Articles

Three-Dimensional Structure−Activity Relationships of Nonsteroidal Ligands in Complex with Androgen Receptor Ligand-Binding Domain

Annu A. Söderholm,[†,‡] Pekka T. Lehtovuori,[†] and Tommi H. Nyrönen*[,†]

CSC—Scientific Computing Ltd., P.O. Box 405, 02101 Espoo, Finland, and Department of Biochemistry and Pharmacy, Åbo Akademi University, P.O. Box 66, 20521 Turku, Finland

Received May 31, 2004

We studied the three-dimensional quantitative structure−activity relationships (3D QSAR) of 70 structurally and functionally diverse androgen receptor (AR) binding compounds using the comparative molecular similarity indices analysis (CoMSIA) method. The compound set contained 67 nonsteroidal analogues of flutamide, nilutamide, and bicalutamide whose binding mode to AR was unknown. Docking was used to identify the preferred binding modes for the nonsteroidal compounds within the AR ligand-binding pocket (LBP) and to generate the ligand alignment for the 3D QSAR analysis. The alignment produced a statistically significant and predictive model, validated by random group cross-validation and external test sets (q^2_{LOO} = 0.656, SDEP = 0.576, r^2 = 0.911, SEE = 0.293; q^2_{10} = 0.612, q^2_5 = 0.571; pred-r^2 = 0.800). Additional model validation comes from the CoMSIA maps that were interpreted with respect to the LBP structure. The model takes into account and links the AR LBP structure, docked ligand structures, and the experimental binding activities. The results provide valuable information on intermolecular interactions between nonsteroidal ligands and the AR LBP.

Introduction

Endogenous androgens testosterone (T) and 5α-dihydrotestosterone (DHT) are essential steroid hormones for the development, maintenance, and regulation of the male phenotype. Their androgenic and anabolic actions are elicited via androgen receptors (ARs), which function as ligand-dependent transcription factors in the regulation of AR target gene expression.[1] AR belongs to the superfamily of nuclear receptors (NRs) and shares a common modular three-domain structure of the NRs. The transcriptional activation of the ARs, like other NRs, is regulated through agonist and antagonist binding resulting in conformational changes in the ligand-binding domain (LBD) and subsequent recruitment of coregulators.[1−3]

In the agonist-bound form of AR the LBD folds into a compact structure with the ligand-binding pocket (LBP) enclosed by the carboxyl-terminal helix 12 (H12) (Figure 1).[4−6] The conformational changes occurring upon antagonist binding are unknown for AR. In related NRs, e.g. estrogen receptor (ER), the active antagonists displace the H12 of the ER LBD into a coactivator recognition surface known as activation function 2 (AF2).[7,8] The displaced H12 thus prevents coactivators from binding to the receptor.[7,9] Antagonist-induced displacement is possible in ER because the H12 contains the same recognition sequence (LXXLL) as ER coactivators that interact with the AF2.[7] However, no evidence has been reported for a similar antagonistic mechanism in AR.

Exogenous AR ligands have for long been used in the treatment of conditions resulting from altered androgen levels or responsiveness.[10] AR targeted pharmaceuticals can structurally be classified into steroidal or nonsteroidal compounds and functionally into androgenic or antiandrogenic compounds. Both androgenic (e.g. testosterone and its derivatives) and antiandrogenic (e.g. cyproterone acetate) functional activities of the steroidal structures are used clinically. Only antiandrogenic compounds of the nonsteroidal class of AR ligands (e.g. bicalutamide, flutamide, and nilutamide) are currently in medicinal use. However, several research groups have recently reported discovery of nonsteroidal androgens.[11−16]

From the pharmacological point of view, nonsteroidal compounds are superior to the steroidal ones in terms of receptor selectivity and pharmacokinetic properties.[17,18] These properties make them more attractive alternatives as novel therapeutic agents. Recently, interest has especially been directed toward the development of selective androgen receptor modulators (SARMs) with tissue-specific agonist and antagonist activities.[10,19]

In the past years a number of experimental structure−activity relationship (SAR) studies have been conducted on a few structural classes of nonsteroidal AR binding compounds.[11−16,20−28] In this study we applied the 3D QSAR analysis based on the comparative molecular similarity indices analysis (CoMSIA)[29] method to gain insight into the structural and chemical features influencing the binding affinity of a panel of 70 AR binding

* To whom correspondence should be addressed. Phone: +358-9-4572235. Fax: +358-9-4572302. E-mail: tommi.nyronen@csc.fi.
† CSC—Scientific Computing Ltd.
‡ Åbo Akademi University.

10.1021/jm0495879 CCC: $30.25
Published on Web 01/26/2005

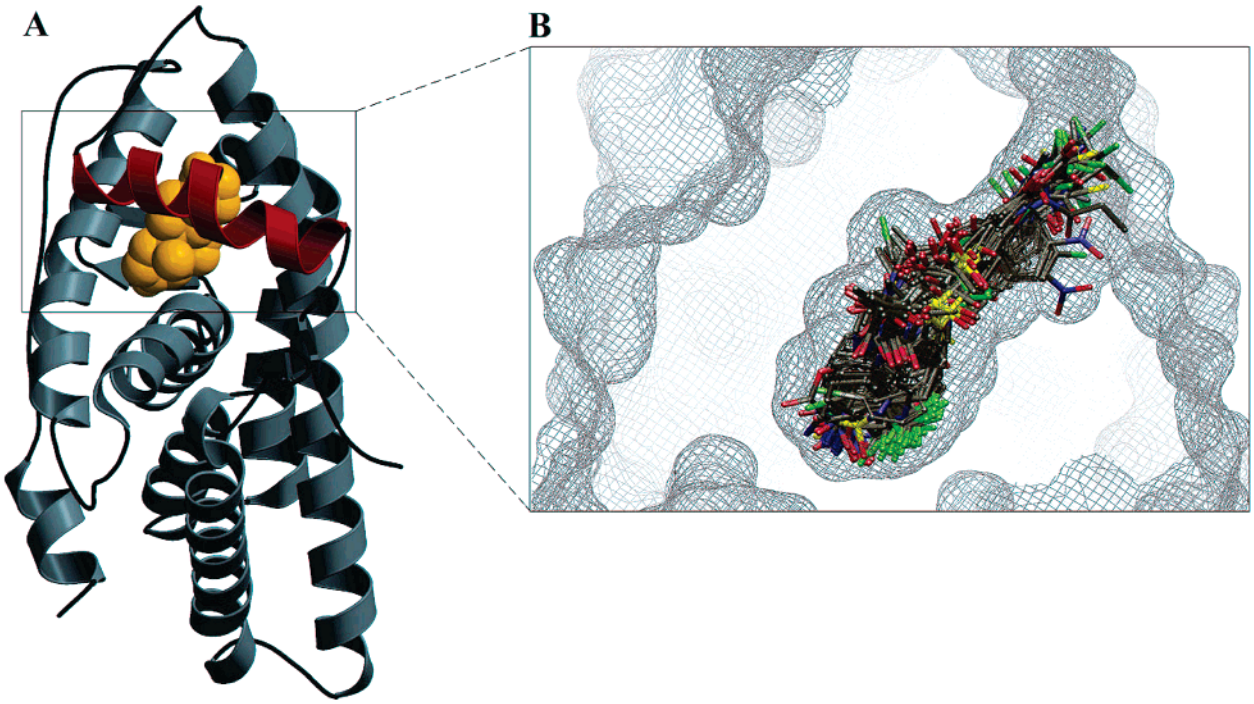


Figure 1. (A) The androgen receptor ligand-binding domain structure presented in the agonist-bound conformation in complex with the natural androgen dihydrotestosterone. Helix 12 is shown in red. (B) A close-up view of the structural alignment of ligands generated by molecular docking. The superposition is shown inside the Connolly solvent accessible surface of the ligand-binding pocket.

compounds containing 67 nonsteroids obtained from the literature.[11,24−27] Since the structure of AR LBD in complex with nonsteroidal ligands has not yet been experimentally resolved, we used molecular docking to predict the biologically active conformations of the nonsteroids and to create the structural alignment of AR ligands for model building. The generation of a reliable 3D QSAR model requires that the ligands be aligned in a way to maximize the 3D overlap of their structural and functional features. To the best of our knowledge, this is the first report within AR research where molecular docking is combined with 3D QSAR analysis. Docking provides a powerful way to screen the conformational space of the ligands in search of the preferred binding conformation, while taking into account the structure and the chemical environment of the AR LBP. Combining docking and the 3D QSAR analysis have been shown to yield predictive 3D QSAR models.[30−32] This approach allows also the visualization and interpretation of the CoMSIA maps to be made within the AR LBP, thus enlightening the interactions that are beneficial or detrimental for the binding affinity of the compounds in their proposed binding mode.

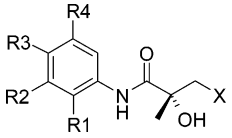
Results and Discussion

Ligand Alignment. The structures and pharmacological data for 70 AR binding compounds used in the 3D QSAR analysis were obtained from five previous publications reported by two laboratories.[11,24−27] The results from the two laboratories were merged in our study, as the laboratories follow similar experimental methods in binding affinity measurements. The steroid reference compound DHT is included in the articles of both laboratories and is reported to have affinity values close to each other thus indicating that the laboratories produce comparable affinity data. The compound series was divided into a training set of 61 compounds and a test set of 9 compounds. The test set was selected to contain compounds spanning a wide activity range and the main structural elements integrated within the training set.

The most crucial and challenging step in 3D QSAR analysis is the generation of the structural alignment of the compounds under study. The aim is to identify an alignment, among numerous possible ones, that represents the biologically active conformations of the compounds. At many instances this task is facilitated if the binding geometry for a compound from the series under investigation has been experimentally observed.[30,32] This was not the situation in our case, however, as the experimentally determined binding mode of the nonsteroidal AR ligands has not been reported to date. Another challenge in the superpositioning was the structural and functional diversity of the compound set. Structurally the set can be divided into six chemical families (Tables 1−6). Primarily, the compounds are derivatives of the nonsteroidal structures flutamide, nilutamide, and bicalutamide. The steroidal structures of T, DHT, and mibolerone, which are known to bind AR LBP, were included in the published binding affinity data sets and incorporated in the examined set of compounds. Functionally the compounds represent the entire spectrum of pharmacological activities ranging from agonists to partial agonists and antagonists. For all of the compounds the functional activity has not been reported. The competitive binding assays indicate, however, that all the compounds do bind to AR LBP, as they are able to displace the high-affinity radioligand mibolerone from binding to the receptor.

Inclusion of compounds with different functionalities can be a potential source of errors in this study, since

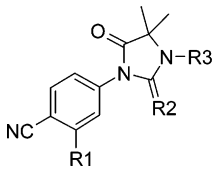
Table 1. Structures of Flutamide Derivatives



compd	R_1	R_2	R_3	R_4	X	pK_{iExp}[b]	pK_{iPred}[c]
1[a,d]	H	CF_3	CN	H	Br	8.10	7.86
2[d]	H	CF_3	NO_2	H	Br	9.52	9.19
3[d]	H	CF_3	CN	H	I	8.55	8.09
4[d]	H	CF_3	NO_2	H	I	9.07	8.80
5[e]	H	H	NO_2	H	Br	7.48	7.82
6[e]	NO_2	H	NO_2	H	Br	5.93	5.63
7[e]	H	NO_2	H	NO_2	Br	7.09	6.67
8[e]	H	CF_3	NH_2	H	Br	7.09	7.17

[a] Compound that belongs to the test set. [b] Experimental binding affinity (pK_i). [c] Predicted binding affinity (pK_i). [d] Reference 11. [e] Reference 27.

Table 2. Structures of Nilutamide Derivatives



compd	R_1	R_2	R_3	pK_{iExp}[b]	pK_{iPred}[c]
9[d]	CF_3	O	CH_3	7.43	7.17
10[d]	CF_3	O	$(CH_2)_2OCH_3$	6.92	7.10
11[d]	CF_3	O	$(CH_2)_2F$	6.85	7.04
12[d]	CF_3	O	$(CH_2)_3F$	6.49	7.04
13[d]	CF_3	O	$(CH_2)_4F$	6.67	7.08
14[d]	CF_3	O	CH_2(*m*-I-Ph)	6.91	6.49
15[d]	CF_3	O	CH_2(*p*-I-Ph)	6.82	6.46
16[d]	CF_3	O	$(CH_2)_2$(*m*-I-Ph)	6.09	5.87
17[a,d]	CF_3	O	$(CH_2)_2$(*p*-I-Ph)	6.38	7.01
18[d]	CF_3	O	$(CH_2)_3$(*m*-I-Ph)	6.44	6.47
19[d]	CF_3	O	$(CH_2)_3$(*p*-I-Ph)	7.19	7.01
20[d]	CF_3	O	$(CH_2)_4OH$	7.59	7.28
21[e]	CF_3	S	$(CH_2)_4OH$	8.65	8.42
22[e]	I	S	$(CH_2)_4OH$	9.15	8.96

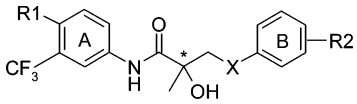
[a–c] See Table 1. [d] Reference 26. [e] Reference 25.

the molecular mechanism for AR antagonism and the resulting conformational changes are not known. It is possible that AR displays an antagonistic mechanism similar to the antagonistic mechanism of ERs that involves displacement of H12 into the cofactor-binding cleft on the surface of LBD.[7,9] However, H12 of AR LBD does not possess the same recognition sequence as the recognition sequences proposed for cofactors and AR amino-terminal sequences that target AR LBD.[33,34] Also, AR antagonists in this study do not have an extension or "arm", like typical ER antagonists.[8] Thus, we expect the conformational changes induced by AR antagonists in our ligand set to cause less violent changes in AR LBD than those caused by ER antagonists in ER LBD. For these reasons we did not model the antagonist-bound structure of AR LBD on the basis of known antagonist forms of other steroid receptors. Instead, we aligned all the ligands using a protein model based on the agonist structure of AR LBD (PDB ID: 1gs4).[6]

Docking Simulations. We used the docking program GOLD[35] to identify the most favorable conformations for the nonsteroidal compounds within the AR LBP and to study the ligand−receptor interactions. The structural water molecule, which is found in most of the ligand-complexed NR LBD structures (like ER, GR, PR, and AR), was kept as part of the protein structure during the docking.

Crystal structures of AR LBD complexed to steroidal agonists, both natural and synthetic, have revealed the common binding mode and the key interactions contributing to the binding of steroids.[4−6] Besides the van der Waals interactions between the steroid core and hydrophobic residues of the LBP, the steroids form three hydrogen bonds to the receptor in the wild-type AR. Hydrogen bonds are formed to Arg752 at the carbonyl

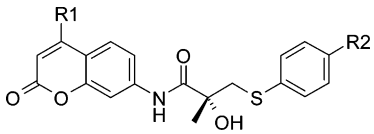
Table 3. Structures of Bicalutamide Derivatives



compd	R_1	R_2	X	isomer	pK_{iExp}[b]	pK_{iPred}[c]
23[d]	CN	*p*-F	SO_2	*R*	7.96	7.32
24[d]	CN	*p*-F	SO_2	*S*	6.44	6.45
25[a,d]	CN	*p*-NH_2	S	*R*	7.05	6.88
26[a,d]	CN	*p*-NH_2	S	*S*	6.10	6.70
27[d]	CN	*m*-NH_2	S	*R*	7.19	7.25
28[d]	CN	*p*-NCS	S	*S*	6.37	6.63
29[d]	CN	*m*-NCS	S	*R*	6.64	7.03
30[d]	CN	*m*-NCS	S	*S*	6.89	6.94
31[d]	CN	*p*-NCS	SO_2	*R*	7.39	7.34
32[d]	CN	*m*-NCS	SO_2	*R*	6.85	7.28
33[d]	CN	*p*-$NHCOCH_2Cl$	S	*R*	8.78	8.42
34[d]	CN	*p*-$NHCOCH_2Cl$	S	*S*	6.70	6.84
35[d]	CN	*p*-$NHCOCH_2Br$	S	*R*	7.00	7.38
36[d]	CN	*p*-$NHCOCH_2Cl$	SO_2	*R*	7.97	7.75
37[a,d]	CN	*m*-$NHCOCH_2Cl$	SO_2	*R*	6.89	7.06
38[d]	CN	*p*-$NHCOCH_2Br$	SO_2	*R*	6.44	6.92
39[d]	CN	*p*-NH_2	SO_2	*R*	7.49	7.51
40[d]	CN	*p*-$NHCOCH_3$	S	*R*	8.31	8.26
41[d]	CN	*p*-$NHCOCH_3$	S	*S*	5.86	6.25
42[d]	CN	*p*-$N(COCH_3)_2$	S	*R*	6.98	6.90
43[d]	CN	*p*-$NHCOCH_3$	SO_2	*R*	7.80	8.07
44[d]	CN	*p*-$NHCOCH_2CH_3$	S	*R*	7.57	7.52
45[d]	CN	*p*-$N(COCH_2CH_3)_2$	S	*R*	6.18	6.59
46[a,e]	NO_2	*p*-NH_2	S	*R*	7.24	7.86
47[a,e]	NO_2	*p*-NH_2	S	*S*	6.28	6.19
48[e]	NO_2	*p*-$NHCOCH_3$	S	*R*	8.54	8.68
49[e]	NO_2	*p*-$NHCOCH_3$	S	*S*	6.91	6.83
50[e]	NO_2	*p*-$NHCOCH_3$	SO_2	*R*	8.03	8.39
51[e]	NO_2	*p*-$NHCOCH_3$	SO_2	*S*	6.23	6.09
52[e]	NO_2	*p*-$NHCOCF_3$	S	*R*	8.59	8.82
53[a,e]	NO_2	*p*-$NHCOCF_3$	SO_2	*R*	8.21	8.65
54[e]	NO_2	*p*-$NHCOCH_2Cl$	S	*R*	8.48	8.63
55[e]	NO_2	*p*-$NHCOCH_2Cl$	S	*S*	7.01	7.23
56[e]	NO_2	*p*-$NHCOCH_2Cl$	SO_2	*R*	8.16	7.73
57[e]	NO_2	*p*-$NHCOCH_2Cl$	SO_2	*S*	6.63	6.12
58[e]	NO_2	*p*-$NHSO_2CH_3$	S	*R*	7.30	7.34
59[e]	NO_2	*p*-$NHSO_2CH_3$	SO_2	*R*	7.82	8.19

[a–c] See Table 1. [d] Reference 24. [e] Reference 27.

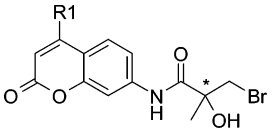
Table 4. Structures of Bicalutamide Derivatives Bearing a Coumarine Ring



compd	R_1	R_2	pK_{iExp}[b]	pK_{iPred}[c]
60[d]	CH_3	F	7.04	7.16
61[d]	CF_3	H	6.88	6.80
62[a,d]	CF_3	F	6.47	6.81
63[d]	CF_3	NCS	6.87	6.77
64[d]	CF_3	NH_2	6.54	6.65

[a–c] See Table 1. [d] Reference 27.

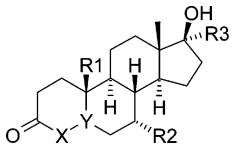
Table 5. Structures of Hydroxyflutamide Derivatives Bearing a Coumarine Ring



compd[a]	R_1	isomer	pK_{iExp}[b]	pK_{iPred}[c]
65[d]	CH_3	*R*	6.56	6.53
66[d]	CF_3	*R*	6.88	6.97
67[d]	CF_3	*S*	6.61	6.61

[a] No compounds in this table belong to the test set. [b,c] See Table 1. [d] Reference 27.

Table 6. Structures of Steroids



compd[a]	X−Y	R_1	R_2	R_3	pK_{iExp}[b]	pK_{iPred}[c]
68[d]	CH_2−CH	CH_3	H	H	9.55	9.36
69[e]	CH−C	CH_3	H	H	8.85	8.97
70[e]	CH−C	H	CH_3	CH_3	9.12	9.14

[a] No compounds in this table belong to the test set. [b,c] See Table 1. [d] Reference 11. [e] Reference 25.

oxygen end of the steroid and to Asn705 and Thr877 at the hydroxyl group end of the steroid (Table 6).[4,5] Hydrogen bonding is considered essential also for nonsteroidal ligand binding.[20,21] The significance of Asn705 for the activity and binding of nonsteroidal antagonists hydroxyflutamide and bicalutamide (casodex) has been demonstrated in a mutation study.[36]

Docking simulations were started using a modified AR LBP structure in which two mutations (His701Leu and Ala877Thr) restoring the wild-type LBD sequence were made. Upon docking into this LBP many of the large nonsteroidal compounds with several aromatic rings adopt "sandwich-like" conformations where the rings lie on top of each other. These conformations are not likely the bioactive conformations but merely the result of forcing the ligands to fit into the volume and shape of the agonist-bound conformation of AR LBP. These docking poses do not satisfy important ligand−receptor interactions, and the alignment of such poses does not provide an explanation for the variance of the biological data. To emphasize the hydrogen-bonding interactions of the steroidal and nonsteroidal ligands to Asn705, DHT was used to guide docking, although it is not structurally an optimal template for the large and flexible nonsteroids. Additionally, another mutation (Phe876Ala) was made to allow rational alignment of the bulky nonsteroidal ligands in the LBP. The enlarged volume generated by this mutation corresponds to a region surrounding the five-ring of the steroidal scaffold in their bioactive conformation, i.e. the region on the opposite side of the LBP from the structural water molecule. This mutation resulted in extended binding conformations of nonsteroidal ligands and enabled key interactions between ligand and receptor to be formed. The Phe876Ala mutation could partly account for the structural changes that we assume to take place in the LBD upon binding of large nonsteroidal compounds, which would otherwise have problems to fit in the AR binding cavity.

The aim of docking was to generate the ligand alignment as automatically and objectively as possible. Unfortunately, the selection of top-scored conformations based on the scoring functions of GOLD, XSCORE,[37] or CSCORE[38] (including the individual scoring functions implemented in CSCORE) did not produce alignments resulting in statistically significant models. To obtain a well-superposed set of ligands, the alignment based on GOLD scoring was improved by manually selecting docking poses for ligands that deviated from the alignment. In the set of best-ranked conformations according to GOLD there were 10 ligands (**12**, **14**, **22**, **32**, **36**, **44**, **48**, **50**, **52**, **54**) whose structural features were not superimposed with the rest of the docked ligand set. For these ligands the aromatic A-ring and its substitutions (Table 3) were not aligned with the rest of the docked ligands and did not produce statistical correlation with experimental binding affinities in 3D QSAR analysis using the CoMSIA method. Also the amide group next to the A-ring was often not oriented similarly as in the other docked ligands. For these 10 problematic ligands, GOLD simulation did generate conformations whose structural and functional elements were aligned with the majority of the ligands but which were not top-ranked in GOLD scoring. An average decrease of 2.0 in the GOLD fitness score was detected between the best-ranked conformation and the conformations where the structural and functional elements for these 10 ligands were aligned with the rest of the ligand set. These differences in the GOLD score can be considered to be of minor importance. The molecular alignment resulting from docking of the training set compounds is shown in Figure 1 within the AR LBP visualized with the Connolly solvent accessible surface calculated in BODIL.[39]

3D QSAR Model. The molecular alignment derived from docking simulations was used for the generation of the 3D QSAR model based on the CoMSIA method.[29] This approach allowed us to study the physicochemical characteristics that contribute to the binding affinity of the investigated set of AR ligands. Hydrophobic and hydrogen bond acceptor fields and five PLS components in model building provided the best explanation for the variation in the binding affinity data. The hydrophobic fields contribute 51% to the information content of the final model, while the hydrogen bond acceptor fields represent the remaining 49%.

The leave-one-out (LOO) cross-validation used to determine the optimum number of PLS components for model building produced a correlation coefficient q^2_{LOO} of 0.656 and standard error of prediction (SDEP) of 0.576. Internal correlation coefficient values for 10 random group cross-validation yielded an average q^2_{10} of 0.612 and SDEP of 0.612, and those for five random group cross-validation yielded an average q^2_5 of 0.571 and SDEP of 0.643. The conventional correlation coefficient r^2 obtained by the nonvalidated analysis gave 0.911 with a standard error of estimate (SEE) of 0.293. The q^2 and r^2 values of our 3D QSAR model indicate a statistically significant and stable model.

The r^2 value signifies that approximately 91% of the variance in ligand binding of the present series of

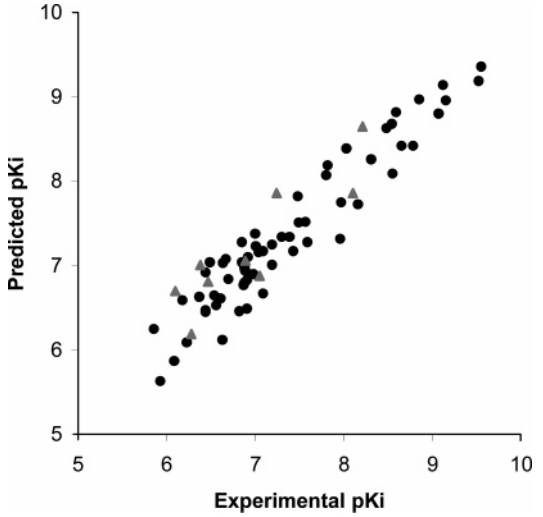


Figure 2. Correlation between the experimental and predicted activities (pK_i) for the training set compounds (●) and test set compounds (▲).

compounds can be explained with the model. The standard error value of the model means that if a compound were predicted to have a 10 nM (K_i) affinity, its actual affinity would fall into a range between 5 and 20 nM ($10^{-8\pm0.3}$ M). The error of model can never be less than the experimental error of the data that was used to train the model. The experimental and predicted binding affinity values of the training set ligands are listed in Tables 1−6 and graphed in Figure 2.

Six compounds from the publications where the binding affinity data was collected were excluded from this study (compounds 16, 39, 47, 58 of Yin et al.,[27] compound 13 of van Dort et al.,[26] and compound (*R*)-7 of Kirkovsky et al.[24]). They turned out as outliers in 3D QSAR model building and were thus removed from the final model. Compounds 16, 39, 47, and 58 from Yin et al.[27] were excluded from the ligand set because their structures caused problems in molecular modeling. Compound 16 has a metal-containing B-ring substituent (*p*-Sn$(CH_3)_3$) and compound 39 has an A-ring substitution (*p*-NCS) that are not present in any other investigated compound. Compound 47, on the other hand, is the only structure of the bicalutamide derivatives bearing a coumarine ring (Table 4) with a sulfone ($-SO_2-$) linkage, and compound 58 is the only nilutamide derivative bearing a coumarine ring in the entire compound set. Compounds with singular substituents are likely to cause structural changes in the AR LBP that we were unable to predict with the current modeling programs. Compound 13 from van Dort et al.[26] has an *N*-iodopropenyl side chain with a highly reactive double bond. Due to the apparent reactivity we decided to remove it from the training set. For the removal of compound (*R*)-7 from Kirkovsky et al.[24] we were unable to find any structural explanation. This compound repeatedly turned out as an outlier in all CoMSIA models we attempted to generate. This result led us to suspect that perhaps the reported experimental value might be erroneous.

Besides internal validation with the training set, we used an external test set of 9 compounds to validate the predictivity of the model. The external validation process can be considered the most valuable validation method, as these compounds are completely excluded during the training of the model. Prior to prediction, the test set compounds were processed identically to the training set compounds, as described in the Methods section. Visual observation of the top-ranked docking poses by GOLD showed the conformations to be well structurally aligned with the training set compounds and thus were used for the prediction of the binding affinity. All of the test set compounds, which represent the different structural properties incorporated within the training set, are well predicted without any apparent outliers. Low- and high-affinity compounds are clearly separated in the prediction. We chose not to include compounds from the highest affinity end of the scale into the test set as, after all, there are only a few sub-nanomolar binders and they are valuable for the 3D QSAR model building. The chosen external test set yielded a predictive r^2 of 0.800 with a SEE value of 0.367. These values indicate a good predictive power and are in agreement with the statistical values from the internal validation procedures. The predicted versus experimental affinity values for the test set compounds are listed and plotted with the training set compounds in Tables 1−6 (marked with a) and Figure 2, respectively.

While our study was underway, Marhefka et al.[40] reported a new series of bicalutamide-like compounds in which the metabolically susceptible sulfur linkage was replaced with either oxygen or nitrogen, and new substitutions were introduced into the B-ring. We did not include this high-affinity compound series in the 3D QSAR model building because the training set already had 31 structurally similar compounds (Table 3), but rather used them to further validate our model. We produced 10 possible poses of each of these compounds using the docking procedure described in the Methods section. We then used our 3D QSAR model to predict the binding affinity of each of these poses and picked the one that was predicted to be the best. The docked conformations that give the highest predictions in the 3D QSAR model are well aligned with the ligand set used in this study. The predicted vs the experimental binding affinity value yields a SEE of 0.6. This value corresponds to a situation where a compound predicted to have a 10 nM (K_i) affinity in reality shows an affinity in the range between 2 and 40 nM ($10^{-8\pm0.6}$ M). The binding mode inside AR LBP of these novel high-affinity AR ligands can thus be identified with our 3D QSAR model. A similar approach can be used in the identification of active compounds from a molecular library in virtual screening, too. The compound is docked several times to the receptor model, and the representative docking pose for a novel compound is then picked using the best prediction according to the 3D QSAR model. If the selected docking pose is structurally aligned with the training set compounds in this study, the predicted affinity gives a good estimate of the actual binding affinity and the ligand conformation takes into account the chemical and structural features of the AR LBP. The limitation is that the model can only predict K_i of compounds that are structurally described by the training set.

Visualization of the 3D QSAR Model. The statistically relevant results from the 3D QSAR analysis are

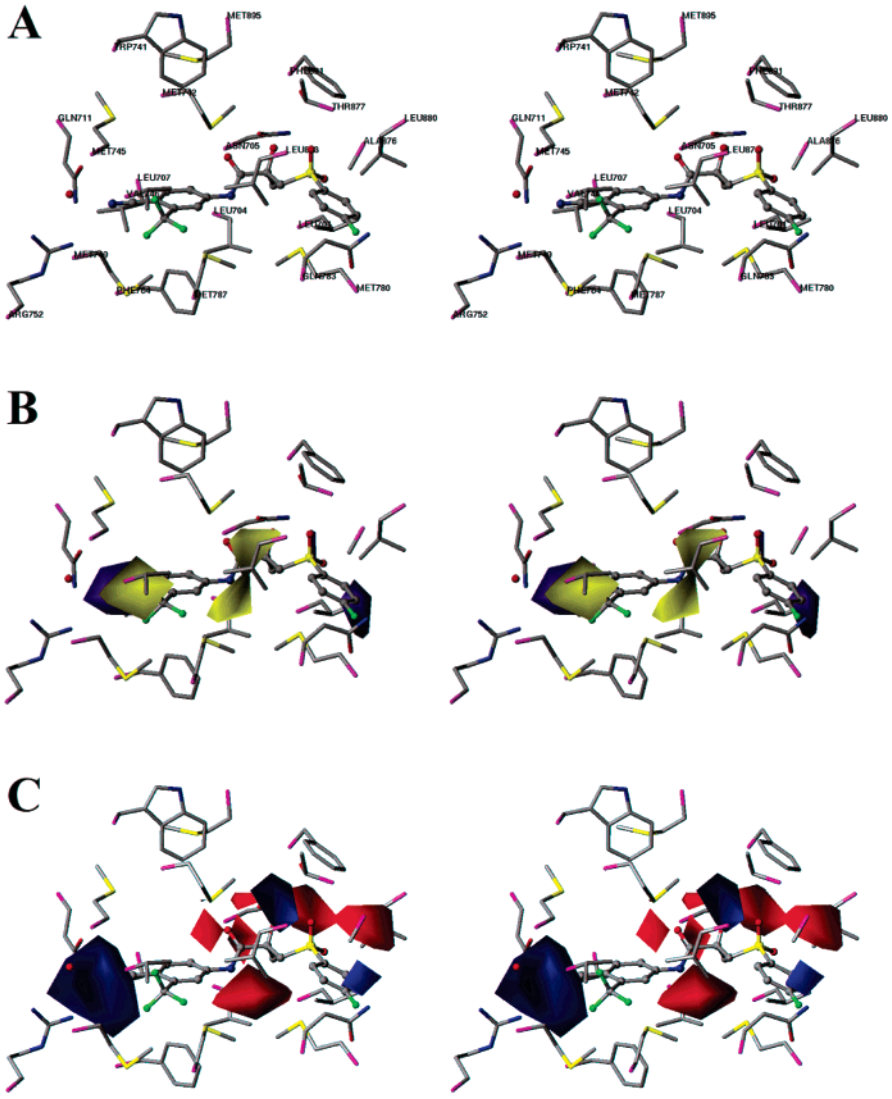


Figure 3. Stereoviews of the LBP and the 3D contour maps of the CoMSIA model represented as stdev*coeff plots. The favorable and unfavorable maps are shown with contribution levels of 80% and 20%, respectively. (A) The LBP residues (sticks) with the docked conformation of compound 23 ((*R*)-bicalutamide) (ball-and-stick). All hydrogens are omitted for clarity, and C^{α}-carbons are colored in purple. The structural water is shown as a red sphere. (B) The yellow contours represent regions in ligands where an increase in hydrophobicity is connected with enhanced affinity; purple contours represent regions where hydrophobicity is detrimental to affinity, thus favoring hydrophilicity. (C) The blue contours indicate volumes where hydrogen bond acceptors in ligands increase binding affinity; red contours indicate areas where hydrogen bond acceptors decrease binding affinity.

visualized as 3D contour maps. The contour maps of our CoMSIA model based on hydrogen bond acceptor and hydrophobic fields are displayed as PLS stdev*coeff maps in Figure 3. Since the superpositioning of ligands for the analysis was done using the receptor structure and ligand docking, the CoMSIA maps can be drawn inside the AR binding cavity. The maps of the 3D QSAR model based on the chemical properties and molecular interaction fields of ligands should correlate with the features of the LBP. Also, the ligand interactions with the residues of the binding cavity should provide an explanation of the variation of the experimental binding affinity. In Figure 3, the contour maps of our CoMSIA model are displayed with the binding site residues and with the docked conformation of ligand **23** ((*R*)-bicalutamide) as a reference structure.

Interpretation of the 3D QSAR Maps with Respect to AR LBP. For the analysis of the CoMSIA maps we divided the LBP roughly into three sections: the outer (solvent accessible) part, the center part, and the inner part. The contours of the favorable hydrophobic and hydrogen bond acceptor fields used in the analysis are drawn at a contribution level of 80%, while the corresponding unfavorable fields are contoured at a contribution level of 20% (Figure 3).

The Outer Part. There is a small favored volume for acceptor interactions close to the solvent accessible surface, located near the amide group of residue Gln783 and the backbone carbonyl oxygen of Arg779 (not shown in Figure 3). The volume next to the side chain of Leu880 and the carbonyl oxygen of Phe876 (Ala876 in the receptor model used for docking) in the outer part of the binding site is unfavorable for acceptor interactions. There is a small hydrophilic volume located above the carbonyl oxygen of Ser778 and close to the volume that is unfavorable for acceptor interactions. Consistent with the aqueous surface, this part of the binding site has no favored volumes for hydrophobic interactions. Receptor structure and the 3D QSAR model consistently indicate that hydrophobic parts of

the ligands that come close to the solvent surface are not beneficial for binding affinity. It is sensible that hydrophilicity is favored in this volume, since polar water molecules can form interactions with the ligands that reach the outer part of the binding site.

The Center Part. According to the 3D QSAR model a volume in the center of the LBP, lined by the main chain carbonyl of Leu873, the aliphatic carbons of side chain Thr877, and the terminal methyl group of Met742, is defined as favorable for acceptor interactions. This result is rather surprising, because at first glance the interactions provided by these side chains seem hydrophobic. Our first impression was that the contour could be an artifact. However, Superstar[41,42] calculations (data not shown) with water oxygen probe indicate that the volume could accommodate acceptor interactions even though these interactions do not seem very strong. The favored acceptor contour in this volume can partially explain the high binding affinity of e.g. compounds **36**, **50**, and **56**. In their docked conformations a sulfone linkage ($-SO_2-$) is positioned close to this volume. Although a general observation from the experimental results is that in most cases bicalutamide derivatives with a sulfide linkage show higher affinity than ones with a sulfone linkage, the effect of the linkage also largely depends on the substituent and its position in the aromatic B-ring.[24,27]

There are two volumes in the center of the LBP where visualized CoMSIA maps indicate unfavorable acceptor interactions. These volumes are placed on the opposite sides of the superimposed ligands. The first unfavorable volume is located in the vicinity of the side chain carbonyl of Asn705 and the side chain aliphatic carbon of Leu704, and on top of the plane of the peptide bond between Leu704 and Asn705. This volume is clearly not a favorable site for acceptor interactions. The unfavorable acceptor interaction volume extends toward the side chain hydroxyl group of Thr877 and the hydrophobic residues Leu701, Leu880, and Phe891. While the hydroxyl group of Thr877 could provide an interaction that is constructive with acceptors, the other residues make the volume mostly hydrophobic. The second unfavorable volume for acceptor interactions is also indicated as favorable for hydrophobic interactions. Residues Met742, Met780, Met787, Leu873, and one edge of the aromatic ring of Phe764 form the borders of this volume in the center of the binding cavity. Thus, both the unfavorable acceptor and the favorable hydrophobic CoMSIA contours are in accordance with the receptor structure. On the whole, the CoMSIA maps described above show that the center of the LBP favors binding of hydrophobic moieties. Moreover, the center part lacks volumes defined as unfavorable for hydrophobic interactions. This result is in agreement with the observed interactions between the endogenous steroid ligand core and the AR LBP.[5]

The Inner Part. The inner part of the LBP, next to the side chains of Gln711 and Arg752 and the structural water molecule, displays a large favored acceptor volume that partly overlaps with a favored hydrophilic volume. Both maps agree with each other and with the receptor structure, since the hydrogen-bonding network between side chains of Gln711 and Arg752 and the water molecule are most likely able to donate a hydrogen bond. The inner part of the binding site displays a CoMSIA volume favoring hydrophobicity in a region above the plane of the peptide bond between Met745 and Val746 and close to the side chains of Met745, Val746, and Met749. In this part of the LBP the CoMSIA model is devoid of areas that are unfavorable for acceptor interactions.

Conclusions

In this paper we report a 3D QSAR model (CoMSIA) for a series of 70 AR binding compounds containing 67 nonsteroids. The set contains six structurally distinct scaffolds and a variety of pharmacological activities. Because the binding mode for nonsteroidal AR ligands was unknown, we used molecular docking to identify the bioactive conformations, and to superpose the structurally and functionally diverse ligands. Docking screens the conformational space of the ligands in search of the preferred binding conformation, while taking the chemical and structural features of the AR LBP into account.

The superposition of the ligands made with docking produced a statistically significant 3D QSAR model that has been carefully validated. Further validation for the 3D QSAR model comes from the interpretation of the contour maps explaining variation of the binding affinity of the ligands. The 3D QSAR model is compatible with the protein environment in the binding site, as the interpretation of the contour maps can be reflected to the amino acids of the AR LBP. These results indicate that the superposition is likely to represent the biologically active conformations of the nonsteroidal ligands.

The steps in the model building process are interdependent. The docking simulations depend on the structure of the AR LBP, and the 3D QSAR analysis in turn depends on the superposition of the ligands produced with the docking simulations. The experimentally observed changes in the binding affinity were mapped back to the structural features of the ligands in the 3D QSAR analysis, and the interpretation of the 3D QSAR model fits the chemical environment of AR LBP. Analysis of the maps is also consistent with the possible interactions formed between the nonsteroidal ligands and the AR LBP residues in the docking simulations. As a whole, there is a chain of dependencies between AR LBP, docking simulations, ligand alignment, statistical analysis, and experimental binding activities. The statistical validation thus confirms that there is a link between the different steps in the model building process and the reported experimental binding affinities. Together, the results provide valuable information on how nonsteroidal ligands bind and interact with the AR LBP. Nevertheless, it would be interesting to see whether similar binding geometries as presented here would be detected by experimental structure determination methods.

The procedure described in the paper can be used for, e.g., automated virtual identification of high-affinity AR ligands from chemical databases and structure-based optimization of AR ligands.

Methods

The Protein Data. The crystal structure of androgen receptor LBD in complex with an agonist 9α-fluorocortisol (PDB ID: 1gs4)[6] used in docking simulations was retrieved

from the Protein Data Bank.[43] 1gs4 is a mutant structure with two mutations in the active site (Leu701His, Thr877Ala). Prior to docking, His701Leu and Ala877Thr mutations restoring the wild-type AR sequence were made. Side chain conformations corresponding to ones observed in the crystal structure of AR LBD from PDB ID 1e3g[4] were fixed using the rotamer library in the BODIL software.[39] An additional mutation (Phe876Ala) was made to allow more space to the solvent surface side of the binding cavity. 9α-Fluorocortisol and all water molecules except the structural water located between residues Gln711 and Arg752 (Wat957 in 1gs4) were removed from the crystal structure. All hydrogens of the protein structure and of the water molecule were added using Sybyl 6.9.1.[44]

The Ligand Data. The structural and pharmacological data for 70 AR binding compounds were obtained from five publications reported by two laboratories.[11,24−27] Compounds lacking defined stereochemistry or exact affinity values were excluded from the data set. The results reported by the two laboratories were combined into our study, as the experimental procedures for binding affinity measurements are highly similar. The affinity measurements have been performed using competitive binding assay with cytosolic AR from rat ventral prostate. ^{3}H-Mibolerone was used as the high-affinity ligand in the measurements and triamcinolone acetonide to block interaction of ^{3}H-mibolerone with glucocorticoid and progesterone receptors. Moreover, the incubation times and temperatures in binding assays were the same and the hydroxylapatite precipitation was used as a method to determine the protein bound radioactivity from free radioactivity. The binding affinities reported as K_i (nM) were converted to pK_i values ($-\log K_i$) for the derivation of the 3D QSAR model.

The ligands were converted into 3D structures for docking using CORINA, version 2.6.[45,46] The maximum number of ring conformations for one compound was restricted to three while using an energy window of 20 kJ/mol between the best and the worst conformations. This setting yielded a single conformation for each ligand, except testosterone (2 conformations) and mibolerone (3 conformations). Gasteiger−Hückel atom charges[47,48] used for 3D QSAR analysis were calculated in Sybyl 6.9.1.

Docking Simulations. The docking program GOLD, version 2.0,[35] was used to predict the bioactive conformations and binding modes of the ligands within AR LBP. The binding pocket was defined as a cavity within a radius of 15 Å from atom C9 in 9α-fluorocortisol in the crystal structure of 1gs4. The docking procedure was repeated 10 times for each of the CORINA-generated ligand conformations using the standard docking parameters in the GOLD manual. Docking of a single ligand conformation was allowed to terminate if the three top-scoring solutions were within a 1.5 Å root-mean-squared deviation (rmsd) of each other.

Docking of the nonsteroidal ligands was biased to emphasize similar hydrogen-bonding patterns as seen in the experimentally determined binding mode of the natural ligand DHT. The crystal structure of AR LBD with DHT (PDB ID: 1i37)[5] was superimposed with our receptor model used in docking. The bound DHT was then extracted and merged into the receptor model with BODIL to be used as a template for docking. The template similarity constraint option of GOLD was applied by evaluating the overlap of all donor atoms and all acceptor atoms as well as shape overlap between DHT and the ligand being docked, with constraint weights 5.0, 5.0, and 10.0, respectively.

Selection of Docking Poses for 3D QSAR Analysis. Attempts to pick the top-scoring ligand poses using merely GOLD, XSCORE,[37] or CSCORE[38] scoring functions did not produce statistically significant alignments in the 3D QSAR analysis. As a consequence, manual selection of docking poses that we based on GOLD scoring was included into the alignment generation. For 10 of the ligands the top-ranked docking pose was not properly superimposed with the rest of the docked ligand set. For these problematic cases representatives that are consistent with the alignment produced by the majority of the ligands were manually selected among the docking poses generated by GOLD simulation.

3D QSAR Analysis. The molecular alignment of the 61 docked ligands in the training set was used to build the 3D QSAR model with the CoMSIA method[29] to explore physicochemical properties contributing to binding affinity. The CoMSIA molecular descriptor fields, expressed as steric, electrostatic, hydrophobic, and hydrogen-bonding properties, were calculated using the default settings in Sybyl and correlated with the variations in the binding affinity data using the statistical method of partial least squares (PLS).[49] Molecular field descriptors with an energy variance less than 2.0 kcal/mol were filtered out from the PLS analysis. The 3D QSAR model was built using hydrophobic and hydrogen bond acceptor fields. The LOO cross-validation method was applied to determine the optimum number of PLS components, corresponding to the highest q^2 value and to the lowest SDEP value. The optimum of five components derived from the LOO cross-validation was used in the development and further validation of the 3D QSAR model.

The predictive value of the 3D QSAR model was validated first with internal cross-validation using 10 and five random groups. Due to the structural variability and the relatively small size of the training set, we did not apply random group cross-validation with a smaller number of groups. Each random group cross-validation procedure was repeated 25 times to calculate the mean q^2 values. Further validation of the predictivity of the model was done with an external test set of 9 compounds not included in 3D QSAR model building. The correlation between the experimental and the predicted activities for the test set compounds for the 3D QSAR model is represented as the predictive r^2 value.

Acknowledgment. This work was funded by ISB—Graduate School for Informational and Structural Biology and the Finnish Technology Agency (TEKES). Computational resources and program licenses were supplied by CSC—Scientific Computing Ltd. and Fatman Bioinformational Designs Ltd. We thank Hormos Medical Corp. for helpful discussions regarding SARMs.

References

(1) Gobinet, J.; Poujol, N.; Sultan, C. Molecular action of androgens. *Mol. Cell. Endocrinol.* **2002**, *198*, 15−24.
(2) Bourguet, W.; Germain, P.; Gronemeyer, H. Nuclear receptor ligand-binding domains: three-dimensional structures, molecular interactions and pharmacological implications. *Trends Pharmacol. Sci.* **2000**, *20*, 381−388.
(3) Berrevoets, C. A.; Umar, A.; Brinkmann, A. O. Antiandrogens: selective androgen receptor modulators. *Moll. Cell. Endocrinol.* **2002**, *198*, 97−103.
(4) Matias, P. M.; Donner, P.; Coelho, R.; Thomaz, M.; Peixoto, C.; et al. Structural evidence for ligand specificity in the binding domain of the human androgen receptor. *J. Biol. Chem.* **2000**, *275*, 26164−26171.
(5) Sack, J. S.; Kish, K. F.; Wang, C.; Attar, R. M.; Kiefer, S. E.; et al. Crystallographic structures of the ligand-binding domains of the androgen receptor and its T877A mutant complexed with the natural agonist dihydrotestosterone. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 4904−4909.
(6) Matias, P. M.; Carrondo, M. A.; Coelho, R.; Thomaz, M.; Zhao, X.-Y.; et al. Structural basis for the glucocorticoid response in a mutant human androgen receptor (Ar^{ccr}) derived from an androgen-independent prostate cancer. *J. Med. Chem.* **2002**, *45*, 1439−1446.
(7) Shiau, A. K.; Barstad, D.; Loria, P. M.; Cheng, L.; Kushner, P. J.; et al. The structural basis of estrogen receptor/coactivator recognition and the antagonism of this interaction by tamoxifen. *Cell* **1998**, *95*, 927−937.
(8) Shiau, A. K.; Barstad, D.; Radek, J. T.; Meyers, M. J.; Nettles, K. W.; et al. Structural characterization of a subtype-selective ligand reveals a novel mode of estrogen receptor antagonism. *Nat. Struct. Biol.* **2002**, *9*, 359−364.
(9) Brzozowski, A. M.; Pike, A. C.; Dauter, Z.; Hubbard, R. E.; Bonn, T.; et al. Molecular basis of agonism and antagonism in the oestrogen receptor. *Nature* **1997**, *389*, 753−758.
(10) Zhi, L.; Martinborough, E. Selective androgen receptor modulators (SARMs). *Annu. Rep. Med. Chem.* **2001**, *36*, 169−180.

(11) Dalton, J. T.; Mukherjee, A.; Zhu, Z.; Kirkovsky, L.; Miller, D. D. Discovery of nonsteroidal androgens. *Biochem. Biophys. Res. Commun.* **1998**, *244*, 1−4.
(12) Edwards, J. P.; West, S. J.; Pooley, C. L. F.; Marschke, K. B.; Farmer, L. J.; et al. New nonsteroidal androgen receptor modulators based on 4-(trifluoromethyl)-2(1H)-pyrrolidino[3,2-g] quinolinone. *Bioorg. Med. Chem. Lett.* **1998**, *8*, 745−750.
(13) Hamann, L. G.; Mani, N. S.; Davis, R. L.; Wang, X. N.; Marschke, K. B.; et al. Discovery of a potent, orally active, nonsteroidal androgen receptor agonist: 4-ethyl-1,2,3,4-tetrahydro-6-(trifluoromethyl)-8-pyridono[5,6-*g*]- quinoline (LG121071). *J. Med. Chem.* **1999**, *42*, 210−212.
(14) Edwards, J. P.; Higuchi, R. I.; Winn, D. T.; Pooley, C. L. F.; Caferro, T. R.; et al. Nonsteroidal androgen receptor agonists based on 4-(trifluoromethyl)-2H-pyrano[3,2-g]quinolin-2-one. *Bioorg. Med. Chem. Lett.* **1999**, *9*, 1003−1008.
(15) Zhi, L.; Tegley, M. T.; Marschke, K. B.; Jones, T. K. Switching androgen receptor antagonists to agonists by modifying C-ring substituents on piperidino[3,2-g]quinolinone. *Bioorg. Med. Chem. Lett.* **1999**, *9*, 1009−1012.
(16) Higuchi, R. I.; Edwards, J. P.; Caferro, T. R.; Ringgenberg, J. D.; Kong, J. W.; et al. 4-Alkyl- and 3,4-dialkyl-1,2,3,4-tetrahydro-8-pyridono[5,6-g]quinolines: potent, nonsteroidal androgen receptor agonists. *Bioorg. Med. Chem. Lett.* **1999**, *9*, 1335−1340.
(17) Cockshott, I. D.; Cooper, K. J.; Sweetmore, D. S.; Blacklock, N. J.; Denis, L. The pharmacokinetics of casodex in prostate cancer patients after single and during multiple dosing. *Eur. Urol.* **1990**, *18*, 10−17.
(18) Teutsch, G.; Goubet, F.; Battmann, T.; Bonfils, A.; Bouchoux, F.; et al. Non-steroidal antiandrogens: synthesis and biological profile of high-affinity ligands for the androgen receptor. *J. Steroid Biochem. Mol. Biol.* **1994**, *48*, 111−119.
(19) Negro-Vilar, A. Selective androgen receptor modulators (SARMs): a novel approach to androgen therapy for the new millennium. *J. Clin. Endocrinol. Metab.* **1999**, *84*, 3459−3462.
(20) Tucker, H.; Crook, J. W.; Chesterson, G. J. Nonsteroidal antiandrogens. Synthesis and structure−activity relationships of 3-substituted derivatives of 2-hydroxypropionanilides. *J. Med. Chem.* **1988**, *31*, 954−959.
(21) Morris, J. J.; Hughes, L. R.; Glen, A. T.; Taylor, P. J. Nonsteroidal antiandrogens. Design of novel compounds based on an infrared study of the dominant conformation and hydrogen-bonding properties of a series of anilide antiandrogens. *J. Med. Chem.* **1991**, *34*, 447−455.
(22) Hamann, L. G.; Higuchi, R. I.; Zhi, L.; Edwards, J. P.; Wang, X. N.; et al. Synthesis and biological activity of a novel series of nonsteroidal, peripherally selective androgen receptor antagonists derived from 1,2-dihydropyridono[5,6-*g*]quinolines. *J. Med. Chem.* **1998**, *41*, 623−639.
(23) Burden, P. M.; Ai, T. H.; Lin, H. Q.; Akinci, M.; Costandi, M.; et al. Chiral derivatives of 2-cyclohexylideneperhydro-4,7-methanoindenes, a novel class of nonsteroidal androgen receptor ligand: synthesis, X-ray analysis, and biological activity. *J. Med. Chem.* **2000**, *43*, 4629−4635.
(24) Kirkovsky, L.; Mukherjee, A.; Yin, D.; Dalton, J. T.; Miller, D. D. Chiral nonsteroidal affinity ligands for the androgen receptor. 1. Bicalutamide analogues bearing electrophilic groups in the B aromatic ring. *J. Med. Chem.* **2000**, *43*, 581−590.
(25) Van Dort, M. E.; Robins, D. M.; Wayburn, B. Design, synthesis, and pharmacological characterization of 4-[4,4-dimethyl-3-(4-hydroxybutyl)-5-oxo-2-thioxo-1-imidazolidinyl]-2-iodobenzonitrile as a high-affinity nonsteroidal androgen receptor ligand. *J. Med. Chem.* **2000**, *43*, 3344−3347.
(26) Van Dort, M. E.; Jung, Y. W. Synthesis and structure-activity studies of side-chain derivatized arylhydantoins for investigation as androgen receptor radioligands. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 1045−1047.
(27) Yin, D.; He, Y.; Perera, M. A.; Hong, S. S.; Marhefka, C.; et al. Key structural features of nonsteroidal ligands for binding and activation of the androgen receptor. *Mol. Pharmacol.* **2003**, *63*, 211−223.
(28) He, Y.; Yin, D.; Perera, M.; Kirkovsky, L.; Stourman, N.; et al. Novel nonsteroidal ligands with high binding affinity and potent functional activity for the androgen receptor. *Eur. J. Med. Chem.* **2002**, *37*, 619−634.
(29) Klebe, G.; Abraham, U.; Mietzner, T. Molecular similarity indices in a comparative analysis (CoMSIA) of drug molecules to correlate and predict their biological activity. *J. Med. Chem.* **1994**, *37*, 4130−4146.
(30) Sippl, W. Binding affinity prediction of novel estrogen receptor ligands using receptor-based 3-D QSAR methods. *Bioorg. Med. Chem.* **2002**, *10*, 3741−3755.
(31) Wolohan, P.; Reichert, D. E. CoMFA and docking study of novel estrogen receptor subtype selective ligands. *J. Comput.-Aided Mol. Des.* **2003**, *17*, 313−328.
(32) Tervo, A. J.; Nyrönen, T. H.; Rönkkö, T.; Poso, A. Comparing the quality and predictiveness between 3D QSAR models obtained from manual and automated alignment. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 807−816.
(33) He, B.; Kemppainen, J. A.; Wilson, E. M. FXXLF and WXXLF sequences mediate the NH2-terminal interaction with the ligand binding domain of the androgen receptor. *J. Biol. Chem.* **2000**, *275*, 22986−22994.
(34) He, B.; Minges, J. T.; Lee, L. W.; Wilson, E. M. The FXXLF motif mediates androgen receptor-specific interactions with coregulators. *J. Biol. Chem.* **2002**, *277*, 10226−10235.
(35) Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.* **1997**, *267*, 727−748.
(36) Poujol, N.; Wurtz, J.-M.; Tahiri, B.; Lumbroso, S.; Nicolas, J.-C.; et al. Specific recognition of androgens by their nuclear receptor. *J. Biol. Chem.* **2000**, *275*, 24022−24031.
(37) Wang, R.; Lai, L.; Wang, S. Further development and validation of empirical scoring functions for structure-based binding affinity prediction. *J. Comput.-Aided Mol. Des.* **2002**, *16*, 11−26.
(38) Clark, R. D.; Strizhev, A.; Leonard, J. F.; Matthew, J. B. Consensus scoring for ligand/protein interactions. *J. Mol. Graphics Modell.* **2002**, *20*, 281−295.
(39) Lehtonen, J. V.; Still, D.-J.; Rantanen, V.-V.; Ekholm, J.; Björklund, D.; et al. BODIL: a molecular modeling environment for structure-function analysis and drug design, *J. Comput.-Aided Mol. Des.* **2004**, *18*, 401−419.
(40) Marhefka, C. A.; Gao, W.; Chung, K.; Kim, J.; He, Y.; et al. Design, synthesis, and biological characterization of metabolically stable selective androgen receptor modulators. *J. Med. Chem.* **2004**, *47*, 993−998.
(41) Verdonk, M. L.; Cole, J. C.; Taylor, R. SuperStar: a knowledge-based approach for identifying interaction sites in proteins. *J. Mol. Biol.* **1999**, *289*, 1093−1108.
(42) Verdonk, M. L.; Cole, J. C.; Watson, P.; Gillet, V.; Willett, P. SuperStar: improved knowledge-based interaction fields for protein binding sites. *J. Mol. Biol.* **2001**, *307*, 841−859.
(43) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; et al. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235−242.
(44) *Sybyl 6.9.1*; Tripos Inc.: St. Louis, MO, 2003.
(45) Sadowski, J.; Gasteiger, J. From atoms and bonds to three-dimensional atomic coordinates: model builders. *Chem. Rev.* **1993**, *93*, 2567−2581.
(46) Sadowski, J.; Gasteiger, J.; Klebe, G. Comparison of automatic three-dimensional model builders using 639 X-ray structures. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 1000−1008.
(47) Purcell, W. P.; Singer, J. A. A brief review and table of semiempirical parameters used in the Hückel molecular orbital method. *J. Chem. Eng. Data* **1967**, *12*, 235−246.
(48) Gasteiger, J.; Marsili, M. Iterative partial equalization of orbital electronegativity: a rapid access to atomic charges. *Tetrahedron* **1980**, *36*, 3219−3222.
(49) Cramer, R. D., III; Bunce, J. D.; Patterson, D. E. Crossvalidation, bootstrapping, and partial least squares compared with multiple regression in conventional QSAR studies. *Quant. Struct.-Act. Relat.* **1988**, *7*, 18−25.

JM0495879