

J-CAMD 177

## A comparison of progestin and androgen receptor binding using the CoMFA technique

Deborah A. Loughney\* and Charles F. Schwender

*Department of Medicinal Chemistry, R.W. Johnson Pharmaceutical Research Institute, P.O. Box 300, Route 202, Raritan, NJ 08869, U.S.A.*

Received 31 January 1992

Accepted 20 July 1992

*Key words:* QSAR; CoMFA; Androgen; Progestin

---

### SUMMARY

A series of 48 steroids has been studied with the SYBYL QSAR module using Relative Binding Affinities (RBAs) to progesterone and androgen receptors obtained from the literature. Models for the progesterone and androgen data were developed. Both models show regions where sterics and electrostatics correlate to binding affinity but are different for androgen and progesterone which suggests differences possibly important for receptor selectivity. The progesterone model is more predictive than the androgen (predictive  $r^2$  of 0.725 vs. 0.545 for progesterone and androgen, respectively).

---

### INTRODUCTION

The mechanism of action of steroid hormones has been the topic of much interest over the years. Of particular interest has been the elucidation of the mode of binding of steroids of their receptors. Using in vitro receptor binding screens for the progesterone, estrogen and androgen receptors amongst others, subtle changes in the steroid skeleton have been used to develop SARs (Structure–Activity Relationships) to understand the mode of binding of steroidal ligands to each receptor type [1].

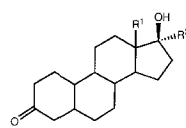
The receptor binding affinity of 48 steroids in the estrogen, progesterone, androgen, mineralocorticoid and glucocorticoid receptor binding assays suggested that a relationship between conformational flexibility and receptor affinity was important [2]. We have incorporated this data set into a QSAR model. The activities chosen for study were progesterone and androgen Relative Binding Affinities (RBAs). Evaluation by the CoMFA technique has led to a better understanding of the molecular requirements for optimal drug–receptor interactions for these steroidal receptors.

---

\*To whom correspondence should be addressed.

The SYBYL QSAR [3] module was used in this study. It incorporates a relatively new technique called Comparative Molecular Field Analysis [4] or CoMFA. The SYBYL QSAR module was used to generate a large table of physical properties which incorporated both steric and electrostatic data (as well as parameters such as logP, molar refractivity, etc.) which was then correlated to the desired activity. The goal of this study was not necessarily to be able to predict androgen or progestin binding but rather to gain additional insight into the structural requirements for binding to both of these receptors which could not be as easily ascertained using traditional modelling or QSAR techniques.

TABLE 1  
RELATIVE BINDING AFFINITIES (RBA) OF A SERIES OF SUBSTITUTED STEROIDS

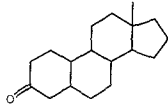
	$\Delta$	R <sup>1</sup>	R <sup>2</sup>	RBA progestin	RBA androgen
Progesterone				100	5.5 ± 0.6
Testosterone				1.0 ± 0.3	100
Nortestosterone	4	CH <sub>3</sub>	H	20 ± 3	154 ± 20
RU-3118	4,9	CH <sub>3</sub>	H	17 ± 2	134 ± 27
RU-2341	4,9,11	CH <sub>3</sub>	H	74 ± 16	197 ± 197
RU-3366	4	C <sub>2</sub> H <sub>5</sub>	H	34 ± 4	126 ± 14
RU-3470	4,9	C <sub>2</sub> H <sub>5</sub>	H	26 ± 2	93 ± 26
RU-4458	4,9,11	C <sub>2</sub> H <sub>5</sub>	H	86 ± 15	172 ± 9
RU-1285	4	C <sub>3</sub> H <sub>7</sub>	H	4.5 ± 0.8	108 ± 22
RU-3484	4,9	C <sub>3</sub> H <sub>7</sub>	H	4.6 ± 1.0	42 ± 9
RU-25923	4,9,11	C <sub>3</sub> H <sub>7</sub>	H	38 ± 7	105 ± 16
RU-598	4	CH <sub>3</sub>	CH <sub>3</sub>	100 ± 10	146 ± 23
RU-3467	4,9	CH <sub>3</sub>	CH <sub>3</sub>	71 ± 8	64 ± 21
RU-1881	4,9,11	CH <sub>3</sub>	CH <sub>3</sub>	208 ± 18	204 ± 5
RU-2309	4,9,11	C <sub>2</sub> H <sub>5</sub>	CH <sub>3</sub>	230 ± 40	143
U-10997 7 $\alpha$ -CH <sub>3</sub>	4	CH <sub>3</sub>	CH <sub>3</sub>	214 ± 33	108 ± 35
RU- 2788 7 $\alpha$ -CH <sub>3</sub>	4,9	CH <sub>3</sub>	CH <sub>3</sub>	198 ± 59	122 ± 7
RU-2420 7 $\alpha$ -CH <sub>3</sub>	4,9,11	CH <sub>3</sub>	CH <sub>3</sub>	306 ± 58	180 ± 23
RU-4841 7 $\alpha$ -CH <sub>3</sub>	4,9,11	C <sub>2</sub> H <sub>5</sub>	CH <sub>3</sub>	236 ± 48	124
Norethindrone	4	CH <sub>3</sub>	C $\equiv$ CH	156 ± 18	43 ± 3
RU-3097	4,9	CH <sub>3</sub>	C $\equiv$ CH	42 ± 6	19 ± 3
RU-2010	4,9,11	CH <sub>3</sub>	C $\equiv$ CH	63 ± 5	70 ± 18
Norgestrel	4	C <sub>2</sub> H <sub>5</sub>	C $\equiv$ CH	170 ± 12	84 ± 18
RU-3714	4,9	C <sub>2</sub> H <sub>5</sub>	C $\equiv$ CH	68 ± 7	41 ± 13
RU-2323	4,9,11	C <sub>2</sub> H <sub>5</sub>	C $\equiv$ CH	76 ± 7	83 ± 3
RU-1364	4	C <sub>3</sub> H <sub>7</sub>	C $\equiv$ CH	73 ± 7	44 ± 3
RU-1475	4,9	C <sub>3</sub> H <sub>7</sub>	C $\equiv$ CH	11 ± 2	10 ± 3
RU-2715	4,9,11	C <sub>3</sub> H <sub>7</sub>	C $\equiv$ CH	61 ± 4	66 ± 6

## METHODS

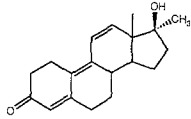
*Relative binding affinities*

The compounds chosen for study are shown in Table 1 and were obtained from the literature [2]. There were 3 different backbone types represented, as well as varying substitution at C-13 and -17 and modifications to the A ring. The progesterone RBAs (progesterone = 100) ranged from 0.6 to 376 while the androgen (testosterone = 100) RBAs ranged from 0.6 to 204. RBA is defined as the  $IC_{50}$  of the ligand divided by the  $IC_{50}$  of the standard (progesterone for the progestins, testosterone for the androgens) multiplied by 100.

TABLE 1 (continued)

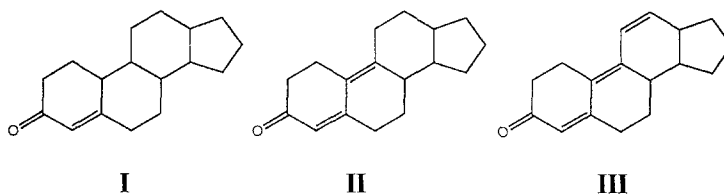
	$\Delta$	R <sup>1</sup>	R <sup>2</sup>	Progestin	Androgen
RU-801	4			190 $\pm$ 34	37 $\pm$ 4
RU-22848	4,9			218 $\pm$ 5	29 $\pm$ 4
RU-2060	4,9,11			274	138 $\pm$ 18
RU-22779	4			252 $\pm$ 50	3.5 $\pm$ 1.3
RU-23747	4,9			226 $\pm$ 47	2.3 $\pm$ 0.6
RU-25051	4,9,11			376 $\pm$ 48	14 $\pm$ 1
Norprogesterone	4	COCH <sub>3</sub>	H	230 $\pm$ 19	6.4 $\pm$ 0.8
H-3163	4,9	COCH <sub>3</sub>	H	181 $\pm$ 18	8.8 $\pm$ 2.5
RU-2061	4,9,11	COCH <sub>3</sub>	H	230 $\pm$ 16	16 $\pm$ 2
H-3510	4	COCH <sub>3</sub>	CH <sub>3</sub>	317 $\pm$ 60	5.5 $\pm$ 1.5
RU-2453	4,9	COCH <sub>3</sub>	CH <sub>3</sub>	230 $\pm$ 18	1.1 $\pm$ 0.4
RU-2236	4,9,11	COCH <sub>3</sub>	CH <sub>3</sub>	220 $\pm$ 27	1.9 $\pm$ 0.2

	Structural modification	Progestin	Androgen
RU-3773	3-deoxo	0.7	0.6
RU-4814	no substituent at C-17	9.9 $\pm$ 2.4	8.5 $\pm$ 4.1
RU-2065	C = O at C-17	0.6	0.6
RU-2956	2-gem-dimethyl	1.0 $\pm$ 0.2	14 $\pm$ 2
RU-4743	6-gem-dimethyl	69 $\pm$ 16	178 $\pm$
RU-2922	2 $\beta$ -methyl	13 $\pm$ 2	66 $\pm$ 8
RU-4089	4-methyl	27 $\pm$ 9	89 $\pm$ 3
RU-2992	A-nor	229 $\pm$ 63	166 $\pm$ 26
RU-2999	2-oxo	262 $\pm$ 20	158 $\pm$ 14

### Model construction

The first and probably most important step in this QSAR study was choosing the conformation and alignment (or fit) of each molecule. The 3 molecules shown below were searched in the Cambridge Structural Database [5]. This was done to determine which backbone conformation was most prevalent in each backbone type. Substructure searches on compound **I** (4-monoene) netted over 100 hits, compound **II** (4,9-diene) netted 19 hits and compound **III** (4,9,11-triene) netted 7 hits.



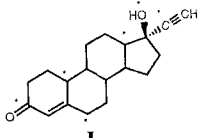
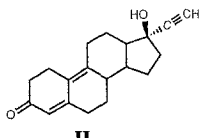
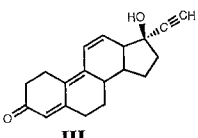
The resulting hits were then narrowed down by substituent. For example, in the monoene series only those compounds with either a 17 $\beta$ -acetyl or a 17 $\beta$ -OH, 17 $\alpha$ -ethynyl substituent were considered. This resulted in the definition of backbone conformations which were used for the rest of the study. Once a representative backbone conformation was found for each structural type its geometry was optimized using the MAXIMIN algorithm in SYBYL.

The next step was the Multifit procedure. The atoms used in the Multifit are marked (\*) below. A spring constant of 20 was used in the flexifit. The energies both before and after the Multifit as well as the RMS (Root-Mean-Square) fits of the Multifit structure to the starting structure are listed in Table 2.

As can be seen by the data in Table 2, two of the compounds decreased in energy after the Multifit while one (the triene) had to invest 0.6 kcal/mol to achieve the fit. The triene also had the least movement of atoms which is in agreement with this slight investment of energy. The energies of all compounds are essentially the same before and after the Multifit and the resulting overlap can be seen in Fig. 1.

Once representative conformers were found for each backbone type and the alignment rule was

TABLE 2  
RESULTS OF MULTIFIT

Compound			
Energy before Multifit (kcal/mol)	1.722	7.413	7.973
Energy after Multifit (kcal/mol)	1.412	6.842	8.547
RMS (Å) (before vs. after)	0.146	0.211	0.075

set, substituents were added to create the 48 steroids included in the analysis. The ethyl and propyl side chains at position 13 were added such that they reproduced the torsions found in the crystal structures. The orientation of the 17 $\beta$ -acetyl group in progesterone used was also that found in the crystal. The hydroxyl hydrogens were oriented such that the sp<sup>3</sup>C-sp<sup>3</sup>C-sp<sup>3</sup>O-H torsion angle was the same as the sp<sup>3</sup>C-sp<sup>3</sup>C-sp<sup>2</sup>C-sp<sup>3</sup>C torsion of the 17 $\beta$ -acetyl group of the other compounds in the series.

#### *Generation of the QSAR table*

The 48 compounds were entered into a QSAR database. Each molecule in the study became a row in the QSAR table. Charges were calculated using MNDO (MOPAC 4.0). A region was defined which encompassed all of the molecules in the study. The coordinates of the lower corner of the region were -10, -8, -9. The coordinates of the upper corner were 10, 8, 11. A grid size of 2 Å was used. An sp<sup>3</sup>C probe atom with a + 1 charge was used to generate the CoMFA columns. The maximum electrostatic and steric energies used were 30 kcal/mol.

Columns were also added for dipole moment (total and X, Y and Z components), logP and molar refractivity (MR). The logP and molar refractivity were calculated by the Medchem software [6]. Finally, columns were added for progesterone RBA and androgen RBA.

#### *3D QSAR analysis*

The table resulting from a CoMFA study is unusual in that it contains many more columns than rows. Therefore, the partial least squares (PLS) method of Wold et al. [7] was used to extract a QSAR. PLS components are extracted as long as the crossvalidated, or predictive  $r^2$  increases. In the crossvalidation technique, the analysis is repeated with a randomly chosen subset of the compound-rows excluded (the number of compounds excluded is determined by the crossvalidation group parameter), and the resulting model is used to 'predict' the biological property value of interest for the excluded compound (see below). This procedure is repeated until every such property value has been 'predicted' by a model from whose derivation it was excluded. The number of crossvalidation groups used in these studies was 48. That is, each compound was predicted from the model that used all other compounds.

It may be apparent that the amount of calculation involved could become very large depending on the parameters used. The minimum sigma value determines which values of the CoMFA columns will be included in the analysis. The default setting is 0, which essentially includes all values.

TABLE 3  
ANALYSIS RUN USING PROGESTERONE RBA AND CoMFA, LOG P, DIPOLE (X, Y, Z) AND MR  
Test on minimum sigma.

Minimum sigma	Crossvalidation groups	Components	No. of lattice points	$r^2$	'Press' <sup>a</sup>	Optimal components
0	48	8	2178	0.718	62.061	7
2	48	8	85	0.719	61.912	7

<sup>a</sup> Predictive sum of squares. The sum, over all compounds, of the squared differences between the actual and 'predicted' biological properties.

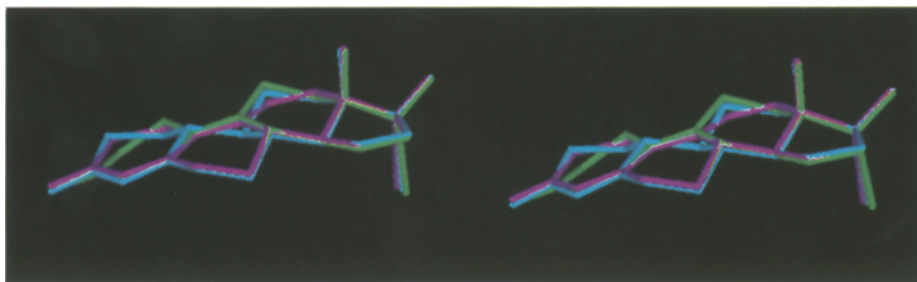


Fig. 1. Overlap of monoene (blue), diene (green) and triene (magenta) structures.

For these crossvalidated runs the minimum sigma was set to 2. That is, any region point column having a variance less than 2 was omitted from the analysis. In order to test whether this option seriously affected the results 2 test runs were performed, first setting the minimum sigma to 0 (the default value) and then to 2.0. The results are tabulated in Table 3.

The CPU time (VAX 8650) for run 1 (minimum sigma equals 0) was 118159.29 seconds. Run 2 (min. sigma = 2), on the other hand, took 392.31 seconds. Using a minimum sigma value of 2.0 gave a 300-fold increase in speed without an appreciable effect on the results, i.e. number of optimal components and  $r^2$ . Therefore, the minimum sigma was set to 2.0 for all of the crossvalidated runs.

TABLE 4  
PROGESTERONE CROSSVALIDATED RUNS

Rows	Columns	$r^2$	'Press'	Optimal components
1:48	CoMFA	0.725	61.214	7
1:48	CoMFA, logP, MR	0.725	61.214	7
1:48	CoMFA, logP, MR, dipole, X, Y, Z	0.725	61.214	7
1:48	CoMFA, dipole, X, Y, Z	0.725	61.214	7
1:48	logP, MR, X, Y, Z, dipole	0.358	91.249	2
1:48	steric	0.733	60.383	7
1:48	electrostatic	0.661	67.993	8
1:48	MR	0.054	105	1
1:48	dipole	-0.081	112	1
1:48	logP	-0.033	109	1

## RESULTS AND DISCUSSION

*Crossvalidated analyses*

Multiple crossvalidated runs were performed using various combinations of columns. This was done for both the progesterone and androgen series. The results are listed in Tables 4 and 5. In all the analyses 48 crossvalidation groups were used and minimum sigma was set to 2.0.

When CoMFA data are used in the analysis, the addition of other factors (such as logP) has no effect on the results. An analysis was done weighting logP by a factor of 10 which also had no effect on the results. This may be due to the fact that the number of CoMFA columns is so large that any effect the other columns would have on the results is masked.

CoMFA data vs. relative binding affinity gives a very good crossvalidated  $r^2$  in the progesterone case (0.725). (Using log RBA for progesterone gave a crossvalidated  $r^2$  of 0.48.) The androgen crossvalidated  $r^2$  is not as high (0.525). Interestingly, if one performs an analysis using sterics and electrostatics independently, one can see that both correlated very nicely in the progesterone model (0.733 and 0.661, respectively), whereas electrostatics correlated better in the androgen case (0.736 vs. 0.491). However, the contours of the standard deviation times the coefficient for the electrostatic field, taken from a non-crossvalidated run using androgen binding and electrostatics alone, were very difficult to interpret as some of the contour lines actually ran through the molecule. This is due to the fact that when one uses pure electrostatics the steric cutoffs are not turned on. Since we could not easily interpret such a result, the use of electrostatics alone was not pursued any further.

TABLE 5  
ANDROGEN CROSSVALIDATED RUNS

Rows	Columns	$r^2$	'Press'	Optimal components
1:48	CoMFA	0.545	50.722	9
1:48	CoMFA, logP, MR	0.545	50.722	9
1:48	CoMFA, logP, MR, dipole, X, Y, Z	0.545	50.722	9
1:48	CoMFA, dipole, X, Y, Z	0.545	50.722	9
1:48	logP, MR, X, Y, Z, dipole	0.074	66.857	1
1:48	steric	0.491	50.850	6
1:48	electrostatic	0.736	36.639	8
1:48	MR	0.074	63.117	1
1:48	dipole	-0.052	67.276	1
1:48	logP	0.076	63.072	1

TABLE 6  
PROGESTERONE AND ANDROGEN CROSSVALIDATED RUN USING RBA AND CoMFA

Progesterone			Androgen		
No. of compounds	Standard error	r <sup>2</sup>	No. of compounds	Standard error	r <sup>2</sup>
1	92.000	0.379	1	63.002	0.298
2	80.626	0.523	2	59.806	0.368
3	73.775	0.601	3	60.242	0.358
4	69.246	0.648	4	57.526	0.415
5	63.959	0.700	5	57.968	0.406
6	62.180	0.716	6	53.604	0.492
7 <sup>a</sup>	61.213	0.725	7	52.210	0.518
8	62.725	0.711	8	51.471	0.532
			9 <sup>a</sup>	50.722	0.545
			10	50.725	0.545
			11	52.394	0.515
			12	52.415	0.514

<sup>a</sup> Optimal number of components, i.e. standard error at a minimum.

The numbers for the standard error were somewhat high; approximately 60 in the progesterone model and 50 in the androgen. However, examination of the original dataset shows that the range in the standard error of the mean (for those compounds with  $n > 1$ ) for the progesterone data is from 0.2 to 63. The range for the androgen data is 0.2 to 30. The average standard error for the progesterone receptor binding data is 20.82, and the standard error for the androgen receptor is 12.41.

Another point to be made about the results is the high number of components calculated as necessary to explain the data. The large numbers (7 and 9 for progesterone and androgen, respectively) indicate that the model developed is a very complex one. Models composed of fewer components generally have better predictive capability. Therefore, the results were analyzed to determine whether a smaller number of components could be used. Table 6 lists the number of components, standard errors and  $r^2$  of the crossvalidated runs for both progesterone and androgen using binding data vs. CoMFA data.

As can be seen in Table 6, the progesterone model is fairly predictive after 3 components ( $r^2 = 0.601$ ), while the androgen model only reaches an  $r^2$  of 0.545 after 9 components.

One of the purposes of the crossvalidated runs is the determination of the optimal number of components to use in the non-crossvalidated runs. While it would have been permissible to calculate the non-crossvalidated model for progesterone using as little as 3 or 4 components, 7 components were used in order to be consistent with the androgen model.

Attempts were made to improve upon the predictive  $r^2$  of the androgen model. If 3 compounds (RU-3773, RU-2922, RU-4089) were omitted, the predictive  $r^2$  could be improved to 0.631 at 7 components. However, for the sake of comparison, all compounds were included and the non-crossvalidated model for androgen was generated using 9 components. That is, for both progesterone and androgen the number of components used in the non-crossvalidated runs was the



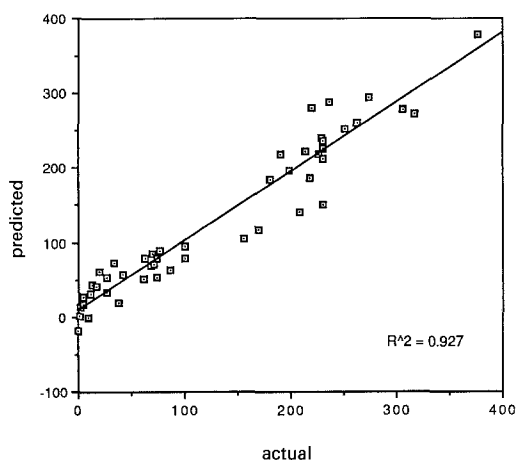


Fig. 2. Progesterone non-crossvalidated model. Relationship between actual and predicted relative binding affinities.

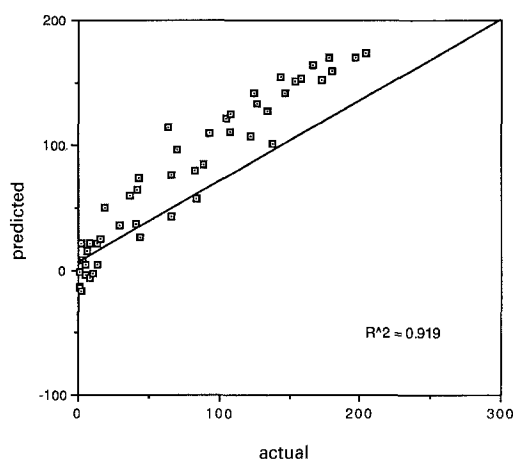


Fig. 3. Androgen non-crossvalidated model. Relationship between actual and predicted relative binding affinities.

number given as optimal in the crossvalidation analysis. Plots of the actual vs. predicted RBAs are shown in Figs. 2 and 3.

#### *Non-crossvalidated analyses*

The QSAR equation derived using CoMFA contains a potentially non-zero coefficient for each column in the data table, which corresponds to 2 coefficients for each lattice point. Because of the correspondence between a coefficient and a spatial position the QSAR equation can be contoured in 3D space to give 'coefficient contour maps'.

Listed in Table 7 are the data from the non-crossvalidated analyses. The field retrieved in these analyses is the standard deviation times coefficient field (STDEV\*COEFF). Graphs of these fields are shown in Figs. 4, 5, 6 and 7.

An important point to keep in mind when interpreting the results of the models developed is that the progesterone model is a more predictive and perhaps a more meaningful model than the androgen model. Recall that the progesterone model had a very acceptable predictive  $r^2$  (0.601) after 3 components, while the androgen model only achieved an  $r^2$  of 0.545 even after 9 compo-

TABLE 7  
NON-CROSSVALIDATED ANALYSES

RBA	$r^2$	Standard error	Components	Field range
Progesterone	0.927	30.859	7	-43.209 to 27.852 <sup>s</sup> -1.865 to 0.612 <sup>e</sup>
Androgen	0.919	20.301	9	-31.364 to 21.452 <sup>s</sup> -6.013 to 0.768 <sup>e</sup>

<sup>s</sup> Steric field.

<sup>e</sup> Electrostatic field.

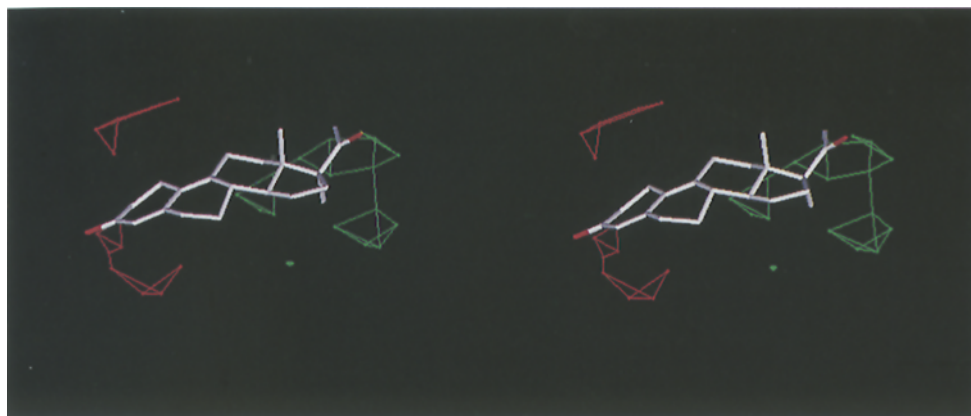


Fig. 4. Contour map of the progestin steric field (standard deviation times coefficient) with RU-2453. The red contour level is at  $-10$  and the green contour level is  $10$ .

nents. While prediction of potencies was not the goal of this work, much information can be gained in terms of model feasibility and scope by the prediction of compounds not included in the model. A close structural analog of RU-2453, RU-5020, (4,9-diene, 13-methyl, 17 $\alpha$ -methyl, 17 $\beta$ -propionyl, RBA = 223 for progesterone and 1.2 for androgen) was available which was not included in the original study. Its RBA was predicted by both the androgen and progesterone models. The progesterone model comprised of 7 components predicted the progesterone RBA to be 215. Another progesterone model comprised of 4 components predicted the RBA to be 213. The androgen model comprised of 9 components predicted the androgen RBA to be  $-13$ . The prediction of additional compounds will help to better define the optimal number of components which should be used in each model.

One can compare the steric and electrostatic contours calculated for the progesterone and androgen models. The steric contours for both the androgen and progesterone models (Figs. 4 and

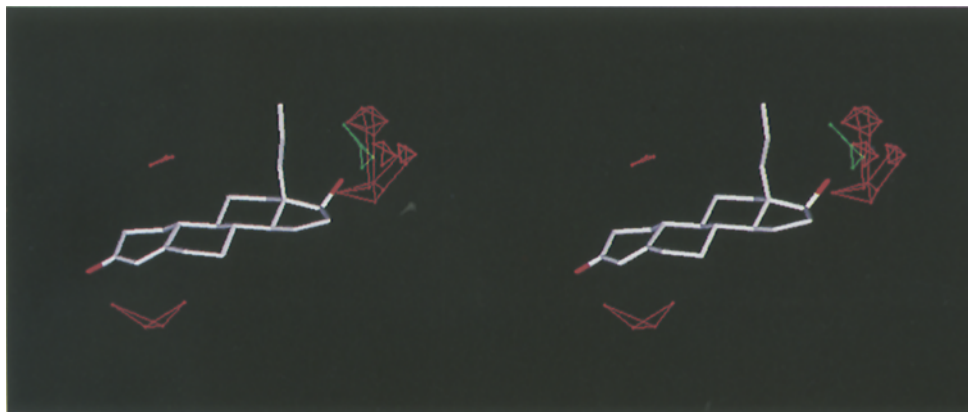


Fig. 5. Contour map of the androgen steric field (standard deviation times coefficient) with RU-1285. The red contour level is at  $-10$  and the green contour level is  $10$ .

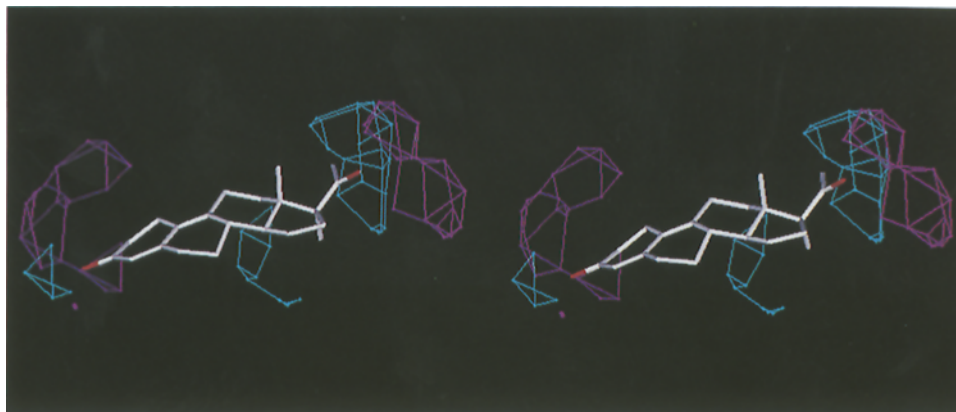


Fig. 6. Contour map of the progesterin electrostatic field (standard deviation times coefficient) with RU-2453. The magenta contour level is at  $-0.5$  and the blue contour level is  $0.5$ .

5) show a negative correlation ( $-10$  coefficient in red) in the region of the A ring. Substitution either above or below the plane of this ring is detrimental to activity. These contours are more widespread in the progesterone case, perhaps indicating that the progesterone receptor is more sensitive to steric bulk in the region of the A ring than is the androgen receptor.

The steric contour for the progesterone model shows a positive correlation to steric bulk in the region of the  $17\alpha$  position (i.e.  $+10$  coefficient in green). In the  $11\beta$ -substituted series of Roussel Uclaf this position has been substituted with large, somewhat linear groups such as substituted ethynyl groups, and progesterone binding potency has been retained or increased relative to H substitution [8]. In the androgen model, however, the 17 position is a region which correlates negatively to sterics (red contours =  $-10$ ). This observation is corroborated by the fact that many steroids which bind to the androgen receptor have relatively small substituents at the 17 position, that is, a hydroxy at the  $17\beta$  position and a H or  $\text{CH}_3$  in the  $17\alpha$  position [9]. Substitution

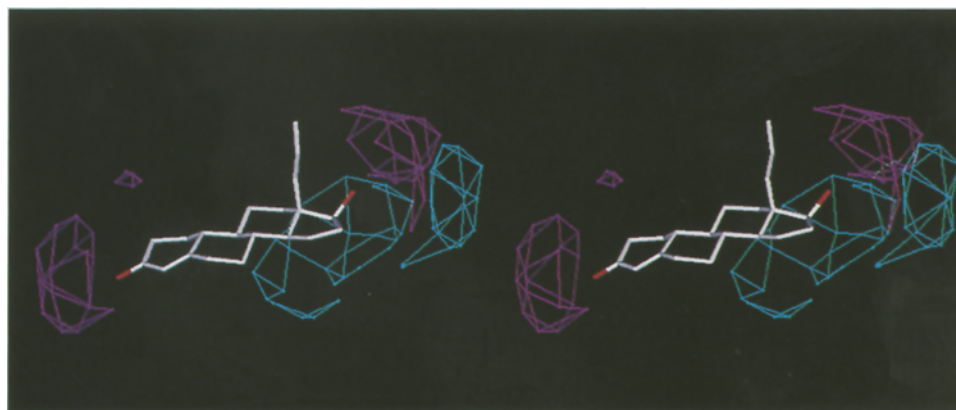


Fig. 7. Contour map of the androgen electrostatic field (standard deviation times coefficient) with RU-1285. The magenta contour level is at  $-0.5$  and the blue contour level is  $0.5$ .

of these steroids with an ethynyl group (i.e. nortestosterone or RU-598 vs. norethindrone) consistently results in a loss in potency. The androgen and progesterone RBAs of compounds with various substituents in the 17 $\alpha$  position would help determine whether or not this observed trend is based entirely on sterics.

The electrostatic contours are more difficult to interpret. Both the androgen and progesterone models have contours (Figs. 6 and 7) corresponding to a favorable interaction (i.e. magenta contours) with negative charge in the area of the D ring. Most progestones and androgens have substituents on the 17 position which have partial negative charges such as acetyl and hydroxy groups. Many hypotheses have been put forth suggesting that the binding of progestones (and androgens) involves the formation of a hydrogen bond at the C-20 substituent [10].

The androgen model has a region corresponding to negative charge which was not found in the progesterone model. This may be an indication that androgen binding is indeed more correlated to electrostatics than is progesterone. (See Table 5 – electrostatic  $r^2$  for androgen = 0.736, overall  $r^2$  = 0.545) The fact that this additional contour is in such close proximity to the molecule may suggest a spatial requirement for negative charge and that this charge should be located near the steroid 17 position for optimal androgen binding. Evidence for this may be found in the fact that the androgen RBA for nortestosterone (17 $\beta$ -OH) is 154, while the RBA for norprogesterone (17 $\beta$ -COCH<sub>3</sub>) is 6.4. The negatively charged carbonyl O in norprogesterone is further removed from the steroid backbone than the hydroxy O in nortestosterone. However, one should also consider the fact that any hydrogen bond donating function at the 17 position is lost in the substitution of a hydroxy with an acetyl which may also account for the loss in potency. It would be interesting to explore differences in electrostatic contours between hydrogen bond donors and acceptors in a separate study to get a better grasp on the interpretation of the electrostatic contours.

It should also be noted that due to the nature of the CoMFA technique, contours will be found in regions where changes in structure are correlated to changes in activity. Therefore, the absence of a contour does not imply that a region is unimportant, rather, it indicates that this region is constant in the data set used.

## CONCLUSION

In conclusion, based on the crossvalidated  $r^2$ , the progesterone model is more predictive than the androgen model. However, it appears that the steric components of both models can be readily interpreted in terms of what is known of the SAR in the literature. The electrostatic components show some interesting differences in the area of the 17 position. This may suggest that the 17 position plays an important role in defining receptor selectivity. However, it cannot be determined from the present study whether the differences found arise from receptor requirements for charge distribution at this position or whether different H-bonding characteristics are implicated. The fact that differences between the androgen and progesterone models were seen at all is encouraging because the structural changes in the steroids used in this study were somewhat subtle.

The prediction of more steroid compounds, the addition of these compounds to the models generated and the inclusion of fields which take into account explicit H-bonding terms will perhaps help to define the present models better. The use of the SYBYL QSAR module has provided us with some insight into the relative relationships between steric and electronic factors in binding

to progesterone and androgen receptors. This information could have potential utility in the design of more receptor-selective binding agents.

## REFERENCES

- 1 Wolff, M.E., In Topliss, J.G. (Ed.) Quantitative Structure-Activity Relationships of Drugs, Academic Press, New York, N.Y., 1983, pp. 351-392.
- 2 Delettre, J., Mornon, J.P., Lepicard, G., Ojasoo, T. and Raynaud, J.P., *J. Steroid Biochem.*, 13 (1980) 45.
- 3 Tripos Associates, St. Louis, Missouri, USA.
- 4 Cramer III, R.D., Patterson, D.E. and Bunce, J.D., *J. Amer. Chem. Soc.*, 110 (1988) 5959.
- 5 Cambridge Crystallographic Data Centre, University Chemical Laboratory, Lensfield Road, Cambridge CB2 1EW, U.K.
- 6 Daylight Chemical Informations Systems, Inc. 18500 Von Karman Ave., Irvine, Ca. 92715, USA.
- 7 Lindberg, W., Persson, J.-A. and Wold, S., *Anal. Chem.*, 55 (1983) 643.
- 8 Teutsch, G., In Baulieu, E. and Segal, S.J. (Eds.) *The Antiprogesterin Steroid RU 486 and Human Fertility Control*, Plenum Press, New York, 1985, pp. 27-47.
- 9 Ringold, H.J., In Villée, C.A. and Engel, I.I. (Eds.) *Mechanism of Action of Steroid Hormones*, Pergamon Press, New York, 1961, pp. 200-234.
- 10 a. Seeley, D.H., Wang, W. and Salhanick, H.A., *J. Biological Chemistry*, 257 (1982) 13359.  
b. Ojasoo, T., Delettre, J., Mornon, J.P., Turpin-VanDyke, C. and Raynaud, J.P., *J. Steroid Biochem.*, 27 (1987) 255.