Biochemical Pharmacology 78 (2009) 406–413

Contents lists available at ScienceDirect

Biochemical Pharmacology

journal homepage: www.elsevier.com/locate/biochempharm

Evaluation of the binding orientations of testosterone in the active site of homology models for CYP2C11 and CYP2C13

Hongwu Wang, Jemmie D. Cheng, Diana Montgomery, K.-C. Cheng *

Schering-Plough Research Institute, 2015 Galloping Hill Road, Kenilworth, NJ 07033, USA

ARTICLE INFO

Article history:
Received 4 March 2009
Accepted 15 April 2009

Keywords:
CYP2C
Homology modeling
Testosterone metabolism
Substrate docking
CYP binding site

ABSTRACT

Cytochromes P-450 2C11 and 2C13 are the major CYPs in rat liver microsomes. Despite a high degree of sequence identity, these two isozymes display different positional and regio-specific metabolism of steroid hormones, such as testosterone. CYP2C11 converts testosterone to 2α-hydroxyl and 16α-hydroxyl metabolites, while CYP2C13 produces primarily the 6β-hydroxyl metabolite. Using a human CYP2C9 crystal structure as the template, homology models were generated for CYP2C11 and CYP2C13. Despite similar volume of the binding pockets for CYP2C11 and CYP2C13, the models for these two CYPs showed a substantial difference in the shape of the substrate-binding sites. Substrate docking using rigid and induced-fit methods showed that testosterone fits into the substrate-binding sites of both CYP2C11 and CYP2C13 without the need of added constraints. These docking exercises appear to support testosterone binding in both CYP2C11 and CYP2C13. A constrained docking using energy minimization is required to position testosterone for more precise positional and regio-specificity in supporting the observed metabolism. These results demonstrate the complexity of using modeling for understanding the binding of substrate to CYPs, and suggest that, as a complement to the metabolism data, modeling and docking may yield reliable structural information for the molecular interaction between the substrate and the CYPs.

1. Introduction

Liver microsomal cytochromes P-450 (CYP) constitute a large and diverse gene superfamily. These hemoproteins are the enzymes responsible for the metabolism of a large variety of xenobiotics, such as drugs, carcinogens, pesticides, and pollutants [1,2]. They also play an essential role in maintaining the homeostasis of many physiological substrates, such as steroid hormones and bile acids [3]. Among the most common endogenous steroid hormones are androgens, progestins and estrogens. The metabolic pathways responsible for the metabolism of steroid hormones in both animals and humans have been well elucidated. For example, human CYP3A4 converts testosterone primarily to 6β-hydroxyl-testosterone [4]. CYP2C9, on the other hand, showed low catalytic activity in converting testosterone to multiple hydroxylated metabolites and androstenedione [5]. The major isoform in rat liver, CYP2C11, produces 2α-hydroxyl and 16α-hydroxyl-testosterone [6], and CYP2C13, produces 6β-hydroxyl-testosterone. Rabbit CYP2C5 converts progesterone to either 16α or 6β-hydroxyl metabolite [7]. Rats are commonly used in research laboratories as well as in the pharmaceutical industry as the primary rodent species for metabolism and toxicology studies. In rat liver, CYP2C11, similar to CYP3A4 in humans, accounts for about 50% of the total content of rat liver microsomal cytochrome P-450 [8]. The diverse region- and positional specificity of the cytochromes P-450 represents one of the most interesting aspects of the enzymology and structure–function relationship of this group of enzymes. Despite the sequence homology between CYP2C11 and CYP2C13, their region- and positional specificities for testosterone appear to be quiet different.

Recent progress in X-ray crystallography has generated a wealth of information of the 3-dimensional (3-D) structures of mammalian cytochromes P-450 [9–12]. The structures of two members of the CYP2C family, human CYP2C8 and CYP2C9, have been elucidated [13,14]. Since these two enzymes are closely related to the rat CYP2C11 and CYP2C13, they represent suitable templates for homology modeling. These 3-D structures represent snap shots of the conformation of various CYPs. How to relate these snap shots to the complex reaction mechanism of CYPs remains an active and challenging research topic. During the reaction cycle, the substrate binding to the CYP represents the initial step, but not

Abbreviations: CYP, cytochrome P-450; 3-D, 3-dimensional; IFD, induced-fit docking.

* Corresponding author at: Schering-Plough Research Institute, K-15-2-2700, 2015 Galloping Hill Rd, Kenilworth, NJ 07033, USA. Tel.: +1 908 740 4056; fax: +1 908 740 2916.
E-mail address: kuo-chi.cheng@spcorp.com (K.-C. Cheng).

0006-2952/$ – see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.bcp.2009.04.020

necessarily the rate-limiting step for the catalysis [15]. Nevertheless, the binding interaction most likely dictates the location of the oxidation site on the substrate.

Several hypotheses have been proposed with respect to the positional and regio-specificity. For example, multiple binding orientations may provide the opportunity for the production of multiple oxidized metabolites, such as in the case of hydroxylation of progesterone by CYP2C5. This concept only involves a single-step of each binding mode which evokes some induce-fit conformation changes. The other concept involves the multiple-step substrate-binding model, in which the initial substrate binding to the peripheral site followed by subsequent movement toward the active site and interaction with the heme [16]. The central question is whether the X-ray crystal structural information may allow computational examination for the distinct positional and regio-specificity. In this study, we use CYP2C11 and CYP2C13 as models to examine the binding mechanism. The results based on rigid docking, induced-fit docking and energy minimization may shed some light on the positional and regio-specificity of a given substrate.

2. Methods

2.1. Template selection

The amino acid sequences of rat CYP2C11 (UniProt entry P08683) and 2C13 (UniProt entry P20814) were retrieved from the UniProt website [17] and used to search against the PDB database [18] using the Blast method [19]. The pair-wise sequence identities were calculated for 2C11, 2C13, and their homologs. Cyp2C9 was identified as the most homologs to both rat enzymes. Three structures of human CYP2C9 are available from PDB, an apo form (PDB ID: 1og2) and two ligand-bound complexes (PDB ID: 1og5 and 1r9o). The crystal structure with the highest resolution (1r9o) was selected as the template for building the homology model.

2.2. Homology modeling

Homology models of CYP2C11 and CYP2C13 were constructed using the Prime comparative modeling method [20]. The heme and five water molecules were preserved during the model construction. The five water molecules are: 600, 618, 622, 696 and 819 (numbered as in the PDB entry 1r9o) [9]. Based on the 1r9o crystal structure, water 600 forms the sixth coordinate bond with the heme. Waters 618, 622 and 696 form the H-bonding networks around the heme molecule to maintain the structural integrity of the binding pocket. Water 819 forms H-bonds with backbone atoms from residues in the I-helix and induces a kink in the middle of the long helix. This kink is present in all mammalian P450 enzymes [21]. Keeping these water molecules is likely to reduce modeling artifacts introduced in energy minimizations. The Prime loop prediction algorithm was used to generate the conformations in the homology models for the loop regions (aa38–42 and aa214–220) that were absent in the 1r9o structure [22]. Three segments around the missing loops were included for loop prediction: aa36–43 (Loop1), aa69–72 (Loop2) and aa213–221 (Loop3). First, the Serial loop prediction approach with "extend medium" setting was used for the prediction of the conformation of Loop2. Sidechain conformation of amino acid residues within 7.5 Å of the loop was allowed to optimize during the loop prediction. Next, structures of the two missing loops were predicted using the cooperative loop sampling algorithm since their conformations are interdependent. This approach allowed the simultaneous optimization of both loops. Again, sidechain conformation of the residues within 7.5 Å of the two loops was allowed to change. The best loop conformation ranked by Prime energy was retained for further optimization. The model was further refined with sidechain optimization. Finally, Ramachandran plots were generated to verify the quality of the homology model.

The volume of the substrate-binding site was calculated using VOIDOO software [23]. The volume was calculated by rolling a probe with 1.4 Å radius (water) in the active site using the "Probe occupied" option. The heme was treated as part of the protein structure, whereas water molecules were excluded from the calculation.

2.3. Compound preparation

The 3-D structures of testosterone and its metabolites were generated from its two-dimensional structure using the Concord program (Concord software, CA). The structures were then subjected to energy minimization using MacroModel based on Merck Molecular Force Field (MMFFs) and a dielectric constant of 1.0. Energy minimization was terminated with a gradient cut-off of 0.05. The minimized structure was used for the docking exercises.

2.4. Rigid docking

Rigid dockings of testosterone to the CYP2C11 and 2C13 models were carried out using Glide in XP mode [24]. The binding site of each CYP was represented by energy grids using a cubic box centered on the centroid of the cavity. Dimension for the box within which the centroid of a docked pose falls was set to be 12 Å × 12 Å × 12 Å. No geometric or hydrogen-bonding constraint was introduced for substrate docking. A maximum of 10 poses were evaluated for each docking calculation.

2.5. Induced-fit docking (IFD)

The IFD was performed according to the following three consecutive steps [25,26]. First, the ligand was docked into a rigid-receptor model with scaled-down vdW radii (Glide SP mode). A vdW scaling of 0.5 was used for non-polar atoms of both CYP and testosterone. Up to twenty initial ligand poses were retained for subsequent CYP structural refinement. Next, Prime was used to generate the induced-fit protein–substrate complexes. Each of the structures from the previous step was subjected to sidechain and backbone refinement. Any amino acid residue within 5.0 Å of testosterone was included in the refinement. The refined complexes were then ranked by Prime energy. The structures within 30 kcal/mol of the structure with minimal energy were selected for a second round of Glide docking and scoring. Finally, testosterone was re-docked into the refined lower energy CYP structures. An IFD score that accounts for both the protein–substrate interaction energy and the total energy of the system was calculated for ranking the IFD poses. All energy calculations were carried out using the OPLS-2001 force field with implicit solvation model.

2.6. Constrained energy minimization

Testosterone was first manually docked into the binding site of each model. The compound was oriented such that the oxidized hydrogen atom was close to the heme iron while the rest of the entire molecule was fit in the general shape of the active site. A stepwise energy minimization procedure was used to generate the docked structure of each complex. First, a force constant of 100 was applied to constrain the distance between the oxidized hydrogen atom (C2α for CYP2C11 or C6β for CYP2C13) and the ferric atom in the range of 3 ± 1 Å in order to reduce the vdW contacts in the initial manual docking poses. All other non-hydrogen protein atoms were frozen in their initial positions. Minimization was terminated either after 500 steps (PRCG) or when gradient dropped below 0.05. A