

Molecular Shape and QSAR Analyses of a Family of Substituted Dichlorodiphenyl Aromatase Inhibitors

P. I. Nagy,[†] John Tokarski, and A. J. Hopfinger*

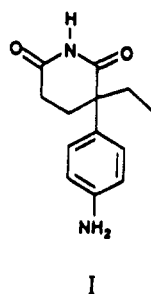
Laboratory of Molecular Modeling and Design, and the Department of Medicinal Chemistry and Pharmacognosy (M/C 781), College of Pharmacy, The University of Illinois at Chicago, Chicago, Illinois 60612

Received February 24, 1994*

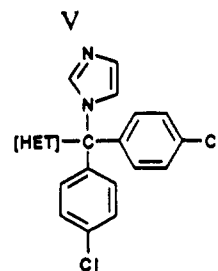
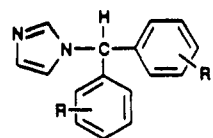
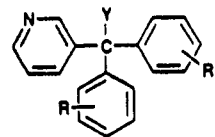
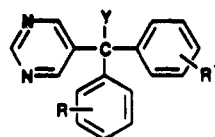
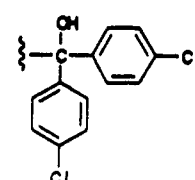
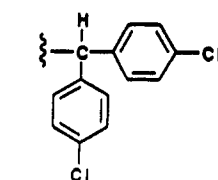
Conformational analyses of three families of substituted dichlorodiphenyl aromatase inhibitors indicated that both potent and weak inhibitors adopt a common global minimum energy conformation. Further, this global minimum energy conformation is the only meaningful intramolecular conformer state that can be energetically realized and is virtually identical to the crystal structure of one of the analogs. Quantitative structure–activity relationships, QSARs, were separately, and jointly, developed for two series of inhibitors. The distance, D , of a nitrogen atom in the variable heterocycle from the core C_c atom is the most important activity descriptor. The optimum distance between the nitrogen and C_c to maximize inhibitor potency is about 3.6 Å for both classes of analogs. Integrated potential energy field difference calculations were also carried out using a proton probe and some of the variable heterocycles. The field calculations coupled with the QSAR studies suggest that the nitrogen 3.6 Å from C_c acts as a hydrogen bond acceptor. Two possible three-dimensional pharmacophores are proposed for effective aromatase inhibitors.

INTRODUCTION

Enzymatic aromatization of androgens is the terminal step in the biosynthesis of estrogens. The inhibition of aromatase has become a pharmacological endpoint in the control of estrogen-dependent diseases. Much of this interest in aromatase inhibition has come from the success of aminoglutethimide, I, in the treatment of breast cancer through the inhibition of aromatase.^{1–4}



Jones et al.⁵ have synthesized and determined the *in vitro* aromatase inhibitory activity of a family of heterocyclic (4,4'-dichlorodiphenyl)methanes and methanols. The analog classes within this family are given by II–VII. Many of the compounds made and tested by Jones et al.⁵ exhibit EC_{50} potencies for aromatase inhibition at the nanomolar level. These workers have also derived a structural model for their most active compounds based upon the observed structure–activity relationship (SAR). A key feature of the structural model has the organic moiety attached to the heterocycle at a position β to the N atom. The crystal structure of 2-[bis(4-chlorophenyl)methyl]-2H-tetrazole, compound 11 of Table 2, which is a moderately potent aromatase inhibitor ($-\log(EC_{50}) = 6.83$, see Table 2) is also reported by Jones et al.⁵



It occurred to us that this family of aromatase inhibitors might be a good data set to apply molecular shape analysis, MSA,⁶ in order to generate a three-dimensional quantitative structure–activity relationship, 3D-QSAR.⁷ Moreover, three specific questions arise from a survey of the available data: (1) Is the crystal conformation of compound 11 the “active” conformation? (2) What physicochemical properties of the heterocycle govern activity? (3) Is enough structure–activity information available to construct a 3D-QSAR and corresponding 3D-pharmacophore?

METHODS

A. Conformational Analysis. Systematic fixed valence geometry conformational scans were carried out for the

* To whom correspondence should be addressed.

[†] On leave from Chemical Works of Gedeon Richter Ltd., Budapest, Hungary. Current address: College of Pharmacy, University of Toledo, Toledo, OH 43606.

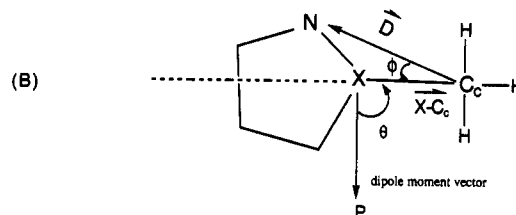
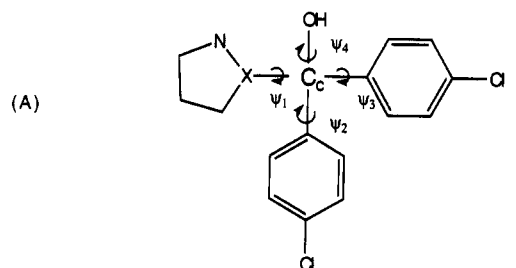
* Abstract published in *Advance ACS Abstracts*, August 1, 1994.

Table 1. Ab Initio, Experimental, and AM1 Dipole Moments (in debyes) for Five-Membered Azoles

	ab initio unsubstituted	exp ^a unsubstituted	AM1	
			unsubstituted	substituted
pyrrole	1.85 ^a	1.74	1.95	2-Me 1.95 3-Me 1.83
imidazole	3.84 ^a	3.67	3.61	2-Me 3.52 4-Me 3.39 5-Me 3.69
pyrazole	2.42 ^a	2.21	2.10	1-Me 2.29 3-Me 1.86 4-Me 2.08 5-Me 2.41
triazole-1,2,3	4.6 ^b 4.66 ^c		4.11	1-Me 4.33 3-Me 3.99 5-Me 4.45
triazole-1,2,4	2.90 ^a 2.9 ^b	2.72	2.73	1-Me 3.21 3-Me 2.32 5-Me 3.00
triazole-1,2,5	0.5 ^b 0.06 ^c		0.12	1-Me 0.55 3-Me 0.49
triazole-1,3,4	5.99 ^b 6.0 ^c		5.32	1-Me 5.77 2-Me 5.35
tetrazole-1,2,3,4	5.63 ^a	5.30	5.13	5-Me 5.47
tetrazole-1,2,3,5	2.23 ^a	2.19	1.99	1-Me 2.54

^a Mo, O; de Paz, J. L. G.; Yanez, M. *J. Phys. Chem.* **1986**, *90*, 5597–5604. ^b Cox, J. R.; Woodcock, S.; Hillier, I. H.; Vincent, M. A. *J. Phys. Chem.* **1990**, *94*, 5499–5501. ^c Tornkvist, C.; Bergman, J.; Liedberg, B. *J. Phys. Chem.* **1991**, *95*, 3123–3128.

5-pyrimidinyl and 2-pyridyl analogs in each of the II(H), III(OH), and VII(Im) series. The reason for doing these conformational searches was to first establish the set of stable free space conformer states available to each analog series and then to ascertain which of the stable conformations are common to active analogs of all three series. The four torsion



$$\phi = \angle \text{NC}_c\text{X}$$

$$\theta = \angle \text{PXC}_c$$

$$\vec{D} = \vec{C}_c\text{N}$$

Figure 1. (A) Definition of the torsion angles ψ_1 , ψ_2 , ψ_3 , and ψ_4 and (B) schematic definitions of the descriptors D , θ , and ϕ .

angles searched are defined in Figure 1A. For the analogs in series II, only ψ_1 , ψ_2 , and ψ_3 are explored.

The six molecules were built using standard bond lengths and angles, and the structures optimized using the MMFF

Table 2. Heterosubstituted (Dichlorodiphenyl) Methanes, II

no.	HET	D (Å)	θ (deg)	ϕ (deg)	$\log P$	$-\log(\text{EC}_{50})$ obsd	$-\log(\text{EC}_{50})$ pred. (eq 2, Table 6)
1		3.76	0.0	150.0	0.04	7.26	7.64
2		4.31	0.0	179.7	1.16	6.80	6.18
3		2.47	108.0	120.0	1.16	<5.30	5.34
4		2.53	62.7	123.4	1.26	<5.30	5.63
5		4.26	98.3	154.3	-0.10	5.71	6.37
6		2.47	50.2	124.8	0.65	5.70	5.34
7		3.63	15.2	159.5	-0.75	8.06	7.77
8		3.63	0.00	162.8	-0.91	8.03	7.77
9		3.55	46.1	160.2	-0.53	7.73	7.81
10		2.47	0.0	123.0	-0.04	<5.30	5.34
11 ^a		3.54	28.9	159.0		6.83	^a

^a Outlier—not used to derive eq 2 of Table 6.

option of the molecular modeling package CHEMLAB-II.⁸ MMFF is an extended version, in terms of parameterization and force field representation, of Allinger's MM2 program.⁹ Atomic charges were computed for the minimum energy conformation of each molecule using the AM1 semiempirical quantum chemical method.¹⁰ The structures were then reoptimized in MMFF including the monopole electrostatic potential when computing the conformational energies. The valence geometries were next frozen, and systematic conformational searching was performed at 30° increments for each of the four possible torsion angles. The nonbonded portion of the Allinger force field, including the monopole electrostatic potential term, was used to compute the conformational energies. Each of the apparent minima identified in the conformational scans was rigorously minimized again using MMFF. The resulting optimized structures were retained for further inspection.

B. QSAR Descriptors. An inspection of the structure-activity data generated by Jones et al.⁵ suggested that the electronic structure of the variable heterocyclic ring might be important to inhibitor potency. Moreover, AM1 calculations on the six compounds considered in the conformational analysis indicated that the electronic properties of each of their heterocyclic rings are virtually identical to those of the corresponding unsubstituted heterocyclic ring. Consequently, the corresponding unsubstituted ring for each unique heterocycle contained in the composite set of II, III, and IV analog series was optimized using AM1. The dipole moments of the unsubstituted heterocyclic rings computed using AM1 are in good agreement with the experimental and *ab initio* values, see Table 1. Some methyl analogs were also considered to evaluate how sensitive the AM1 dipoles are to aliphatic substitutions (see the last column of Table 1). Methylation has minimal effect on the dipole moment of the heterocycle. Thus, both unsubstituted and methylated heterocyclic rings are reasonable electronic structure models for this unit when it is part of the complete inhibitor. The AM1 values tend to underestimate, while the *ab initio* overestimate, the experimental values. Where measured values are not available in Table 1, the difference in *ab initio* and AM1 values does not exceed 10%. Overall, the AM1 dipoles constitute a reasonably accurate, self-consistent set of hetero-ring properties.

Electronic properties derived from the dipole moment and heterocyclic ring geometry were explored as possible QSAR descriptors. In particular, three descriptors derived from the net electronic structure, which are defined in Figure 1B, were considered. *D* is the N-C_c distance, based upon the optimized geometry, where N is a heterocycle nitrogen in any position by X, see Figure 1. If there is more than one nitrogen in such a position in the heterocycle, *D* refers to the N-C_c distance for the β nitrogen. If there are two β nitrogens in the heterocycle, *D* is the average value of the two N(β)-C_c distances. Two directional descriptors, θ and ϕ , are the angles which the total heterocycle dipole moment and the N-atom position vector, *D*, make, respectively, with the X-C_c vector (X is the heterocycle atom bonded to C_c). Other electronic properties that were also explored as QSAR descriptors in the study, but not found to be of statistical significance, include the *x*, *y*, and *z* components of the total dipole moment, the highest occupied, HOMO, and lowest unoccupied, LUMO, molecular orbital energies of the heterocyclic ring and also the distribution of partial atomic charges in the heterocyclic ring.

The integrated potential field difference, ΔF ,¹¹ using the heterocycle of analog 7 in Table 2 as the reference comparison

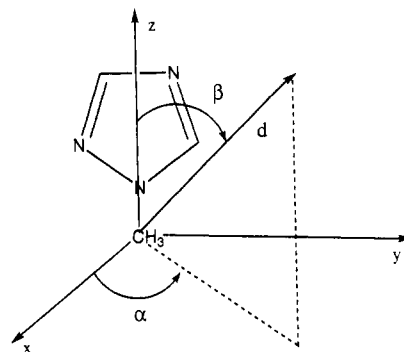
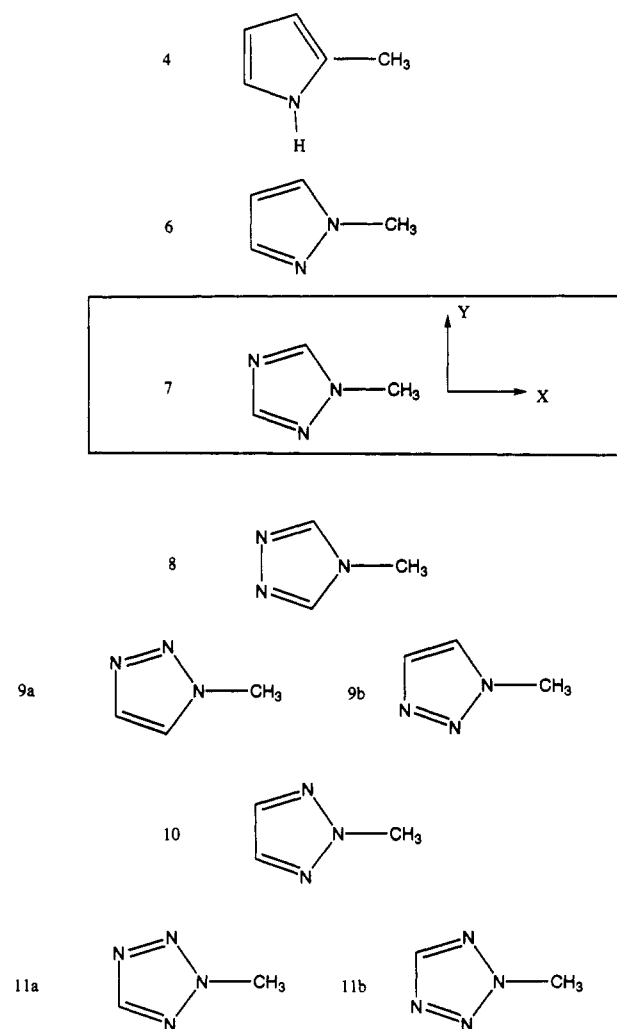


Figure 2. Spatial orientation for potential field calculations using the reference comparison structure 7 of Table 3 as an example.














Table 3. Methyl Heterocycles Considered in the Integrated Potential Field Difference Study^a



^a Structure 7 is the reference comparison structure, and the relative alignments used for the other structures are shown relative to 7.

structure was computed for each of the heterocycles of seven analogs in Table 2. These seven analogs were chosen because they are identical in ring topology. Therefore differences in activity should be of electronic origin. For each of the seven analogs three ΔF 's, corresponding to an oxygen anion, a methyl group, and a proton probe, were estimated. The seven structures are reported in Table 3, and the orientation geometry used to carry out the field difference integration is shown in Figure 2. The alignment criterion was to match the CH₃-X-Y unit where X = (-C-) or (-N-) and Y = (-C=) or (=N-) of the heterocycle. In two cases, the heterocycles of

Table 4. Heterosubstituted (Dichlorophenyl) Methanols, III

no.	HET	<i>D</i> (Å)	θ (deg)	ϕ (deg)	log <i>P</i>	log (EC ₅₀) obsd	-log (EC ₅₀) pred (eq 4, Table 6)
1		3.76	0.0	150.0	0.04	7.12	7.43
2		4.27	49.0	178.4	0.04	7.00	7.12
3		4.31	0.0	179.7	1.16	7.08	7.07
4		2.47	108.0	120.0	1.16	<5.30	5.57
5		3.76	6.5	146.3	0.04	7.03	7.43
6		2.55	79.8	125.5	0.22	6.13	5.80
7		3.68	50.2	162.5	0.22	7.55	7.43
8		3.60	3.1	161.5	0.59	7.05	7.41
9		3.69	78.2	163.0	0.59	7.52	7.43
10		3.67	66.0	157.2	-0.65	<5.60	^a
11		3.64	92.0	162.5	0.12	5.70	^a
12		3.78	41.0	160.1	0.71	8.02	7.43
13		3.92	46.6	164.1	0.76	7.72	7.40

^a Outliers not used to derive eq 4 of Table 6.

analogs 9 and 11, two different alignments were considered as shown in Table 3. Different spatial sampling about superimposed pairs of heterocycles, in terms of *d*, α and β , see Figure 2, were considered.

Overall, the objective of the field calculations was to gain insight about the electronic environment of the receptor. The integrated potential field difference, in a given region of space about a pair of superimposed analogs, yields a quantitative measure of how differently a receptor "sees" these two analogs in the region of space. Of course, this mimic of the receptor view is biased by the choice of the probe. Nevertheless, ΔF has been found to be a significant 3D-QSAR descriptor in past studies.^{6,7,11}

The relative lipophilicity of each heterocycle, log *P*, was determined using the CLOGP program.¹² Experimental log *P* values are available for six compounds (2, 3, and 4 of Table 2 and 2, 5, and 7 of Table 4). The differences in calculated and measured log *P* values varies from 0.04 to 0.19 suggesting the calculated values are reliable over the entire data set.

QSARs were constructed for series II and series III analogs both separately and jointly for compounds with both common and distinct heterocycles in each series. The log *P* values of analogs in series III, for the joint QSAR, were adjusted by a constant value of -1.20 to reflect the constant substituent difference of OH (III) and H (II).

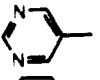
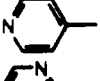

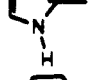
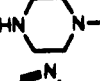
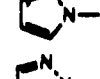
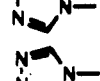

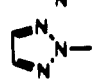
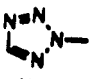
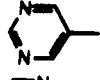

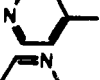
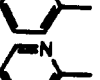

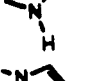
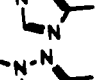
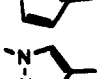
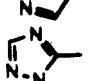
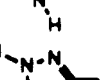
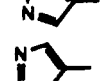

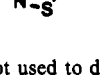

Molecular shape descriptors, namely the common overlap steric volume and the difference in steric volumes, were also computed and considered in the development of QSAR models.

The QSARs were determined by performing multidimensional linear regression analysis using the SAS software¹³ for all combinations of the physicochemical descriptor sets. Outliers were defined as analogs whose activities were incorrectly predicted by more than 25% of the range in the observed -log (EC₅₀) values for the series. Cross-correlation coefficients were used to eliminate collinear descriptors. No collinearity greater than 0.4 was permitted for descriptors in a specific QSAR. Reliability of a QSAR was estimated using the *leave-one-out* cross-validation method.¹⁴

RESULTS

The analogs in the II(H) series of aromatase inhibitors considered in the QSAR analysis are listed in Table 2; those of the III series are listed in Table 4 (OH), and the combined II and III series are listed in Table 5. Also, given in each table are the *D*, θ , ϕ , and log *P* descriptor measures as well as -log (EC₅₀), where EC₅₀ is the molar concentration of analog necessary for 50% *in vitro* inhibition of aromatase activity. The optimum QSARs, for the complete set of analogs, and for selective deletion of outliers, are given in Table 6 for each of the three series. The QSARs are described in terms of the

Table 5. Heterosubstituted (Dichlorophenyl) Methanes II and Methanols III

no.	HET	R	D (Å)	θ (deg)	ϕ (deg)	log <i>P</i>	log (EC ₅₀) obsd	-log (EC ₅₀) pred (eq 10, Table 5)
1		H	3.76	0.0	150.0	0.04	7.26	7.50
2		H	4.31	0.0	179.7	1.16	6.80	6.96
3		H	2.47	108.0	120.0	1.16	<5.30	5.42
4		H	2.53	62.7	123.4	1.26	<5.30	5.62
5		H	4.26	98.3	154.3	-0.10	5.71	<i>a</i>
6		H	2.47	50.2	124.8	0.65	5.70	5.42
7		H	3.63	15.2	159.5	-0.75	8.06	7.50
8		H	3.63	0.0	162.8	-0.91	8.03	7.50
9		H	3.55	46.1	160.2	-0.53	7.73	7.48
10		H	2.47	0.0	123.0	-0.04	<5.30	5.42
11		H	3.54	28.9	159.0		6.83	7.48
12		O H	3.76	0.0	150.0	0.04	7.12	7.50
13		O H	4.27	49.0	178.4	0.04	7.00	7.03
14		O H H	4.31	0.0	179.7	1.16	7.08	6.96
15		O H H	2.47	108.0	120.0	1.16	<5.30	5.42
16		O H	3.76	6.5	146.3	0.04	7.03	7.50
17		O H	2.55	79.8	125.5	0.22	6.13	5.69
18		O H	3.68	50.2	162.5	0.22	7.55	7.51
19		O H	3.60	3.1	161.5	0.59	7.05	7.50
20		O H	3.69	78.2	163.0	0.59	7.52	7.51
21		O H	3.67	66.6	157.2	-0.65	<5.60	<i>a</i>
22		O H	3.64	92.0	162.5	0.12	5.70	<i>a</i>
23		O H	3.78	41.0	160.1	0.71	8.02	7.49
24		O H	3.92	46.6	164.1	0.76	7.72	7.43

^a Outliers not used to derive eq 6 of Table 6.

Table 6. Optimized QSAR Equations Represented in Terms of Regression Coefficients of the Key Descriptors and the Statistics of Fit and Optimum Values of D , D_{opt}^a

eq no.	D^2 (Å ²)	D (Å)	log P	const	n	R^2	F	se	D_{opt} (Å)	outliers	cross-validated R^2
$R = H$											
1	-2.11 (±0.45)	14.81 (±2.95)		-18.35	11	0.83	19.8	0.51	3.51	11	
2	-2.41 (±0.41)	16.82 (±2.72)		-21.48	10	0.89	28.7	0.44	3.49		0.66
$R = OH$											
3		1.00 (±0.38)	-0.53 (±0.41)	3.01	13	0.44	3.9	0.71		3, 4, 7, 10, 11, 12	
4	-1.14 (±0.35)	8.53 (±2.32)		-8.56	11	0.82	17.8	0.36	3.74	10, 11 not used to construct QSAR	0.66
$R = H$											
5	-1.27 (±0.42)	9.27 (±2.78)		-9.75	24	0.56	13.3	0.68	3.65	$R = H$: 5, 7, 8 $R = OH$: 10, 11, 12	
6	-1.41 (±0.25)	10.38 (±1.66)		-11.63	21	0.86	57.3	0.37	3.68	$R = H$: 5 $R = OH$: 10, 11 not used to construct QSAR	0.83

^a n is the number of compounds, se is the standard deviation of fit, the statistical ratio F is used to determine the probability of a linear relationship between the dependent variable and the set of independent variables and the constant, and R^2 is the regression coefficient.

regression coefficients, and the standard deviations of fit of the regression coefficients are given in parentheses for each significant descriptor. The absence of a regression coefficient indicates the descriptor was not significant in the regression fit. The number of compounds considered, n , the square of the correlation coefficient, R^2 , F -statistic, F , standard error, se , the value of D which maximizes $-\log(EC_{50})$, D_{opt} , the outlier(s), and the cross-validated R^2 are given in the columns to the right of the regression coefficients in Table 6. The value of D_{opt} remains nearly constant at 3.6 ± 0.1 Å for five of the six regression equations in Table 6. Equation 3 is not very significant, and only has D as a linear term.

Each of the three "best" equations (2, 4, and 6 of Table 6) for each of the three structure classes are quite significant as measured by R^2 and the cross-validated R^2 . However, there does not appear to be any structural pattern, or physicochemical property, that is common to outliers within, or between, analog classes. In other words, it is not possible to postulate why the outliers do not fit the QSAR.

The dominance of D in the QSARs and the lack of other physicochemical property descriptors is a surprise. In particular, descriptors related to the dipole moment were anticipated to be important in specification of inhibition potency for both the $R = H$ and $R = OH$ analog series.

The conformational analyses of the 5-pyrimidinyl and 2-pyridyl analogs of the $R = H$ (II), OH (III) and Im (VII) series indicate a common global minimum energy conformer state in terms of ψ_1 to ψ_4 . This conformation is virtually identical in overall molecular shape to the crystal structure of compound II of Table 2.⁵ The steric bulk of the substituents about the C_c center markedly limits conformational flexibility so that the global minimum energy conformation is, energetically, the only plausible conformation which can be considered an "active" conformation. Figure 3 specifically displays the superposition of the global minima of compounds 1 and 3 of Table 2, 1 and 4 of Table 4, and the same two corresponding heterocycle analogs in series VII. These additional compounds were considered in the conformational analyses to further establish the existence of a single, restricted conformer state common to series II, III, and VII.

It is important to point out that the global minimum conformer state can adopt different subgeometries. That is, the heterocycle nitrogens and/or protons can assume different

locations and orientations, with respect to the remainder of the molecule, for the same bulk occupancy of the constituent rings in space. Most often, these subgeometries arise from 180° ring-flips with respect to ψ_1 . Subgeometries of the global minimum can also arise through ψ_4 rotations for OH and



the energy differences among these subgeometry states does not exceed 2 kcal/mol in overall range for the six analogs studied.

When the conformational data is combined with the significance of D in the QSARs in Table 6, two well-defined 3D-pharmacophores can be postulated. The two aromatic rings define a rigid, nonpolar molecular shape in space. A negative charge center (the heterocycle nitrogen), which is capable of being involved in a hydrogen bond, is located at one of two specific sites in space relative to C_c and the aromatic rings. In addition, the OH of the methanol series of analogs and the



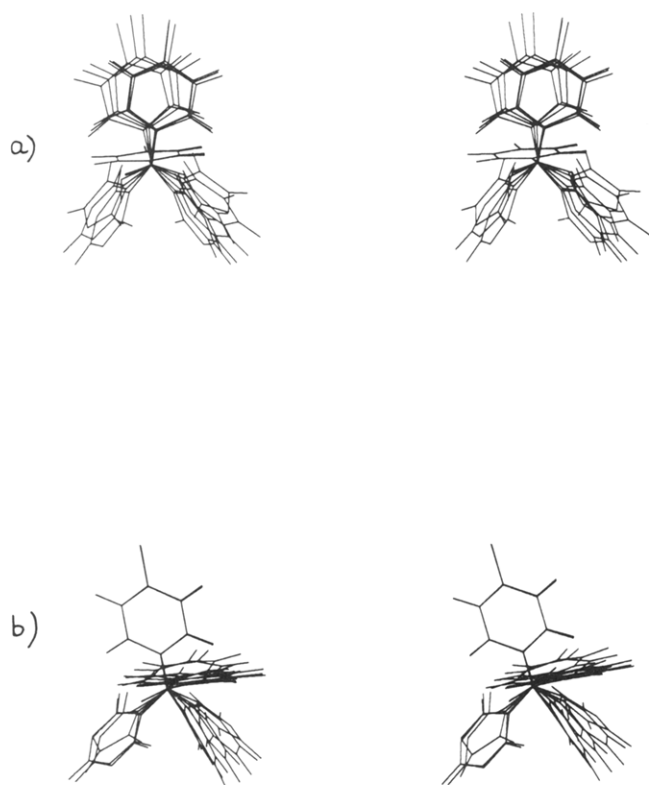
of the VII series indicates an additional site of ligand allowed space. Moreover, a comparison of the inhibition potencies of $R = H$ and $R = OH$ analog pairs having the same, or nearly equivalent heterocycles indicates the $-OH$ has little, if any, effect on activity. This, in turn, suggests there is little interaction between the receptor and the $-OH$ group at this site in the receptor.

Potential energy field difference calculations, as described in the METHODS section, were carried out for the seven methylated heterocycles listed in Table 3. These methylated heterocycles were taken as local potential field models for their corresponding complete class II and class III analogs. Different regions of space about these model compounds, defined in terms of α and β , were explored for d ranging from the ligand steric contact point to 20 Å from the exocyclic carbon shown in Figure 2. The increments in the field probes

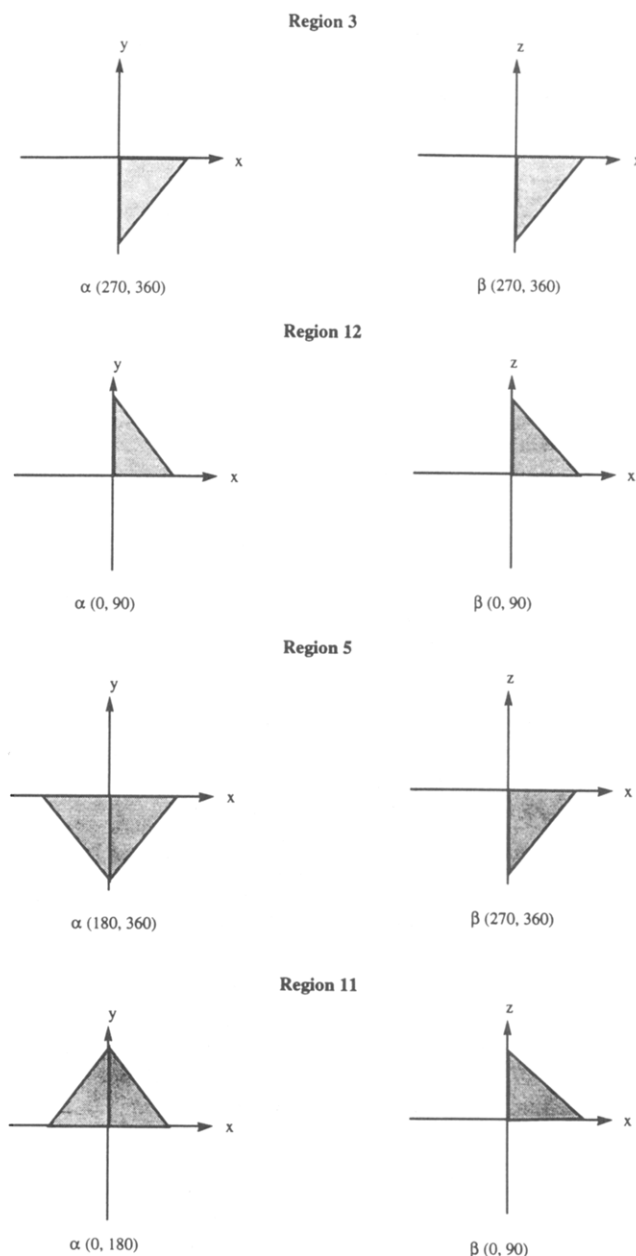
Table 7. Region of Space, in Terms of α and β and Corresponding R^2 Values Sampled in the Integrated Potential Energy Field Correlation Studies

region	α_i	α_f	β_i	β_f	R^2_a	R^2_b
1	0	360	0	360	0.47	0.22
2	180	360	180	360	0.40	0.36
3	270	360	270	360	0.78	0.85
4	270	360	180	360	0.37	0.31
5	180	360	270	360	0.76	0.85
6	180	360	315	360	0.59	0.44
7	0	180	0	360	0.36	0.36
8	0	90	0	360	0.51	0.51
9	0	180	0	180	0.40	0.36
10	0	90	0	180	0.42	0.40
11	0	180	0	90	0.76	0.85
12	0	90	0	90	0.70	0.86

^a R^2_a is the regression coefficient between $-\log(EC_{50})$ and the linear and quadratic terms for ΔF where the values of ΔF were calculated using a particular alignment of structures 9 and 11, namely 9a and 11a as shown in Table 3. ^b R^2_b is identical to R^2_a except alignment of 9b and 11b was used in the calculations.

**Figure 3.** Molecular superposition, based upon matching of equivalent phenyl rings, of the global minimum energy conformations of compounds 1 and 3 of Table 2, compounds 1 and 4 of Table 4, and the same heterocycle analogs in series VII: (a) along the plane of the matched rings and (b) perpendicular to the plane of the matched rings.

were $\Delta\alpha = 5^\circ$, $\Delta\beta = 5^\circ$, and $\Delta d = 0.5 \text{ \AA}$. The results of the different regional probes are given in Table 7 for the two alignments considered. Twelve different regions of space were considered for each alignment and are defined in terms of α_i , α_f , β_i , and β_f . Figure 2 can be used to map out these regions of space relative to the methylated heterocycle of analog 7 of Table 2 which is also the field reference structure. The integrated potential field difference, ΔF , using an oxygen anion, methyl group, and hydrogen ion probe in each of three ΔF calculations, was then correlated with $-\log(EC_{50})$ for the corresponding seven class II analogs. The proton probe gave superior results compared to the other probes, and, therefore, only results using the former are presented. The correlation

**Figure 4.** Two-dimensional Cartesian projections of the (α, β) angles defining the regions of space for which ΔF correlates with $-\log(EC_{50})$. The dark triangles represent the critical wedged-shaped regions of space.

coefficient squared, R^2 , using both linear and quadratic terms for ΔF are reported in Table 7.

An inspection of Table 7 reveals that only four regions of space, (3, 5, 11, and 12) for each of the alignment schemes, yield significant correlations as measured by R^2 . Figure 4 illustrates the two-dimensional projections of the α and β angles into Cartesian space for these four regions of space. It is also clear from Figure 4 that region 5 includes region 3, and that region 11 encompasses region 12. However, the R^2 of regions 3 and 5 and regions 11 and 12 are about the same for both alignments. Therefore, one concludes that the correlations between ΔF and $-\log(EC_{50})$ arise from the field behavior occurring in region 3 and in region 12. Regions 3 and 12 are complementary in space to the two possible locations of the negative charge center identified in the combined study of the QSARs and conformational behavior of the analogs. It is also clear from Table 7 that R^2 is quite sensitive to the region of space in which ΔF is computed.

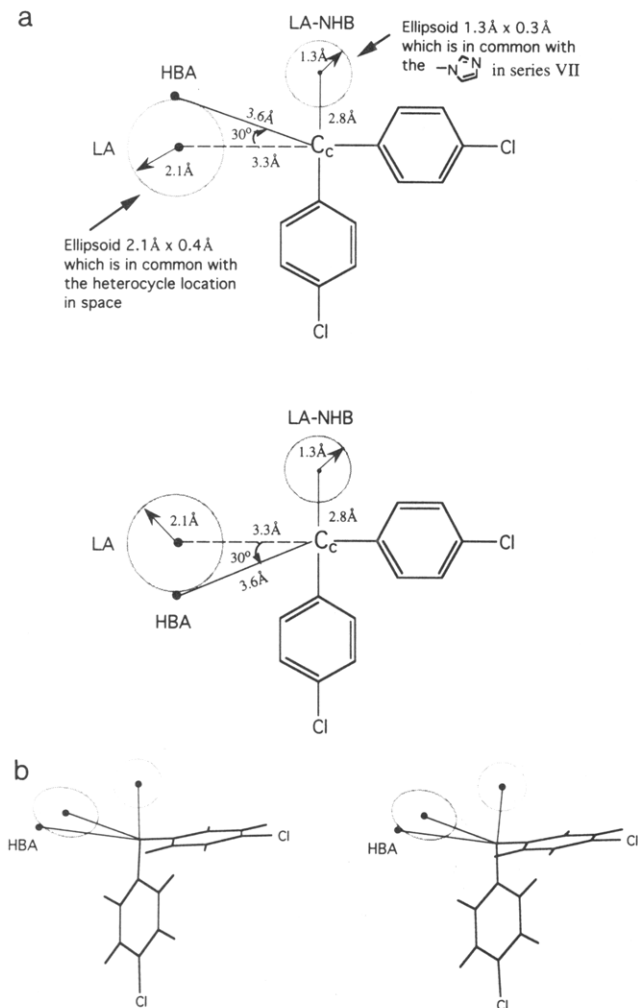


Figure 5. (a) Schematic representation of the two 3D-pharmacophores that are consistent with the QSARs, conformational analyses, and potential field calculations. LA refers to ligand allowed space, HBA defines a hydrogen bond acceptor, and NHB defines a no hydrogen bonding receptor site. (b) Spatial representation of one of the two 3D-pharmacophores in stereo.

DISCUSSION

The findings that D in eqs 1–6 of Table 6 and the ΔF of specific regions in space, see Table 7, correlate with $-\log(EC_{50})$ suggests an intermolecular hydrogen bond (or strong electrostatic interaction) between a nitrogen in the inhibitor heterocycle and the receptor. Both D and ΔF , by the way they are defined, relate a property of a particular site in space to inhibition potency. This interpretation of the QSAR, conformational, and field analyses data is also indirectly supported by the finding that dipole properties of the heterocycle do not correlate with $-\log(EC_{50})$. Dipole descriptors of the heterocycle are not associated with specific locations on the heterocycle. Rather, they reflect the gross spatial and electronic properties of the entire heterocycle.

If the assumption is made that all analogs in Tables 2 and 4 having $-\log(EC_{50}) > 7.0$ sterically fit into the receptor site,

then it is possible to map out a ligand allowed (LA) receptor pocket near the hydrogen bond acceptor (HBA) site defined by D and ΔF . The ellipsoid formed by the union of the heterocycles of these active analogs, in their global minimum energy conformations, contains all centers of the non-hydrogen atoms within $a = 2.1$ Å, $b = 2.1$ Å, and $c = 0.4$ Å of the geometric center of the ellipsoid which is 3.3 Å from C_c .

In like fashion the



of the VII series analogs defines a ligand allowed receptor pocket which is an ellipsoid having $a = 1.3$ Å, $b = 1.3$ Å, and $c = 0.3$ Å and centered about 2.8 Å from C_c along the C_c –N bond vector. Since equivalent II($R = H$) and III($R = OH$) analogs have the same potency, this is a site in the receptor pocket where intermolecular hydrogen bonding does not likely (NHB) occur.

The two candidate 3D-pharmacophores derived in this study for an effective aromatase inhibitor are shown in Figure 5. The structure–activity data does not contain information to resolve where to position the HBA. Nevertheless, LA regions of space are sufficient to size, number, and resolution to permit the construction of many new structural classes that fit one, or both, of the 3D-pharmacophores.

ACKNOWLEDGMENT

This work was carried out using resources of the Laboratory of Molecular Modeling and Design at UIC. J.T. is the recipient of an American Foundation for Pharmaceutical Education Fellowship.

REFERENCES AND NOTES

- (1) *Cancer Res. Suppl.* **1982**, *42*, 3389s–3467s.
- (2) Harris, A. L. *Exp. Cell. Biol.* **1985**, *53*, 1–8.
- (3) Sutherland, C. M.; Muchmore, J. H.; Carter, R. D. *South Med. J.* **1985**, *78*, 987–989.
- (4) Brufman, G.; Biran, S. *Eur. J. Surg. Oncol.* **1985**, *11*, 27–31.
- (5) Jones, C. D.; Winter, M. A.; Hirsch, K. S.; Stamm, N.; Tayler, H. M.; Holden, H. E.; Davenport, J. D.; Krumkalns, E. V.; Suhr, R. G. *J. Med. Chem.* **1990**, *33*, 416–429.
- (6) Hopfinger, A. J.; Burke, B. J. In *Concepts and Applications of Molecular Similarity*; Johnson, M. A., and Maggiora, G. M., Eds.; Wiley Interscience, John Wiley & Sons Inc.: New York, 1990; pp 173–209.
- (7) *3D-QSAR in Drug Design*; Kubinyi, H., Ed.; ESCOM Sci. Pub.: Leiden, The Netherlands, 1993.
- (8) Chemlab-II, version 12.0, Molecular Simulations Inc.: Waltham, MA 1992.
- (9) Burkner, U.; Allinger, N. *Molecular Mechanics*; American Chemical Society: Washington, DC, 1982.
- (10) (a) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902–3909. (b) Stewart, J. J. P. MOPAC molecular orbital package, version 4.0, QCPE Nr. 455, Bloomington, IN.
- (11) Hopfinger, A. J. *J. Med. Chem.* **1983**, *26*, 990.
- (12) MedChem Software Manual, Daylight Inc.: Release 3.51, 1987.
- (13) SAS Institute, Inc. *SAS Users Guide: Basics* SAS Release 5.18, Cary, NC, 1986.
- (14) (a) Principal Data Components, *Users Guide*, 2505 Shepard Blvd, Columbia, MO, 1989. (b) Cramer, R. D., III; Bunce, J. D.; Patterson, D. E. *Quant. Struct. Relat.* **1988**, *7*, 18–25.