estrogen	+
dienestrol	estrogen	+
estriol	estrogen	+
estrone	estrogen	+
hexestrol	estrogen	+
megestrol	estrogen	-
norgestrol	estrogen	-
norlestrin (isomer)	estrogen	+
phenol red	xenoestrogen	-

values were transformed into CASE activity units using Equation 1.

Chemicals were also assigned to activity groups. Chemicals reported by the authors to be non-carcinogenic in mice were assigned 10 CASE units together with chemicals exhibiting TD_{50} values in excess of 51 mmol/kg/day. The chemicals in the range of 10–19 CASE units were designated as inactive. Chemicals with activities in the range of 20–29 CASE units were designated marginally active carcinogens, and chemicals in the range of 30–99 CASE units were considered carcinogenic. Overall, the mouse CPDB consisted of 636 chemicals, 291 of which were active, 11 marginal, and 334 noncarcinogenic.

$$\text{CASE activity} = 14.133 \times \log (1/TD_{50}) + 44.133 \quad \text{(Equation 1)}$$

Results and Discussion

The 6Å biophore was identified as being associated with carcinogenicity in the mouse CPDB. Among the chemicals in the database that possessed the 6Å biophore were estradiol and related chemicals (Fig. 1), suggesting that this biophore may be related to the estrogenicity as well as the carcinogenicity of chemicals that contain it (21). In our original study, a series of 42 chemicals reported to be endowed with estrogenic activity were tested for the presence of the 6Å biophore; of these, about one-half contained the 6Å distance descriptor (see Table I for an abbreviated list).

The therapeutic antiestrogen tamoxifen lacked the 6Å biophore; however, its active antiestrogenic metabolite, 4-hydroxytamoxifen, possessed it. Oxidative metabolism is often required for chemicals to gain their estrogenicity as is the case for tamoxifen (3). Polychlorinated biphenyls and methoxychlor have also been shown to require oxidative metabolism to exhibit estrogenicity (32), and, in fact, hydroxylated metabolites of polychlorinated biphenyls display the 6Å descriptor. Moreover, the antiestrogens ICI 164,384 and ICI 182,780 contain the 6Å biophore whereas LY 117018, another antiestrogen, lacks it. It has been suggested that LY 117018 and tamoxifen (32, 34) and LY 117018, ICI 182,780, and 164,384 (35) have different bases for their antiestrogenicity. Thus possession or lack of the 6Å biophore may be indicative of a dichotomy that exists among estrogens.

The present exercise entailed analysis for the 6Å biophore in a group of 54 phytoestrogens (Table I). As mentioned above, oxidative metabolism may be needed to transform estrogens (i.e., proestrogens) into active congeners. Many of the chemicals in this set were hydroxylated and methoxylated congeners of flavonoids. Thus if oxidative metabolism was involved in the conversion of these chemicals to carcinogens or estrogens, this set of chemicals may contain active congeners. However, only indenestrol A and 4,4′-dihydroxystilbene, an analog of diethylstilbestrol, pos-

sessed the 6Å descriptor (Table I). No flavonoids or other phytoestrogens tested contained the 6Å biophore.

Conclusion

We postulated that the 6Å descriptor is associated not only with murine carcinogenicity, but also with an estrogen-receptor ligand. This suggests that the phenomena of estrogenicity and carcinogenicity may result from a similar mechanism as is possibly the case for the reported carcinogenicity of tamoxifen. However, a large population of chemicals endowed with estrogenic activity, particularly phytoestrogens, are devoid of this biophore, suggesting that a structurally based dichotomy may exist in the estrogenic response.

In fact Baker (36) suggests that the estrogenicity of phytoestrogens is derived from their inherent ability to interact with mammalian enzymes involved in the regulation and production of endogenous estrogens. Thus phytoestrogens may exert their estrogenicity through a mechanism that is not involved with carcinogenicity as indicated by the absence of the 6Å descriptor associated with carcinogenicity in mice.

1. Davis DL, Bradlow HL, Wolff M, Woodruff T, Hoel DG, Anton-Culver H. Medical hypothesis: Xenoestrogens as preventable causes of breast cancer. Environ Health Prospect **101:**372–377, 1993.
2. Sheehan DM. Introduction: The case for expanded phytoestrogen research. Proc Soc Exp Biol Med **208:**3–5, 1995.
3. Lerner LJ, Jordan VC. Development of antiestrogens and their use in breast cancer. Eighth Cain Memorial Award Lecture. Cancer Res **50:**4177–4189, 1990.
4. Miller JA, Miller EC. Ultimate chemical carcinogens as reactive mutagenic electrophiles. In: Hiatt HH, Watson JD, Winsten JA, Eds. Origins of Human Cancer Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press, pp605–627, 1977.
5. Ashby J, Tennant RW. Definitive relationships among chemical structure, carcinogenicity, and mutagenicity of 301 chemicals tested by the U.S. National Toxicological Program. Mutation Res **257:**229–306, 1991.
6. Ashby J, Paton D. The influence of chemical structure on the extent and sites of carcinogenesis for 522 rodent carcinogens and 55 different human carcinogen exposures. Mutation Res **286:**3–74, 1993.
7. Ennever FK, Noonan TJ, Rosenkranz HS. The predictivity of animal bioassays and short-term genotoxic tests for carcinogenicity and non-carcinogenicity to humans. Mutagenesis **2:**73–78, 1987.
8. Shelby MD. The genotoxicity of human carcinogens and its implications. Mutat Res **204:**3–15, 1988.
9. Bartsch H, Malaveille C. Prevalence of genotoxic chemicals among animal and human carcinogens evaluated in the IARC Monograph Series. Cell Biol Toxicol **5:**115–127, 1989.
10. Barrett JC. Mechanisms of action of known human carcinogens. In: Vainio H, Magee PN, McGregor DB, McMichael AJ, Eds. Mechanism of Carcinogenesis in Risk Identification, IARC Scientific Publication No. 116. Lyon, France: International Agency for Research on Cancer, pp115–134, 1992.
11. Lucier GW. Receptor-mediated carcinogenesis. In: Vainio H, Magee PN, McGregor DB, McMichael AJ, Eds. Mechanism of Carcinogenesis in Risk Identification, IARC Scientific Publication No. 116. Lyon, France: International Agency for Research on Cancer, pp87–112, 1992.
12. Gladek A, Liehr JG. Mechanism of genotoxicity of diethylstilbestrol *in vivo*. J Biol Chem **264:**16847–16852, 1989.
13. Williams GM, Iatropoulos MJ, Djordjevic MV, Kaltenberg OP. The triphenylethylene drug tamoxifen is a strong liver carcinogen in the rat. Carcinogenesis **14:**315–317, 1993.
14. Han X, Liehr JG. Induction of covalent DNA adducts in rodents by tamoxifen. Cancer Res **52:**1360–1363, 1992.
15. White INH, DeMatteis F, Davies A, Smith LL, Crofton-Sleigh C, Venitt S, Hewer A, Phillips DH. Genotoxic potential of tamoxifen and analogs in female Fischer F344/n rats, DBA/2 and C57BL/6 mice and in human MCL-5 cells. Carcinogenesis **13:**2197–2203, 1992.
16. Hard GC, Iatropoulos MJ, Jordan K, Radi L, Kaltenberg OP, Imondi AR, Williams GM. Major differences in the heptocarcinogenicity and DNA adduct forming ability between toremifene and tamoxifen in female Crl:CD(BR) rats. Cancer Res **53:**4534–4541, 1993.
17. Montandon F, Williams GM. Comparison of DNA reactivity of the polyphenylethylene hormonal agents diethylstilbestrol, tamoxifen, and toremifene in rat and hamster liver. Arch Toxicol **68:**272–275, 1994.
18. Vainio H, Magee PN, McGregor DB, McMichael AJ, Eds. Mechanism of Carcinogenesis in Risk Identification, IARC Scientific Publication No. 116. Lyon, France: International Agency for Research on Cancer, pp 9–54, 1992.
19. Ashby J. Use of short-term tests in determining the genotoxicity or nongenotoxicity of chemicals. In: Vainio H, Magee PN, McGregor DB, McMichael AJ, Eds. Mechanism of Carcinogenesis in Risk Identification, IARC Scientific Publication No. 116. Lyon: International Agency for Research on Cancer, pp135–164, 1992.
20. Ashby J. Two million rodent carcinogens? The role of SAR and QSAR in their detection. Mutat Res **305:**3–12, 1994.
21. Rosenkranz HS. Structure-activity relationships for carcinogens with different modes of action. In: Vainio H, Magee PN, McGregor DB, McMichael AJ, Eds. Mechanism of Carcinogenesis in Risk Identification, IARC Scientific Publication No. 116. Lyon: International Agency for Research on Cancer, pp271–277, 1992.
22. Cunningham A, Klopman G, Rosenkranz HS. Study of the structural basis of the carcinogenicity of tamoxifen, toremifene, and their metabolites. Mutat Res **304:**85–95, 1996.
23. Rosenkranz HS, Cunningham A, Klopman G. Identification of a 2-D geometric descriptor associated with nongenotoxic carcinogens and some estrogens and antiestrogens. Mutagenesis **11:**95–100, 1996.
24. Klopman G. MULTICASE 1. A hierarchical Computer Automated Structure Evaluation program. Quantitative Struct Activity Relationships **11:**176–184, 1992.
25. Klopman G, Rosenkranz HS. Prediction of carcinogenicity/mutagenicity using MULTICASE. Mutat Res **305:**33–46, 1994.
26. Gold LS, Sawyer CB, Magaw R, Backman GM, deVeciana M, Levinson R, Hooper NK, Havender WR, Berstein L, Peto, R, Pike MC, Ames BN. A carcinogenic potency database of the standardized results of animal bioassays. Environ Health Perspect **58:**9–319, 1984.
27. Gold LS, deVeciana M, Backman GM, Lopipero M, Smith M, Blumenthal R, Levinson R, Berstein L, Ames BN. Chronological supplement to the carcinogenic potency database: Standardized results of animal bioassays published through December 1982. Environ Health Perspect **67:**161–200, 1986.
28. Gold LS, Slone TH, Backman GM, Magaw R, DaCosta M, Lopipero P, Blumenthal M, Ames BN. Second chronological supplement to the carcinogenic potency database: Standardized results of animal bioassays published through December 1984 and by the National Toxicology Program through May 1986. Environ Health Perspect **74:**237–329, 1987.
29. Gold LS, Slone TH, Backman GM, Eisenberg S, DaCosta M, Wong M, Manley NB, Rohrbach L, Ames BN. Third chronological supplement to the carcinogenic potency database: Standardized results of animal bioassays published through December 1986 and by the National Toxicology Program through June 1987. Environ Health Perspect **84:**215–286, 1990.
30. Gold LS, Manley NB, Slone TH, Garfinkle TH, Rohrbach L, Ames

BN. Fifth plot of the carcinogenic potency database: Results of animal bioassays published in the general literature through 1986 and by the National Toxicology Program through 1989. Environ Health Perspect **100:**65–135, 1993.

31. Peto R, Pike MC, Bernstein L, Gold LS, Ames BN. The TD_{50}: A proposed general convention for the numerical description of the carcinogenic potency of chemicals in chronic-exposure animal experiments. Environ Health Perspect **58:**1–8, 1984.
32. Korach KS, Sarver P, Chae K, McLachlan JA, McKinney JD. Estrogen receptor-binding activity of polychlorinated hydroxybiphenyls: Conformationally restricted structural probes. Mol Pharmacol **33:**120–126, 1998.
33. Black LJ, Goode RL. Evidence for biological action of the antiestrogens LY 117018 and tamoxifen by different mechanisms. Endocrinology **109:**987–989, 1981.
34. Scholl SM, Huff KF, Lippman ME. Antiestrogenic effects of LY 117018 in MCF-7 cells. Endocrinology **113:**611–617, 1983.
35. Coradini D, Biffi A, Cappelletti V, DiFronzo G. Activity of tamoxifen and new antiestrogens on estrogen receptor positive and negative breast cancer cells. Anticancer Res **14:**1059–1064, 1994.
36. Baker ME. Endocrine activity of plant-derived compounds: An evolutionary perspective. Proc Soc Exp Biol Med **208:**131–138, 1995.