Med Chem Res (2013) 22:4568–4580
DOI 10.1007/s00044-012-0456-5

MEDICINAL CHEMISTRY RESEARCH

ORIGINAL RESEARCH

Design of potent human steroid 5α-reductase inhibitors: 3D-QSAR CoMFA, CoMSIA and docking studies

Rajnish Kumar · Priyanka Malla · Abhilasha Verma · Manoj Kumar

Received: 22 August 2012 / Accepted: 27 December 2012 / Published online: 13 January 2013

Abstract 3D-QSAR CoMFA, CoMSIA and docking studies were performed on a set of 4-azasteroidal human steroid 5α-reductase inhibitors. The models developed using maximal common substructure-based alignment was found to be reliable and significant with good predictive r^2 value. CoMSIA model developed using combination of steric, electrostatic, hydrophobic, hydrogen bond donor and hydrogen bond acceptor features has shown $r^2_{cv} = 0.564$ with six optimum components, $r^2_{ncv} = 0.945$, *F value* = 101.196, $r^2_{Pred} = 0.693$ and SEE = 0.209. The contour plots obtained has shown a favourable effect of bulkier groups at C-17 position. Docking studies indicates the importance of bulkier groups at C-17 position for favourable activity. The study further helps in design of potent inhibitors of the enzyme.

Keywords 5α-Reductase · Docking · Dihydrotestosterone · CoMFA · CoMSIA

Introduction

Testosterone is the most abundant androgen in serum, synthesised by Leydig cells of the testes under the control of the hypothalamus and anterior pituitary gland. It has a variety of effects including growth of male genitalia, building and maintaining muscle mass, development of libido and initiation of spermatogenesis at puberty (Azzouni *et al*., 2012). Intracellularly, testosterone (T) is converted to its more active metabolite dihydrotestosterone (DHT) by NADPH dependent; membrane bound (microsomal) human 5α-reductase (5α-R) (EC 1.3.99.5) enzyme (Fig. 1). The DHT is responsible for in utero differentiation and growth of prostate gland, male external genitalia and pubertal growth of facial and body hair. It plays an important role in several human diseases like acne, hirsutism, male pattern baldness, benign prostatic hyperplasia (BPH) and prostate cancer (PCa) (Cilotti *et al*., 2001). Three isozymes of 5α-R have been identified (5α-R1-3). Type 1 is found in the periphery, type 2 is mainly located in the prostate tissue, (Andersson and Russell, 1990) and recently type 3 has been identified in hormone refractory prostate cancer cells (HRPC), which is located on SRD5A3 gene and converts testosterone into DHT in a similar way to type 1 enzyme (Aggarwal *et al*., 2010c; Uemura *et al*., 2008). Isozymes 1 and 2 are expressed in normal prostate tissue but overexpressed in BPH where they are responsible for the hyperplasia of stromal and epithelial cells in the transition zone and periurethral glands of the prostate that finally results in prostate gland enlargement. Currently, two structurally similar drugs are available for the clinical management of BPH, i.e. finasteride and dutasteride (Fig. 2). Dutasteride is a dual 5α-reductase inhibitor that inhibit both type 1 and type 2 isozymes while finasteride is type 2 inhibitor. Dutasteride induces a more profound reduction of serum DHT in the range of 90–95 % compared with 70–75 % for finasteride due to its dual inhibition pattern (Gravas and Oelke, 2010).

The aim behind development of 5α-R inhibitors is to obtain potent binding with the enzyme 5α-R with little or no affinity to androgen or other steroid receptors. Several steroidal and non-steroidal inhibitors have been synthesised and tested against 5α-R in last three decades (Aggarwal *et al*., 2010c). However, only finasteride and dutasteride (Fig. 2) belonging to 4-azasteroid category are approved

R. Kumar · P. Malla · A. Verma · M. Kumar (✉)
University Institute of Pharmaceutical Sciences, Panjab University, Chandigarh 160014, India
e-mail: mkgw4pu@yahoo.com; manoj_uips@pu.ac.in

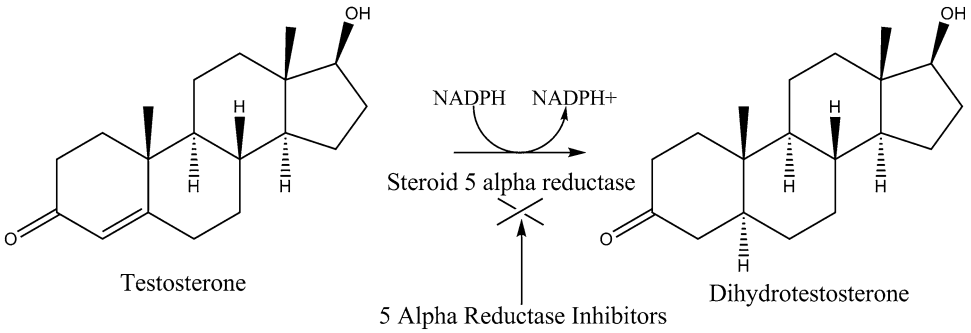


Fig. 1 Inhibition of enzyme 5α-reductase which catalyses the conversion of testosterone into the more active dihydrotestosterone using NADPH as a cofactor

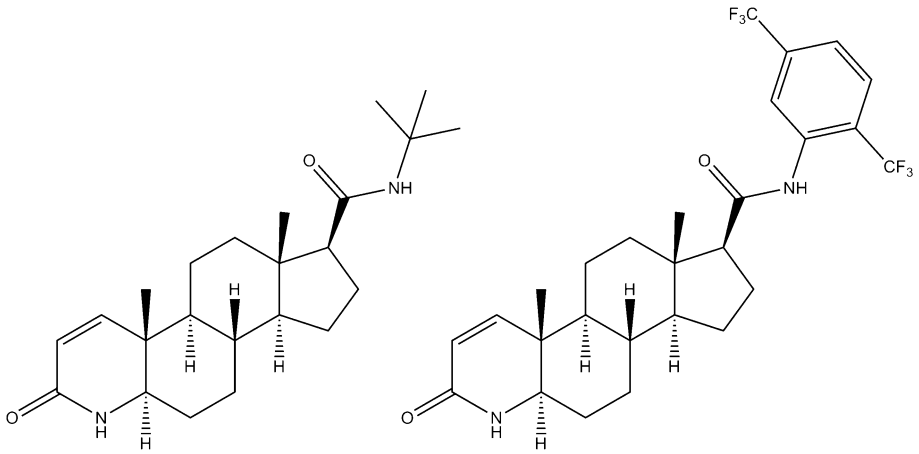


Fig. 2 Clinically available 5α-reductase inhibitors used for the management of BPH

for use in clinical management of BPH and other 5α-R dependent diseases. Finasteride is slow acting and produces side effects affecting sexual functions (Faragalla *et al*., 2003), therefore, it is needed to design and synthesise novel molecules exhibiting potent 5α-R inhibition.

Structure activity relationship study is the basis for perception of structural features of ligand as well as receptor for exhibiting the desired activity. Ligand-based 3D-QSAR studies and receptor-based molecular docking studies are known to be complimentary to each other and are important tools in the design and development of potent inhibitors. In order to optimise the structural features required for potent inhibition of 5α-R, several 3D-QSAR studies have been reported from our laboratory involving use of Comparative Molecular Field Analysis (CoMFA), Comparative Molecular Similarity Indices Analysis (CoMSIA) and Self Organising Molecular Filed Analysis (SOMFA) techniques (Kumar and Kumar, 2013; Aggarwal *et al*., 2010a, b; Thareja *et al*., 2009). Lack of knowledge about crystal structure of the type 1 and type 2 5α-R is the major hurdle in the discovery of novel 5α-reductase inhibitors (5ARIs). They have not been yet isolated and purified from tissues or cells, and the only information available is the primary sequence estimated from their c-DNAs. However, several crystal structures of 5β reductase are available binding with different steroids and same cofactors (NADP+ or NADPH). In the absence of structural information on target protein, various receptor mapping technologies were developed allowing constructing 3D surrogate of the binding pocket. In a QSAR context, those can act as substitutes for the structure of the true biological receptor (Vedani *et al*., 2007). Recently Yao *et al*., (2011) have reported docking studies using crystal structure of 5β reductase (AKR1D1. NADP+.finasteride) as a surrogate receptor.

In the present work, 3D-QSAR studies have been performed on a series of substituted 4-azasteroidal inhibitors of 5α-R. CoMFA and CoMSIA methodologies were used to obtain 3D QSAR derived contour maps as output which gives insight into the topological features of the binding pocket (Potshangbam *et al*., 2011). The molecular docking study of 4-azasteroids into finsasteride binding site of 5β-reductase was also carried out using Glide to identify the binding orientation and the receptor ligand interactions responsible for exhibited activity. The free energy of binding of ligands was calculated using Prime MMGBSA approach. The information obtained from 3D-QSAR and docking studies could be helpful in optimisation of structural requirements for potent inhibition of human 5α-reductase enzyme.

Methods

Dataset collection

A set of 52 substituted 4-azasteroidal inhibitors of the enzyme human 5α-reductase were selected from the literature (Rasmusson *et al*., 1986) on the basis of diversity in structural motif and variation in the biological activity. The biological activity was reported for human 5α-reductase, which showed some assay to assay variability, so the inhibition is expressed as a ratio of the inhibitor IC_{50} value to the IC_{50} value of the compound **44**. Where IC_{50} is the molar concentration in nanomoles of the inhibitors producing 50 % inhibition of human 5α-reductase enzyme. The activity was converted into negative logarithm. The dataset was divided into a training set of 42 molecules and a test set of ten molecules which were used to assess the predictively of the models. The test set was judiciously chosen so that it covered almost entire range of biological activity and structural diversity. The general structure of the test and training set molecules along with the obtained and predicted activity have been presented in Tables 1 and 2.

Computational approach and molecular alignment

Molecular modelling, CoMFA and CoMSIA analysis were performed using SYBYL 7.0 software (SYBYL Molecular Modeling System and version, 2003). The other calculations were done using Schrödinger and New York, 2012 suite installed on Linux based Intel® Core™2 Quad CPU Q8400@2.66 GHz platform. Since the crystal structure of enzyme 5α-reductase in complex with 4-azasteroid inhibitor is not reported, the most active compound from the dataset was subjected to conformational search using OPLS_2005 forcefield using water as a solvent in MacroModel (MacroModel, 2012). The global minimum conformation obtained from the search was used as a template to build the other molecules. The partial charges for all of the compounds were calculated using Gasteiger-Huckel method. The geometry of the molecules was optimised using Tripos force field with a distance dependent dielectric function and energy convergence criterion of 0.001 kcal/mol Å with standard SYBYL settings and keeping maximum 1,000 iterations. Alignment of molecules is important and the method which maintains the bioactive conformation of the molecules is used to obtain a better alignment of molecules (Dessalew *et al*., 2007). The positioning of the molecular model within the fixed lattice is an important input variable in CoMFA, since the relative interaction energies depend strongly on the relative molecular positions (Cramer *et al*., 1988). The optimised structures of the 42 selected molecules were aligned using maximum common substructure (Fig. 3) methodology implemented in SYBYL. Superimposition of all the molecules under study on the template has been shown in Fig. 4. The test set molecules were also processed in a similar fashion as training set molecules.

CoMFA methodology

The steric (Lennard–Jonnes potential) and electrostatic fields (Coulombic potentials) were calculated at each lattice intersection for the aligned molecules kept in a 3D cubic lattice with a grid spacing of 2.0 Å in x, y and z coordinates. The van der Waals potentials and coulombic terms representing the steric and electrostatic fields, respectively, were calculated using standard Tripos force fields. A sp^3 carbon atom having a charge of +1 and a radius of 1.52 Å was used as a probe to calculate the steric and electrostatic fields. The steric and electrostatic fields were truncated at 0.3 kcal/mol. The other parameters were used as default because, the effect of altering the lattice spacing, column filtering and energy cutoff values in the CoMFA process appears to be minimal, and, thus, the use of the default settings for these parameters seems appropriate (Mittal *et al*., 2009).

CoMSIA interaction energy calculation

In CoMSIA interaction energy calculation, the steric, electrostatic, hydrophobic, hydrogen bond donor and hydrogen bond acceptor potential fields were calculated at each lattice intersection of the same lattice box used for CoMFA calculations (Klebe *et al*., 1994). In the present study, standard settings of CoMSIA were used to calculate the steric, electrostatic, hydrophobic, donor and acceptor fields. The equation used to calculate the similarity indices is as follows:

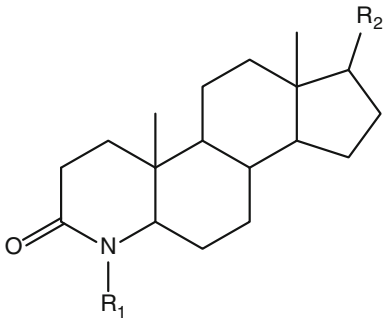
$$A^{q}_{F,k}(j) = \sum_{i} w_{\text{probe},k} w_{ik} e^{-\alpha r^{2}_{iq}}$$

where A is the similarity index at grid point q, summed over all atom i of the molecule j under investigation. $w_{\text{probe},k}$ is the probe atom with radius 1 Å, charge +1, hydrophobicity +1, hydrogen bond donating +1 and hydrogen bond accepting +1. w_{ik} is the actual value of physicochemical property k of atom i. r_{iq} is the mutual distance between the probe atom at grid point q and atom i of the test molecule. α is the attenuation factor.

Partial least square (PLS) analysis

PLS method is a general and powerful tool for deriving multiple linear relationships among columns of data and shown to be much less susceptible to chance correlation (Puntambekar *et al*., 2006). The models were developed using PLS methodology to correlate 5α-reductase inhibitory

Table 1 Compounds (**1**–**43**) used in the study



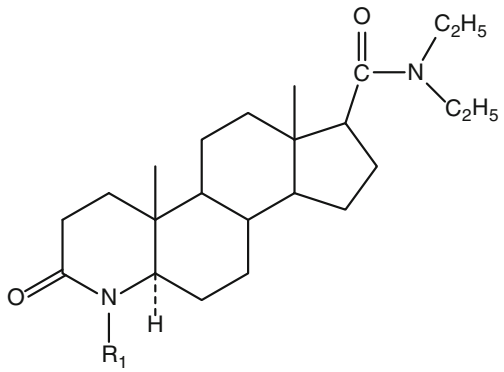
Compound no.	R_1	R_2	Human 5α-R IC_{50}/IC_{50} 44	PIC IC_{50}/IC_{50} 44	Predicted	Residual
01	CH_3	β-$COCH_3$	27	6.569	6.219	0.350
02	CH_3	β-$CHOHCH_3$	17	6.770	5.918	0.852
03	CH_3	β-$CH(CH_3)CH_2OH$	15	6.824	6.439	0.385
04	H	β-$COOCH_3$	42	6.377	6.337	0.040
05[a]	CH_3	β-$COOCH_3$	17	6.770	6.646	0.124
06	H	β-$CONH_2$	150	5.824	5.901	−0.077
07	CH_3	β-$CONHCH_3$	28	6.553	6.735	−0.182
08	CH_3	β-$CONHC_2H_5$	9.4	7.027	7.167	−0.140
09[a]	CH_3	β-$CONHCH(CH_3)_2$	4.0	7.398	7.588	−0.190
10	H	β-$CONHC(CH_3)_3$	1.7	7.770	7.43	0.340
11	CH_3	β-$CONHC(CH_3)_3$	1.1	7.959	7.815	0.144
12	CH_3	β-$CONH(CH_2)_4CH_3$	2.4	7.620	7.362	0.258
13[a]	H	β-$CONH(CH_2)_7CH_3$	2.4	7.620	7.272	0.348
14	CH_3	β-$CONH(CH_2)_7CH_3$	4.0	7.398	7.675	−0.277
15	H	β-$CONHC(CH_3)_2CH_2C(CH_3)_3$	0.56	8.252	7.924	0.328
16	CH_3	β-$CONHC(CH_3)_2CH_2C(CH_3)_3$	0.52	8.284	8.335	−0.051
17	CH_3	β-$CONHC(CH_3)_2CH_2OH$	2	7.699	7.52	0.179
18[a]	CH_3	β-$CONHCH_2CON(C_2H_5)_2$	11	6.959	6.991	−0.032
19	CH_3	β-$CONHCH_2CH(OCH_3)_2$	20	6.699	6.834	−0.135
20	H	β-$CON(C_2H_5)_2$	1.2	7.921	7.739	0.182
21	CH_3	β-$CON(C_2H_5)_2$	1	8.000	8.165	−0.165
22	H	β-$CON(i\text{-}C_3H_7)_2$	1.5	7.824	7.835	−0.011
23	CH_3	β-$CON(i\text{-}C_3H_7)_2$	0.82	8.086	8.251	−0.165
24[a]	CH_3	β-$CON(C_2H_5)CH_2CH_2OH$	3	7.523	8.219	−0.696
25	CH_3	β-$CON(C_2H_4)_2O$	5.6	7.252	7.516	−0.264
26	CH_3	β-$CH_2COOC_2H_5$	29	6.538	6.914	−0.376
27	CH_3	β-$CH_2CON(C_2H_5)_2$	2	7.699	7.849	−0.150
28[a]	CH_3	=$CHCON(C_2H_5)_2$	1.5	7.824	7.796	0.028
29	CH_3	β-$CH(CH_3)COOCH_3$	143	5.845	6.226	−0.381
30	CH_3	β-$CH(CH_3)CON(C_2H_5)_2$	390	5.409	6.722	−1.313
31	CH_3	β-$CH(CH_3)(CH_2)_2CON(C_2H_5)_2$	27	6.569	6.435	0.134
32	CH_3	β-CN	83	6.081	6.087	−0.006
33[a]	CH_3	β-(2-Oxazolinyl)	11	6.959	6.741	0.218
34	CH_3	β-(4,4-Dimethyl-2-oxazolinyl)	125	5.903	6.091	−0.188
35	H	α-$CH_2CH_2CH_2O$-β	735	5.134	5.655	−0.521
36	CH_3	α-$CH_2CH_2CH_2O$-β	144	5.842	5.915	−0.073
37	H	β-$COCH_2CH(CH_3)_2$	4.2	7.377	7.214	0.163

Table 1 continued

Compound no.	R_1	R_2	Human 5α-R IC_{50}/IC_{50} 44	PIC IC_{50}/IC_{50} 44	Predicted	Residual
38	CH_3	β-$COCH_2CH(CH_3)_2$	2.7	7.569	7.635	−0.066
39[a]	H	β-$COCH(CH_3)C_2H_5$	4.0	7.398	7.302	0.096
40	H	β-CO(2-pyrrolyl)	3.6	7.444	7.406	0.038
41	CH_3	β-CO(2-pyrrolyl)	2.0	7.699	7.731	−0.032
42	H	β-$COOC(CH_3)_3$	8.6	7.066	7.164	−0.098
43	CH_3	β-$COSC(CH_3)_3$	47	6.328	6.503	−0.175

[a] Test set molecule

Table 2 Compounds (**44**–**52**) used in the study



Compound no.	R_1	Others	Human 5α-R IC_{50}/IC_{50} 10X	P(IC_{50}/IC_{50} 10X)	Predicted	Residual
44	CH_3	–	1	8.000	7.668	0.332
45	H	–	1.2	7.921	7.240	0.681
46[a]	CH_3	1-α-CH_3	7	7.155	7.390	−0.235
47	CH_3	2-β-CH_3	3.6	7.444	7.277	0.167
48	CH_3	1-α, 2-α-CH_2	1.3	7.886	7.623	0.263
49[a]	CH_3	1-α, 2-α-O	63	6.201	7.857	−1.656
50	CH_3	2-β-F	5.0	7.301	7.369	−0.068
51	CH_3	2-β-OCH_3	15	6.824	6.845	−0.021
52	CH_3	2-$(OCH_3)_2$	28	6.553	6.375	0.178

[a] Test set molecule

activity with the CoMFA and CoMSIA fields. Cross-validation analysis was performed internally using leave one out (LOO) method in which one compound is removed from the dataset and its activity is predicted using the model derived from the rest of the dataset (Kumar *et al*., 2011). The cross-validation r^2 that resulted in optimum numbers of components and lowest standard error of prediction were considered for further analysis. The external validation of various models was performed using a test set of 10 molecules. The analysis was carried out with the column filtering value of 2.0 kcal/mol to speed up the calculation and reduce the noise. Final analysis was performed to calculate non-cross-validated r^2 using the optimum number of components. The cross-validation r^2, Fischer's statistic (F test), Standard error of estimate and predicted r^2were calculated.

Docking studies

Docking studies were carried out using standard Glide (Glide, 2012) molecular docking protocol implemented in Maestro molecular modelling suite of Schrödinger, LLC, New York, NY, 2012. As the receptor for 5α-reductase is not available, recently a docking study has been reported using 5β-reductase as a surrogate to the receptor on the basis that the substrate for both of the enzymes is same,

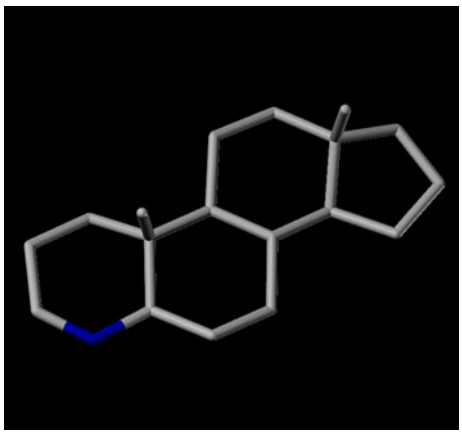


Fig. 3 Maximum common substructure used for alignment

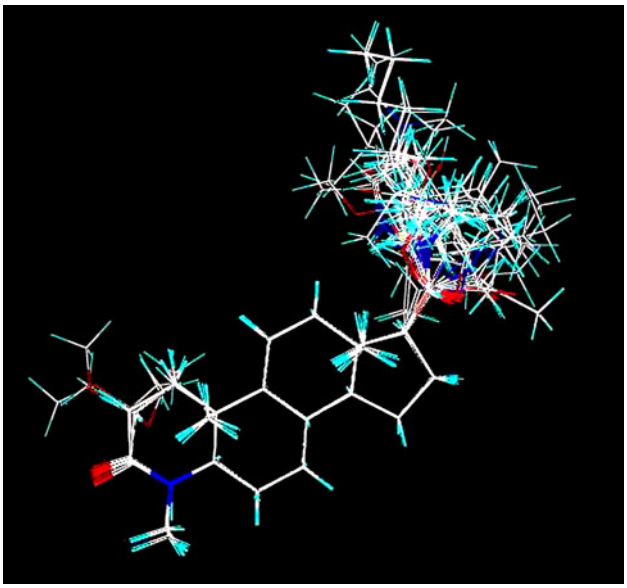


Fig. 4 Maximum common substructure-based alignment of dataset compounds used for CoMFA and CoMSIA analysis

thus, they have same or at least similar enzymatic functions when metabolising steroid hormones (Yao *et al*., 2011). The protein structure PDB code: 3G1R (Drury *et al*., 2009) was retrieved from Protein Data Bank (www.rcsb.com) and used for docking studies. The monomer B was kept and treated using protein preparation wizard of the Schrödinger suite. Following hydrogen bonding assignment optimisation, removal of water molecules and the protein ligand complex were energy minimised until an rmsd of 0.30 Å. The docking studies were carried out in the presence of cofactor NADPH to elucidate its potential role. The docking was performed using Glide extra precision mode (XP). The receptor grid was generated using the centroid of the co-crystallized ligand and a maximum size of 20 Å. To mimic the flexibility of the protein structure, a scaling van der Waals radii for non-polar parts in the binding site and the ligand was carried out.

Further Prime MMGBSA approach (Prime, 2012) was used to predict the free energy of binding for the receptor inhibitor complex (Lyne *et al*., 2006). The docked complexes were taken from pose viewer files of initial Prime XP docking. The MMGBSA approach employs molecular mechanics, the generalised Born model and solvent accessibility method to elicit free energies from structural information circumventing the computational complexity of free energy simulations (Dill, 1997). The binding free energy of each ligand was calculated using following equation:

$$\Delta G_{\text{bin}} = \Delta E_{\text{mm}} + \Delta G_{\text{sol}} + \Delta G_{\text{SA}}$$

Where ΔE_{mm} is the difference in the minimised energies between the receptor ligand complex and the sum of energies of unliganded receptor and ligands. ΔG_{sol} is the difference in the GBSA solvation energy of the receptor ligand complex and the sum of energies of unliganded receptor and ligands. ΔG_{SA} is the difference in surface area energies for the receptor ligand complex and the sum of energies of unliganded receptor and ligands.

Results and discussion

CoMFA and CoMSIA analysis were used to develop QSAR models on a set of fifty-two 4-azasteroidal 5ARIs. The lowest energy conformation of the most active compound was used as a template to build all the molecules. The dataset was subjected to maximum common substructure alignment methodology. A number of predictive models were developed using CoMFA and CoMSIA fields. The developed models were predicted externally using a test set of 10 molecules. The models with significant predictive and correlative power are reported in Table 3. The predicted and residual activity of the compounds obtained using the best model is given in Tables 1 and 2. The plot between observed and predicted activities of training set and test set is shown in Fig. 5.

CoMFA analysis

The negative logarithm of $IC_{50}/IC_{50}44$ (p $IC_{50}/IC_{50}44$) was used as the dependent variable and is given in Table 1 and 2. The CoMFA models were built using a training set of forty-two compounds and tested against a test set of ten compounds. The selection of training set and test set was done by considering the structural diversity and the range of biological activity. The energy-minimised structures of all the compounds were aligned using maximum common structure-based alignment methodology. PLS analysis was carried out using column filtering value of 1.0 kcal/mol.

Table 3 Summary of PLS statistics of the best CoMFA and CoMSIA models based on the maximum common substructure-based alignment

Parameter	CoMFA	COMSIA					
		SEHDA	SEHD	SEHA	SEH	SEDA	SEA
r^2_{cv}	0.522	0.564	0.522	0.550	0.564	0.608	0.619
ONC	5	6	3	7	4	7	7
r^2_{ncv}	0.902	0.945	0.813	0.936	0.894	0.904	0.938
SEE	0.256	0.209	0.372	0.223	0.283	0.274	0.220
F value	69.080	101.196	54.936	105.519	98.129	87.577	108.079
r^2_{Pred}	0.583	0.693	0.513	0.618	0.572	0.608	0.599
Contribution (%)							
Steric (S)	58.8	20.8	22.7	30.9	33.2	25.6	62.9
Electrostatic (E)	41.2	12.9	23.2	17.3	26.6	20.9	26.0
Hydrophobic (H)	–	27.2	32.2	33.6	40.2	–	–
Donor (D)	–	19.9	21.9	–	–	30.9	–
Acceptor (A)	–	19.1	–	18.2	–	22.7	11.1

r^2_{cv} cross-validated correlation coefficient, *ONC* optimum number of components from PLS analysis, r^2_{ncv} non-cross-validated correlation coefficient, *SEE* standard error of estimate, *F* Fischer statistic, r^2_{Pred} predictive correlation coefficient

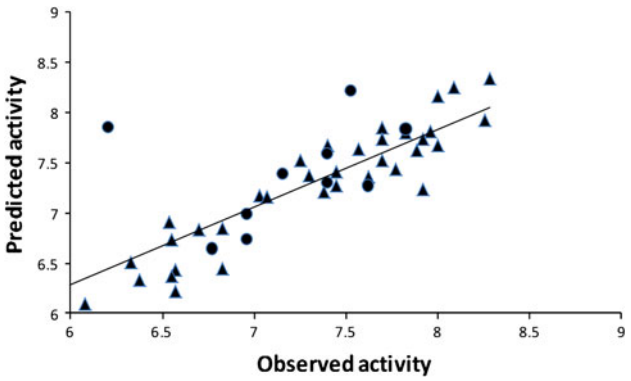


Fig. 5 *Graph* between observed and predicted activity for training set and test set molecules from the best predictive CoMSIA model (*Closed triangle* and *closed circle* represents training and test set molecules, respectively)

In CoMFA and CoMSIA study, predicted r^2 value of >0.3 is considered as statistically significant but a value of >0.5 can be considered as statistically more significant (Bohm *et al*., 1999). The CoMFA model showed cross-validation $r^2 = 0.522$ with 5 optimum components, non-cross-validation $r^2 = 0.902$, *F value* = 69.080 predicted $r^2 = 0.583$, standard error of estimate (SEE) = 0.256; the steric and electrostatic contributions were 58.8 and 41.2 %, respectively. The developed model was found to be statistically significant towards describing the 5α-reductase inhibitory activity due to its high non-cross-validation r^2 and low standard error of estimate.

CoMSIA analysis

The CoMSIA analysis was performed using steric, electrostatic, hydrophobic, hydrogen bond donor and hydrogen bond acceptor fields. Various models were developed using a combination of different fields and the statistically significant models are reported in Table 3. The CoMSIA model with combination of all the fields yielded a cross-validation $r^2 = 0.564$ with 6 optimum components, non-cross-validation $r^2 = 0.945$, *F value* = 101.196, predicted $r^2 = 0.693$, SEE = 0.209; the steric, electrostatic, hydrophobic, hydrogen bond donor and hydrogen bond acceptor fields contribution were 20.8, 12.9, 27.2, 19.9 and 19.1 %, respectively. Combination of steric, electrostatic, hydrophobic and hydrogen bond donor fields yielded a CoMSIA model with cross-validation $r^2 = 0.552$ with 3 optimum components, non-cross-validation $r^2 = 0.813$, *F value* = 54.936, predicted $r^2 = 0.513$, SEE = 0.372; the steric, electrostatic, hydrophobic and hydrogen bond donor fields contribution were 22.7, 23.2, 32.3 and 21.9 %, respectively. Combination of steric, electrostatic, hydrophobic and hydrogen bond acceptor fields yielded a CoMSIA model with cross-validation $r^2 = 0.550$ with 7 optimum components, non-cross-validation $r^2 = 0.936$, *F value* = 105.519, predicted $r^2 = 0.618$, SEE = 0.223; the steric, electrostatic, hydrophobic and hydrogen bond donor fields contribution were 30.9, 17.3, 33.6 and 18.2 %, respectively. Combination of steric, electrostatic and hydrophobic fields yielded a CoMSIA model with cross-validation $r^2 = 0.564$ with 4 optimum components, non-cross-validation $r^2 = 0.894$, *F value* = 98.129, predicted $r^2 = 0.572$, SEE = 0.283; the steric, electrostatic and hydrophobic fields contribution were 33.2, 26.6 and 40.2 %, respectively. Combination of steric, electrostatic, hydrogen bond donor and hydrogen bond acceptor fields yielded a CoMSIA model with cross-validation $r^2 = 0.608$ with 7 optimum components, non-cross-validation $r^2 = 0.904$, *F value* = 87.577, predicted $r^2 = 0.608$, SEE = 0.274; the steric, electrostatic,

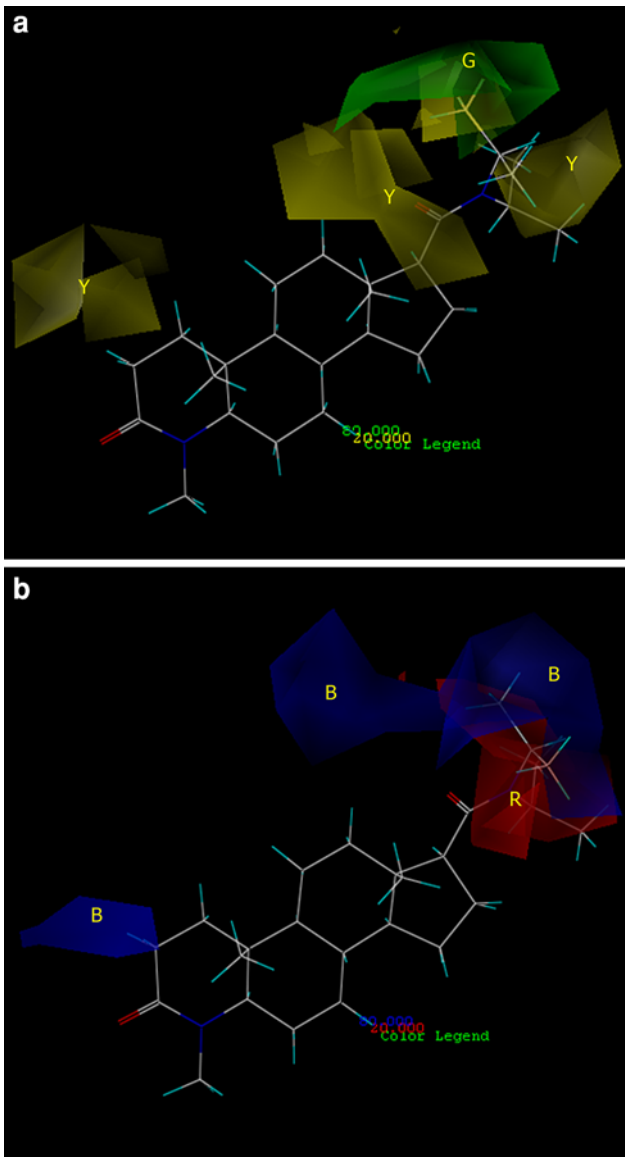


Fig. 6 The CoMFA STDEV*COEFF *contour plots* obtained from PLS analysis mapped over most active compound (**23**). **a** Steric, sterically favoured areas (contribution level 80 %) are represented by *green polyhedron*. Sterically disfavoured areas (contribution level 20 %) are represented by *yellow polyhedron*. **b** Electrostatic, Positively charged favoured areas (contribution level 80 %) are represented by *blue polyhedron*. Negatively charged favoured areas (contribution level 20 %) are represented by *red polyhedron*

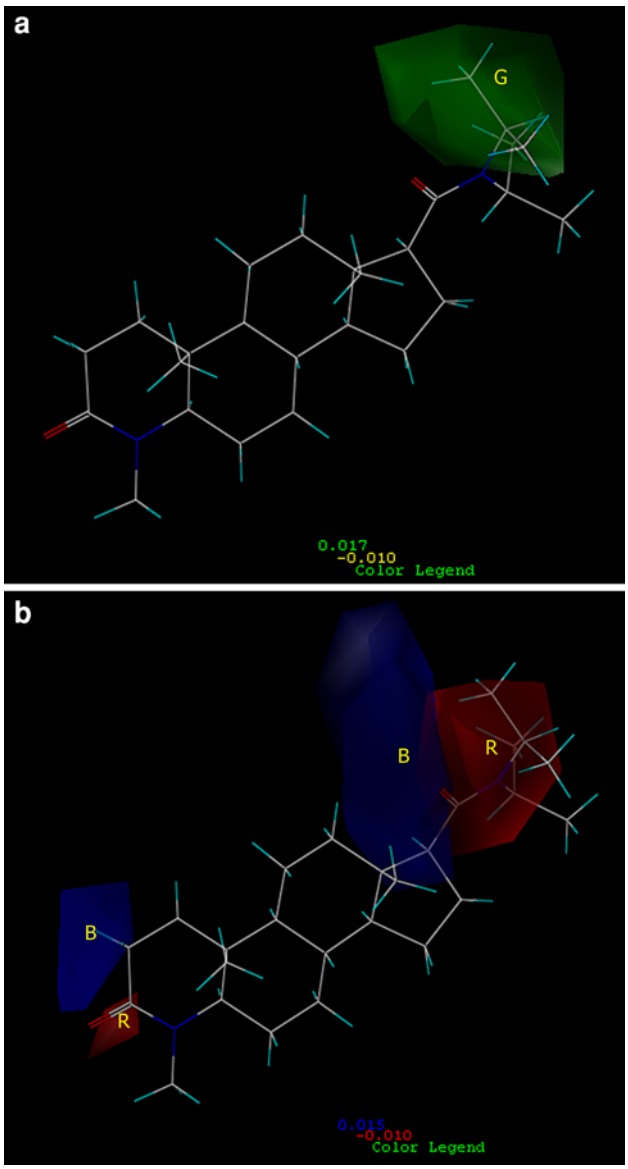


Fig. 7 The CoMSIA STDEV*COEFF *contour plots* of obtained from PLS analysis mapped over most active compound (**23**). **a** Steric, **b** Electrostatic. *Yellow* and *green contours* show region where steric group is favoured and disfavoured the activity, respectively. *Blue* and *red contours* show region where positive charged substituent is favoured and negatively charged substituent is favoured the activity, respectively

hydrogen bond donor and hydrogen bond acceptor fields contribution were 25.6, 20.9, 30.9 and 22.7 %, respectively. Combination of steric, electrostatic and hydrogen bond acceptor fields yielded a CoMSIA model with cross-validation $r^2 = 0.619$ with 7 optimum components, non-cross-validation $r^2 = 0.938$, F *value* $= 108.079$, predicted $r^2 = 0.599$, SEE $= 0.220$; the steric, electrostatic, and hydrogen bond acceptor fields contribution were 62.9, 26.0 and 11.1 %, respectively. The developed models were shown to be highly predictive as indicated by their robust statistical parameters as given in Table 3. The major contributing field in CoMFA model was steric and in the case of CoMSIA, all the fields were contributing well. The predicted and residual activities determined using the best model (SEHDA) are given in Tables 1 and 2.

CoMFA and CoMSIA contour maps

The CoMFA steric and electrostatic maps are shown in Fig. 6a, b respectively. In steric contour plots, the sterically

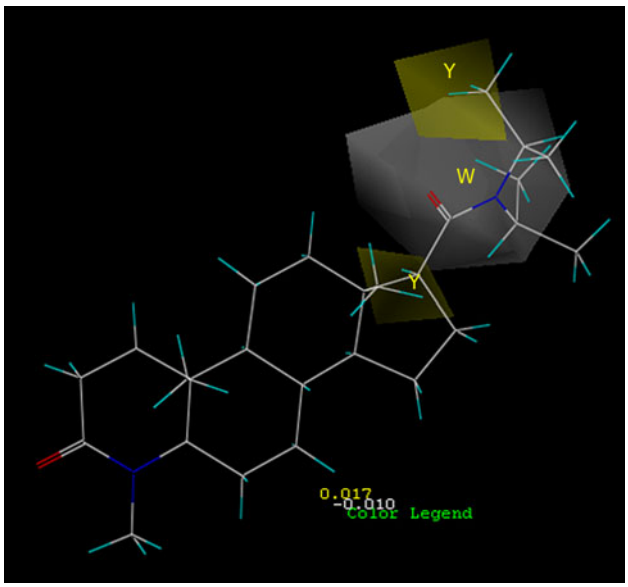


Fig. 8 CoMSIA STDEV*COEFF hydrophobic *contour plots* obtained from PLS analysis mapped over most active compound (**23**). *Yellow* and *white contours* show region where hydrophobic group is favoured and disfavoured the activity, respectively

favoured areas are represented by green polyhedron. Sterically disfavoured areas are represented by yellow polyhedron. The green polyhedron near C-17 substituted position indicates that the steric field is favourable at this position for activity while yellow colour at C-2 position and around C-20 carbonyl indicated that steric bulk at this position disfavour the activity. In the case of electrostatic contour map (Fig. 6b), blue colour region show area where electropositive groups enhances the activity, while red region represents the area where electronegative charge groups enhances the activity. It was found from the contour plot that substitution of electropositive substituents at C-17 and C-2 positions leads to enhanced activity. The CoMSIA steric and electrostatic contour plots mapped over most active compound **44** from the dataset are shown in Fig. 7: (a) steric, (b) electrostatic. In steric contour plots, the sterically favoured areas are represented by green polyhedron. Sterically disfavoured areas are represented by yellow polyhedron. The green polyhedron near C-17 substituted position indicates that the steric field is favourable at this position for activity. In the case of electrostatic contour map (Fig. 7b), blue colour region show area where electropositive groups enhances the activity, while red region represents the area where electronegative charge groups enhance the activity. It was found to be similar to CoMFA electrostatic contour map. Fig. 8 shows the contour plots mapped over compound **44** for hydrophobic field. Yellow and white contours show region where hydrophobic group is favoured and disfavoured the activity, respectively. It was found from the hydrophobic contour map that

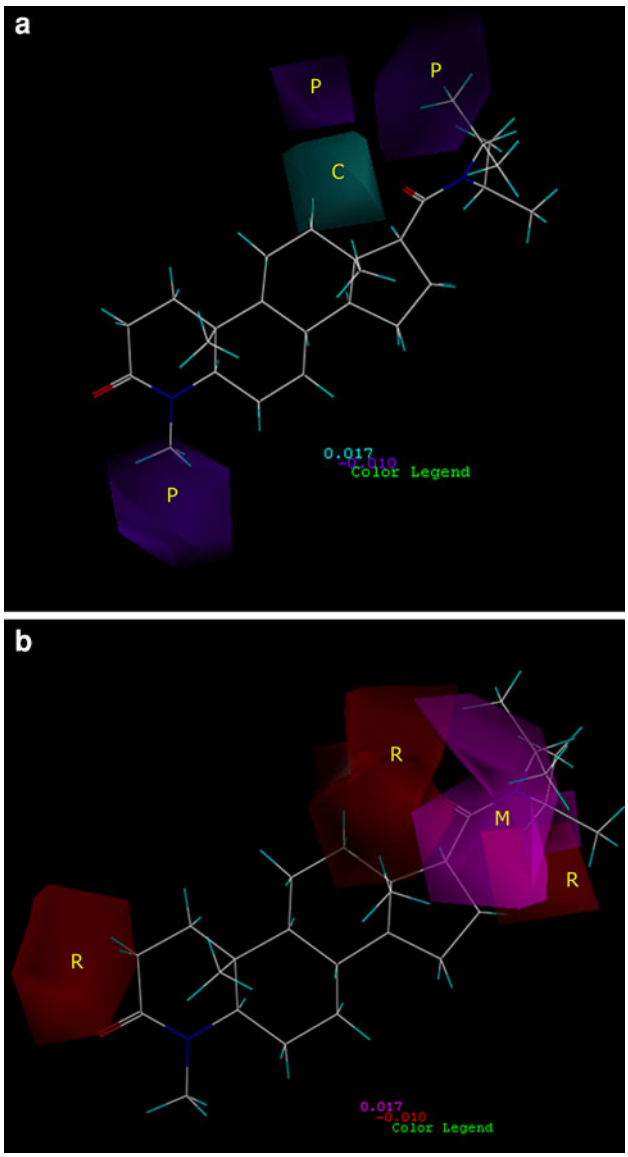


Fig. 9 The CoMSIA STDEV*COEFF *contour plots* obtained from PLS analysis mapped over most active compound (**23**). **a** H-bond donor, *Cyan* and *purple contours* show region where H-bond donors are favoured and disfavoured the activity, respectively. **b** H-bond acceptor, *Magenta* and *red contours* show region where H-bond donors are favoured and disfavoured the activity, respectively

substitution at C-17 position with hydrophobic groups leads to increase in the activity, which indicates that the C-17 position substituents are responsible for hydrophobic interaction with the active site of the enzyme. It is in agreement with previously reported study which suggests that the presence of hydrophobic groups at C-17 position of steroidal skeleton is mandatory for inhibitory activity (Kurup *et al*., 2000). The CoMSIA STDEV*COEFF contour plots obtained from PLS analysis mapped over most active compound (**44**) is shown in Fig. 9: (a) H-bond donor, Cyan and purple contours show region where H-bond donors are favoured and disfavoured the activity,

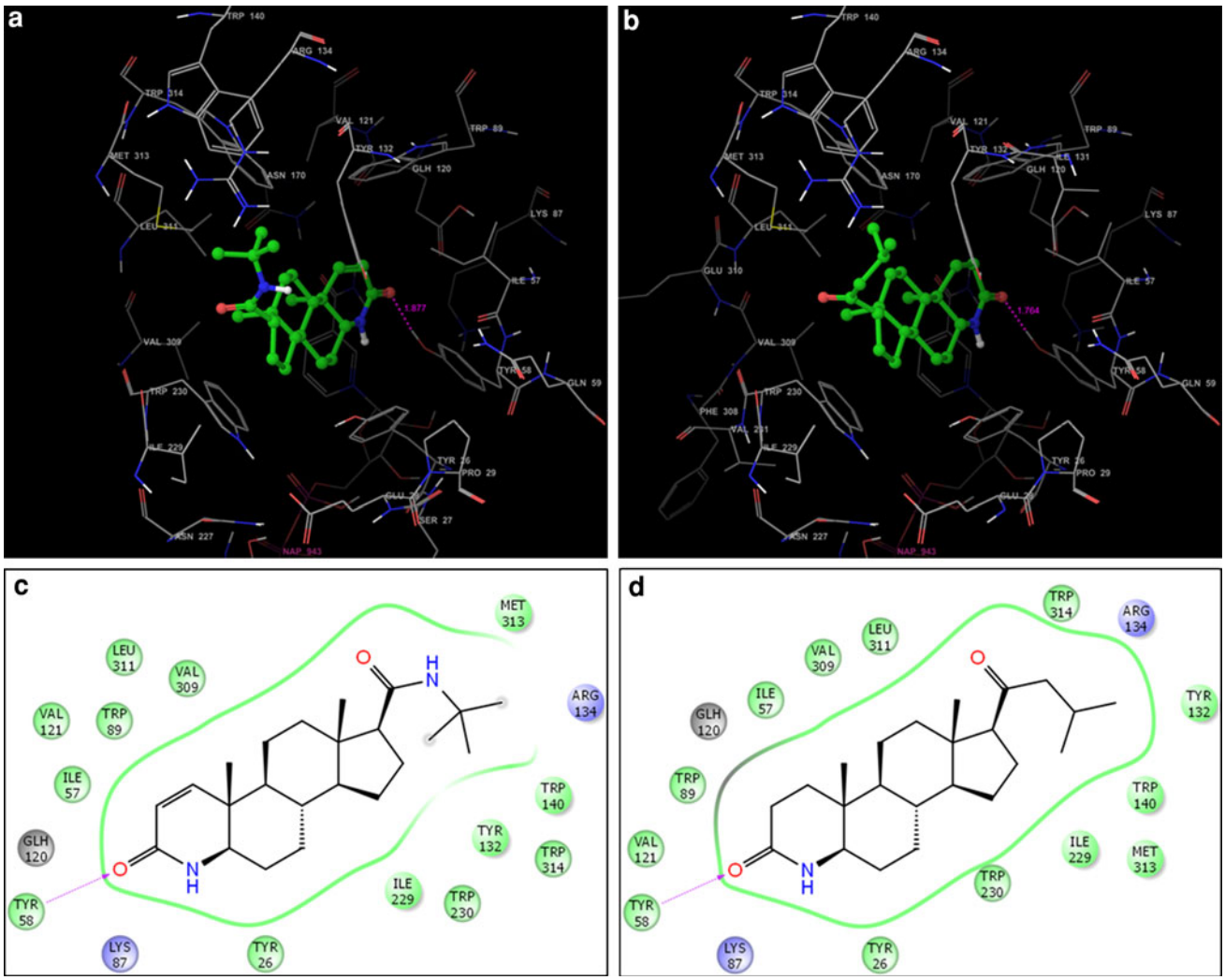


Fig. 10 Key interactions of the docked finasteride (**a** and **c**) and compound **17** (**b** and **d**) into the active site of 5-β reductase

respectively. The cyan polyhedron near C-20 carbonyl shows that the presence of hydrogen bond donor leads to increased activity while the presence of purple polyhedron near substitution at C-4 shows the presence of hydrogen bond donor leads to decrease in activity. (b) H-bond acceptor, Magenta and red contours show region where H-bond donors are favoured and disfavoured the activity, respectively. The presence of magenta polyhedron near C-17 substituents indicated that hydrogen bond acceptor is favourable for activity at this substitution while the presence of red polyhedron near C-2 indicated decreased preference for hydrogen bond acceptor. These details were used for further design of novel inhibitors of the enzyme 5AR.

Docking studies

The docking studies with enzyme 5β-reductase were carried out using the Glide XP methodology. The finasteride was also considered for docking evaluation as it is most widely used clinically available drug as 5ARI. The docked compounds interact with the enzyme 5β-reductase by hydrophobic as well as hydrogen bonds. The hydrophobic substituent at C-17 position of finasteride has shown interaction with a hydrophobic pocket composed of Met313, Arg134, trp140, Tyr132, Trp314, Ile229 and Trp230. The site around C-5 which is supposed to be catalytic site for reduction is composed of Tyr26, Lys87, Tyr58, Glh120, Ile57, Trp89, Val121, Trp89, Leu311 and Val309. The C-3 keto acting as a hydrogen bond acceptor to Tyr58 with a bond length of 1.877 Å. It has shown a docking score (XP GScore) of −10.019. The dataset compounds have also shown similar kind of interactions with the active site residues and the top scored compound (**17**) among dataset that has shown a docking score (XP GScore) of −9.845. The C-3 keto is forming a hydrogen bond with Tyr58 with a length of 1.764 Å. The major interactions of finasteride and top score ligand from dataset

Table 4 Docking score and energy of binding of receptor inhibitor complex calculated using Prime MMGBSA method

Compound no.	Docking score (XP GScore)	Prime MMGBSA DG bind	MMGBSA dG bind Coulomb	MMGBSA dG bind covalent	MMGBSA dG bind vdW	MMGBSA dG bind Solv GB
Finasteride	−10.018503	−87.859688	9.960327	5.999644	−42.733933	2.359468
1	−7.973819	−91.239223	3.607656	4.049858	−45.583692	10.008923
2	−9.165698	−88.679819	13.297103	1.109717	−43.628535	3.779395
3	−9.84664	−88.798037	7.927461	−1.926553	−42.232923	5.689026
4	−9.524238	−96.546	2.851867	−2.907145	−39.977236	0.758775
5	−9.787208	−81.39632	22.58573	−4.262472	−40.4008	1.173214
6	−9.192957	−89.748171	16.781104	−1.385704	−45.465341	2.071019
7	−9.324031	−96.342717	8.904561	−0.182597	−40.083127	−2.232718
8	−9.661474	−87.525945	10.662818	4.519418	−42.184531	−0.435427
9	−9.453767	−82.92257	15.047871	−2.290514	−44.668426	8.005508
10	−9.42741	−83.736598	23.024542	1.057421	−43.298872	−0.774863
11	−9.541476	−73.74172	33.611011	−3.329565	−47.050998	−2.804115
12	−9.704006	−76.443589	139.055596	0.557908	−38.358884	−120.500507
13	−9.756585	−71.923403	137.583643	5.898713	−41.084567	−121.411581
14	−9.848105	−89.810289	10.831061	2.746606	−40.917087	−0.829834
15	−9.439031	−95.146683	13.468115	−0.349854	−40.227845	−4.57803
16	−9.849299	−89.114126	3.590487	4.193917	−39.091219	4.739484
17	−9.184434	−93.800427	13.798652	−0.15029	−39.317679	−4.742056
18	−9.288659	−66.827939	29.246578	−5.526845	−41.293125	−0.61745
19	−9.334523	−97.316801	13.32399	1.643934	−47.602691	−1.520586
20	−9.20397	−88.853522	11.551093	0.841713	−39.949947	−2.839316
21	−8.830986	−87.036877	−0.984167	4.223438	−43.972228	9.738005
22	−9.379841	−105.499758	−0.106462	5.247049	−52.488349	18.74902
23	−9.544745	−92.895047	6.530258	2.25849	−43.676731	6.015347
24	−8.495176	−93.797267	5.049959	4.448976	−51.450948	10.236216
25	−9.267337	−89.014198	4.216711	1.735271	−43.494338	9.406928
26	−9.252118	−105.716646	−14.778726	8.005601	−48.117782	24.318182
27	−9.126283	−85.827991	8.630974	2.064539	−43.780574	6.738237
28	−9.284998	−98.135961	10.252	−0.853204	−43.054507	−0.576413
29	−9.328559	−91.040094	10.676023	−2.096823	−41.607064	0.244939
30	−9.271078	−90.507554	12.30791	3.596532	−45.741513	2.194284
31	−9.758375	−90.926614	7.32146	3.72632	−45.25025	4.014972
32	−9.420349	−93.873985	12.12075	−0.769747	−41.34117	−1.857371
33	−9.525062	−87.660464	7.603173	0.725881	−42.07771	2.034013
34	−8.924793	−88.723043	9.482864	−0.262163	−42.460211	1.346273
35	−8.613479	−100.709886	−2.716151	9.681352	−60.128732	18.761503
36	−8.536865	−93.614435	11.336846	11.333685	−53.050065	2.14422
37	−8.962538	−99.586655	3.685199	2.954378	−46.675344	7.847616
38	−9.228962	−87.721992	8.017987	2.168414	−47.182534	6.855098
39	−9.634012	−116.210802	5.345136	−0.998961	−54.112859	7.003205
40	−8.697722	−100.28545	7.180056	10.542426	−48.637135	2.75002
41	−8.794481	−110.562879	6.887629	5.052888	−57.52481	5.569424
42	−9.123153	−90.310616	7.083512	1.315786	−41.117475	0.259413
43	−9.690708	−85.581175	13.389067	8.955381	−44.388836	−1.284021
44	−9.18053	−92.593898	8.344491	5.057608	−44.805355	−0.484626
45	−9.308603	−86.206399	11.983134	−0.347521	−39.687885	−2.426924
46	−8.78899	−86.663537	2.666339	6.896304	−45.753575	5.801478
47	−9.122766	−76.189667	−0.407226	3.549797	−37.566629	7.769055

Table 4 continued

Compound no.	Docking score (XP GScore)	Prime MMGBSA DG bind	MMGBSA dG bind Coulomb	MMGBSA dG bind covalent	MMGBSA dG bind vdW	MMGBSA dG bind Solv GB
48	−8.605821	−92.631054	5.686294	4.285569	−51.207434	7.86236
49	−9.388033	−81.929099	12.405195	1.701641	−38.574966	−4.168391
50	−9.230797	−89.578194	9.227302	4.259991	−42.154862	2.277227
51	−9.194778	−99.430183	−1.793843	5.089545	−45.576205	8.89159
52	−9.264837	−81.07553	11.096361	0.687301	−40.795182	−0.950016

are shown in Fig. 10. Further, the Prime MMGBSA method was used to calculate the drug receptor binding energy calculation. The binding energies of the different compound receptor complexes were in accordance to the Glide score (Table 4). The docking studies indicated that the binding orientation of the compounds are similar to that of finasteride and the bulk present at C-17 position is mainly responsible for the potent binding of the inhibitors to the enzyme active site.

Conclusion

In the present study, fifty-two 4-azasteroidal inhibitors of the enzyme human steroid 5α-reductase were selected for the development of 3D-QSAR CoMFA and CoMSIA models. The selected molecules were subjected to maximum common substructure alignment method. The CoMFA model suggested high contribution of steric field for inhibition of 5AR. CoMSIA model developed with a combination of steric, electrostatic, hydrophobic and hydrogen bond donor fields showed better predictive ability for a test set of 10 compounds. Further, the contribution of hydrophobic and steric field was found to be more thus indicating, steric and hydrophobic interaction of the inhibitors with the active site of enzyme. Docking studies were also carried out to determine the potential binding interaction between receptor and the inhibitors. Glide XP docking was performed followed by receptor inhibitor free energy of binding calculation using Prime MMGBSA approach. The higher Glide score and free energy of binding indicated the major contribution of bulkiness at C-17 position which is responsible for activity. In conclusion, various 3D-QSAR models were developed using CoMFA and CoMSIA methodologies which could be used for design and optimisation of leads as 5α-reductase inhibitors. The studies were supported by docking studies leading to insight into interaction of inhibitors of the enzyme with the receptor.

Acknowledgments Authors are thankful to University Grant Commission (UGC), New Delhi, India for financial support (File no. 37-319/2009) (SR) and to MSU of Baroda for necessary help regarding computational work.

References

Aggarwal S, Thareja S, Bhardwaj TR, Kumar M (2010a) 3D-QSAR studies on unsaturated 4-azasteroids as human 5alpha-reductase inhibitors: a self organizing molecular field analysis approach. Eur J Med Chem 45(2):476–481.

Aggarwal S, Thareja S, Bhardwaj TR, Kumar M (2010b) Self-organizing molecular field analysis on pregnane derivatives as human steroidal 5alpha-reductase inhibitors. Steroids 75(6): 411–418

Aggarwal S, Thareja S, Verma A, Bhardwaj TR, Kumar M (2010c) An overview on 5alpha-reductase inhibitors. Steroids 75(2): 109–153

Andersson S, Russell DW (1990) Structural and biochemical properties of cloned and expressed human and rat steroid 5 alpha-reductases. Proc Natl Acad Sci U S A 87(10):3640–3644

Azzouni F, Godoy A, Li Y, Mohler J (2012) The 5 alpha-reductase isozyme family: a review of basic biology and their role in human diseases. Adv Urol 2012:530121. doi:10.1155/2012/530121

Bohm M, St rzebecher J, Klebe G (1999) Three-dimensional quantitative structure-activity relationship analyses using comparative molecular field analysis and comparative molecular similarity indices analysis to elucidate selectivity differences of inhibitors binding to trypsin, thrombin, and factor Xa. J Med Chem 42(3):458–477. doi:10.1021/jm981062r

Cilotti A, Danza G, Serio M (2001) Clinical application of 5alpha-reductase inhibitors. J Endocrinol Invest 24(3):199–203

Cramer RD, Patterson DE, Bunce JD (1988) Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins. J Am Chem Soc 110(18):5959–5967. doi:10.1021/ja00226a005

Dessalew N, Patel DS, Bharatam PV (2007) 3D-QSAR and molecular docking studies on pyrazolopyrimidine derivatives as glycogen synthase kinase-3beta inhibitors. J Mol Graph Model 25(6): 885–895

Dill KA (1997) Additivity principles in biochemistry. J Biol Chem 272(2):701–704

Drury JE, Di Costanzo L, Penning TM, Christianson DW (2009) Inhibition of human steroid 5beta-reductase (AKR1D1) by finasteride and structure of the enzyme-inhibitor complex. J Biol Chem 284(30):19786–19790

Faragalla J, Bremner J, Brown D, Griffith R, Heaton A (2003) Comparative pharmacophore development for inhibitors of human and rat 5-alpha-reductase. J Mol Graph Model 22(1): 83–92

Glide version 5.8, Schrödinger LLC. New York (2012)

Gravas S, Oelke M (2010) Current status of 5alpha-reductase inhibitors in the management of lower urinary tract symptoms and BPH. World J Urol 28(1):9–15. doi:10.1007/s00345-009-0493-y

Klebe G, Abraham U, Mietzner T (1994) Molecular similarity indices in a comparative analysis (CoMSIA) of drug molecules to correlate and predict their biological activity. J Med Chem 37(24):4130–4146

Kumar R, Kumar M (2012) 3D-QSAR CoMFA and CoMSIA studies for design of potent human steroid 5α-reductase inhibitors. Med Chem Res 22:1–10. doi:10.1007/s00044-012-0006-1

Kumar R, Kumar A, Jain S, Kaushik D (2011) Synthesis, antibacterial evaluation and QSAR studies of 7-[4-(5-aryl-1,3,4-oxadiazole-2-yl)piperazinyl] quinolone derivatives. Eur J Med Chem 46(9): 3543–3550

Kurup A, Garg R, Hansch C (2000) Comparative QSAR analysis of 5alpha-reductase inhibitors. Chem Rev 100(3):909–924

Lyne PD, Lamb ML, Saeh JC (2006) Accurate prediction of the relative potencies of members of a series of kinase inhibitors using molecular docking and MM-GBSA scoring. J Med Chem 49(16):4805–4808. doi:10.1021/jm060522a

MacroModel version 9.9, Schrodinger LLC. New York (2012)

Mittal RR, McKinnon RA, Sorich MJ (2009) The effect of molecular fields, lattice spacing and analysis options on CoMFA predictive ability. QSAR Comb Sci 28(6–7):637–644. doi:10.1002/qsar.200860128

Prime version 3.1, Schrödinger LLC. New York (2012)

Potshangbam AM, Tanneeru K, Reddy BM, Guruprasad L (2011) 3D-QSAR and molecular docking studies of 2-pyrimidinecarbonitrile derivatives as inhibitors against falcipain-3. Bioorg Med Chem Lett 21(23):7219–7223

Puntambekar D, Giridhar R, Yadav MR (2006) 3D-QSAR studies of farnesyltransferase inhibitors: a comparative molecular field analysis approach. Bioorg Med Chem Lett 16(7):1821–1827

Rasmusson GH, Reynolds GF, Steinberg NG, Walton E, Patel GF, Liang T, Cascieri MA, Cheung AH, Brooks JR, Berman C (1986) Azasteroids: structure-activity relationships for inhibition of 5 alpha-reductase and of androgen receptor binding. J Med Chem 29(11):2298–2315

SYBYL Molecular Modeling System version 7.0 (2003)

Thareja S, Aggarwal S, Bhardwaj TR, Kumar M (2009) Self organizing molecular field analysis on a series of human 5alpha-reductase inhibitors: unsaturated 3-carboxysteroid. Eur J Med Chem 44(12):4920–4925

Uemura M, Tamura K, Chung S, Honma S, Okuyama A, Nakamura Y, Nakagawa H (2008) Novel 5 alpha-steroid reductase (SRD5A3, type-3) is overexpressed in hormone-refractory prostate cancer. Cancer Sci 99(1):81–86

Vedani A, Lill MA, Dobler M (2007) Predicting the toxic potential of drugs and chemicals in silico. ALTEX 24:63–66

Yao Z, Xu Y, Zhang M, Jiang S, Nicklaus MC, Liao C (2011) Discovery of a novel hybrid from finasteride and epristeride as 5alpha-reductase inhibitor. Bioorg Med Chem Lett 21(1): 475–478