*Bioorganic & Medicinal Chemistry Letters*, Vol.3, No.6, pp. 1157-1162, 1993
Printed in Great Britain

0960-894X/93 $6.00 + .00


# QSAR STUDY OF BENZOQUINOLINONES AS INHIBITORS OF HUMAN TYPE 1 5-α-REDUCTASE.

James. H. Wikel*, Kerry G. Bemis, James E. Audia, Loretta A. McQuaid, Charles. D. Jones, Pamela A. Pennington, David E. Lawhorn, Kenneth R. Hirsch, and Nancy B. Stamm
Lilly Research Laboratories, Eli Lilly & Company, Indianapolis, Indiana 46285

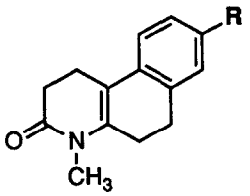
(*Received in USA* 12 *January* 1993)

**Abstract.**
QSAR models have been developed with regression analysis that related a specific lipophilic feature of the substituent of the molecule with biological activity. A second important feature of the compounds, the energy of the HOMO, was revealed using a simple visualization technique.

## Introduction

Testosterone and dihydrotestosterone have substantially distinct physiological roles as principle androgens in humans and other mammalian species.[1] Testosterone possesses anabolic activities leading to increase in bone and muscle mass, as well as effects on sexual differentiation leading to the development of the vas deferens, epididymis and seminal vessicles; whereas dihydrotestosterone mediates the androgenic effects on formation of external genitalia and the prostate.[2] In most species of mammals, dihydrotestosterone is produced from testosterone in an NADPH-dependent reduction catalyzed by steroid 5-α-reductase (EC 1.3.99.5) by local synthesis in the target cells for androgen action. Recently, two human 5-α-reductase isozymes which differ in structure, biochemical properties, patterns of expression and pharmacology, have been identified.[3-5] Although the precise roles of each 5-α-reductase isozyme have yet to be delineated, preliminary findings indicate that both forms are produced in human prostrate.[6] In scalp tissue one form predominates and has been identified as Type 1 5-α-reductase.[7]

The human genital skin fibroblast cell line Hs68 has been shown to express the Type 1 steroid 5-α-reductase isozyme, and has provided an assay system used to identify a series of novel, non-steriodal inhibitors of this enzyme.[6] Herein we discribe a quantitative structure activity (QSAR) study related to this series of benzoquinolinones and their ability to inhibit Type 1 steroid 5-α-reductase in Hs68 cells.




**I, R=H**

The compound dataset consisted of 14 compounds represented by structure I with variations in the phenyl ring substituent (Table I). The compounds were divided into two groups designated Series A and Series B. Series A represented the unreduced 4a,10b ring fusion and consisted of 5 compounds (1-5). Series B represented the reduced ring fusion and was comprised of 9 compounds (6-14). Three of these compounds (12-14) had angular methyl groups at the 10b position while the remainder had a angular hydrogen at this site.

Biological activity was defined as the ability of the compound to inhibit the Type 1 5-α-reductase enzyme in cultured Hs68 human foreskin fibroblast cells. Results were reported as the micromolar concentration required for 50% inhibition (IC50) of enzyme activity. For the QSAR analysis, the activity was expressed as the logarithm of inverse of the IC50 (log 1/IC50) (Table I). Quantum chemistry calculations were performed on a DEC VAX 8600 computer with the MOPAC program Version 4.0 using the AM1 Hamilitonian.[8] All of the structures for MOPAC calculations were geometry preoptimized using the SYBYL 5.3 molecular mechanics program[9]. Molecular mechanics calculations were performed on DEC VAX 8600 attached to an Evans and Sutherland PS390 Graphics terminal. The QM parameters included the LUMO ($E_{LUMO}$)and HOMO ($E_{HOMO}$) energies. In addition to these data, the mathematical sum and difference of the $E_{HOMO}$ and $E_{LUMO}$ values were included representing an excitation energy and an electronegativity value for the molecule. The CLOGP program provided a value for the octanol/water partition coefficient (clogP).[10] Published substituent constants were used for the groups on the phenyl ring.[11] These parameters included values for sigma, pi, F, R, and substituent MR obtained from a database of parameters.[12] The biological activity data and the complete set of parameters for QSAR are given in Table I. The correlation coefficient matrix of pertinent parameters is shown in Table II. Regression analysis was performed on a Apple Macintosh IIci using the program JMP, a statistical visualization program, from the SAS Institute.[13] The size of the dataset required the application of regression analysis using each descriptor listed in Table I singularily and in combination with the squared term. The use of an indicator variable for each of the series did not appreciably alter the models. In the QSAR equations the numbers in parenthesis are the standard errors, n is the number of observations, r is the correlation coefficient, and F is a measure for significance of the equation.

Regression analysis yielded both linear and quadratic correlations between the biological response (log 1/IC50) and the substituent descriptor pi (Equations 1-6) representing the lipophilic property of the substituent. Equations 1 and 4 are for both series together, Equations 2 and 5 represent Series A, and Equations 3 and 6 represent Series B. The entire dataset and Series B can be represented by either the linear or quadratic model. Series A is not well represented by either model.

*QSAR Equation 1:* Predicted $\log(1/IC_{50}) = 6.56(\pm0.19) + 1.21\ (\pm0.37)\ \mathbf{pi}$
(Series A & B) n=14 r=0.69 F=10.75 p=0.01

*QSAR Equation 2:* Predicted $\log(1/IC_{50}) = 6.37(\pm0.35) + 1.08\ (\pm0.63)\ \mathbf{pi}$
(Series A) n=5 r=0.70 F=2.94 p=0.19

*QSAR Equation 3:* Predicted $\log(1/IC_{50}) = 6.64(\pm0.22) + 1.42\ (\pm0.45)\ \mathbf{pi}$
(Series B) n=9 r=0.77 F=10.08 p=0.02

*QSAR Equation 4:* Predicted $\log(1/IC_{50}) = 6.41(\pm0.19) + 4.07\ (\pm1.53)\ \mathbf{pi} - 3.57\ (\pm1.86)\ \mathbf{pi^2}$
(Series A & B) n=14 r=0.78 F=8.42 p=0.01

*QSAR Equation 5:* Predicted $\log(1/IC_{50}) = 6.14(\pm0.37) + 4.14\ (\pm2.51)\ \mathbf{pi} - 3.66\ (\pm2.93)\ \mathbf{pi^2}$
(Series A) n=5 r=0.85 F=2.53 p=0.28

*QSAR Equation 6:* Predicted $\log(1/IC_{50}) = 6.50(\pm0.22) + 4.30\ (\pm1.93)\ \mathbf{pi} - 3.68\ (\pm2.41)\ \mathbf{pi^2}$
(Series B) n=9 r=0.84 F=7.17 p=0.03

Inspection of the *in vitro* biological data clearly indicates that compounds in Series B representing the compounds with reduced ring fusions were more active than their Series A counterpart. Interest was concentrated in Series B type compounds and subsquent to the development of these models, two additional compounds (15 and 16, Table 3) were prepared and evaluated for their *in vitro* biological activity. Equations 1 and 4 were used to predict the biological activity of these compounds (Table 3). A comparision of the validity of the two models was then possible. Table 3 reports the difference between predicted and experimental in vitro response indicating the improved accuracy of the quadratic model represented by Equation 4. In addition, the regression analysis was repeated for the expanded dataset (Equations 7 and 9) and for Series B alone (Equations 8 and 10). A comparision of the linear models for the entire dataset (Equation 1 vs Equation 7) and for Series B alone (Equation 3 vs Equation 8) indicates a modest dimunition in the significance of the linear model. On the other hand the quadratic models remain stable (Equation 4 vs Equation 9 and Equation 6 vs Equation 10).

*QSAR Equation 7.* Predicted $\log(1/IC_{50}) = 6.63(\pm0.19) + 0.87\ (\pm0.32)\ \mathbf{pi}$
(Series A & B) n=16 r=0.59 F=7.46 p=0.02

*QSAR Equation 8:* Predicted $\log(1/IC_{50}) = 6.74(\pm0.24) + 0.81\ (\pm0.39)\ \mathbf{pi}$
(Series B) n=11 r=0.57 F=4.25 p=0.07

*QSAR Equation 9:* Predicted $\log(1/IC_{50}) = 6.44(\pm0.17) + 3.28\ (\pm0.97)\ \mathbf{pi} - 2.58\ (\pm1.00)\ \mathbf{pi^2}$
(Series A & B) n=16 r=0.75 F=8.59 p=0.00

*QSAR Equation 10:* Predicted $\log(1/IC_{50}) = 6.53(\pm0.20) + 3.64\ (\pm1.11)\ \mathbf{pi} - 2.94\ (\pm1.10)\ \mathbf{pi^2}$
(Series B) n=11 r=0.80 F=7.12 p=0.02

Prior to regression analysis in our QSAR study, we visualized each variable's relationship to the biological activity using 2-D plots as shown in Figure 1. This visualization procedure facilitated the rapid identification of an intriguing relationship between the energy of

the highest occupied molecular orbital ($E_{HOMO}$) and the biological activity. All of the compounds in the more active Series B have a higher $E_{HOMO}$ than the corresponding compound in Series A. The increase in energy spanned a narrow range (0.90 to 1.17 ev) and was accompanied by a more active compound *in vitro*. The difference in biological response between the two groups (5 pairs) was significant ($p=0.05$) based upon a paired t-test.

The results of this QSAR study also suggest the importance of the lipophilic effects of the aromatic ring substituent. A quadratic model was found to best describe this relationship. Since we were unable to define a similar relationship with the calculated octanol/water partition coefficient (clogP) describing the lipophilic property of the whole molecule, we would postulate that the substituent interacts with a local lipophilic pocket in the enzyme. This descriptor of the lipophilic property of the substituent can be useful from a predictive point since the values were obtained from a table in a textbook or within a database. Therefore, it is possible to use the QSAR equation to suggest or prioritize possible substituents. Equation 10 was used for this purpose using the values for the descriptor pi. Approximately 100 substituents were screened using QSAR Equation 10 and the results for the predicted biological activity were sorted. Examination of the data suggest that the optimum activity may reside in the property space around the chlorine substituent. Other substituents appeared in the list and may be worth consideration for reasons other than intrinsic potency described by the present QSAR.

A second important feature was identified by comparing the two series in a visual inspection of the 2-D relationship of the $E_{HOMO}$ and the biological activity. The importance of the HOMO is warranted in this series since this molecular orbital would be involved in a reduction reaction as performed by this enzyme if these agents served as substrates. The prior chemical reduction of the double bond in Series A may remove the most readily accessible orbital making the reduction more difficult by approximately 23 kcal/mol (1 ev). Alternatively the location of the HOMO in Series B may not be properly located for rapid reduction by the enzyme. It remains to be determined as to the ability of these compounds to function as substrates as well as inhibitors.

In conclusion, we have identified two molecular features in a series of benzoquinolinones that were correlated with their *in vitro* ability to inhibit the Type 1 5-α-reductase enzyme in clutured Hs68 human foreskin fibroblast cells. One feature, the substituent lipophilic property, was identified by regression analysis and describes a localized feature of the structure around the aromatic ring substituent. A second feature, the energy of the HOMO, was revealed from a visualization technique and was descriptive a global molecular feature that differentiated the two structural subtypes in this study. This knowledge provides a basis for continuing studies.

**Table 1. Biological activity data and the parameters used for QSAR.**

| ID | Series | Substituent | BR[a] | clogP[b] | F | cMR | R | pi | $E_{HOMO}$ | E LUMO | HL Gap[c] |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | A | 8-F | 6.82 | n a | 0.43 | 0.92 | -0.34 | 0.14 | 8.3 | -0.41 | 8.72 |
| 2 | A | 8-CH3 | 7.49 | n a | -0.04 | 5.65 | -0.13 | 0.56 | 8.15 | -0.18 | 8.33 |
| 3 | A | 8-Br | 7.22 | n a | 0.44 | 8.88 | -0.17 | 0.86 | 8.38 | -0.49 | 8.86 |
| 4 | A | 8-H | 6.00 | n a | 0.00 | 1.03 | 0.00 | 0.00 | 8.23 | -0.18 | 8.41 |
| 5 | A | 8-Cl | 6.76 | n a | 0.41 | 6.03 | -0.15 | 0.71 | 8.32 | -0.43 | 8.75 |
| 6 | B | 8-Cl ang H | 8.10 | 2.3 | 0.41 | 6.03 | -0.15 | 0.71 | 9.49 | -0.09 | 9.58 |
| 7 | B | 8-CH3 ang H | 7.49 | 2.6 | -0.04 | 5.65 | -0.13 | 0.56 | 9.27 | 0.23 | 9.04 |
| 8 | B | 8-H ang H | 6.25 | 1.9 | 0.00 | 1.03 | 0.00 | 0.00 | 9.47 | 0.25 | 9.22 |
| 9 | B | 8 Br ang H | 7.46 | 2.8 | 0.44 | 8.88 | -0.17 | 0.86 | 9.52 | 0.18 | 9.34 |
| 10 | B | 8-F ang H | 7.46 | 2.1 | 0.43 | 0.92 | -0.34 | 0.14 | 9.45 | -0.09 | 9.54 |
| 11 | B | 8-OCH3 ang H | 6.94 | 1.8 | 0.26 | 7.87 | -0.51 | -0.02 | 9.02 | 0.23 | 8.79 |
| 12 | B | 8-F ang CH3 | 6.94 | 2.6 | 0.43 | 0.92 | -0.34 | 0.14 | 9.47 | -0.06 | 9.53 |
| 13 | B | 8-Cl ang CH3 | 7.49 | 3.2 | 0.41 | 6.03 | -0.15 | 0.71 | 9.46 | -0.07 | 9.53 |
| 14 | B | 8-H ang CH3 | 6.00 | 2.4 | 0.00 | 1.03 | 0.00 | 0.00 | 9.41 | 0.28 | 9.14 |

a) Biological Response expressed as log (1/IC50) where IC50 was the micromolar concentration of the compound required for 50% inhibition of Type I 5-a-reductase enzyme activity in cultured Hs68 human foreskin fibroblast cells. b) na indicates the value was not available from the program. c) HL GAP = $E_{HOMO}$ - E LUMO

**Table II. Correlation matrix of QSAR parameters for Compounds 1-14.**

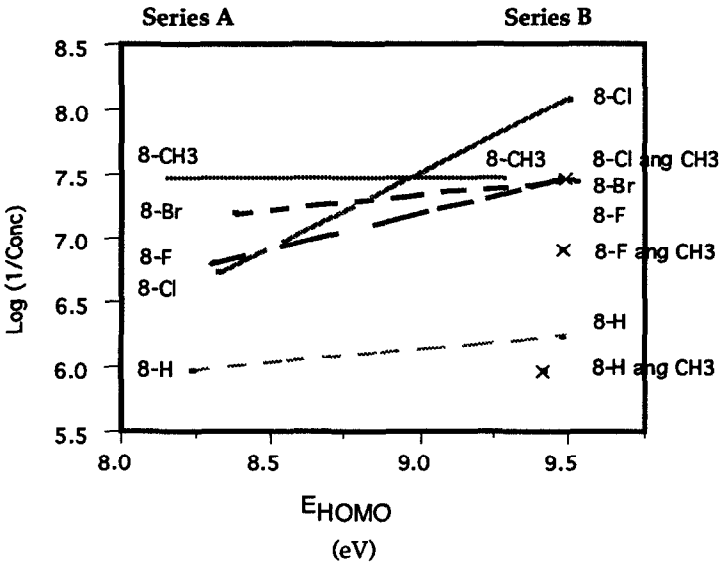
| Variable | pi | E HOMO | E LUMO | HL Gap | F | R | cMR |
|---|---|---|---|---|---|---|---|
| pi | 1.00 | -0.06 | -0.35 | 0.13 | 0.39 | 0.15 | 0.77 |
| E (HOMO) | | 1.00 | 0.72 | 0.90 | 0.13 | -0.04 | -0.08 |
| E (LUMO) | | | 1.00 | -0.35 | -0.47 | 0.11 | -0.07 |
| HL Gap | | | | 1.00 | 0.47 | -0.12 | -0.06 |
| F | | | | | 1.00 | -0.56 | 0.28 |
| R | | | | | | 1.00 | -0.15 |
| cMR | | | | | | | 1.00 |

**Table 3. Predicted and experimental biological activity data.**

| ID | Series | Substituent | BR[a] | Predicted BR (Eq 1) | Error[b] | Predicted BR (Eq 4) | Error[b] | pi |
|---|---|---|---|---|---|---|---|---|
| 15 | B | 8-I ang H | 7.10 | 7.92 | 0.82 | 6.49 | 0.58 | 1.12 |
| 16 | B | 8-CF3 ang H | 7.05 | 7.63 | 0.58 | 7.22 | 0.17 | 0.88 |

a) Biological Response expressed as log (1/IC50) where IC50 was the micromolar concentration of the compound required for 50% inhibition of Type I 5-a-reductase enzyme activity in cultured Hs68 human foreskin fibroblast cells; b) error difference between actual BR and predicted BR.

**Figure 1.**




**References and Notes**

1. Wilson, J. D. *Handbook of Physiology: Endocrinology;* Greep, R. O.; Astwood, E. B. Eds. Am. Physiol. Soc:Washington D. C., 1975; pp. 491-508.

2. Griffen, J. E.; Wilson, J. D.; *The Metabolic Basis of Inherited Diseases;* Seriver, C. R.; Beaudet, A. L.; Sly, W. S.; Valle, D., Eds.; McGraw-Hill, New York, NY, 1989; pp. 1919-1944.

3. Andersson, S.; Rissell, D. W. *Proc. Natl. Acad. Sci. USA* **1990,** 87, 3640.

4. Andersson, S.; Berman, D. M.; Jenkins, E. P.; Russell, D. W. *Nature* **1991,** 354, 159.

5. Jenkins, E. P.; Andersson, S.; Imperato-McGinley, J.; Wilson, J. D.; Russell, D. W. *J. Clin. Invest.* **1992,** 89, 293.

6. Hirsch, K. S.; Jones, C. D.; Audia, J. E.; Andersson, S.; McQuaid, L.; Stamm, N. B.; Pennington, P.; Russell, D. W., Newbauer, B. L., Toomey, R. E. *Proc. Natl. Acad. Sci. USA,* in press.

7. Itami, S.; Kurata, S.; Sonoda, T.; Takayasu, S. *J. Invest. Dermatol.* **1991,** 96, 57.

8 J. J. P. Stewart, *QCPE Bull.* **1989,**9, 10. QCPE Program 255, MOPAC Version 4.0.

9. SYBYL Software, Tripos Associates, St. Louis, MO USA.

10. a) A. Leo, Environmental Health Perspectives **1985,** 61, 275. b) J. Pharm. Sci. 1987, **76,** 166; Pomona College Medchem Software Manual, Version 3, 1984

11. Hansch, C.; Leo, A. *Substituents Constants for Correlation Analysis in Chemistry and Biology.* John Wiley & Sons, New York, 1979.

12. Boyd, D. B.; Seward, C. M. *QSAR: Rational Approaches to the Design of Bioactive Compounds.*; Silipo, C.; Vittoria, A., Ed.; Elsevier Science Publishers B. V.: Amsterdam, 1991.; pp. 167-170.

13. JMP, Statistical Visualization Software, Version 2.0.2, SAS Institute, Cary, NC, USA.