Journal of Enzyme Inhibition and Medicinal Chemistry, December 2005; 20(6): 581–585

Three-dimensional model of cytochrome P450 human aromatase

CEDRIC LOGE[1], MARC LE BORGNE[1], PASCAL MARCHAND[1], JEAN-MICHEL ROBERT[1], GUILLAUME LE BAUT[1], MARTINA PALZER[2], & ROLF W. HARTMANN[2]

[1]*Département de Pharmacochimie, Biomolécules et Cibles Thérapeutiques, UPRES EA1155, UFR Sciences Pharmaceutiques, 1 rue Gaston Veil, F-44035 Nantes cedex 1, France, and* [2]*Pharmazeutische und Medizinische Chemie, Universität des Saarlandes, P.O. Box 151150, D-660041 Saarbrücken, Deutschland*

(Received 5 January 2005; accepted 3 March 2005)

Abstract

A three-dimensional (3-D) structure of human aromatase (CYP19) was modeled on the basis of the crystal structure of rabbit CYP2C5, the first solved X-ray structure of an eukaryotic cytochrome P450 and was evaluated by docking *S*-fadrozole and the steroidal competitive inhibitor (19*R*)-10-thiiranylestr-4-ene-3,17-dione, into the enzyme active site. According to a previous pharmacophoric hypothesis described in the literature, the cyano group of *S*-fadrozole partially mimics the steroid backbone C(17) carbonyl group of (19*R*)-10-thiiranylestr-4-ene-3,17-dione, and was oriented in a favorable position for H-bonding with the newly identified positively charged residues Lys119 and Arg435. In addition, this model is consistent with the recent combined mutagenesis/modeling studies already published concerning the roles of Asp309 and His480 in the aromatization of the steroid A ring.

Keywords: *Breast cancer, CYP19, aromatase, homology modeling, active site, inhibitors, S-fadrozole*

Introduction

Breast cancer is the most frequent cancer amongst women. Laboratory work in this therapeutic field aims at developing structures able to inhibit human cytochrome P450 aromatase, also called CYP19. This enzyme is a member of the heme-containing cytochrome P450 superfamily which requires 3 moles of oxygen and 3 moles of NADPH to function for every mole of C(19) steroid metabolized [1,2]. CYP19 catalyses the bioconversions of androgens (androstenedione, testosterone) into estrogens (estrone, estradiol), through the aromatization of the A ring (Figure 1). In this process, the 19-methyl is removed. Taking into account the known proliferative effects of estrogens, CYP19 constitutes an important target in the treatment strategy of hormone-dependent breast cancers in postmenopausal women. Aromatase inhibitors may be divided into two major classes, the steroidal and the non-steroidal compounds (Figure 2). The steroidal compounds such as formestane and exemestane—also referred to as type I agents or inactivators—are analogues of androstenedione and bind competitively but irreversibly to the substrate-binding site of the enzyme [3,4]. The non-steroidal inhibitors, previously termed type II inhibitors, interfere competitively but reversibly with the heme Fe^{2+} ion of the cytochrome P450 moiety of aromatase. Marketed non-steroidal inhibitors include the prototype aminoglutethimide (first-generation) and azole derivatives such as fadrozole (second-generation), letrozole and anastrazole (third-generation) [5,6]. Recent efforts are being carried out both to identify new molecules of therapeutic interest and to clarify the mechanism of action that could be improved by a better knowledge of the enzyme's active site. However,

Correspondence: C. Loge, Département de Pharmacochimie, Biomolécules et Cibles Thérapeutiques, UPRES EA1155, UFR Sciences Pharmaceutiques, 1 rue Gaston Veil, F-44035 Nantes cedex 1, France. Tel: 33 240 411 108. Fax: 22 240 412 876. E-mail: cedric.loge@univ-nantes.fr

ISSN 1475-6366 print/ISSN 1475-6374 online © 2005 Taylor & Francis
DOI: 10.1080/14756360500220574

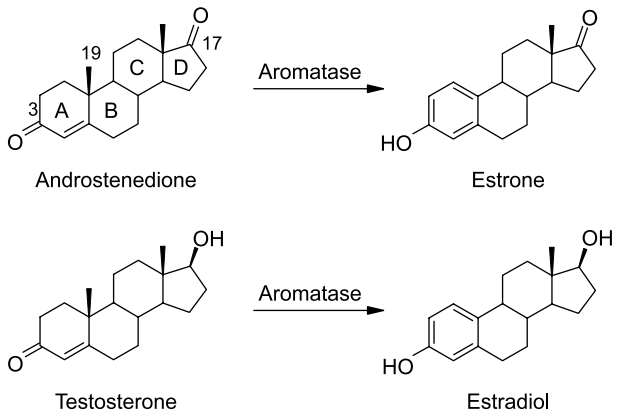

Figure 1. Aromatization of androstenedione and testosterone to estrone and estradiol, respectively.

crystallization of the human membrane-bound enzyme has been difficult to accomplish, and only this result would allow an X-ray analysis of the structure. So far, numerous molecular models of aromatase have been published [7–12], but all are based on the crystallographic data of P450cam (CYP101), terp (CYP108) and bm3 (CYP102), coming from bacterial sources whose percent sequence identity with aromatase ranges from 13–18%. In 2000, the crystal structure of rabbit CYP2C5 was reported as the first solved X-ray structure of a mammalian microsomal cytochrome P450 [13]. This prompted us to construct a 3-D structure of human CYP19, by the homology modeling technique, using the structure of CYP2C5 complexed with a relatively general inhibitor of human 2C enzymes as a template [14].

Materials and methods

Molecular modeling studies were performed using Sybyl software version 6.9.1† running on a Silicon Graphics Octane workstation. The crystallographic structure of rabbit P450 2C5 complexed with a substrate, 4-methyl-*N*-methyl-*N*-(2-phenyl-2*H*-pyrazol-3-yl)benzene sulfonamide (DMZ), at 2.3 Å resolution (pdb code: 1N6B) was used as a template [14]. The amino acid sequence of the human aromatase was aligned to that of the rabbit P450 2C5 using ClustalW [15]. This alignment was further checked by comparing a secondary structure elements prediction for aromatase, obtained through the PSIPRED protein structure prediction server [16], with the experimental secondary structure assignments for CYP2C5 deduced from the PDB file. The 3-D model of the human aromatase was then constructed by the Nest program from the protein structure modeling package JACKAL [17]. The resulting model was subjected to an energy minimization using Powell's method

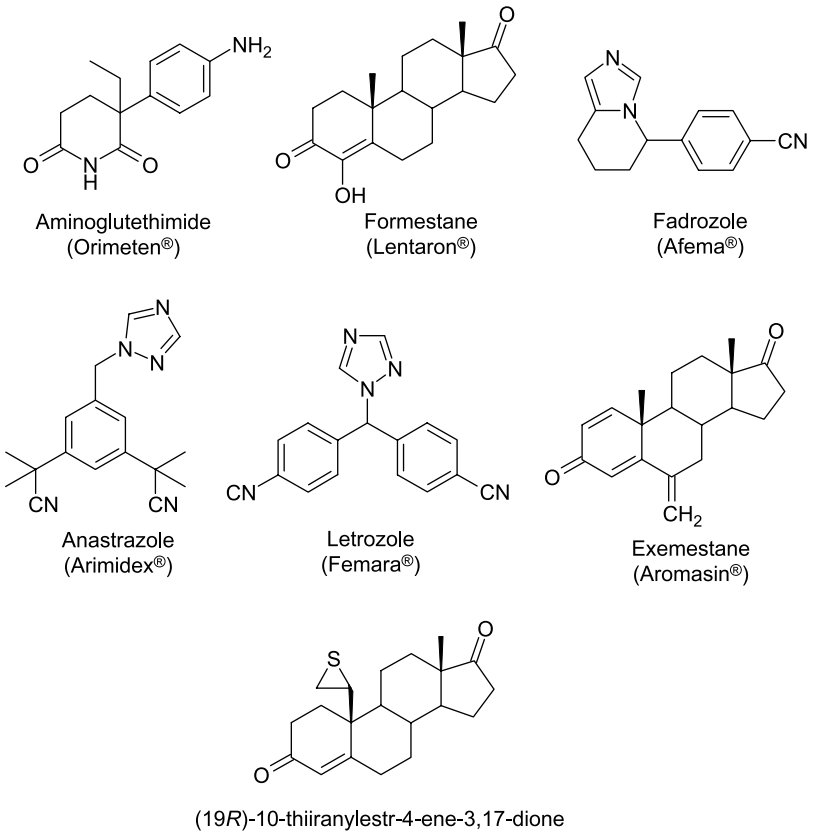

Figure 2. Chemical structures of some aromatase inhibitors.

available in Maximin2 procedure with the Tripos force field [18] and a dielectric constant of 4.0 until the gradient value reached 0.1 kcal/mol Å. Energy minimization was started with the core side chains, then the core main chains. The structure was periodically checked by the Sybyl ProTable module. If present, improper geometries were manually corrected and the structure minimized again with the above procedure. Heme extracted from the rabbit P450 2C5 structure was added to the human aromatase model, and residues neighboring heme were minimized to avoid steric conflicts.

The 3-D structures of *S*-fadrozole and (19*R*)-10-thiiranylestr-4-ene-3,17-dione were constructed using the standard sketch procedure of Sybyl and their geometries were subsequently optimized using the Tripos force field including the electrostatic term calculated from Gasteiger and Hückel atomic charges. Powell's method available in Maximin2 procedure was used for energy minimization until the gradient value was smaller than 0.001 kcal/mol Å. Flexible docking of these compounds into the enzyme active site was performed using GOLD software [19]. For each compound, the most stable docking model was selected according to the best scored conformation predicted by the GoldScore scoring function. The complexes were energy-minimized using Powell's method available in Maximin2 procedure with the Tripos force field and a dielectric constant of 4.0, until the gradient value reached 0.05 kcal/mol Å.

Results and discussion

Homology modeling of aromatase

We constructed the 3-D structure of human CYP19 using the crystal structure of the rabbit CYP2C5 complexed with a dimethyl derivative of sulfaphenazole as a template [14]. As a result of the increased resolution (2.3 vs. 3.0 Å), the overall geometry of CYP2C5 is improved relative to the previously published structure [13]. Sequence alignment was performed as described in Materials and methods section, and the results are shown in Figure 3. The N-terminal residues Met1-Ile47 were not constructed because the coordinates of the corresponding region of CYP2C5 were not determined. Sequence identity between CYP2C5 and aromatase is about 20%.

Highly conserved and functionally important residues in the CYP superfamily are: (i) three absolutely conserved residues, ExxR in the K helix and C just before the L helix (for example Glu362, Arg365 and Cys437 in CYP19); (ii) the consensus sequence (A/G)xx(E/D)T in the center of the I helix around the conserved Thr residue (Thr310 in Cyp19); (iii) the consensus sequence F(G/S)xGx (R/H)xCxGxx(I/L/F)A containing the cysteine (Cys437 in CYP19) responsible for heme binding [20]. As shown in Figure 3, these important sequences of both CYPs are correctly aligned with each other. In addition, from Figure 3, it results that the secondary structure element predictions for aromatase ($H =$ helix; $E =$ strand) match reasonably well with the secondary structure elements determined in CYP2C5 by X-ray analysis. This correspondence, despite the low degree of homology between the two sequences, might confer an acceptable feasibility to the aromatase model.

The heme is sandwiched between the L helix including its N-terminal loop and the I helix. The sulfide of Cys437 provides the axial ligand at the fifth coordination site of the heme iron (Figure 4). Coordination of the heme iron by a cysteine plays a central role in the capacity of P450s to catalyze the scission of dioxygen bound to the sixth coordination site of the heme iron. The structural conservation of the heme binding site is likely to reflect structural features necessary to preserve the capacity of P450s to catalyze this reaction [13]. The propionate side chains of the heme interact with Arg435 before the L helix, Trp141 and Arg145 in the C helix, three highly conserved residues across mammalian P450s, and with Lys376 in a β-strand (Figure 4).

Amino acids whose side chains face the active site pocket are the hydrophobic residues Leu122, Ile125, Phe134, Trp224, Ile229, Ala306, Ala307, Val369, Val370, Leu372, Val373, Met374, Leu477 and Leu479, and the hydrophilic residues Lys119, Asp309, Thr310, Arg375, Lys376, Arg435, Ser478 and His480 (Figure 5).

Binding site evaluation

In order to assess the validity of our model, *S*-fadrozole, responsible for the high aromatase inhibitory activity, and the steroidal competitive inhibitor (19*R*)-10-thiiranylestr-4-ene-3,17-dione (Figure 2) were docked into the active site as described in Materials and methods section, and their positions together with some surrounding residues are shown in Figure 6. Both inhibitors coordinate to the iron atom of the heme group present at the active site of the enzyme by their imidazole and thiirane rings. According to previous studies [2,21], the 4-cyanobenzyl moiety present in *S*-fadrozole partially mimics the steroid backbone C(17) carbonyl group of (19*R*)-10-thiiranylestr-4-ene-3,17-dione, assuming that the binding mode is representative of that of the natural substrate androstenedione [21,22]. In addition, as suggested by S. Chen et al. [23], the C(17) carbonyl binding region of the active site could anchor the D ring of the steroid or participate in binding various electronegative groups of inhibitors by acting as hydrogen bond donor with the positively charged residues Lys119 and Arg435.

```
                                                                              HH
CYP2C5      ---------------MAKKTSSKGKLP--------------------PGPTPFPIIGNIL  (43)
aromatase   MVLEMLNPIHYNITSIVPEAMPAATMPVLLLTGLFLLVWNYEGTSSIPGPGYCMGIGPLI  (60)
                HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH                 HH

                          A                            B
            H         HHHHHHHHHHHH  EEEEE    EEEEEE HHHHHHHH     HHHH EE
CYP2C5      QID---AKDISKSLTKFSECYGPVFTVYLGMKPTVVLHGYEAVKEALVD--LGEEFAGRG  (98)
aromatase   SHGRFLWMGIGSACNYYNRVYGEFMRVWISGEETLIISKSSSMFHIMKHNHYSSRFGSKL  (120)
            HHHHHH    HHHHHHHHHHHH  EEEEEE  EEEEEE  HHHHHHHHHH

                                        C                         D
              HHHHHH          HHHHHHHHHHHHHHH          HHHHHHHHHHHHHHHHH
CYP2C5      SVPILEKVSKGLGIAFSNAKTWKEMRRFSLMTLRNFGMGKRSIEDRIQEEARCLVEELRK  (158)
aromatase   GLQCIGMHEKGI-IFNNNPELWKTTRPFFMKALS--GPGLVRMVTVCAESLKTHLDRLEE  (177)
             HHH       EE EEE   HHHHHHHHHHHH      HHHHHHHHHHHHHHHHHHHHHHHHH

                          E                        F
                      HHHHHHHHHHHHHHH        HHHHHHHHHHHHHHHHH       HHHH
CYP2C5      TNASPCDPTFILGCAPCNVICSVIFHNRFDYKDEEFLKLMESLHENVELLGTPWLQVYNN  (218)
aromatase   VTNESGYVDVLTLLRRVMLDTSNTLFLRIPLDESAIVVKIQGYFDAWQAL----LIKPDI  (233)
            HHH      HHHHHHHHHHHHHHHHHHH        HHHHHHHHHHHHHHHHH     H    H

                             G                             H
              HHHHH HHHHHHHHHHHHHHHHHHHHHHHHHHH            HHHHHHHH
CYP2C5      FPALLDYFPGIHKTLLKNADYIKNFIMEKVKEHQKLLDVNNPRDFIDCFLIKMEQENNLE  (278)
aromatase   FFKISWLYKKYEKSVKDLKDAIEVLIAEKRRRISTEEKLEECMDFATELILAEKRG---D  (290)
            HHH      HHHHHHHHHHHHHHHHHHHHHHHHH         HHHHHHHHHHHHH

                          I                        J
              HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH HHHHHHHHHHHHHH         HH
CYP2C5      FTLESLVIAVSDLFGAGTETTSTTLRYSLLLLLKHPEVAARVQEEIERVIGRHRSPCMQD  (338)
aromatase   LTRENVNQCILEMLIAAPDTMSVSLFFMLFLIAKHPNVEEAIIKEIQTVIG-ERDIKIDD  (349)
              HHHHHHHHHHHHH      HHHHHHHHHHHHHH    HHHHHHHHHHHH         HHH

                   K
            HH  HHHHHHHHHHHHHH        EE     EEE   EEE     EEEEEHHHHH
CYP2C5      RSRMPYTDAVIHEIQRFIDLLPTNLPHAVTRDVRFRNYFIPKGTDIITSLTSVLHDEKAF  (398)
aromatase   IQKLKVMENFIYESMRYQPVVDLVMRKALEDDVIDG-YPVKKGTNIILNIG-RMHRLEFF  (407)
            HH  HHHHHHHHHHHH                 EEEE EEEEE

                                                        L
                 HHHH                              HHHHHHHHHHHHHHHHHHHHEEEE
CYP2C5      PNPKVFDPGHFLDESGNFKKSDYFMPFSAGKRMCVGEGLARMELFLFLTSILQNFKLQSL  (458)
aromatase   PKPNEFTLENFAKN----VPYRYFQPFGFGPRGCAGKYIAMVMMKAILVTLLRRFHVKTL  (463)
                    HHHHH          EEE           HHHHHHHHHHHHHHHHHHHH EEE

              HHH      EE     E    E       EEEE
CYP2C5      VEPKDLDITAVVNGFVS---VPPSYQLCFIPI     (487)
aromatase   QGQCVESIQKIHDLSLHPDETKNMLEMIFTPR     (495)
                       EEE        EEEEEEE
```

Figure 3. Alignment of the primary sequences of rabbit P450 2C5 and aromatase. Secondary structure elements retrieved from the PDB file (for CYP2C5) and predicted (for aromatase) are shown above and below the CYP2C5 and the aromatase sequences, respectively (H = helix; E = strand). The A–L helices are labeled as defined by Wester et al. [14]. Highly conserved and functionally important residues in the CYP superfamily are depicted by red letters.

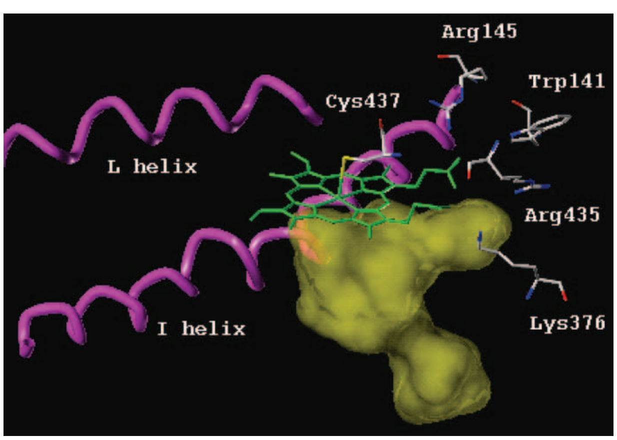

Figure 4. Interaction of the heme with some amino acid residues. In yellow, a view of the active site pocket (MOLCAD surface; program Sybyl 6.9.1).

Moreover, recent site-directed mutagenesis studies suggested that Ser478 and His480 probably play an important role in the catalysis and are likely to be closer to the active site than previously predicted [24]. In particularly, it is proposed that Asp309 and His480 are involved in the aromatization of the 3-keto androgen. It is thought that Asp309 facilitates the enolization by abstracting a hydrogen atom from C-2. His480 donates its hydrogen to the 3-keto group and receives one hydrogen from Asp309 [24]. Our results are in-line with these earlier studies. The three residues Asp309, Ser478 and His480 are positioned in front of the A ring of the steroid (Figure 6) and could form a catalytic triad consistent with the reported mechanism.

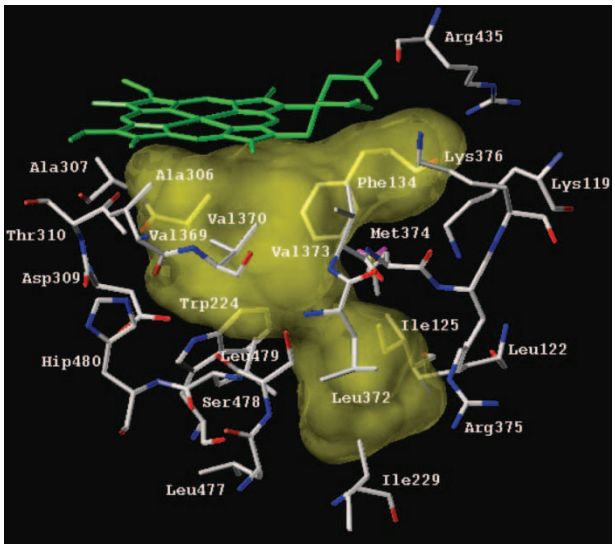

Figure 5. Heme and amino acid residues constituting the active site pocket.

Conclusion

In summary, we have developed a new model of aromatase based on the crystal structure of rabbit CYP2C5, and studied the docking of *S*-fadrozole and (19*R*)-10-thiiranylestr-4-ene-3,17-dione into the enzyme active site.

According to the pharmacophoric hypothesis described by P. Furet et al. [21], the cyano group of *S*-fadrozole partially mimics the steroid backbone C(17) carbonyl group of (19*R*)-10-thiiranylestr-4-ene-3,17-dione, and was oriented in a favorable position for H-bonding with the new identified positively charged residues Lys119 and Arg435. In addition, our model is consistent with the recent combined mutagenesis/modeling studies published by Y. C. Kao et al. concerning the roles of Asp309 and His480 in the aromatization of the steroid A ring [24].

This new model will help us to formulate strategies for design of new aromatase inhibitors. Mutagenesis studies on the newly identified amino acids Lys119 and Arg435 should confirm or disagree with our hypotheses.

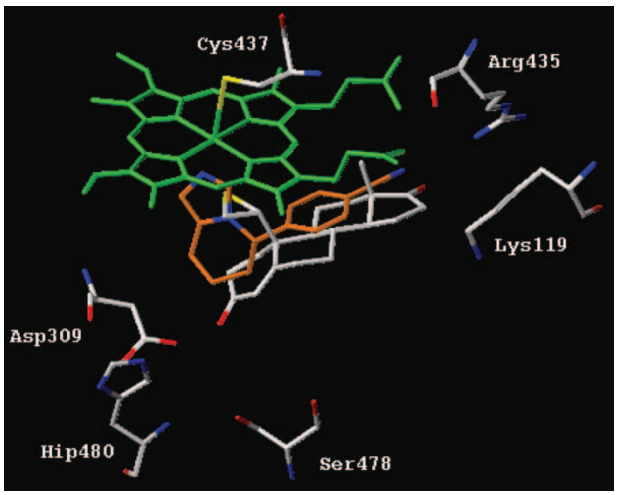

Figure 6. Best docking solutions of *S*-fadrozole and (19*R*)-10-thiiranylestr-4-ene-3,17-dione into the enzyme active site.

Note

[†]Sybyl 6.9.1, Tripos Associates, Inc. 1699 South Hanley Road, St. Louis, MO 63144, USA.

References

[1] Thompson EA, Siiteri PK. J Biol Chem 1974;249:5373.
[2] Ahmed S. Drug Des Discov 1998;15:239.
[3] Johnston SR, Dowsett M. Nat Rev Cancer 2003;3:821.
[4] Lonning PE. Endocr Relat Cancer 2004;11:179.
[5] Mokbel K. Int J Clin Oncol 2002;7:279.
[6] Miller WR. Semin Oncol 2003;30:3.
[7] Auvray P, Nativelle C, Bureau R, Dallemagne P, Séralini GE, Sourdaine P. Eur J Biochem 2002;269:1393.
[8] Laughton CA, Zvelebil MJ, Neidle SJ. Steroid Biochem Mol Biol 1993;44:399.
[9] Koymans LMH, Moereels H, Vanden Bossche H. Steroid Biochem Mol Biol 1995;53:191.
[10] Graham-Lorence S, Amarneh B, White RE, Peterson JA, Simpson ER. Protein Sci 1995;4:1065.
[11] Zhou D, Cam LL, Laughton CA, Korzekwa KR, Chen S. J Biol Chem 1994;269:19501.
[12] Cavalli A, Greco G, Novellino E, Recanatini M. Bioorg Med Chem 2000;8:2771.
[13] Williams PA, Cosme V, Sridhar EF, Johnson DE, McRee DE. J Inorg Biochem 2000;81:183.
[14] Wester MR, Johnson EF, Marques-Soares C, Dansette PM, Mansuy D, Stout CD. Biochemistry 2003;42:6370.
[15] Thompson JD, Higgins DG, Gibson TJ. Nucleic Acids Res 1994;22:4673.
[16] McGuffin LJ, Bryson K, Jones DT. Bioinformatics 2000;16:404, http://bioinf.cs.ucl.ac.uk/psipred/.
[17] Xiang Z, Honig B. J Mol Biol 2001;421, http://trantor.bioc.columbia.edu/programs/jackal/.
[18] Clarck M, Cramer III, RD, Van Opdenbosch N. J Comput Chem 1989;10:982.
[19] Jones G, Willett P, Glen RC. J Mol Biol 1995;245:43.
[20] Graham-Lorence S, Peterson JA. Methods Enzymol 1996;272:315.
[21] Furet P, Batzl C, Bhatnagar A, Francotte E, Rihs G, Lang M. J Med Chem 1993;36:1393.
[22] Kellis TJ, Childers WE, Robinson CH, Vickery LE. J Biol Chem 1987;262:4421.
[23] Chen S, Zhang F, Sherman MA, Kijima I, Cho M, Yuan YC, Toma Y, Osawa Y, Zhou D, Eng ET. J Steroid Biochem Mol Biol 2003;86:231.
[24] Kao YC, Korzekwa KR, Laughton CA, Chen S. Eur J Biochem 2001;268:243.