European Journal of Medicinal Chemistry 44 (2009) 4920–4925

Contents lists available at ScienceDirect

European Journal of Medicinal Chemistry

journal homepage: http://www.elsevier.com/locate/ejmech

Original article

Self organizing molecular field analysis on a series of human 5α-reductase inhibitors: Unsaturated 3-carboxysteroid

Suresh Thareja [a], Saurabh Aggarwal [a], T.R. Bhardwaj [a,b], Manoj Kumar [a,*]

[a] University Institute of Pharmaceutical Sciences, Panjab University, Chandigarh 160 014, India
[b] I.S.F College of Pharmacy, Ferozepur Road, Moga 142 021, India

ARTICLE INFO

Article history:
Received 21 July 2009
Accepted 10 August 2009
Available online 14 August 2009

Keywords:
BPH
3D-QSAR
SOMFA
Steroid
Testosterone

ABSTRACT

Steroidal 5α-reductase is a NADPH dependent enzyme that catalyzes the irreversible conversion of 4-en-3-oxo-steroid testosterone to the corresponding 5α-H-3-oxo-steroid dihydrotestosterone thus involved in Benign Prostatic Hyperplasia (BPH). As the crystal structure of target enzyme is not available; we have carried out ligand based designing using Self Organizing Molecular Field Analysis (SOMFA). SOMFA, a novel 3D-QSAR methodology used in present case to study the correlation between molecular properties and human 5α-reductase inhibitory activities of a series of unsaturated 3-carboxysteroid. The statistical results, good cross-validated r^2_{cv} (0.693) and non cross-validated r^2 (0.732), showed satisfied predictive ability. All analysis of SOMFA model may provide some useful information in the design of human steroidal 5α-reductase inhibitors with better spectrum of activity.

© 2009 Elsevier Masson SAS. All rights reserved.

1. Introduction

Benign prostatic hyperplasia (BPH) is a non-malignant enlargement of the prostate gland. It is caused by increase in the number of stromal and epithelial cells, resulting in the obstruction of proximal urethra, thus causes urinary flow disturbances [1]. Testosterone (**A**), an androgen plays a major role in prostate growth. Within the prostate, testosterone gets converted to a more powerful androgen, dihydrotestosterone (DHT) (**B**). DHT stimulates cell growth in the tissue (the glandular epithelium) that lines the prostate gland and is the major cause of the rapid prostate enlargement that occurs between puberty and young adulthood. DHT is a prime suspect in prostate enlargement in later adulthood [2]. Steroid 5α-reductase enzyme (3-oxo-steroid-4-ene dehydrogenase) is a membrane bound NADPH-dependent enzyme responsible for the conversion of testicular testosterone into dihydrotestosterone [3]. Thus 5α-reductase dictates the cellular availability of dihydrotestosterone to prostatic epithelial cells and consequently modulates its growth. Therefore, inhibition of 5α-reductase is a logical treatment for benign prostatic hyperplasia [4].

5α-Reductase inhibitors suppress the dihydrotestosterone concentration by blocking the enzyme (Fig. 1) thus shrinking the size of prostate and ultimately provide relief from the symptoms related to the static mechanical obstruction caused by BPH. Further, the rationale for use of 5α-reductase inhibitors is rooted in the observation that these are more specific to DHT androgens action without affecting/lowering testosterone level, thus capable of decreasing long term side effect of castration due to loss of testosterone without compromising the efficacy of hormonal therapy [5,6].

Hundreds of steroidal and non-steroidal inhibitors ranging from classical, reversible and irreversible inhibitors, and transition state analogues to mechanism-based analogues have been synthesized during last two decades (Fig. 2). Finasteride (MK-906) (**C**) was the first 5α-reductase inhibitor approved in U.S. for the treatment of benign prostatic hyperplasia [7]. Dutasteride (**D**) is another related drug approved by U.S. FDA in 2002, for the symptomatic treatment of BPH [8]. Epristeride (SK&F 105657) (**E**), a novel 5α-reductase inhibitor, is an interesting drug in the treatment of BPH. It belongs to class of carboxysteroid [9]. It has been shown to be an uncompetitive inhibitor against both testosterone and NADPH. Its inhibitory action results from a preferential association to an enzyme binary complex containing NADP and hence increases in testosterone concentration does not overcome its inhibition. It is a specific inhibitor of type 2 5α-reductase isoenzyme. It also attenuates the growth rate of some androgen responsive prostate cancers [10].

Recently, Holt et al. [11] reported a series of unsaturated-3-carboxysteroids (**1**–**23**) those were designed because of presumably favorable electrostatic interaction between the carboxylate and the positively charged oxidized cofactor, the steroidal acrylates

* Corresponding author. Tel.: +91 172 2534115; fax: +91 172 2541142.
E-mail address: manoj_uips@pu.ac.in (M. Kumar).

0223-5234/$ – see front matter © 2009 Elsevier Masson SAS. All rights reserved.
doi:10.1016/j.ejmech.2009.08.004

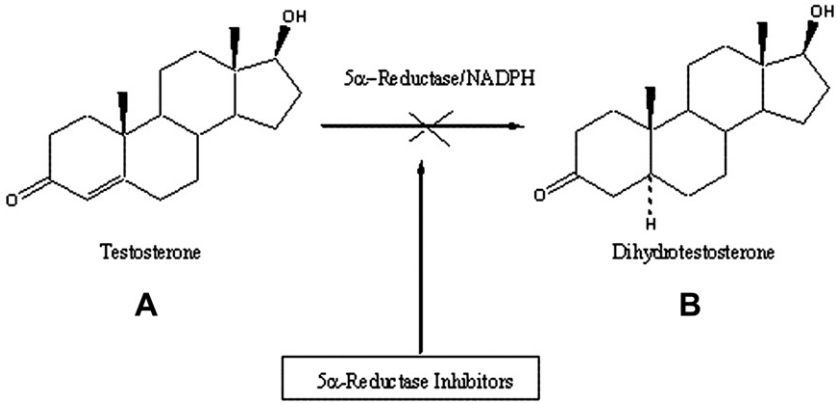

Fig. 1. Site of action of 5α-reductase Inhibitors.

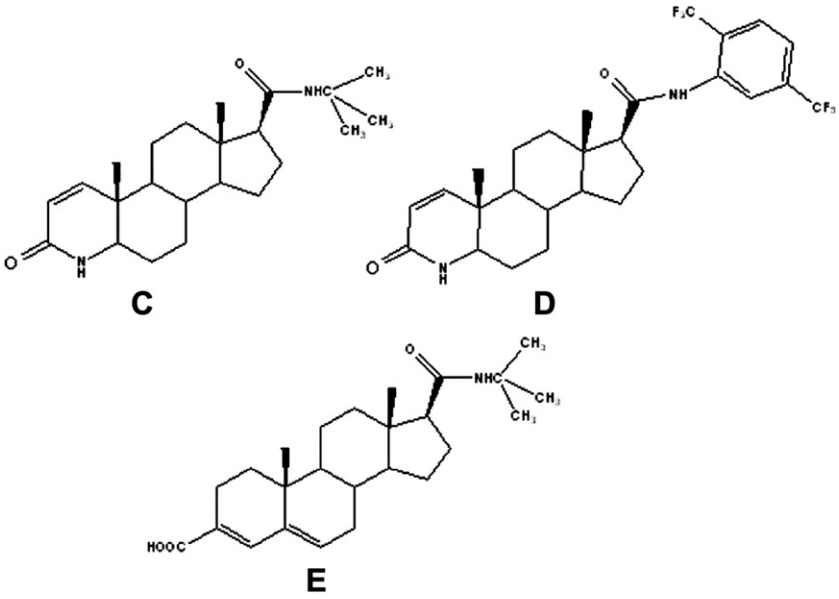

Fig. 2. Clinically used 5α-reductase Inhibitors.

preferentially binds in a ternary complex with enzyme and $NADP^+$, which leads to the observed uncompetitive kinetics. As the crystal structure of target enzyme, i.e. human 5α-reductase is not available therefore we have used SOMFA technique on this series of unsaturated 3-carboxysteriod in order to design new and effective inhibitors of human 5α-reductase.

Robinson et al. has reported a novel three-dimensional quantitative structure activity relationship (3D-QSAR) technique called Self-organizing molecular field analysis (SOMFA) [12]. This method has similarities to both comparative molecular field analysis (CoMFA) and molecular similarity studies [13]. Like CoMFA, it is also grid-based approach however, no probe interaction energies are required to be calculated. Like the similarity methods it generates intrinsic molecular properties, such as the molecular shape and electrostatic potential, which are used to develop QSAR models [14].

A successful 3D-QSAR model not only helps in better understanding of the structure–activity relationships of any class of compounds, but also provides researcher an insight at molecular level about the lead compounds for further developments. A SOMFA model could suggest a method of tackling the all important alignment, which all 3D-QSAR methods have faced. The inherent simplicity of this method allows the possibility of aligning the training compounds as an integral part of the model derivation process and of aligning prediction compounds to optimize their predicted activities [15].

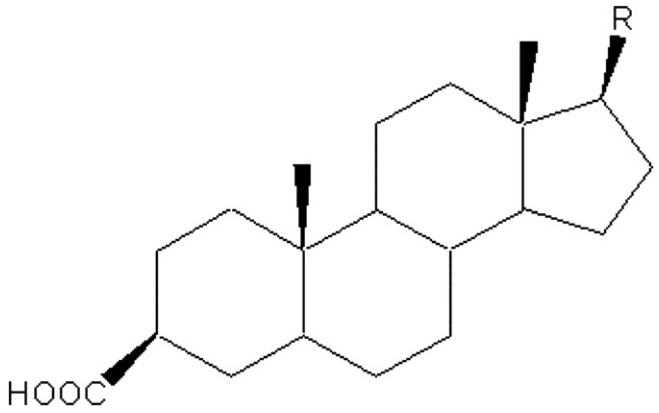

Fig. 3. General Structure of unsaturated 3-carboxysteroid derivatives.

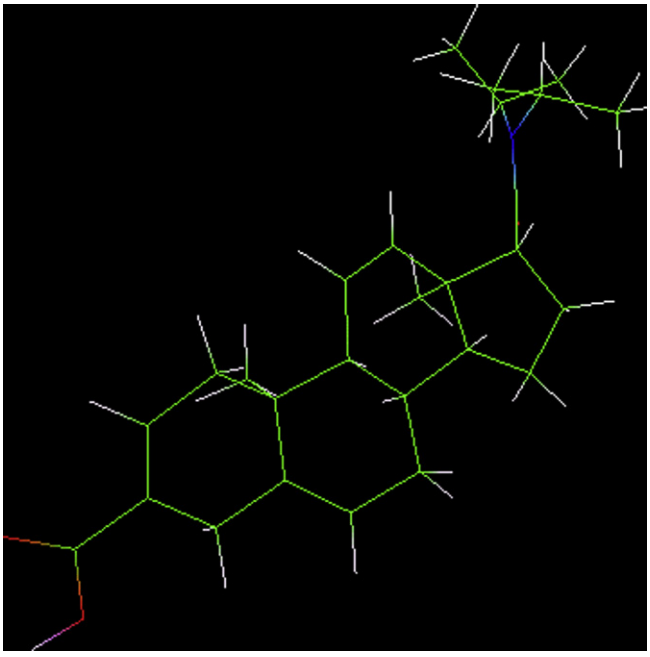

Fig. 4. Template used for alignment.

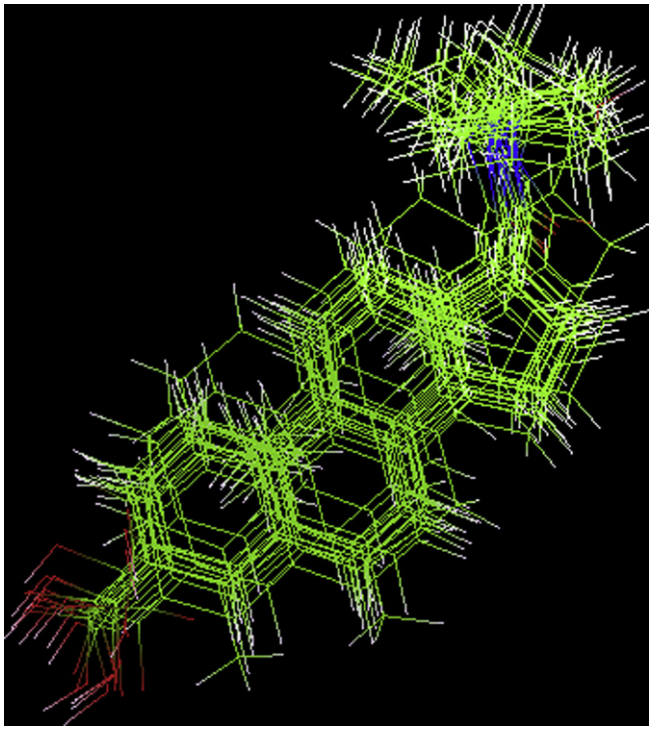

Fig. 5. Superimposition of compounds on template.

The main objective of present work was to carry out SOMFA analysis on a series of unsaturated 3-carboxysteroids to investigate optimal molecular architecture required for designing new specific inhibitors of human 5α-reductase for the treatment of benign prostate hyperplasia.

2. Results and discussion

QSAR analysis in computational research is responsible for the generation of models to correlate biological activity and physico-chemical properties of a series of compounds. The underlying assumption is that the variations of biological activity within a series can be correlated with changes in measured or computed molecular features of the molecules. In the present studies SOMFA was employed for the analysis with the training set composed of 18 compounds whose biological activities are known to find out molecular features responsible for optimum biological activities. Statistical results of SOMFA models obtained by PLS analysis i.e. cross-validated value r^2_{cv} (q^2) serves as a quantitative measure of the predictability of the SOMFA [16].

It is evident from Table 2 that the results are less sensitive to resolution of grid. The best SOMFA (Model I) obtained using template-based alignment shows good cross-validated correlation coefficient r^2_{cv} (q^2) (0.693) and non cross-validated correlation

Table 1
Actual and SOMFA predicted activities of training and test set.

Compound number	Unsaturation	Substitution	R	Observed activity	Predicted activity	Residual activity
1	3–4	–	–CON(iPr)$_2$	−1.477	−1.619	0.142
2^T	3–4	–	–CONH(tBu)	−2.041	−1.84	−0.201
3^T	3–4,5–6	–	–CON(iPr)$_2$	−1.097	−1.868	0.771
4	3–4,5–6	–	–CONH(tBu)	−1.519	−1.819	0.3
5	2–3	–	–CON(iPr)$_2$	−1.929	−1.568	−0.361
6	–	(3β)	–CON(iPr)$_2$	−3.342	−2.736	−0.606
7	2–3,4–5	–	–CON(iPr)$_2$	−1.716	−1.9	0.184
8	4–5	(3α)	–CON(iPr)$_2$	−2.756	−2.983	0.227
9	4–5	(3β)	–CON(iPr)$_2$	−2.301	−2.979	0.678
10	4–5	(3β)–OH	–CON(iPr)$_2$	−2.342	−2.294	−0.048
11	4–5	(3α)–OH	–CON(iPr)$_2$	−2.505	−2.146	−0.359
12^T	1–2,3–4	–	–CON(iPr)$_2$	−1.792	−1.819	0.027
13	1–2,3–4,5–6	–	–CON(iPr)$_2$	−1.778	−1.861	0.083
14	2–3,4–5,6–7	–	–CON(iPr)$_2$	−0.978	−1.664	0.686
15	3–4,5–6,11–12	–	–CON(iPr)$_2$	−0.845	−1.309	0.464
16^T	3–4	4-F	–CON(iPr)$_2$	−1.415	−1.558	0.143
17^T	3–4,5–6	6-F	–CON(iPr)$_2$	−1.505	−1.619	0.114
18	3–4,5–6	4–CH_3	–CON(iPr)$_2$	−1.544	−1.53	−0.014
19	3–4,5–6	6–CH_3	–CON(iPr)$_2$	−2.23	−1.785	−0.445
20	3–4,5–10	19-nor	–CON(iPr)$_2$	−2.041	−1.739	−0.302
21	3–4,5–6	19-nor	–CON(iPr)$_2$	−1.699	−1.468	−0.231
22	3–4	–	−20(*S*)CH_3CHCH_2OH	−3.699	−3.176	−0.523
23	3–4,5–6	–	–CN	−2.898	−3.023	0.125

T-Test set.

Table 2
PLS statistical results of SOMFA.

Parameter	Resolution 1.0 Å (Model I)	Resolution 0.5 Å (Model II)
r^2	0.732	0.731
$r^2_{CV}(q^2)$	0.693	0.691
S	0.403	0.404
F-test	43.816	43.551
r^2_{pred}	0.641	0.644

coefficient r^2 values (0.732) which show significant statistical correlation and satisfied predictive ability.

During the SOMFA investigation, grid spacings of 1 and 0.5 Å were investigated. The 1 Å grid spacing produces a good correlation equal to 0.5 Å grids. This has been improved marginally with the 1 Å spacing used for the results presented here. Further increases in resolution have produced small improvement in model quality but not enough to warrant the extra computational time.

The observed and predicted activities of the training set are reported in Table 1 using best Model 1. Figs. 6a and 7 showed a satisfied linear correlation and moderate difference between actual and predicted values of molecules in the training set.

The best way to validate a 3D-QSAR model is to predict biological activities for some compounds of test set [17]. The SOMFA analysis of the test set composed of 5 compounds is reported in Table 1. Most of the compounds in test set show good correlation between actual and predicted values Figs. 6b and 8.

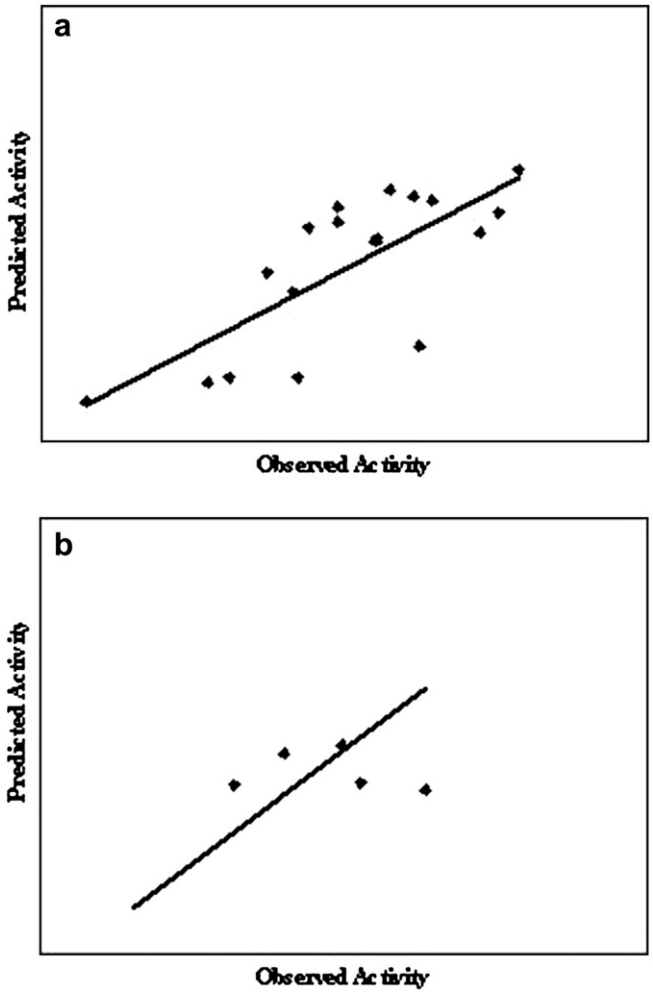

Fig. 6. Graph of Actual vs. Predicted activities for training and test set molecules from the best predictive SOMFA model. (a) Training Set (b) Test Set.

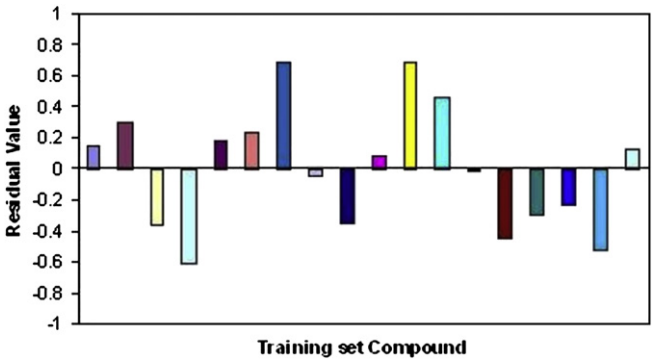

Fig. 7. Histogram of SOMFA residual value for training set.

SOMFA calculation for both shape and electrostatic potentials were performed. The SOMFA shape and electrostatic potential have been presented as master grids (Figs. 9 and 10) at different resolutions of grid.

The master grid maps derived from the best model were used to display the contribution of shape and electrostatic potential. The master grid maps gave a direct visual indication regarding structural features responsible to differentiate the activities of compounds in the training set under study. The master grid also offered an interpretation as to design and synthesize some novel compounds with much higher activities. Each master grid map was colored in two different colors for favorable and unfavorable effects. In other words, the electrostatic features were red (more positive charge increases activity, or more negative charge decreases activity) and blue (more negative charge increases activity, or more positive charge decreases activity), and the shape features are red (more steric bulk increases activity) and blue (more steric bulk decreases activity), respectively.

The SOMFA electrostatic potential map shows some important features. A high density of blue points around the C-3 of steroid indicating presence of electronegative groups whereas high density of red points are also in the neighborhood indicated some electropositive groups where as at substituent 'R' presence of blue points indicating presence of electronegative groups as well few red points indicating presence of electropositive groups are favorable for their optimal activity. Meanwhile, in the map of shape master grid, a high density of red points around the substituent 'R' suggests a favorable steric interaction; simultaneously, we also find few blue points near C-3 of steroid where an

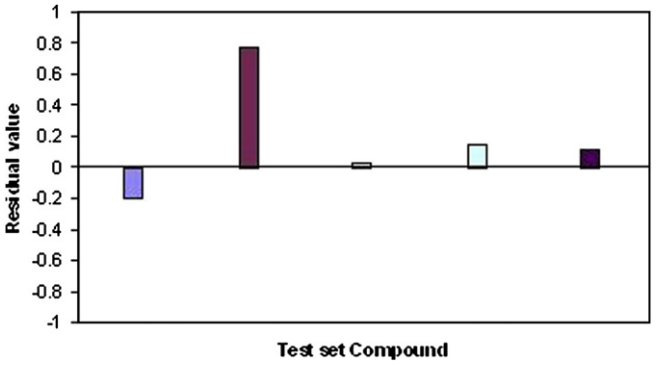

Fig. 8. Histogram of SOMFA residual value for test set.

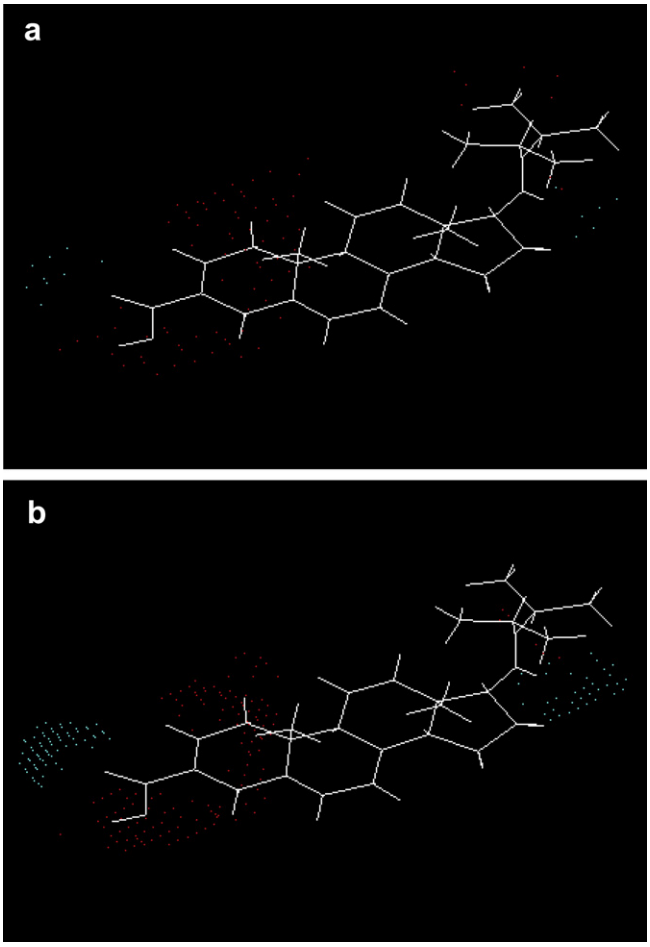

Fig. 9. SOMFA derived Electrostatic grids. The most active compound 15 is displayed in the background. Red and blue indicates region where more electronegative groups or electropositive groups, respectively, will enhance the activity at different resolutions (a) 1 Å (b) 0.5 Å.

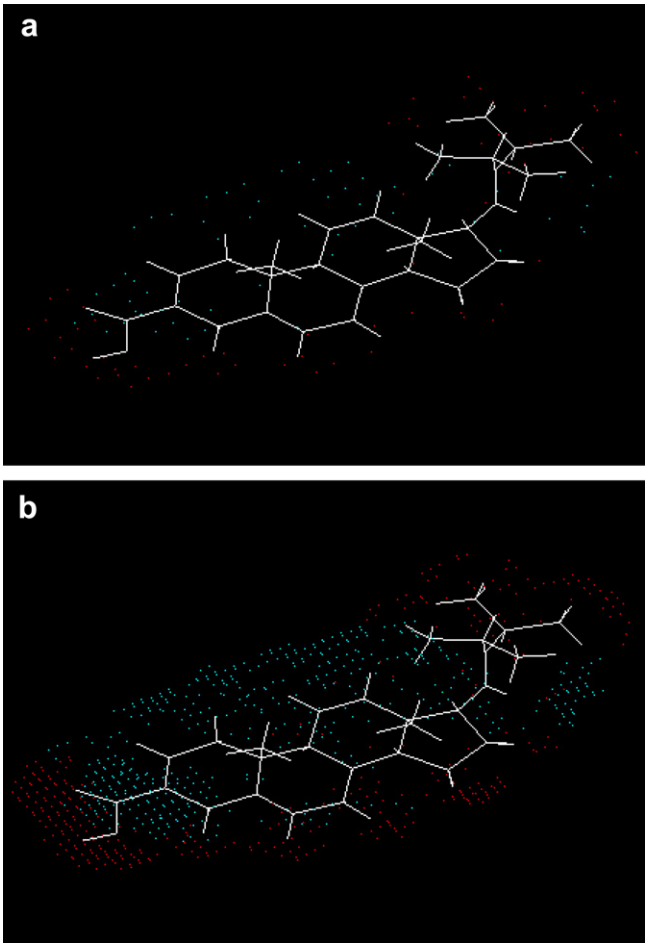

Fig. 10. SOMFA derived Shape grids. The most active compound 15 is displayed in the background. Red and blue indicates region where more steric bulk or less steric bulk, respectively, will enhance the activity at different resolutions (a) 1 Å (b) 0.5 Å.

unfavorable steric interaction may be expected to enhance activities.

3. Conclusion

A predictive SOMFA 3D-QSAR models for unsaturated 3-carboxysteroid derivatives having wide variations in structure and potency profile against human 5α-reductase has been developed. The master grid obtained for the various SOMFA models indicated electrostatic and shape potential contributions can be mapped back onto structural features relating to the trends in activities of the molecules. Shape and electrostatic potential contributions calculations will be helpful in designing novel molecules with improved spectrum of activity.

4. Experimental

4.1. Data set

A dataset of 23 molecules belonging to unsaturated 3-carboxysteroid derivatives as human 5α-reductase inhibitors were taken from the literature and used for SOMFA study [11]. The above reported series of 3-carboxysteroid derivatives showed wide variations in their structure and potency profiles. Different models were generated for this series using a training set of 18 molecules. Predictive powers of the resulting models were evaluated by a test set of 5 molecules with uniformly distributed biological activities. The general structures of the training and test set molecules have been presented in Table 1. Selections of test set molecules were made by considering the fact that test set molecules represent structural features similar to compounds in the training set [18]. Thus the test set compounds are a true representative of the training set.

4.2. Biological activities

The negative logarithm of the measured K_i (nM) against human 5α-reductase enzyme as pK_i (pK_i or log $1/K_i$) was used as dependent variable [19], thus correlating the data linear to the free energy change. Only those compounds which showed significant activity/inhibition were included in the present SOMFA studies.

4.3. Molecular modeling and alignment

The three-dimensional structures of the unsaturated 3-carboxysteroid (Fig. 3) were constructed with the Chemdraw Ultra 8.0 running on an Intel Pentium IV 2.80GHz Processor/Microsoft Win XP Home Edition platform and were subjected to energy minimization using molecular mechanics (MM2). The minimization is continued until the root mean square (RMS) gradient value reaches

a value smaller than 0.001 kcal/mol Å. The Hamiltonian approximations Austin model 1 (AM1) method [20] available in the MOPAC module [21] of Chem3D is adopted for re-optimization until the root mean square (RMS) gradient attains a value smaller than 0.001 kcal/mol Å. Unless otherwise indicated, all parameters were kept default.

The selected template molecule for alignment is typically one of the following:(a) the most active compound; (b) the lead and/or commercial compound; (c) the compound containing the greatest number of functional groups [22,23]. Generally, the low energy conformation of the most active compound is set as reference [24]. In the present study, the compounds were aligned using template based alignment against minimum energy conformation of most active compound **15** selected as reference compound (Fig. 4). Superimposition of all the compounds on template compound **15** has been shown in Fig. 5.

4.4. SOMFA 3D-QSAR models

In the SOMFA study, a $40 \times 40 \times 40$ Å grid originating at $(-20,-20,-20)$ with a resolution of 0.5 and 1 Å respectively, was generated around the aligned compounds [25,26,27,28]. Two different models using different resolution of grid under exploration using template-based alignment has been presented in Table 2.

The partial least square (PLS) algorithm was used in conjugation with leave one out (LOO) cross-validation to develop the final model. In this analysis, one compound was dropped in turn and a model was generated from the remaining compounds. This model was then used to predict the activity of the dropped compound. This procedure was repeated until all the compounds were predicted. This PLS analysis gave the optimum number of components that was used to generate the final models without cross-validation. The result from a cross-validation analysis was expressed as r^2_{cv} (q^2) value, which is defined as

$$r^2_{cv} = 1 - \text{PRESS}/\sum(Y - Y_{mean})^2$$

$$\text{where PRESS} = \sum(Y - Y_{pred})^2$$

The r^2_{cv} (q^2) can take up values in the range from 1, suggesting a perfect model, to less than 0 where errors of prediction are greater than the error from assigning each compound mean activity of the model [29].

Since the final equations are not very useful to represent efficiently the SOMFA models, 3D master grid maps of the best models are displayed by Grid-Visualizer program. These grids represent area in space where steric and electrostatic field interactions are responsible for the observed variations of the biological activity.

Acknowledgements

We gratefully acknowledge Dr Daniel Robinson and the Computational Chemistry Research Group (Oxford University, UK) for the use of the SOMFA software and UGC (University Grant Commission), New Delhi for their financial support.

References

[1] T.L. Bullock, G.L. Andriole, Expert Opin. Emerg. Drugs 11 (2006) 111.
[2] G. Andriole, N. Bruchovsky, L.W. Chung, A.M. Matsumoto, R. Rittmaster, C. Roehrborn, D. Russell, D. Tindall, J. Urol. 172 (2004) 1399.
[3] J.D. Wilson, J.E. Griffin, D.W. Russell, Endocr. Rev. 14 (1993) 577.
[4] S.A. Kaplan, Urology 58 (2001) 65.
[5] A. Kurup, R. Garg, C. Hansch, Chem. Rev. 100 (2000) 909.
[6] P. Sanchez, J.M. Torres, E. Ortega, Neurochem. Res. 30 (2005) 577.
[7] H. Weisser, S. Tunn, M. Debus, M. Krieg, Steroids 59 (1994) 616.
[8] B. Djavan, S. Milani, Y.K. Fong, Expert. Opin. Pharmacother. 6 (2005) 311.
[9] B.W. Metcalf, D.A. Holt, M.A. Levy, J.M. Erb, J.I. Heaslip, M. Brandt, H.J. Oh, Bioorg. Chem. 17 (1989) 372.
[10] M.A. Levy, M. Brandt, R. Heys, D.A. Holt, B.W. Metcalf, Biochemistry 29 (1990) 2815.
[11] D.A. Holt, M.A. Levy, H.J. Oh, J.M. Erb, J.I. Heaslip, M. Brandt, H.-Y.L. Hargest, B.W. Metcalf, J. Med. Chem. 33 (1990) 943.
[12] D.D. Robinson, P.J. Winn, P.D. Lyne, W.G. Richards, J. Med. Chem. 42 (1999) 573.
[13] R.D. Cramer, D.E. Patteerson, J.D. Bunce, J. Am. Chem. Soc. 110 (1988) 5959.
[14] S. Li, Y. Zheng, Int. J. Mol. Sci. 7 (2006) 220.
[15] M. Li, L. Du, B. Wu, L. Xia, Bioorg. Med. Chem. 11 (2003) 3945.
[16] A. Golbraikh, A. Tropsha, J. Mol. Graph. Model. 20 (2002) 269.
[17] D.S. Puntambekar, R. Giridhar, M.R. Yadav, Eur. J. Med. Chem. 43 (2008) 142.
[18] V.S. Murthy, V.M. Kulkarni, Bioorg. Med. Chem. 10 (2002) 2267.
[19] N. Sachan, S.S. Kadam, V.M. Kulkarni, J. Enzyme Inhib. Med. Chem. 22 (2007) 267.
[20] M.J.S. Dewar, E.G. Zoebisch, E.F. Healy, J.J.P. Stewart, J. Am. Chem. Soc. 107 (1985) 3902.
[21] J. Stewart, J. Comput. Aided Mol. Des. 4 (1990) 1.
[22] A. Agarwal, E.W. Taylor, J. Comput. Chem. 14 (1993) 237.
[23] N. Baurin, E. Vangrevelinghe, L.M. Allory, J. Med. Chem. 43 (2000) 1109.
[24] M. Xu, A. Zhang, S. Han, L. Wang, Chemosphere 48 (2002) 707.
[25] A. Pedretti, L. Villa, G. Vistoli, J. Mol. Graph. Model. 21 (2002) 47.
[26] VEGA ZZ Release 2.3.1.1 can be free downloaded from: http://www.ddl.unimi.it/vega/index2.htm.
[27] SOMFA2 v2.0.0 can be downloaded from: http://bellatrix.pcl.ox.ac.uk.
[28] M.Y. Li, H. Fang, L. Xia, Bioorg. Med. Chem. Lett. 15 (2005) 3216.
[29] S.S. Kulkarni, L.K. Gediya, V.M. Kulkarni, Bioorg. Med. Chem. Lett. 7 (1999) 1475.