Structure-biological function study of 17β-hydroxysteroid dehydrogenase type 1 and reductive steroid enzymes: inhibitor design targeting estrogen-dependent diseases

Thèse

Tang Li

Doctorat en médecine moléculaire
Philosophiæ doctor (Ph. D.)

Québec, Canada

© Tang Li, 2019
Résumé

La 17β-HSD1 catalyse l’activation de l’œstrogène le plus actif, l’estradiol, ainsi que la désactivation de la dihydrotestosterone, l’androgène le plus puissant. Cette enzyme est considérée comme une cible prometteuse pour le traitement des maladies dépendantes des œstrogènes. Malgré des décennies de recherches, aucun inhibiteur ciblant la 17β-HSD1 n’a encore atteint le stade clinique. De plus, le mécanisme de l’inhibition du substrat de la 17β-HSD1, qui peut être utilisé pour faciliter la conception d’inhibiteur, n’est toujours pas bien démontré de manière structurelle. Ici, nous avons Co-cristallisé trois inhibiteurs de différence, à savoir l’EM-139, le 2-MeO-CC-156 et le PBRM, avec la 17β-HSD1 et avons résolu ces structures cristallines. L’inhibiteur réversible EM-139 s’est révélé moins stable dans le site de liaison aux stéroides, avec seulement la fraction du noyau stéroidien de l’inhibiteur présentant une densité d’électron définissable. La fraction volumineuse de 7α-alkyle de l’inhibiteur, qui limite son activité anti-œstrogénique, n’est pas définie dans la densité électronique, peut compromettre l’effet inhibiteur de l’inhibiteur sur l’enzyme. Quant à l’inhibiteur réversible, le 2-MeO-CC-156, il interagit de manière similaire que le CC-156 avec l’enzyme. Cependant, avec la présence du groupe 2-MeO, le pouvoir inhibiteur de la 17β-HSD1 est nettement inférieur à celui du CC-156. L’analyse du complexe ternaire PBRM avec la 17β-HSD1 montre clairement la formation d’une liaison covalente entre l’His221 et la chaîne latérale bromoethyl de l’inhibiteur, donnant un aperçu des interactions moléculaires bénéfiques qui favorisent la liaison et l’avènement de N-alkylation ultérieur dans le site catalytique de l’enzyme. En outre, le groupe bromoethyl en position C-3 du PBRM justifie son profil non œstrogénique, ralentit son métabolisme et assure son action spécifique de la 17β-HSD1 par la formation d’une liaison covalente avec Nε du résidu His221. Nous avons aussi Co-cristallisé la 17β-HSD1 avec l’œstrone ainsi qu’avec l’analogue de l’œstrone et du cofacteur NADP+, la structure a révélé un mode de liaison inversé de l’œstrone dans l’enzyme, jamais trouvé dans les complexes d’estradiol. L’analyse structurale a démontré que His221 est le résidu clé responsable de la réorganisation et de la stabilisation de l’œstrone liée de manière inversée, conduisant à la formation d’un complexe sans issue. Ainsi, sur la base du mécanisme d’inhibition du substrat et de l’analyse computationnelle, une nouvelle entité chimique (SX7) est proposée qui peut inhiber la 17β-HSD1 et former un complexe sans issue. De plus, avec un grand nombre d’échantillons cliniques, nous avons démontré la modulation et la corrélation d’expression significative de plusieurs enzymes clés de conversion des stéroides, supportant les 17β-HSD1 et 17β-HSD7 réductrices comme cibles prometteuses et la nouvelle thérapie combinée ciblant les 11β-HSD2 et 17β-HSD7.
Abstract

Human 17β-hydroxysteroid dehydrogenase type 1 (17β-HSD1) catalyzes the activation of the most potent estrogen estradiol as well as the deactivation of the most active androgen dihydrotestosterone, and is considered as a promising target for the treatment of estrogen-dependent diseases such as endometriosis, breast cancer, endometrial cancer and ovarian cancer. Despite decades of research, no inhibitor targeting 17β-HSD1 has yet reached the stage of clinical trials. Moreover, the structure-biological function of the substrate inhibition of 17β-HSD1, which can be used to facilitate the inhibitor design, is still not well demonstrated. Here we co-crystallized three different inhibitors, namely EM-139, 2-MeO-CC-156 and PBRM, with 17β-HSD1 and solved the structures of these complexes. The reversible inhibitor EM-139 showed high mobility in the steroid binding site with only its steroid core moiety could be defined in the electron density. The bulky 7α-alkyl moiety of the inhibitor, which guarantees its anti-estrogenic activity but unable to be defined in the electron density, may compromise the inhibitory effect of the inhibitor on the enzyme. As for the reversible inhibitor 2-MeO-CC-156, it interacts similarly to CC-156 with the enzyme. However, in the presence of the 2-MeO group, it shows much less inhibitory potency to 17β-HSD1 as compared to the CC-156. The analysis of the PBRM ternary complex with 17β-HSD1 clearly shows an unambiguous continuity of electron density from the side chain of His^{221} to the bound PBRM, demonstrating the formation of a covalent bond between the Nε of His^{221} and the C-31 (BrCH₂) of the inhibitor. This result provides insight into beneficial molecular interactions that favor the binding and subsequent N-alkylation event in the enzyme catalytic site. Also, the bromoethyl group at position C-3 of the PBRM warrants its non-estrogenic profile, slows down its metabolism, and secures the specific action of 17β-HSD1 through the formation of a covalent bond with Nε of residue His^{221}. Meanwhile, we co-crystallized 17β-HSD1 with estrone as well as with estrone and cofactor analog NADP⁺, revealed a reversely orientated binding mode of estrone in the enzyme, never found in reported estradiol complexes. Structural analysis demonstrated that His^{221} is the key residue responsible for the reorganization and stabilization of the reversely bound estrone, leading to the formation of a dead-end complex. Thus, based on the substrate inhibition mechanism and computational analysis, a chemical entity (SX7) is proposed that may inhibit 17β-HSD1 and form a dead-end complex. Furthermore, with large number clinical samples, we demonstrated the significant expression modulation and expression correlation of several key steroid-converting enzymes, supporting the reductive 17β-HSD1 and 17β-HSD7 as promising targets and the new combined therapy targeting 11β-HSD2 and 17β-HSD7.
# Table des matières

Résumé ................................................................................................................................. ii
Abstract ................................................................................................................................. iii
Table des matières ............................................................................................................... iv
Liste des figures ..................................................................................................................... vi
Liste des figures, tableaux, illustrations ................................................................................ vii
Liste des abréviations ........................................................................................................... viii
Remerciements ...................................................................................................................... xii
Avant-propos ....................................................................................................................... xiii
Introduction .......................................................................................................................... 1
Chapitre 1 Combined biophysical chemistry reveals a new covalent inhibitor with a low-reactivity alkyl halide 16
  1.1 Résumé ......................................................................................................................... 16
  1.2 Abstract ....................................................................................................................... 16
  1.3 Introduction .................................................................................................................. 17
  1.4 Results and Discussion ............................................................................................... 19
  1.5 Conclusion ................................................................................................................... 19
  1.6 Experimental Procedures ............................................................................................ 22
  1.7 Reference .................................................................................................................... 25
Chapitre 2 Crystal structures of human 17β-hydroxysteroid dehydrogenase type 1 complexed with the dual-site inhibitor EM-139 .................................................................................. 38
  2.1 Résumé ......................................................................................................................... 38
  2.2 Abstract ....................................................................................................................... 38
  2.3 Introduction .................................................................................................................. 39
  2.4 Materials and Methods ............................................................................................... 39
  2.5 Results ........................................................................................................................ 40
  2.6 Discussion ................................................................................................................... 41
  2.7 Conclusion ................................................................................................................... 41
  2.8 Reference .................................................................................................................... 43
Chapitre 3 Crystal structures of human 17β-hydroxysteroid dehydrogenase type 1 complexed with estrone and cofactor reveal the mechanism of substrate inhibition ...................................................................... 53
  3.1 Résumé ......................................................................................................................... 53
  3.2 Abstract ....................................................................................................................... 53
  3.3 Introduction .................................................................................................................. 54
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.4</td>
<td>Materials and Methods</td>
<td>55</td>
</tr>
<tr>
<td>3.5</td>
<td>Results</td>
<td>56</td>
</tr>
<tr>
<td>3.6</td>
<td>Discussion</td>
<td>59</td>
</tr>
<tr>
<td>3.7</td>
<td>Conclusion</td>
<td>60</td>
</tr>
<tr>
<td>3.8</td>
<td>Reference</td>
<td>62</td>
</tr>
<tr>
<td>Chapitre 4</td>
<td>Remarkable steroid-converting enzyme and receptor regulations in large number breast tumor samples: molecular correlation and combined therapies</td>
<td>78</td>
</tr>
<tr>
<td>4.1</td>
<td>Résumé</td>
<td>78</td>
</tr>
<tr>
<td>4.2</td>
<td>Abstract</td>
<td>79</td>
</tr>
<tr>
<td>4.3</td>
<td>Introduction</td>
<td>80</td>
</tr>
<tr>
<td>4.4</td>
<td>Materials and Methods</td>
<td>81</td>
</tr>
<tr>
<td>4.5</td>
<td>Results</td>
<td>84</td>
</tr>
<tr>
<td>4.6</td>
<td>Discussion</td>
<td>86</td>
</tr>
<tr>
<td>4.7</td>
<td>Conclusion</td>
<td>88</td>
</tr>
<tr>
<td>4.8</td>
<td>Reference</td>
<td>88</td>
</tr>
<tr>
<td>Conclusion</td>
<td></td>
<td>101</td>
</tr>
<tr>
<td>Bibliographie</td>
<td></td>
<td>105</td>
</tr>
<tr>
<td>Annexe A</td>
<td>Cold-active extracellular lipase: expression in Sf9 insect cells, homogenization, and catalysis</td>
<td>119</td>
</tr>
<tr>
<td>2.1</td>
<td>Résumé</td>
<td>119</td>
</tr>
<tr>
<td>2.2</td>
<td>Abstract</td>
<td>120</td>
</tr>
<tr>
<td>2.3</td>
<td>Introduction</td>
<td>120</td>
</tr>
<tr>
<td>2.4</td>
<td>Materials and Methods</td>
<td>120</td>
</tr>
<tr>
<td>2.5</td>
<td>Results and Discussion</td>
<td>125</td>
</tr>
<tr>
<td>2.6</td>
<td>Conclusion</td>
<td>129</td>
</tr>
<tr>
<td>2.7</td>
<td>Reference</td>
<td>131</td>
</tr>
</tbody>
</table>
Liste des figures

Figure 1. Ten leading cancer types for the estimated new cancer cases and deaths in women in United States, 2017

Figure 2. The time course of the estimated new BC cases and deaths in women in the United States

Figure 3. Model of the multistep carcinogenesis in BC

Figure 4. Schematic representations of sex hormones synthesis regulations in pre- (A) and postmenopausal (B) women

Figure 5. Human steroidogenic and steroid-inactivating enzymes in peripheral intracrine tissues

Figure 6. Stereo ribbon presentation of human 17β-HSD1 structure

Figure 7. Two possible stepwise catalytic mechanisms for 17β-HSD1

Figure 8. Key inhibitors of 17β-HSD1 from different Series
Liste des figures, tableaux, illustrations

Table 1. Previously published 17β-HSD1 structures

Table 2. The ratio of kinetic constants of Human 17β-HSD1 variants vs. that of wild type enzyme
# Liste des abréviations

<table>
<thead>
<tr>
<th>Abbréviation (fr.)</th>
<th>Abbréviation (en.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-MeO-CC-156</td>
<td>2methoxy-16β-(m-carbamoylbenzyl)-E2</td>
</tr>
<tr>
<td>3β-diol</td>
<td>5α-androstan-3β,17β-diol</td>
</tr>
<tr>
<td>4-dione</td>
<td>androstenedione</td>
</tr>
<tr>
<td>5-diol</td>
<td>5-androstenediol; androst-5-ene-3β,17β-diol</td>
</tr>
<tr>
<td>5-diol-FA</td>
<td>5-diol fatty acid</td>
</tr>
<tr>
<td>5-diol-S</td>
<td>5-diol sulfate</td>
</tr>
<tr>
<td>ACTH</td>
<td>adrenocorticotropic hormone</td>
</tr>
<tr>
<td>A-dione</td>
<td>5α-androstan-3,17-dione</td>
</tr>
<tr>
<td>ADT</td>
<td>androsterone</td>
</tr>
<tr>
<td>AIs</td>
<td>aromatase inhibitors</td>
</tr>
<tr>
<td>AKR</td>
<td>aldo-ketoreductase</td>
</tr>
<tr>
<td>AR</td>
<td>androgen receptor</td>
</tr>
<tr>
<td>BC</td>
<td>breast cancer</td>
</tr>
<tr>
<td>C₁₂₂₃₈</td>
<td>octaethylene glycol monododecyl ether</td>
</tr>
<tr>
<td>CC-156</td>
<td>16β-m-carbamoylbenzyl-E2</td>
</tr>
<tr>
<td>CRH</td>
<td>corticotropin releasing hormone</td>
</tr>
<tr>
<td>DHEA</td>
<td>dehydroepiandrosterone</td>
</tr>
<tr>
<td>DHEAS</td>
<td>dehydroepiandrosterone sulfate</td>
</tr>
<tr>
<td>DHT</td>
<td>dihydrotestosterone</td>
</tr>
<tr>
<td>DMF</td>
<td>Dimethylformamide</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>E1</td>
<td>Estrone</td>
</tr>
<tr>
<td>E1S</td>
<td>estrogen sulfate</td>
</tr>
<tr>
<td>E2</td>
<td>Estradiol</td>
</tr>
<tr>
<td>EDD</td>
<td>estrogen-dependent disease</td>
</tr>
<tr>
<td>EM-139</td>
<td>N-n-Butyl-N-methyl-ll-(16'α-chloro3',17'β-dihydroxyestra-1',3',5'(10')-trien-7'α-yl)undecanamide</td>
</tr>
<tr>
<td>epi-ADT</td>
<td>epiandrosterone</td>
</tr>
<tr>
<td>ER</td>
<td>estrogen receptor</td>
</tr>
<tr>
<td>EREs</td>
<td>estrogen responsive elements</td>
</tr>
<tr>
<td>FSH</td>
<td>follicle-stimulating hormone</td>
</tr>
<tr>
<td>GnRH</td>
<td>gonadotropin-releasing hormone</td>
</tr>
<tr>
<td>GSC</td>
<td>Genome Sequencing Centers</td>
</tr>
<tr>
<td>HSD</td>
<td>hydroxysteroid dehydrogenase</td>
</tr>
<tr>
<td>HTS</td>
<td>high throughput sequencing</td>
</tr>
<tr>
<td>LH</td>
<td>luteinizing hormone</td>
</tr>
<tr>
<td>mg</td>
<td>microgram</td>
</tr>
<tr>
<td>ml</td>
<td>microliter</td>
</tr>
<tr>
<td>NAD+</td>
<td>nicotinamide adenine dinucleotide</td>
</tr>
<tr>
<td>NADP+</td>
<td>nicotinamide adenine dinucleotide phosphate</td>
</tr>
<tr>
<td>NIH</td>
<td>National Institute of Health</td>
</tr>
<tr>
<td>nm</td>
<td>nanometer</td>
</tr>
</tbody>
</table>
nM  nanomolar
OD  optical density
PAGE  polyacrylamide gel electrophoresis
PBRM  3-(2-bromoethyl)-16β-(m-carbamoylbenzyl)-17β-hydroxy-1,3,5(10)-estratriene
PDB  protein data bank
pNPA  p-nitro phenyl acetate
pNPB  p-nitro phenyl butyrate
pNPD  p-nitro phenyl decanoate
pNPL  p-nitro phenyl dodecanoate
pNPM  p-nitro phenyl myristate
pNPP  p-nitro phenyl palmitate
P_{PH}  polyhedron promoter
pro-S  prochiral S configuration
RhB  rhodamine B
RhB-OOe  RhB-olive oil
RNA-seq  RNA sequencing
RoDH-1  Ro dehydrogenase 1
SDR  short chain dehydrogenase/reductase
SDS  sodium dodecyl sulfate
SG  space group
Sult2B1  sulfotransferase 2B1
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>testosterone</td>
</tr>
<tr>
<td>TCGA</td>
<td>The Cancer Genome Atlas</td>
</tr>
<tr>
<td>Testo</td>
<td>testosterone</td>
</tr>
<tr>
<td>UGT1A1</td>
<td>uridine glucuronosyl transferase 1A1</td>
</tr>
<tr>
<td>UGT2B28</td>
<td>uridine glucuronosyl transferase 2B28</td>
</tr>
<tr>
<td>UV</td>
<td>ultra-violet</td>
</tr>
<tr>
<td>β-DDM</td>
<td>n-Dodecyl-β-D-Maltoside</td>
</tr>
<tr>
<td>β-ME</td>
<td>β-mercaptoethanol</td>
</tr>
<tr>
<td>β-OG</td>
<td>n-octyl-β-D-glucoside</td>
</tr>
<tr>
<td>μM</td>
<td>micromolar</td>
</tr>
</tbody>
</table>
Remerciements

I would like to convey my immeasurable gratitude to my director of research, professor Sheng-Xiang Lin, for his meticulous guidance and enlightening discussions that helped me overcome all the difficulty and enabled me to present this thesis. I greatly appreciate his trustiness and supports for giving me this inspiring and challenging project. His diligent directions and interactive concept helped me greatly during my study, and I will certainly benefit from that in my future career. I also sincerely appreciate the following person and organisations for supporting my doctoral study to obtain the degree of Doctor of Philosophy (Ph.D.). The Ph.D. study will light up my future career in scientific research and practise.

I would like to thank all the members in Dr. Lin’s Lab. I express my gratitude to Dre. Ming Zhou for her help in protein purification and crystallization; to Mr. Jean-François Theriault for his help in enzyme kinetics; to Mr. Jian Song for his help in binding study and growing rLcn6 crystals; to Dr. Preyesh Stephen for his help in Crif1-CDK2 project; to Miss. Xiaoye San and Ruixuan Wang for their inspiriting discussion in diagnosis and treatment of breast and ovarian cancers. I would like to thank Dr. Dao-Wei Zhu for his help in familiar with the surroundings and his advices in protein purification and crystallization. As well as Dr. Xiaoqiang Wang, Dre. Dan Xu and Dre. Juliette Adjo Aka for their advices and discussion in cell experiments.

I would like to thank Dr. Donald Poirier. I sincerely appreciate the knowledge from him in the field of medicinal chemistry, especially for the inhibitor design. I would like to thank Dr. Rong Shi for his help in collecting an X-ray diffraction dataset at the CLS synchrotron. I would like to thank Dr. Alexandre Brunet in the Laboratory of Flow-cytometry for his advice in preparing the sample and analysing the results. I also appreciated Dre. Sylvie Bourassa and Dre. Florence Roux-Dalvai for their advice and discussion in mass spectrometry experiment design and sample preparation. I would like to thank Mr. Martin Thibault for his help in analysing image of western blot.

I would like to thank all administration staffs in the Research Center of CHU de Quebec (CHUL). Thank Mme Nicole Almeras for taking care of the registration and financial documents. Thank Mme Marianne Roberge for her help in the order of experiment materials and reagents.

Finally, I am very grateful to my family, especially to my beloved son Guanrui Li and wife Juan Liu as well as my parents. It is your love that encouraging and supporting me during my studies.
Avant-propos

This thesis is submitted to the “Faculté des études supérieures de l'Université Laval” for the requirement of a doctor’s degree in science. The thesis is written in English, except for the summary as well as the abstract of each article, which are in French. Two articles have been published by Journal of Physical Chemistry Letters and Health, respectively. The other three are being submitted for publication or in preparation.

In the introductory section, four major estrogen-dependent diseases were reviewed. The biosynthesis of estrogens, mostly estradiol, and the role of 17β-HSD1 in estrogen activation as well as inactivation of androgen are summarized. The structural and kinetic studies as well as the development of 17β-HSD1 inhibitor design are also discussed. The hypothesis and objectives are described in the end of this chapter.

The Chapter I: “Tang Li, René Maltais, Donald Poirier, Sheng-Xiang Lin. Combined Biophysical Chemistry Reveals a New Covalent Inhibitor with a Low-Reactivity Alkyl Halide. Journal of Physical Chemistry Letters (2017 IF: 8.7). 2018 Aug; 9:5275-5280. doi: 10.1021/acs.jpclett.8b02225.” I conducted all the experiments and wrote the manuscript, and I’m the first author of this article. In this chapter, the crystal structures of 17β-HSD1 with two inhibitors (PBRM and 2-MeO-CC-156) were described. This study constructed the first example of N-alkylation between a human enzyme and a low-reactivity alkyl halide derivative, which opens the door to a new design of alkyl halide-based specific covalent inhibitors as potential therapeutic agents.

The Chapter II: “Tang Li, Dao-Wei Zhu, Fernand Labrie and Sheng-Xiang Lin. Crystal structures of human 17β-hydroxysteroid dehydrogenase type 1 complexed with the dual-site inhibitor EM-139. Health. 2018 Aug; 10(8):1079-89. doi: 10.4236/health.2018.108081.” I processed the crystal diffraction data to solve the complex structure and wrote the manuscript, and I’m the first author of this article. In this chapter, the 17β-HSD1 binary complex with the inhibitor EM-139 was described. The interaction between the steroid moiety of the inhibitor and the enzyme was analyzed. The influence of its bulky 7α-alkyl side chain to its inhibitory effect in 17β-HSD1 was also discussed.

The Chapter III: “Tang Li, Preyesh Stephen, Dao-Wei Zhu, Rong Shi, Sheng-Xiang Lin. Crystal structures of human 17β-hydroxysteroid dehydrogenase type 1 complexed with estrone and cofactor reveal the mechanism of substrate inhibition. FEBS Journal. 2019. Doi: 10.1111/febs.14784.” I conducted all the experiments except for the crystallization of the 17β-HSD1-E1 binary complex. I wrote the manuscript, and I’m the first author of this article. In this chapter, the crystal structures of 17β-HSD1 in complex with E1 and with/without cofactor analog NADP+ were described. Based on the E1 binary and ternary complex structures as well as previously published 17β-HSD1 complexes with other ligands, the mechanism of the long observed substrate inhibition of 17β-HSD1 has been discussed.
The Chapter IV: “Tang Li, Zhongjun Li, Sheng-Xiang Lin. Remarkable steroid-converting enzyme and receptor regulations in large number breast tumor samples: molecular correlation and combined therapies (Article under submission).” I conducted the data analysis and wrote the manuscript, and I’m the first author of this article. In this chapter, the cDNA sequencing data from the public cohort The Cancer Genome Atlas Breast Invasive Carcinoma (TCGA-BRCA) was extracted and statistically analyzed, and identified several key steroid-converting enzymes which are significantly up-regulated in cancer samples. Close expression correlations of the enzymes were also found, suggesting combined therapy for breast cancer treatment.

In the conclusion, I interactively discussed 17β-HSD1 structure-function study from inhibitor interactions to the mechanism of enzyme regulation. Besides, I also discussed the use of cDNA sequencing data in breast cancer research.

The references of introduction and conclusion are listed after the conclusion section. References of publications are listed after the text of each article.

In the end of the thesis is the appendix: “Tang Li, Wenfa Zhang, Jianhua Hao, Mi Sun, Sheng-Xiang Lin. Cold-active extracellular lipase: expression in Sf9 insect cells, homogenization, and catalysis. Biotechnol Rep (Amst). 2018; 21:e00295. doi:10.1016/j.btre.2018.e00295.” I conducted all the experiments and wrote the manuscript, and I’m the first author of this article. In this article, I expressed a novel cold-active marine lipase in Sf9 insect cells. After purification, I carefully characterized its enzymatic properties, such as the optimum temperature and pH ranges, substrate specificity, the effects of detergents, organic solvents as well as enzyme inhibitors. These results will facilitate its application in industries.
Introduction

1 Estrogen-dependent disease

1.1 Breast cancer

Breast cancer (BC) is the most commonly diagnosed cancer in women worldwide, and one of the leading cause of cancer death in women\(^1\). BC can also occur in man, but it is rare \(^1\). It has estimated that 268,670 patients will be diagnosed BC in 2018 in the United States, among which 99% were women (Figure 1)\(^2\). The estimated number of death from BC in women is 40,920, ranking the second among all estimated deaths from cancers\(^2\). The incidence of BC is estimated to increase based on the trend of the past ten years (Figure 2)\(^2\)-\(^{12}\). Similar situation was also observed in Canada, about 26,300 female patients will be diagnosed BC, which account for 25% of all cancers in 2017 (Canadian breast cancer statistics 2017, http://www.cancer.ca/en/cancer-information/cancer-type/breast/statistics/). The majority of female patients diagnosed with BC are above 45 years old, and mostly after menopause\(^13\). Among incidences of all BCs, around 60% in premenopausal women and 75% in postmenopausal women are initially estrogen-dependent\(^14\)-\(^{15}\). There is a multistep process involved in the occurrence of BC, which starts from normal cells through hyperplasia, premalignant change, in situ carcinoma, progression of primary BC and to metastasis formation (Figure 3)\(^16\). During this progression process, hormones such as estrogen, progesterone and prolactin, stimulate cell proliferation through their receptors mediated signaling pathways as well as induced genetic damage and mutations\(^16\)-\(^{17}\).

<table>
<thead>
<tr>
<th>Estimated New Cases</th>
<th>Estimated Deaths</th>
</tr>
</thead>
<tbody>
<tr>
<td>Breast</td>
<td>Lung &amp; bronchus</td>
</tr>
<tr>
<td>268,120</td>
<td>70,500</td>
</tr>
<tr>
<td>Lung &amp; bronchus</td>
<td>Breast</td>
</tr>
<tr>
<td>112,350</td>
<td>40,920</td>
</tr>
<tr>
<td>Colon &amp; rectum</td>
<td>Colon &amp; rectum</td>
</tr>
<tr>
<td>64,640</td>
<td>23,240</td>
</tr>
<tr>
<td>Uterine corpus</td>
<td>Pancreas</td>
</tr>
<tr>
<td>63,230</td>
<td>21,310</td>
</tr>
<tr>
<td>Thyroid</td>
<td>Ovary</td>
</tr>
<tr>
<td>40,900</td>
<td>14,070</td>
</tr>
<tr>
<td>Melanoma of the skin</td>
<td>Uterine corpus</td>
</tr>
<tr>
<td>36,120</td>
<td>11,350</td>
</tr>
<tr>
<td>Non-Hodgkin lymphoma</td>
<td>Leukemia</td>
</tr>
<tr>
<td>32,950</td>
<td>10,100</td>
</tr>
<tr>
<td>Pancreas</td>
<td>Liver &amp; intrahepatic bile duct</td>
</tr>
<tr>
<td>26,240</td>
<td>9,660</td>
</tr>
<tr>
<td>Leukemia</td>
<td>Non-Hodgkin lymphoma</td>
</tr>
<tr>
<td>25,270</td>
<td>8,400</td>
</tr>
<tr>
<td>Kidney &amp; renal pelvis</td>
<td>Brain &amp; other nervous system</td>
</tr>
<tr>
<td>22,660</td>
<td>7,340</td>
</tr>
<tr>
<td>All Sites</td>
<td>All Sites</td>
</tr>
<tr>
<td>878,980</td>
<td>286,010</td>
</tr>
</tbody>
</table>

Figure 1 Ten leading cancer types for the estimated new cancer cases and deaths in women in United States, 2018 (Siegel et al., 2018).
Figure 2 The time course of the estimated new BC cases and deaths in women in the United States (Jemal et al., 2008-2010; Siegel et al., 2011-2018).

Estradiol (E2) is the most biologically potent natural estrogen. In estrogen dependent human breast cancers, E2 plays a critical role in the proliferation and development of carcinoma cells and it is actually essential for some of these carcinomas to continue growth\textsuperscript{18}. The primary biological effects of estrogen are mediated by two distinct nuclear receptors, estrogen receptor (ER)\textsubscript{α}\textsuperscript{19} and ER\textsubscript{β}\textsuperscript{20}, which encoded by unique genes and function in the nucleus as ligand-dependent transcription factors. ER\textsubscript{α} is mainly responsible for the effects of estrogens on normal and malignant breast tissues. Its role in promoting proliferation of BC cells is well characterized, through either membrane and cytoplasmic signaling cascades\textsuperscript{21} or transcriptional regulation\textsuperscript{22}. In contrast, the role of ER\textsubscript{β} in BC is not clearly understood but seems to act as an antagonist of ER\textsubscript{α} activity, attenuating the proliferation stimulation effect of estrogen\textsuperscript{23-25}. 

\[\text{Figure 2 The time course of the estimated new BC cases and deaths in women in the United States (Jemal et al., 2008-2010; Siegel et al., 2011-2018).} \]
1.2 Endometrial cancer

Endometrial cancer is one of the most common gynecologic malignancies. It ranks to be the fourth most diagnosed cancers in women after breast, lung, and colorectal cancers, and was expected to have more than 63,000 new cases in US in 2017 (Figure 1). The death rate for endometrial cancer almost doubled during the past two decades. Endometrial cancer is commonly classified into two types based on the dualistic model of endometrial cancer tumorigenesis described by Bokhman. Type I commonly develops in women before menopause in an estrogen-dependent manner. In contrast, type II endometrial cancer majorly develops in postmenopausal women in an estrogen-independent manner. The pathogenesis of type I endometrial cancer is through atypical endometrial hyperplasia, whereas type II endometrial cancer is proposed to be generated directly from normal endometrium. Most patients diagnosed with endometrial adenocarcinoma are between...
the ages of 50 and 60 years, and 90% of women diagnosed with endometrial cancer are after age of 50, mostly after menopause\textsuperscript{26, 29}. About 80% of endometrial cancers are estrogen-dependent\textsuperscript{30} and the most potent estrogen, estradiol (E2), is suggested to play an important role in the pathogenesis of the disease by increasing the mitotic activity of endometrial cells\textsuperscript{31}.

1.3 Endometriosis

Endometriosis is an estrogen-activated gynecological disease characterized by the presence of endometrial-like tissue growing outside the uterine cavity, typically on the pelvic peritoneum, ovaries, and uterosacral ligaments, and in the rectovaginal septum and vesico-uterine fold\textsuperscript{32}. Severe disease may lead to deformation of pelvic anatomy and extensive pelvic adhesions, often associated with pelvic pain and infertility\textsuperscript{32}. Endometriosis is initially considered largely as a benign condition, while the wide opinion nowadays is that endometriosis is a neoplastic condition which can develop into specific type of invasive ovarian cancer\textsuperscript{33-34}. It is estimated that 6 to 10% of diagnosed endometriosis are in premenopausal women, whereas the frequency rises up to 50% of women with infertility\textsuperscript{32}. Endometriosis is a multifactorial disease. Its pathogenesis involves estrogen overexposure, angiogenesis, inflammation, genetic predisposition, and environmental exposure to pollutants\textsuperscript{35-40}. It has been demonstrated that estrogen plays a central role in the development and maintenance of endometriosis by promoting the growth of ectopic tissue\textsuperscript{41}. In premenopausal patient, the depression of E2 levels through gonadotropin-releasing hormone analogues (GnRH-a) leads to the relieving of pains and regression of endometriotic lesions, which relapsed with the recovery of E2 when the therapy discontinues\textsuperscript{42}. While in postmenopausal women, the administration of hormone replacement therapy may lead to the relapse of endometriosis\textsuperscript{43}.

1.4 Ovarian cancer

Ovarian cancer is the fifth most lethal of all gynecological malignancies in western country with more than 14,000 estimated death in 2017 in The United States (Figure 1)\textsuperscript{12}. As more than 80% of all diagnosed ovarian cancers are in women above age 50, it is mainly considered to be a disease of postmenopausal women\textsuperscript{44}. About 90% of malignant ovarian tumors are epithelial ovarian cancer\textsuperscript{45}. Epidemiological data show that estrogen exposure and metabolism are involved in the stimulation and pathogenesis of ovarian cancer, and patients taking estrogen-only hormone replacement therapy have a higher risk of ovarian cancer\textsuperscript{44, 46-48}. Cell studies confirmed that ovarian cancer cells share several estrogen regulation pathways with other estrogen-associated cancers such as endometrial cancer and breast cancer, and anti-estrogen intervention suppresses the proliferation of ovarian cancer cells \textit{in vitro} and \textit{in vivo}\textsuperscript{49-51}. Moreover, estrogen was demonstrated to promote ovarian cancer cell migration and invasion through activating the PIK3/AKT pathway expression and down-regulating nm23-H1 expression\textsuperscript{52}.
2 Origins of estradiol

The origins of E2 in women can be divided into two sources, one is secreted from the ovary, and another is locally biosynthesized from the adrenal precursor dehydroepiandrosterone (DHEA), dehydroepiandrosterone-sulfate (DHEAS) and androstenedione in the peripheral tissues. In premenopausal women, circulating E2 is produced primarily by the ovaries, and DHEAS is produced primarily by the adrenal glands. As for the DHEA, half of it is produced by adrenal glands, 20% originates from the ovaries and the other 30% is converted from DHEAS in peripheral tissues by sulfatase. The production of androstenedione is equally contributed by the adrenals and the ovaries (Figure 4A). After menopause, when the ovaries become atrophied and cease to act, E2 no longer functions as a circulating hormone. Thus, E2 in postmenopausal women is produced only from precursor steroids of the adrenal glands in an intracrine manner to peripheral sites, which include breast, bone, vascular smooth muscle, and various sites in the brain (Figure 4B). Moreover, it is increasingly being recognised in EDDs that these tumor tissues are not just passively dependent on circulating levels of E2 but rather generate it locally from precursors in an active fashion.

Figure 4 Schematic representations of sex hormones synthesis regulations in pre- (A) and postmenopausal (B) women. GnRH, gonadotropin releasing hormone; LH, luteinizing hormone; FSH, follicle-stimulating hormone; CRH, corticotropin releasing hormone; ACTH, adrenocorticotropic hormone; T, testosterone; DHT, dihydrotestosterone; DHEA, dehydroepiandrosterone; E2, estradiol (Labrie 2015).
3 The role of 17β-HSD1

17β-HSD1 belongs to the short-chain dehydrogenase/reductase (SDR) family. The major function of this enzyme is the activation of estrone (E1) to the most potent estrogen E2 (Figure 5), which is known to play a pivotal role in the occurrence and development of estrogen-dependent diseases (EDDs). It can also catalyze the conversion of DHEA into 5-androstene-3β,17β-diol (5-diol), which has been suggested to be the main estrogen after menopause. Beside the ability of activating estrogen, 17β-HSD1 can also inactivate androgens. It has been demonstrated that 17β-HSD1 can transform the most potent androgen dihydrotestosterone (DHT) into a weak estrogen 5α-Androstane-3β,17β-diol (3β-diol), a reaction which has been proposed to become more important after menopause and may be involved in aromatase inhibitor (AI) resistance. 17β-HSD1 is the most active enzyme in terms of the production of E2. The over-expression of 17β-HSD1 as well as the increased estrogen/androgen ratio indicates the pivotal role of the enzyme in breast cancer, endometrial cancer, endometriosis, and ovarian cancer. Thus, inhibition of 17β-HSD1 is considered as a promising therapeutic approach for the treatment of these diseases.

![Chemical structures](image)

Figure 5 Human 17β-HSD1 catalyze the conversion of E1 to E2, DHEA to 5-Diol, and DHT to 3β-Diol (Dumont et al., 1992; Aka et al., 2010).
4 Structural studies of 17β-HSD1

17β-HSD1 is the first human steroid-converting enzyme whose three-dimensional structure has been solved. 17β-HSD1 consists of 328 amino acids with a molecular weight of 34.5kDa. This membrane-associated enzyme is acting as a homodimer and possesses a conserved Tyr-X-X-X-Lys sequence as a SDR family member and a Ser residue at the active site\(^6,73\). The first crystallization of human estrogenic 17β-HSD1 was reported by Zhu and co-workers in 1993\(^74\). The three-dimensional structure of the enzyme was published in 1995\(^75\). Since then, there are 22 17β-HSD1 structures deposited into the protein data bank (PDB), some in complex with substrate or inhibitor, some in complex with cofactor, and some in combination with cofactor and substrate/inhibitor (Table 1). This has led to the atomic level description of the substrate and cofactor binding cavities of the enzyme and a detailed understanding of its mechanism of action, as well as the molecular basis for the estrogen-specificity of the enzyme\(^76-78\).

The core of 17β-HSD1 structure is consisting of seven-stranded parallel β-sheet (βA to βG) surrounded by six parallel α-helices (αB to αG), evenly distributed by the two sides of the β-sheet (Figure 6). The structure of the protein generally forms into two segments: the first segment, βA to βF, is a classic Rossmann fold, responsible for cofactor binding; the second segment, βD to βG, is partly in the Rossmann fold, governs steroid substrate binding\(^75\). The C-terminus of 17β-HSD1 (285-327) cannot be defined in all published structures and residues 190-199 have very poor density or even no density in many structures (1FDS, 1FDU, 1FDV, 1JTV, 1QYV, 1QYW, 1QYX, 3DEY, 3KLM, 3KLP, 3KM0).

Table 1 Previously published 17β-HSD1 structures

<table>
<thead>
<tr>
<th>PDB code</th>
<th>ligand</th>
<th>Cofactor</th>
<th>Resolution(Å)</th>
<th>SG</th>
<th>Other</th>
<th>Author</th>
<th>βFαG'-loop</th>
</tr>
</thead>
<tbody>
<tr>
<td>1A27</td>
<td>EST</td>
<td>NAP</td>
<td>1.9</td>
<td>C2</td>
<td></td>
<td>Mazza</td>
<td>Closed</td>
</tr>
<tr>
<td>1BHS</td>
<td></td>
<td></td>
<td>2.2</td>
<td>C2</td>
<td></td>
<td>Ghosh</td>
<td>Semi-open</td>
</tr>
<tr>
<td>1DHT</td>
<td>DHT</td>
<td></td>
<td>2.24</td>
<td>C2</td>
<td></td>
<td>Han</td>
<td>Opened</td>
</tr>
<tr>
<td>1EQU</td>
<td>EQI</td>
<td>NAP</td>
<td>3.0</td>
<td>P212121</td>
<td>Sawicki</td>
<td>Closed</td>
<td></td>
</tr>
<tr>
<td>1FDS</td>
<td>EST</td>
<td></td>
<td>1.7</td>
<td>C2</td>
<td></td>
<td>Breton</td>
<td>-</td>
</tr>
<tr>
<td>1FDT</td>
<td>EST</td>
<td>NAP</td>
<td>2.2</td>
<td>C2</td>
<td>SO4</td>
<td>Breton</td>
<td>Closed</td>
</tr>
<tr>
<td>1FDU</td>
<td>EST</td>
<td>NAP</td>
<td>2.7</td>
<td>P21</td>
<td>SO4</td>
<td>Mazza</td>
<td>-</td>
</tr>
<tr>
<td>1FDV</td>
<td></td>
<td>NAD</td>
<td>3.1</td>
<td>P21</td>
<td>SO4</td>
<td>Mazza</td>
<td>-</td>
</tr>
<tr>
<td>1FDW</td>
<td>EST</td>
<td></td>
<td>2.7</td>
<td>C2</td>
<td></td>
<td>Mazza</td>
<td>-</td>
</tr>
<tr>
<td>1I5R</td>
<td>HYC</td>
<td></td>
<td>1.6</td>
<td>C2</td>
<td>GOL</td>
<td>Qiu</td>
<td>Closed</td>
</tr>
<tr>
<td>1IOL</td>
<td>EST</td>
<td></td>
<td>2.3</td>
<td>C2</td>
<td>GOL</td>
<td>Azzi</td>
<td>Opened</td>
</tr>
<tr>
<td>1JTV</td>
<td>TES</td>
<td></td>
<td>1.54</td>
<td>C2</td>
<td>GOL</td>
<td>Gangloff</td>
<td>-</td>
</tr>
<tr>
<td>1QYV</td>
<td></td>
<td>NAP</td>
<td>1.81</td>
<td>C2</td>
<td>GOL</td>
<td>Shi</td>
<td>-</td>
</tr>
<tr>
<td>1QYW</td>
<td>5SD</td>
<td>NAP</td>
<td>1.63</td>
<td>C2</td>
<td>GOL</td>
<td>Shi</td>
<td>-</td>
</tr>
<tr>
<td>1QYX</td>
<td>ASD</td>
<td>NAP</td>
<td>1.89</td>
<td>C2</td>
<td>GOL</td>
<td>Shi</td>
<td>-</td>
</tr>
<tr>
<td>3DEY</td>
<td>DHT</td>
<td></td>
<td>1.7</td>
<td>C2</td>
<td>GOL</td>
<td>Mazumdar</td>
<td>-</td>
</tr>
<tr>
<td>3DHE</td>
<td>AND</td>
<td></td>
<td>2.3</td>
<td>C2</td>
<td></td>
<td>Han</td>
<td>Opened</td>
</tr>
</tbody>
</table>
Figure 6 Stereo ribbon presentation of human 17β-HSD1 structure. The α-helices are represented as magenta coils and designated as αB to αH, β-strands are blue arrows and marked as βA to βF, and loops and turns are drawn as gray ropes. The N-terminus and the C-terminus of the protein molecule are indicated (Ghosh et al., 1995).
### Table 2 The ratio of kinetic constants of 17β-HSD1 variants vs. that of wild type enzyme

<table>
<thead>
<tr>
<th>Enzyme variants</th>
<th>Estradiol to Estrone&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Estrone to Estradiol&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Effects on the enzyme</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Specific activity</td>
<td>$K_m$</td>
<td>$V_{max}$ or $K_{cat}$</td>
<td>Specific activity</td>
</tr>
<tr>
<td>H221A</td>
<td>0.12</td>
<td>2.23</td>
<td>0.11</td>
<td>0.18</td>
</tr>
<tr>
<td>H210A</td>
<td>1.17</td>
<td>0.97</td>
<td>1.02</td>
<td>0.90</td>
</tr>
<tr>
<td>H213A</td>
<td>0.92</td>
<td>0.97</td>
<td>1.02</td>
<td>0.90</td>
</tr>
<tr>
<td>H210A+H213A</td>
<td>0.72</td>
<td>1.08</td>
<td>0.50</td>
<td>0.50</td>
</tr>
<tr>
<td>Y155A</td>
<td>2.43E-03</td>
<td>5.32</td>
<td>5.88E-03</td>
<td>4.51E-04</td>
</tr>
<tr>
<td>C54A</td>
<td>0.92</td>
<td>0.97</td>
<td>0.93</td>
<td></td>
</tr>
<tr>
<td>A237V</td>
<td>0.89</td>
<td>0.97</td>
<td>0.95</td>
<td></td>
</tr>
<tr>
<td>S312V</td>
<td>1.12</td>
<td>1.08</td>
<td>1.08</td>
<td></td>
</tr>
<tr>
<td>S134A</td>
<td>1.20</td>
<td>1.07</td>
<td>0.87</td>
<td>0.49</td>
</tr>
<tr>
<td>S142A</td>
<td>2.30</td>
<td>4.50E-03</td>
<td>1.20</td>
<td>5.21E-03</td>
</tr>
<tr>
<td>K159A</td>
<td>2.35</td>
<td>4.15E-03</td>
<td>0.83</td>
<td>2.93E-03</td>
</tr>
<tr>
<td>E282A</td>
<td>0.71</td>
<td>0.74</td>
<td>0.82</td>
<td>1.16</td>
</tr>
<tr>
<td>H221AE282A</td>
<td>1.63</td>
<td>0.10</td>
<td>3.70</td>
<td>0.36</td>
</tr>
<tr>
<td>E282Q</td>
<td>0.79</td>
<td>0.72</td>
<td>0.48</td>
<td>0.73</td>
</tr>
<tr>
<td>H221AE282Q</td>
<td>1.87</td>
<td>0.17</td>
<td>2.72</td>
<td>0.49</td>
</tr>
<tr>
<td>L111EV113F</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>A170E+F172</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>H221L</td>
<td>3.53</td>
<td>0.20</td>
<td>4.50</td>
<td>0.44</td>
</tr>
<tr>
<td>H221Q</td>
<td>3.33</td>
<td>0.67</td>
<td>2.33</td>
<td>0.65</td>
</tr>
<tr>
<td>L149V</td>
<td>2.17</td>
<td>0.05</td>
<td>10.0</td>
<td>0.04</td>
</tr>
<tr>
<td>S12K</td>
<td>8.14</td>
<td>0.35</td>
<td>0.44</td>
<td>0.77</td>
</tr>
<tr>
<td>L36D</td>
<td>7.03</td>
<td>1.85</td>
<td>285.56</td>
<td>2.32</td>
</tr>
<tr>
<td>H221A</td>
<td>6.75</td>
<td>0.40</td>
<td>30</td>
<td>0.47</td>
</tr>
<tr>
<td>E282A</td>
<td>1.25</td>
<td>0.53</td>
<td>0.5</td>
<td>0.59</td>
</tr>
<tr>
<td>S142C</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>S142G</td>
<td>4.5</td>
<td>0.04</td>
<td>145</td>
<td>0.02</td>
</tr>
<tr>
<td>C10S</td>
<td>1.51</td>
<td>1.13</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup> NAD+ was used as cofactor in the kinetic tests. <sup>b</sup> NADPH was used as cofactor in the kinetic tests. ND, undetectable. Some of the 17β-HSD1 kinetic data were reported by Jin et al. <sup>57</sup>
4.1 Substrate recognition

The substrate recognition domain of 17β-HSD1 structure is buried under the flexible loop located between βF and αG', and delimited by the C-terminal region. The tunnel-like substrate binding cavity is composed majorly by hydrophobic residues, such as Leu96, Val143, Met147, Leu149, Pro150, Pro187, Val255, Phe226, Phe259, Leu262, Leu263 and Met279, as well as polar residues Asn152 and Tyr218. The βFαG'-loop acts as a lid covering the entry of the cavity. This segment is highly flexible and unable to be defined in twelve 17β-HSD1 structures. While in the rest ten structures, it shows three possible conformations, including the closed, semi-opened and opened conformation (Table 1). Interestingly, all structures with the presence of cofactor analog NADP+ adopt a closed conformation, whereas structures only with natural steroid ligands exhibit an opened conformation, which suggests the modulation role of cofactor on the conformation of the loop. Moreover, the loop region in structures complexed with inhibitor CC-156 (E2B) and EM-1745 also has a close conformation even without cofactor. Only the apoenzyme has a semi-opened conformation at this flexible loop region. In the close conformation, residue Phe192 from the loop region forms a T-stacking conformation with residue Tyr155, providing extra contacts for stabilizing the bound ligand94. The roles of residues from the active site of 17β-HSD1 have been investigated by mutagenesis and kinetic experiments which are summarized in Table 2. Residue His221 as well as Tyr155/Ser142 are critical for steroid substrate recognition through their hydrogen bonds with the O3 and O17 of the ligand, respectively. Residue Glu282 is supposed to play the same important role as the His221 does since it might also form a hydrogen bond with the O3 of the bound steroid79, as showed in the E2 complex structure76. However, the variant E282A in Huang et al.’s experiment did not show any significant modification in kinetics82. In contrast, residue Leu149 plays an important role for the discrimination of C-19 steroids and estrogens. Steroid ligand is stabilized by hydrogen bonds between O3 and His221/Glu282 at the recognition end, as well as between O17 and Tyr155/Ser142 at the catalytic end of the cavity.

4.2 Catalytic mechanism of 17β-HSD1

The kinetics of 17β-HSD1 follows the common chemical mechanism: a reversible hydride transfer from NADPH to a ketosteroid or a hydride transfer from a hydroxysteroid to NADP+, which is achieved by a proton shift for charge equalization. Based on mutational and structural studies, three conserved amino acids, Tyr155, Lys159 and Ser142 (catalytic triad), and a water molecule have been identified to be essential for the catalytic process73, 75-76. Previous kinetic studies, which was measuring the rate of isotopic exchange between substrate-product pairs while varying concentrations of unlabeled reactants, demonstrated that the binding of substrate and cofactor is random during the reaction85. Therefore, three hypotheses of the catalytic mechanisms of 17β-HSD1 have been proposed: one is a simultaneous transformation of proton and hydride; the other two are stepwise processes which differed in the intermediate presence of either a carbocation or an
oxyanion (Figure 7). The proton relay is mediated by the phenyl ring of Tyr, an electrostatic interaction between the protonated side chain of Lys and a hydrogen-bond network involving Lys, Asn and two water molecules. Phe may also involve in this step by forming a T-shape conformation with Tyr to increase the acidity of the phenol group of Tyr.

Figure 7 Two possible stepwise catalytic mechanisms for 17β-HSD1. (A) In the first step the prochiral S configuration (pro-S) hydride of NADPH is transferred to the α-face of E1 at the planar C17 carbon (A1), resulting in an energetically favorable aromatic system; subsequently the resultant oxyanion is protonated by the acidic OH group of Tyr (A2). (B) In the first step the keto oxygen of E1 is protonated by the acidic OH of Tyr (B1); then the resultant carbocation accepts the pro-S hydride of NADPH at the α-face (B2). Hydrogen bonds are represented in dashed lines (Marchais-Oberwinkler et al., 2011).

4.3 Inhibitors of 17β-HSD1

The development of inhibitors of 17β-HSD1 began in the 1970s and gradually gained momentum thereafter before culminating in the first decade of the 2000s. Despite the number of years of research, no inhibitor has yet reached the stage of clinical trials. The general properties of a good inhibitor should be highly potent and non-estrogenic. Also, it should be selective to 17β-HSD1 over the other 17β-HSD isozymes, especially 17β-HSD2, which catalyzes the reverse reaction (eg. oxidation of estrogens). The development of 17β-HSD1 inhibitor can generally be concluded into four different series (Figure 8). The first series of 17β-HSD1 inhibitors
were E2 derivatives bearing a bromoalkyl side chain at the 16α-position represented by the compound EM-251\textsuperscript{90}. This irreversible competitive inhibitor EM-251 on 17β-HSD1 has an IC\textsubscript{50} of about 320 nM, but was proven to have estrogenic activity on the estrogen sensitive human breast cancer cell line ZR-75-1\textsuperscript{91}. A modification at the C6 position of E2 has led to the development of a second series of reversible inhibitors. These inhibitors have a thiaheptamamide side chain at the 6β-position of E2, and were represented by the compound EM-678 (IC\textsubscript{50}=0.17μM) which was found to be more potent than the substrate E1 itself\textsuperscript{92}. Similar as the first series inhibitors, it also has an estrogen effect\textsuperscript{92-93}. Based on the binding energies of both the cofactor and substrate sites\textsuperscript{94}, as well as the three dimensional-structure of 17β-HSD1\textsuperscript{75-76}, a third series of inhibitors from E2-adenosine hybrids were developed. These molecules are represented by compound EM-1745. This compound has an E2 moiety to interact with the substrate-binding site and an adenosine moiety to interact with the cofactor binding site, which is connected by an eight methylene groups side chain\textsuperscript{95}. Though it has a high inhibitory activity on purified 17β-HSD1 (IC\textsubscript{50}=52nM), there are some major drawbacks such as difficulty to penetrate the cell membrane and weak competition ability against NADPH in intact cells \textsuperscript{96}. Further studies focused on a benzyl group at the 16β-position of E2, which is proven to be efficient in improving the inhibitory activity, yielded the 16β-m-carbamoylbenzyl-E\textsubscript{2} (CC-156), which is the most potent 17β-HSD1 inhibitor by far with an IC\textsubscript{50} value of 44nM for the conversion of E1 into E2\textsuperscript{97}. However, this fourth series of compounds was demonstrated to have estrogenic activity. It stimulated the proliferation of estrogen receptor positive cell line MCF-7 and T-47D cells\textsuperscript{97}. To reduce the unwanted estrogenic activity of CC-156, a series of modification at position 2, 3 and 7 have been made and assessed, yielding the compound 18 (2-MeO-CC-156)\textsuperscript{97} which is less potent (IC\textsubscript{50} of about 230nM) than CC-156 but bearing no estrogenic activity, and a new potent nonestrogenic compound named as 3-(2-bromoethyl)-16β-(m-carbamoylbenzyl)-17β-hydroxy-1,3,5(10)-estratriene (PBRM)\textsuperscript{98-99}. The latter did neither inhibit other 17β-HSDs nor CYP3A4\textsuperscript{100}, and demonstrated to form a covalent bond with 17β-HSD1. A long delay period (i.e. 3-5 days) was required to restore the 17β-HSD1 activity in cells after they had been treated with PBRM\textsuperscript{101}. Moreover, further investigation demonstrated its efficiency in both breast cancer cells and human tumor xenografts in nude mice\textsuperscript{99-100}. 
Figure 8 Key inhibitors of 17β-HSD1 from different Series (Poirier 2011).

Other than the inhibitors with a steroidal scaffold, several classes of non-steroidal 17β-HSD1 inhibitors have also been reported, such as the phytoestrogens\textsuperscript{102-103}, gossypols\textsuperscript{104-105}, thiophenepyrimidinones\textsuperscript{106}, (hydroxyphenyl)naphthols\textsuperscript{107-109}, and bis(hydroxyphenyl)heterocycles\textsuperscript{110-112}. Among these non-steroidal inhibitors, the bicyclic substituted hydroxyphenylmethanones (BSHs) exhibited high inhibitory activity toward the 17β-HSD1 enzyme\textsuperscript{113-114}. The following structural optimized \((5-(3,5\text{-dichloro-4-methoxyphenyl})\text{thiophen-2-yl})(2,6\text{-difluoro-3-hydroxyphenyl})\text{methanone}\) displayed a subnanomolar IC\textsubscript{50} towards the enzyme as well as high selectivity over other enzymes, especially the 17β-HSD2\textsuperscript{89}, and estrogen receptors\textsuperscript{115}, making it a promising candidate for following development as a therapeutic agent.

Beside the traditional 17β-HSD1 inhibitors, a series of E2 derived pure antiestrogens bearing a 7α-alkylamide side chain and a D-ring modification (a halogen atom or a double bond) were reported to exert potent inhibitory effects on 17β-HSD1 activities\textsuperscript{116}. These compounds were defined as dual-site inhibitor which represented by compound EM-139\textsuperscript{116}. Although the inhibition on 17β-HSD1 activities was obtained with this series of inhibitors, the lack of selectivity for other enzymes compromised their potential in clinical utilities\textsuperscript{117}.

5 The role of 17β-HSD7

17β-HSD7 is another important multi function enzyme in the reductive 17β-HSDs. Like 17β-HSD1, it catalyzes the formation of E2 from E1 and performs a more significant role in the inactivation of DHT into 3β-diol\textsuperscript{118-119}. 17β-HSD7 was reported to be primarily involved in cholesterol synthesis\textsuperscript{120-121}, and was suggested to be predominantly involved in cholesterol metabolism rather than in sex steroid synthesis \textsuperscript{122-123}. However, experiment conducted by Mr. Thériault in Prof. Lin’ lab demonstrated that inhibiting E1 to E2 activity of the enzyme by inhibitor is not blocking its zymosterol to zymosterone activity (unpublished data). Moreover, unlike aromatase, which converts testosterone (T) to E2 and is mostly expressed in stromal cells, 17β-HSD7 is
principally expressed and modulated in epithelial cancer cells such as MCF-7 and T47D\textsuperscript{124}. Furthermore, recent in vitro and in vivo experiments demonstrated that inhibition of 17β-HSD7 can induce cell cycle arrest and trigger cell apoptosis in BC cells, and auto-downregulation feedback of the enzyme, leading to significant shrinkage of xenograft tumors\textsuperscript{118, 124}. Furthermore, recent kinetic study showed that 17β-HSD7 has a $K_{\text{m}}$ value of 5.2±0.4 μM which is much higher than the value of 17β-HSD1 (0.03±0.01 μM); while the $k_{\text{cat}}$ value of 17β-HSD7 (2.9±0.4 s\textsuperscript{-1}) is much lower than the value of 17β-HSD1 (0.0063±0.0003 s\textsuperscript{-1})\textsuperscript{67, 125}. As a result, the $K_{\text{cat}}/K_{\text{m}}$ value of 17β-HSD7 is 80,000 times lower than the value of 17β-HSD1, indicating that these two reductive steroid enzymes may responsible of the E1 to E2 conversion at different substrate (E1) levels.

6 Statistical Analysis of RNA sequencing Data in Cancer Research

DNA sequencing technologies have been advanced during recent years due to the development of high throughput sequencing (HTS) technologies which can sequence multiple DNA molecules in parallel\textsuperscript{126}. They enable simultaneous sequencing of millions of DNA molecules and are widely applied on genomics, epigenomics and transcriptomics\textsuperscript{127}. RNA sequencing (RNA-seq) provides a profound advantage over other methods on cancer diagnosis and classification, prediction of response to therapy and prognosis, as well as unveiling the molecular bases of tumorigenesis\textsuperscript{128}. Moreover, transcriptomic profiling through RNA-seq will facilitate the development of personalized treatment for cancer patients through the molecular classification of subtypes\textsuperscript{128}. The Cancer Genome Atlas (TCGA) is a community resource project launched in 2005 by the National Institute of Health (NIH) as a pilot project aiming to discover and catalogue major cancer-causing genome alterations in large cohorts through large-scale genome sequencing and integrated multi-dimensional analyses. The Genome Sequencing Centers (GSCs) of TCGA performed large-scale DNA sequencing on two complementary DNA (cDNA) samples from every TCGA cancer case: one from the tumor specimen and the second from non-malignant tissue to serve as a control. The TCGA database is currently the largest database of cancer genetic information of over 30 kinds of human tumours\textsuperscript{129}. TCGA database provides the most complete clinical information of each patient, and is widely used in many studies\textsuperscript{130-131}.

7 Working Hypothesis and Research Objectives

7.1 Hypothesis

7.1.1 PBRM inhibiting 17β-HSD1 activity would be through the formation of a covalent bond with the enzyme. The interactions of the three inhibitors (PBRRM, 2-MeO-CC-156 and EM-139) with 17β-HSD1 would have significant difference.
7.1.2 The substrate inhibition of 17β-HSD1 would be due to the formation of a dead-end complex which is involving the binding of a reversely oriented E1 and the enzyme.

7.1.3 The analysis of RNA sequencing data would unveil potential new target and combined therapy for breast cancer treatment.

7.2 Objectives

Objective one: To elucidate the structural detail of representative inhibitors interacting with 17β-HSD1, such as EM-139, 2-Meo-CC-156 and PBRM. To achieve this, we have expressed and purified the recombinant 17β-HSD1 protein with Sf9 cells, which then was used in co-crystallization with these inhibitors in the presence or absence of cofactor analog NADP+. The crystal structures of the three complexes were determined and analyzed.

Objective two: To identify the mechanism of the substrate inhibition of 17β-HSD1 and in silico design of inhibitors based on this information. To reach this goal, we have co-crystallized the purified 17β-HSD1 with E1, in the presence or absence of cofactor analog NADP+. After determination of the binary and ternary complex structures, a comparative analysis with previously reported E2/testosterone complexes will be performed to elucidate the substrate inhibition mechanism, followed by computer assisted inhibitor design.

Objective three: To use RNA-seq data from large number clinical samples from TCGA-BRCA cohort to identify novel targets or combined therapy for breast cancer treatment.
Chapitre 1 Combined biophysical chemistry reveals a new covalent inhibitor with a low-reactivity alkyl halide

1.1 Résumé
La 17β-HSD1 joue un rôle central dans la progression des maladies liées aux œstrogènes en raison de son implication dans la biosynthèse des œstrogènes, en particulier de l’estradiol, constituant une cible thérapeutique importante pour le traitement endocrinien. Auparavant, le composé principal 16β-(m-Carbamoylbenzyl)-E2 (CC-156) était décrit comme un puissant inhibiteur de 17β-HSD1 dans la transformation de l’œstrone en estradiol. Cependant, l’activité œstrogénique de l’inhibiteur a compromis son potentiel de développement ultérieur. Une modification à la position C-2 du CC-156 a produit un inhibiteur non œstrogénique, le 2-MeO-CC-156, avec beaucoup moins de puissance d’inhibition que celle d’origine. Des recherches plus poussées à la position C-3 du CC-156 donnent un nouvel inhibiteur irréversible, non œstrogénique, puissant et stéroidien, le 3-(2-bromoéthyl)-16β-(m-carbamoylbenzyl)-17β-hydroxy-1,3,5(10)-éstratriène (PBRM). Dans cette publication, nous rapportons les structures des complexes ternaires de la 17β-HSD1 avec le NADP+ et l’inhibiteur 2-MeO-CC-156 ou le PBRM.

1.2 Abstract
17β-Hydroxysteroid dehydrogenase type 1 (17β-HSD1) plays a pivotal role in the progression of estrogen-related diseases for its involvement in the biosynthesis of estrogens, especially estradiol, constituting a valuable therapeutic target for endocrine treatment. Previously, the lead compound 16β-(m-Carbamoylbenzyl)-E2 (CC-156) was described as a potent 17β-HSD1 inhibitor of the transformation of estrone into estradiol. However, the estrogenic activity of the inhibitor compromised its potential for further development. A modification at the position C-2 of CC-156 produced a non-estrogenic inhibitor 2-MeO-CC-156, with a much less potency as compared with the original one. Further investigation at the position C-3 of CC-156 yield a new potent and steroïdial non-estrogenic irreversible inhibitor 3-(2-bromoethyl)-16β-(m-carbamoylbenzyl)-17β-hydroxy-1,3,5(10)-éstratriène (PBRM). In the present paper, we report structures of the ternary complexes of 17β-HSD1 with NADP+ and inhibitor 2-MeO-CC-156 and PBRM. In the 17β-HSD1-2-MeO-CC-156-NADP+ complex, the presence of a methoxy group at C-2 of the inhibitor significantly reduces its estrogenic effect in estrogen-dependent cancer cells, however it also impedes the essential hydrogen bond at the recognition end of the ligand binding pocket, significantly decreasing its inhibitory activity to the enzyme. For the 17β-HSD1-PBRM-NADP+ complex, the hydrogen bond between O-19 of the inhibitor and Oγ of Ser142 is much weaker as compared with that of CC-156 complex, contributing to its relatively high IC₅₀ to 17β-HSD1 activity. However, the bromoethyl group at position C-3 of the inhibitor warrants its non-estrogenic profile, and secures its
selectivity of 17β-HSD1 through the formation of a covalent bond with Nε of residue His\textsuperscript{221}, suggesting its potential as a therapeutic agent for EDDs.

1.3 Introduction

Covalent inhibitors (CIs) are more beneficial than noncovalent ones because of the reduced risk of drug resistance, extended inhibition effect, increased efficiency with lower doses, and fewer side effects\textsuperscript{1}. However, despite these advantages, toxicity issues encountered with the first generation of CIs related to their high reactivity, low specificity of action, and some immunogenicity response resulted in resistance from the pharmaceutical industry\textsuperscript{2}. Nevertheless, the approval of more specific and safe targeted CIs in the past decade led to a resurgence of interest in the pharmaceutical research field\textsuperscript{3,4}. However, the design of such inhibitors remains a challenge, considering that a high binding affinity for the targeted protein, as well as an inherent reactivity, are two essential elements that must be combined in a single molecular entity to obtain a valuable drug candidate. Even if some covalent drugs have been documented bearing a low-reactivity group that could lead to alkylation in a particular molecular context\textsuperscript{5}, the electrophilic group incorporated into CI is generally highly reactive (α,β-unsaturated ketone, α-haloketone, cyanamide, fluorophosphate, and epoxide), with the inconvenience of increasing the risk of off-target and nonspecific tagging\textsuperscript{6}. The use of less reactive electrophile groups is thus suitable for increasing the level of CI specificity\textsuperscript{2,7-9}.

Most CI drugs are based on the reactivity of cysteine\textsuperscript{10}, the strongest nucleophile among natural amino acids (AAs), allowing the alkylation of a large diversity of electrophiles\textsuperscript{11}. However, because of its low abundance or an inaccessible position in the enzyme catalytic site, other nucleophilic residues have been exploited for covalent inhibition, such as lysine, serine, tyrosine, threonine, aspartate, and glutamate\textsuperscript{12-13}. One uncommon case is the histidine (His) residue, which, despite its good nucleophilicity and its presence at the catalytic site of many enzymes\textsuperscript{14}, has been very rarely exploited in CI design\textsuperscript{15}.

17β-hydroxysteroid dehydrogenase type 1 (17β-HSD1) catalyzes the final step of the transformation of estrone (E1) to estradiol (E2), the most potent estrogen, and is considered a promising therapeutic target for endocrine treatment\textsuperscript{16-21}. This enzyme also catalyzes the reduction of dehydroepiandrosterone (DHEA) into 5-androstene-3β,17β-diol (5-diol) and dihydrotestosterone (DHT) into 5α-androstane-3β,17β-diol (3β-diol), which has been suggested to become more important after menopause, and may be involved in aromatase inhibitor resistance\textsuperscript{16,21-23}. It is well-known that E2 stimulates breast cancer and also plays a crucial role in other estrogen-related diseases such as ovarian cancer, endometriosis, and endometrial cancer\textsuperscript{24-25}. Thus, the blockade of the biosynthesis of E2 is considered to be a valuable therapeutic approach for treating estrogen-dependent diseases\textsuperscript{24-26}.
Previous reports have described 16β-\(m\)-carbamoylbenzyl)-E2 (CC-156) (Figure 1.1) as a potent competitive and reversible inhibitor of 17β-HSD1 with an IC\(_{50}\) value of 44 nM\(^27\). Unfortunately, this compound has an estrogenic activity observed by the proliferation of the stimulation of estrogen receptor (ER) positive cell lines MCF-7 and T-47D\(^27\). To reduce this unwanted estrogenic activity, further development was then engaged to modify the E2 scaffold of CC-156. The addition of a methoxy (MeO) group at position C-2 of CC-156, which produced 2-MeO-CC-156 (Figure 1.1), was efficient in attenuating estrogenic activity but was unfavorable for enzyme inhibition\(^27\). A more promising strategy next focused on the chemical modification of the C-3 phenolic group, which is known to be important for the binding of the E2 scaffold to ER\(^28\), and resulted in the discovery of PBRM (Figure 1.1), the first nonestrogenic irreversible inhibitor of 17β-HSD1\(^29\). Further investigations demonstrated the PBRM efficiency in both breast cancer cells and human breast tumor xenografts in nude mice\(^30-31\), as well as interspecies differences of 17β-HSD1 inhibition\(^32\). Kinetic studies classified PBRM as a competitive and irreversible inhibitor of 17β-HSD1, and a covalent binding of PBRM with 17β-HSD1 was then demonstrated by using a 17α-tritiated derivative of PBRM\(^32\). Furthermore, a molecular modeling study investigating interspecies inhibitory activity of PBRM noted His\(^221\) as a potential key AA involved in the formation of a covalent bond with the bromoethyl side chain. Interestingly, as an indication of the applicability of the bromoethyl group for developing a specific CI drug, PBRM possesses the expected properties of a CI, such as an extended inhibition action and a very low promiscuity rate\(^33\). The bromoethyl side chain also provides a reduced in vitro CYP metabolism in comparison to its phenolic analog (CC-156), which is translated by a higher in vivo bioavailability for PBRM\(^29\).

Despite the indirect evidence of an alkylation between PBRM and 17β-HSD1, the existence and the exact configuration of expected covalent bonds remain to be proven. This was especially significant, considering the predicted low reactivity between a His residue and a primary alkyl halide, even more in a physiological environment\(^32\). Obviously, the demonstration of the capacity of a common and accessible functional group like a primary alkyl halide to act as a reagent for the N-alkylation of an enzyme could demonstrate the viability of such a weak electrophile group for the development of a new type of selective CI. In fact, very few documented examples of an enzyme alkylation by a primary alkyl halide derivative have been reported to date, including a case of O-alkylation from a carboxylate group of Asp\(^106\) residue for haloalkane dehalogenase tagging\(^34\) and a suspected S-alkylation from Met\(^193\) of 16α-bromopropyl-E2 leading to an irreversible inhibition of 17β-HSD1\(^35-36\). Importantly, the primary alkyl halide electrophile group must not be confused with activated alkyl halide units, like the highly reactive N-ethylhalide of “nitrogen mustard” agents, which form a covalent bond via the formation of an intermediate aziridinium very reactive species that reacts with the DNA nitrogenous base\(^37\), or with benzyl halide\(^38-39\) and \(\alpha\)-halo ketone\(^40\) groups, which are not specific, albeit useful in labeling affinity agents for enzyme characterization\(^41-42\).
1.4 Results and Discussion

Analysis of PBRM molecular interactions before the His\(^{221}\) N-alkylation

The inhibitor PBRM bears a bromoethyl side chain at position C-3, making the inhibitor a little longer compared to CC-156’s. However, since the core structure of PBRM, especially the C-16 benzylamide moiety, is the same as that of CC-156, we expected the major conformational modifications during the binding of PBRM had happened at the recognition end (His\(^{221}\), Glu\(^{282}\)) of the steroid binding site\(^{43-44}\). We were therefore interested in investigating the interaction of the bromoethyl chain with His\(^{221}\) before the N-alkylation event. Besides, the Met\(^{279}\) could possibly act as a nucleophile over the bromoethyl chain, considering that the distance between the C-3 of phenolic OH at C-3 in CC-156 and Met\(^{279}\) (3.97 Å) is similar to with His\(^{221}\) (3.45 Å) (data from CC-156 ternary complex\(^{45}\)).

The pseudo PBRM complex structures were visually built from the CC-156 ternary complex structure using SeeSAR software. In the CC-156 complex, the Glu\(^{282}\) side chain faces the binding site to make a hydrogen bond with the inhibitor, leaving no space to build the bromoethyl side chain on CC-156. Since Glu\(^{282}\) is a solvent-exposed flexible residue with a high average B-factor value of 40.6 Å\(^2\), we modeled its conformation to have the side chain exposed to the solvent (as described in the Experimental Section). Moreover, with the existence of side chains from His\(^{221}\) and Met\(^{279}\), the bromide from generated poses of PBRM maintained at least a van der Waals distance from them, which is too long to overcome the force field limitations that do not allow for covalent reactions between the bromoethyl moiety and the side chains. To explore the possible positions of the bromoethyl side chain before the subsequent N-alkylation reaction, residues His\(^{221}\) and Met\(^{279}\) were mutated into Ala, which has a smaller side chain. The best poses with the highest estimated affinity using this binding site conformation are presented in Figure 1.2 A,B. The distance from the CH\(_2\) of the bromoethyl side chain to the NH of the His\(^{221}\) side chain is about 2.5 Å, whereas that distance to the S of Met\(^{279}\) is 2.0 Å. This result urges us to engage co-crystallization experiment for 17β-HSD1-PBRM to clarify the mechanism.

Because no example of N-alkylation between an enzyme and a primary alkyl halide has been reported to date and also to rule out the possibility of the Met\(^{279}\) of 17β-HSD1 to act as nucleophile over the primary alkyl halide (Figure 1.2), we thus seized this opportunity and engaged cocrystallization experiments of PBRM with 17β-HSD1 to prove the capacity of such a weak electrophile to form a covalent bond with the suspected His\(^{221}\) residue, an AA rarely exploited in design of CI drugs\(^\text{15}\).

Structure determination of enzyme-inhibitor complex crystals

The space group identified for all the crystals was \(P2_12_12_1\) with a dimer in one asymmetric unit representing the functional unit of the enzyme\(^{46-47}\). Two ternary complexes, 17β-HSD1–2-MeO-CC-156–NADP\(^+\) and 17β-
HSD1–PBRM–NADP⁺, were refined to 2.1 and 2.2 Å, respectively. The two models show good stereochemistry⁴⁸, and the quality of the final refined models can be accessed from the statistics in Table 1.1. The models of 17β-HSD1 with PBRM (F0D) and 2-MeO-CC-156 (F0A) ternary complexes show very clear electron density for almost all residues, except for the C-terminal end of the protein (residues 286–327) as well as the flexible loop region from Ala¹⁹¹ to Gly¹⁵⁸, as observed in other 17β-HSD1 complexes²¹, ⁴³, ⁴⁹-⁵⁰. The active-site structure of both inhibitor complexes for the A subunit is shown in Figure 1.3.

Comparison of 2-MeO-CC-156 and CC-156 ternary complexes

The presence of a methoxy group at position C-2 in 2-MeO-CC-156 introduces a strong hydrophobic interaction with residue Leu²⁶² with the distance of 3.15 Å between C-32 (CH₃ of MeO) of inhibitor and C8 of the AA residue. This interaction causes the inhibitor to shift 1.04 Å at the O-4 end and to rotate by approximately 4.8° at the steroid core and 3.5° at the benzylamide ring, as compared to the position of the CC-156 complex when superimposing the 2-MeO-CC-156 complex with the previously reported CC-156 ternary complex (PDB ID 3HB5) by Cα atoms (Figure 1.4A)⁴⁵. The side chain of Glu²⁸², used to make a hydrogen bond with the inhibitor in the CC-156 ternary complex, adopts a conformation facing the outside of the protein. Thus, no hydrogen bond can form between the AA residue and 2-MeO-CC-156. Besides, the movement of the O-4 at the end of 2-MeO-CC-156 forces the imidazole side chain of His²²¹ to shift away by 1.49 Å for the Nε as compared with the position of the Cε of His²²¹ in CC-156 complex. The hydrogen bond between the inhibitor and His²²¹, which is important for ligand recognition and orientation⁵¹, is established in the 2-MeO-CC-156 complex with a distance of 2.86 Å (Table 1.2). However, the movement of the side chain of His²²¹ toward the solvent leads to the decrease of its stability (average B-factor of 49.0 Å² of the AA residue as compared with 39.0 Å² of the subunit) compared with its counterpart (average B-factor of 30.7 Å² of the AA residue as compared with 29.8 Å² of the subunit) in the CC-156 complex.⁴⁵ Indeed, when 0.1 μM inhibitor concentration was used, 2-MeO-CC-156 inhibited 37% of the transformation of E1 into E2, whereas CC-156 inhibited 77% of the same reaction²⁷. This is in agreement with the relatively high flexibility of the bound 2-MeO-CC-156 (average B-factor, 54.9 Å) as compared to CC-156 (average B-factor, 35.6 Å). The hydrogen bonding with Ser¹⁴² is conserved in the 2-MeO-CC-156 complex, as in the CC-156 ternary complex.

For the benzylamide ring, the π–π interaction between Tyr¹⁵⁵ and the ring is conserved (Figure 1.4A). The distance between Cε₂ of Tyr¹⁵⁵ and C-23 of 2-MeO-CC-156 is 3.50 Å, and the distance between the centroid of the two phenyl rings is about 4.4 Å, a little bit longer than the distances observed in the CC-156 ternary complex (4.3 Å). Nevertheless, three hydrogen bonds between the carboxamide group of the inhibitor and Leu⁹⁵ and Asn¹⁵² residues are presented (Figure 1.4A and Table 1.2). However, it is more reasonable that the
O-29 of the carboxamide (CON) group of the inhibitor acts as an acceptor forming a hydrogen bond with N of Leu\(^{95}\), whereas the N-30 (CON) acts as a donor forming two hydrogen bonds with O of Leu\(^{95}\) and O\(\delta\) of Asn\(^{152}\) (Table 1.2). Thus, the CON group in 2-MeO-CC-156 adopts a conformation of 180° flip, as compared with that in the CC-156 ternary complex (Figure 1.4A).

**Enzyme interaction with NADP\(^+\) in ternary complexes**

Similar to the previously described model\(^{50}\), only the adenine ring, the ribose and phosphate groups of NADP\(^+\) molecule can be unambiguously identified in the electron densities of the 17\(\beta\)-HSD1–2-MeO-CC-156–NADP\(^+\) ternary structure (Figure 1.3A). The NMN moiety of the NADP\(^+\) molecules missing from the densities was omitted from the final models. It indicates that the major interaction between the NADP\(^+\) and enzyme happens at the ADP part, in agreement with previous the structure-function study\(^{44}\). As compared with the NADP\(^+\) molecule in CC-156 ternary complex, the 2'-phosphate group attached to the adenine ribose in 2-MeO-CC-156 ternary complexes has moved 3.7 Å toward the position of N\(\eta\) of Arg\(^{37}\) in CC-156 complex, and is stabilized by the water bridged hydrogen bond with the N of Arg\(^{37}\) and Asp\(^{38}\) as well as the O\(\gamma\) of Thr\(^{41}\). As a result, the side chain of Arg\(^{37}\) has moved to the protein surface and stabilized by forming a hydrogen bond with O\(\delta\) of Asp\(^{38}\).

Two important hydrogen bonds between the adenine ring and residues Asp\(^{65}\) and Val\(^{66}\) are conserved in 2-MeO-CC-156 ternary complex, as well as the hydrogen bond between the O-3 attached to the adenine ribose and O\(\gamma\) of Ser\(^{12}\). No obvious different interaction was observed at the NADP\(^+\) binding site in the PBRM complex, as compared with that of the 2-MeO-CC-156 complex. Similarly, the electron density map of the nicotinamide and the attached ribose of the NADP\(^+\) molecule are unable to define (Figure 1.3B). The hydrogen bonds with surrounding residues Ser\(^{11}\), Ser\(^{12}\), Asp\(^{65}\) and Val\(^{66}\), as well as the water bridged hydrogen bond with residues Asp\(^{38}\) and Thr\(^{41}\) stabilized the ADP moiety of the NADP\(^+\) molecule. No direct interaction was observed between the bound inhibitor and cofactor molecule in the ternary complex.

**Comparison of PBRM and CC-156 ternary complexes**

In the 17\(\beta\)-HSD1–PBRM–NADP\(^+\) ternary structure, an unambiguous continuity of electron density from the side chain of His\(^{221}\) to the bound PBRM is observed in both subunits, indicating the formation of a covalent bond between the N\(\varepsilon\) of His\(^{221}\) and the C-31 (BrCH\(_2\)) of PBRM (Figure 1.3B). The structure overlay of the complex with CC-156 complex shows a slight shifting at the C-3 end of PBRM (0.66 Å) as well as the imidazole side chain of His\(^{221}\) (0.89 Å) as compared with the positions of their counterparts in the CC-156 complex, indicating the dynamic process favoring the formation of the covalent bond between them. The slight movement of the steroid core of PBRM and side chain of His\(^{221}\) is caused by the formation of their covalent bond (Figure 1.4B). As a result, the distance of the hydrogen bond between O-19 of the inhibitor and O\(\gamma\) of Ser\(^{142}\) increased to 3.22 Å (Table 1.2). The hydrogen bond with Ser\(^{142}\) is one of the three major interactions in
which the potent inhibitor CC-156 interacts with 17β-HSD1\textsuperscript{30}, the increased distance of the bond thus indicating a less favored interaction of the inhibitor with the enzyme.

Similar to CC-156 and 2-MeO-CC-156 complexes, the π–π interaction between the benzylamide ring of PBRM and the side chain of Tyr\textsuperscript{155} is conserved. The distance between C\textsubscript{2} of Tyr\textsuperscript{155} and C-23 in the benzylamide ring of PBRM (3.35 Å) is slightly shorter than that in both the CC-156 (3.45 Å) and 2-MeO-CC-156 (see above) complexes. Besides, the distance of the centroid of the two phenyl rings (4.32 Å) is almost the same as in the CC-156 complex. The carboxamide group of PBRM adopts the same conformation as 2-MeO-CC-156 described previously, making three hydrogen bonds with Leu\textsuperscript{95} and Asn\textsuperscript{152} residues (Figure 1.4B). The distance of the three hydrogen bonds in the PBRM complex is similar to that in the CC-156 complex (Table 1.2), indicating their important role in the inhibitor binding to the enzyme. These molecular interactions are thus sufficiently favorable to bring the bromoethyl side chain of PBRM in proximity to His\textsuperscript{221} and to favor the reaction between these two complementary groups. In fact, such a reaction between an alkyl halide and a relatively poor nucleophile like His is not possible under physiological conditions. Even in the laboratory, excess amounts of imidazole or His were found to be unable to react with PBRM at room temperature\textsuperscript{32}. The proximity effect is thus a crucial factor to allow this unfavorable event, as has been previously demonstrated for low-reactivity electrophile groups in C\textsubscript{I} reactivity\textsuperscript{9}.

1.5 Conclusion

The present study illustrates the structural details of different inhibitory mechanisms of two potent 17β-HSD1 inhibitors, the reversible inhibitor 2-MeO-CC-156 and the irreversible inhibitor PBRM, as compared to CC-156. The results strongly support PBRM as a promising and selective new drug candidate for the adjuvant therapy of estrogen-dependent diseases. All these represent a breakthrough in the long history of the 17β-HSD1 inhibitor search. Also, and in a broader way, this is the first report of a specific N-alkylation between a His residue and a low-reactivity alkyl halide-based inhibitor, which supports the viability of such an approach toward the development of specific CIs.

1.6 Experimental Procedures

Materials. pFastBac\textsuperscript{™}1 vector, DH10Bac\textsuperscript{™} Competent E. coli, Gibco® Spodoptera frugiperda Sf9 cells, Sf-900\textsuperscript{™} III SFM (serum free medium), Sf-900 Medium (1.3X), Cellfectin\textsuperset{®} II Reagent, PureLink\textsuperset{™} HiPure Plasmid Maxiprep Kit, Ni-NTA Agarose were purchased from Thermo Fisher Scientific Corporation. The I-Max serum free medium for insect cells was purchased from Wisent Bioproducts. Albumin standard was purchased from Thermo Scientific. Protease inhibitor cocktail, sodium chloride, NAD\textsuperscript{+}, NADP\textsuperscript{+}, PMSF, β-octylglucoside (β-OG), estrone, trizma base, disodium ethylenediamine tetraacetate (EDTA), glycerol, phenylmethylsulphonyl (PMSF), polyethylene glycol 8000 (PEG-8K), dithiothreitol (DTT) and potassium phosphate monobasic were
obtained from Sigma-Aldrich. MonoQ (HR 5/5) column and Blue Sepharose® 6 Fast Flow resin were obtained from GE Healthcare Life Sciences. Antibiotic such as ampicillin, kanamycin, gentamicin, tetracycline, and penicillin-streptomycin were obtained from Thermo Scientific. Bradford Protein Assay kit and Protein Marker were purchased from Bio-Rad. The DU-80 spectrophotometer was from Beckman Coulter.

**Recombinant virus preparation.** 17β-HSD1 gene (*HSD17B1*) was first subcloned into pFastBac™1 donor plasmid through *RsrII* and *XhoI* double digestion to generate the pFastBac-HSD17B1 recombinant donor plasmid, which has then transformed into DH10Bac™ *E.Coli* competent cell to form the recombinant Bacmid-HSD17B1 shuttle plasmid. The integrity of these recombinant plasmids was confirmed by sequencing, which was provided by the genome sequencing and genotyping platform of the CHU de Québec - Research Center (Québec, QC, Canada).

**17β-HSD1 expression and purification.** Sf9 cells were maintained at 27 °C in stationary T-flasks and were passaged to 150 × 20 mm dishes for protein expression. Cells were infected with virus at a multiplicity of infection (MOI) of 0.1 to 1 pfu to produce virus stocks or at a MOI ≥10 for maximal protein expression. Recombinant 17β-HSD1 was purified by a fast preparation procedure modified from a previously described method\(^{52-53}\). Briefly, enzyme purification consisted in two FPLC steps using Blue-Sepharose affinity and MonoQ anion exchange columns. β-OG was added to the protein fraction thus obtained to stabilize the enzyme\(^{54}\). The protein concentration was measured by the Bradford method and its activity was measured by the oxidation of E2 to E1\(^{52}\).

**Inhibitors 2-MeO-CC-156 and PBRM.** The reversible and irreversible 17β-HSD1 inhibitors 2-MeO-CC-156 and PBRM, respectively, were synthesized from commercially available estrone, as previously reported\(^{27,29-30}\).

**Co-crystallization.** Ternary complex samples were prepared according to the repeated concentration and dilution method of Zhu *et al*\(^{55}\) to saturate 17β-HSD1 in high concentration with hydrophobic steroid. In brief, purified enzyme was subjected to a buffer change procedure via Centricon. The added buffer contains 0.06% (w/v) β-OG, 1 mM NADP\(^+\) and 25 µM of different inhibitors. The obtained complex samples were then concentrated to 20 mg/ml and used for crystal growth. Crystals were obtained using hanging-drop method with 400 µl of well solution consisting in 24% - 29% (w/v) PEG8K, 100 mM Tris buffer (pH 7.5 – 7.8) and 50 mM KH\(_2\)PO\(_4\) at 27 ºC.

**Data collection and structure determination.** Data collection was carried out using MAR CCD 165 mm detector at APS beamline 31-LRL-CAT at 100 K using a wavelength of 0.979 Å. Mineral oil was used as the cryoprotectant. The datasets were intergraded using MOSFLM\(^{56}\) and scaled with SCALA\(^{57}\) from the CCP4 suite\(^{58}\). The structures were solved by molecular replacement with MOLREP\(^{59}\) using the coordinate of 17β-
HSD1, with the highest current resolution (PDB code 1JTV)\textsuperscript{[50]}, as the search model. The structure parameters for the inhibitor 2-MeO-CC-156 and PBRM were generated using the Sketcher from CCP4 suite and were refined using REFMAC\textsuperscript{51}. The complex structures were subjected to multiple rounds of auto-refinement using REFMAC5 and manual refinement using Coot\textsuperscript{62}. The quality of the final models was evaluated with PROCHECK\textsuperscript{63}. The structure figures were prepared with the PyMOL\textsuperscript{64}.

**Manual edition using SeeSAR.** The manual compound edition was performed using SeeSAR\textsuperscript{65} software. The crystal structure of 17β-HSD1 in complex with CC-156 and NADP was taken from PDB code 3HB5\textsuperscript{45}. Inhibitor PBRM shares its core structure with CC-156. We thus chose to use the SeeSAR to build the pseudo PBRM complex structure from CC-156 in the CC-156 complex. The binding poses of PBRM were generated with SeeSAR, its geometry optimized by the Hydrogen bond and Dehydration (HYDE)\textsuperscript{66} as implemented in the software and ranked according to their estimated affinity to the binding site. Before the visual building of PBRM from CC-156, the system's energy was minimized using UCSF Chimera\textsuperscript{67} software after the side chain of Glu\textsuperscript{282} was modified using Dunbrack\textsuperscript{68} backbone-dependent rotamer library in Chimera in which the highest probability conformer not facing the binding site was selected. The mutation of His\textsuperscript{221} and Met\textsuperscript{279} to Ala was done, respectively using the rotamer tool in UCSF Chimera, and no further energy minimization was required.

**Notes**

The authors declare no competing financial interests.

The PDB ID of 17β-HSD1-2-MeO-CC-156-NADP\textsuperscript{+} and 17β-HSD1-PBRM-NADP\textsuperscript{+} are 6CGC and 6CGE, respectively.

**ACKNOWLEDGMENT**

The authors would like to acknowledge the CIHR's support (MOP97917) to Sheng-Xiang Lin, Donald Poirier and Charles Jean Doillon.
1.7 Reference


64. *The PyMOL Molecular Graphics System, Version 2.0* Schrödinger, LLC.

65. SeeSAR version 7.3; BioSolveIT GmbH, Sankt Augustin, Germany, 2018.


Figures and Legends

**Figure 1.1.** Three potent steroidal inhibitors of 17β-HSD1: 16β-\((m\text{-carbamoylbenzyl})\)E2 (CC-156), 2-methoxy-16β-\((m\text{-carbamoylbenzyl})\)E2 (2-MeO-CC-156), and 3-(2-bromoethyl)-16β-\((m\text{-carbamoyl benzyl})\)-17β-hydroxy-1,3,5(10)-estratriene (PBRM).
Figure 1.2. Results from the *in silico* building of PBRM at the binding site of 17β-HSD1. The binding site conformation of CC-156 ternary complex structure is represented (magenta) with Glu^{282} side chain solvent-oriented (labelled and colored in green). His^{221} and Met^{279} residues are labelled and shown in sticks in (A) and (B) respectively. The best pose of manually built PBRM with (A) His^{221} mutated into Ala, and (B) Met^{279} mutated into Ala are represented by pink and blue sticks, respectively. The distance from the CH\_2 of the bromoethyl side chain to (A) the NH of His^{221} side chain, and (B) the S of Met^{279} side chain are labelled.
Figure 1.3. View of the active sites within the A subunit of 2-MeO-CC-156 (A) and PBRM (B) ternary complex structures. Inhibitors 2-MeO-CC-156 (F0A) and PBRM (F0D) and cofactor NADP⁺ are shown in their omit F₀-\(F_c\) and 2\(F_0\)-\(F_c\) electron densities. The side chains of important residues Leu⁹⁹, Ser¹⁴², Asn¹⁵², Tyr¹⁵⁵, His²²¹, and Glu²⁸² are shown in their 2\(F_0\)-\(F_c\) electron densities. 2\(F_0\)-\(F_c\) maps are drawn in gray and contoured at 1σ; \(F_0\)-\(F_c\) maps are drawn in green and contoured at 2.5σ. The backbones of the A subunit in 2-MeO-CC-156 and PBRM complexes are shown in magenta and blue, respectively.
**Figure 1.4.** Superposition of A subunit of 2-MeO-CC-156 (magenta) and PBRM (blue) ternary complexes along with 17β-HSD1–CC-156–NADP⁺ (pink) at the binding sites, showing the inhibitors and important residues. (A) Superposition of 2-MeO-CC-156 and CC-156 complexes at the steroid binding site. (B) Superposition of PBRM and CC-156 complexes at the steroid binding sites. Interacting residues are labeled and shown as sticks. Hydrogen bonds of inhibitor 2-MeO-CC-156 and PBRM with their surrounding residues are presented in green dashed lines. Several important distances are labeled and indicated with black dashed lines.
### Table 1.1 Data collection and refinement statistics

<table>
<thead>
<tr>
<th></th>
<th>17β-HSD1-2-MeO-CC-156-NADP⁺</th>
<th>17β-HSD1-PBRM-NADP⁺</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data Collection</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Space group</td>
<td>P2₁2₁2₁</td>
<td>P2₁2₁2₁</td>
</tr>
<tr>
<td>Unit cell</td>
<td></td>
<td></td>
</tr>
<tr>
<td>a,b,c (Å)</td>
<td>41.75, 107.98, 115.72</td>
<td>42.82, 108.94, 116.36</td>
</tr>
<tr>
<td>α,β,γ (°)</td>
<td>90, 90, 90</td>
<td>90, 90, 90</td>
</tr>
<tr>
<td>Resolution range (Å)</td>
<td>25-2.10 (2.21-2.10)</td>
<td>25-2.2 (2.32-2.20)</td>
</tr>
<tr>
<td>Number of reflections</td>
<td>173088 (23506)</td>
<td>184811 (24512)</td>
</tr>
<tr>
<td>Unique reflections</td>
<td>28809 (3929)</td>
<td>28372 (4006)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>91.9 (87.7)</td>
<td>99.5 (98.1)</td>
</tr>
<tr>
<td>I/σ(I)</td>
<td>7.4 (3.2)</td>
<td>11.0 (3.4)</td>
</tr>
<tr>
<td>Rmean²</td>
<td>0.144 (0.456)</td>
<td>0.108 (0.521)</td>
</tr>
<tr>
<td>Multiplicity</td>
<td>6.0 (6.0)</td>
<td>6.5 (6.1)</td>
</tr>
<tr>
<td>Wilson B-factor (Å²)</td>
<td>27.5</td>
<td>33.0</td>
</tr>
<tr>
<td><strong>Refinement</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-work³</td>
<td>0.21</td>
<td>0.23</td>
</tr>
<tr>
<td>R-free⁴</td>
<td>0.26</td>
<td>0.30</td>
</tr>
<tr>
<td>r.m.s.d</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bond lengths (Å)</td>
<td>0.013</td>
<td>0.011</td>
</tr>
<tr>
<td>Bond angles (°)</td>
<td>1.79</td>
<td>1.68</td>
</tr>
<tr>
<td>Ramachandran plot (°)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Most favored regions</td>
<td>93.9</td>
<td>93.3</td>
</tr>
<tr>
<td>Additional allowed regions</td>
<td>6.1</td>
<td>6.5</td>
</tr>
<tr>
<td>Generously allowed regions</td>
<td>0.0</td>
<td>0.2</td>
</tr>
<tr>
<td>Disallowed regions</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Average B, all atoms (Å²)</td>
<td>39.0</td>
<td>47.0</td>
</tr>
</tbody>
</table>

* Data statistics for the outer shell are given in parentheses.

² The redundancy-independent R_{merge}/R_{sym}, R_{mean} = \frac{\sum_n \sum_{i=1}^n |I_{hkl,i} - \langle I_{hkl} \rangle|}{\sum_n \sum_{i=1}^n I_{hkl,i}}

³ R_{work} = \sum_{hkl} \left( |F_{obs}(hkl) - |F_{calc}(hkl)| \right) / \sum_{hkl} |F_{obs}(hkl)|

⁴ R_{free} = the cross-validation R factor for 5% of reflections

⁵ Calculated with PROCHECK.
Table 1.2. Hydrogen bonds between bound inhibitor and surrounding residues in 17β-HSD1 ternary complexes

<table>
<thead>
<tr>
<th>Complexes</th>
<th>Donor</th>
<th>Acceptor</th>
<th>Length (Å)</th>
</tr>
</thead>
</table>
| 17β-HSD1–2-MeO-CC-156–NADP⁺  
(PDB ID: 6CGC) | Ser¹⁴²: Oγ          | 2-MeO-CC-156: O-19 | 2.73       |
|           | Leu⁹⁵: N            | 2-MeO-CC-156: O-29 | 2.55       |
|           | 2-MeO-CC-156: N-30  | Leu⁹⁵: O           | 2.94       |
|           | 2-MeO-CC-156: N-30  | Asn¹⁵²: Oδ         | 2.52       |
|           | 2-MeO-CC-156: O-4   | His²²¹: Nε         | 2.86       |
| 17β-HSD1–PBRM–NADP⁺  
(PDB ID: 6CGE) | Ser¹⁴²: Oγ          | PBRM: O-19         | 3.22       |
|           | Leu⁹⁵: N            | PBRM: O-29         | 2.82       |
|           | PBRM: N-30          | Leu⁹⁵: O           | 2.88       |
|           | PBRM: N-30          | Asn¹⁵²: Oδ         | 2.61       |
| 17β-HSD1–CC-156–NADP⁺  
(PDB ID: 3HB5) | CC-156: O-4         | Glu²⁸²: Oε         | 2.61       |
|           | Ser¹⁴²: Oγ          | CC-156: O-19       | 2.72       |
|           | Leu⁹⁵: N            | CC-156: N-30       | 2.77       |
|           | CC-156: O-29        | Leu⁹⁵: O           | 3.06       |
|           | CC-156: O-29        | Asn¹⁵²: Oδ         | 2.65       |
Chapitre 2 Crystal structures of human 17β-hydroxysteroid dehydrogenase type 1 complexed with the dual-site inhibitor EM-139

2.1 Résumé

2.2 Abstract
Human 17β-hydroxysteroid dehydrogenase type 1 (17β-HSD1) catalyzes the biosynthesis of 17β-estradiol (E2) from estrone (E1), playing a pivotal role in the progression of estrogen-dependent diseases. N-n-Butyl-N-methyl-11-(16’α-chloro-3’,17’β-dihydroxyestra-1’,3’,5’(10’)-tri-en-7’α-yl)undecanamide (EM-139) was previously described as a dual-site inhibitor that can inhibit estrogen receptor as well as 17β-HSD1 activity. In the present study, we report the crystal structure of the 17β-HSD1-EM-139 binary complex. Interestingly, the EM-139 binary complex crystal grown under similar condition as native crystal has a space group of I121 never observed in other 17β-HSD1 crystals before. The structural analysis showed that the steroidal moiety of the bound EM-139 molecule has a binding pattern similar to E2 in the E2 binary complex, with the O-3 of the inhibitor interacts with residues His221 and Glu282, and the O-17 of the inhibitor makes hydrogen bonds with Ser142 and Tyr155. As for the long 7α-alkyl moiety of the inhibitor, which is essential for its anti-estrogenic activity, may compromise the inhibitory effect of the inhibitor to 17β-HSD1. Moreover, no obvious interaction is observed between the 16α-Cl atom and the surrounding residues. The atomic level understanding of the inhibitory mechanism of EM-139 provides important information for the inhibitor design of 17β-HSD1, which will facilitate future development of more potent and selective inhibitors of the enzyme for clinical purposes.
2.3 Introduction

Seventeen β-hydroxysteroid dehydrogenase type 1 (17β-HSD1, EC. 1.1.1.62) catalyzes the NAD(P)H dependent conversion of estrone (E1) to the most potent estrogen, 17β-estradiol (E2). E2 is well known to play a crucial role in the progression and development of several estrogen-dependent diseases (EDD). Increased E2 levels as well as up-regulated 17β-HSD1 expression indicate the involvement of the enzyme in EDDs, such as breast cancer\(^2-3\), endometrial cancer\(^4-5\), endometriosis\(^6-8\), and ovarian cancer\(^9\). Moreover, patients with tumors that have high mRNA levels of 17β-HSD1 have significantly shortened disease-free and overall survival\(^10-12\). Therefore, blocking the production of E2 through the specific inhibition of 17β-HSD1 activity is considered to be of therapeutic benefit in the treatment of EDDs.

Over the past decades, major efforts from many different laboratories have been devoted to developing highly selective inhibitors of the key steroidogenic enzyme 17β-HSD1, yielding several lead compounds with significant inhibitory activity\(^13-14\). However, due to the lack of specificity, especially for the presence of undesired estrogenic activity, no inhibitor has yet reached the stage of clinical trials\(^15-18\). N-n-Butyl-N-methyl-lI-(16α-chloro-3',17β-dihydroxyestra-1',3',5(10')-tien-7α-yl) undecanamide (EM-139) is a 7α-alkyl, 16α-halo estradiol derivative which was first synthesized as a pure antiestrogen (Figure 2.1)\(^19\). Following experiments demonstrated its inhibitory effect on 17β-HSD1 activity with a K\(_i\) of 6 μM\(^20\). Thus the compound was defined as a dual-site inhibitor which possesses inhibitory effect on estrogen receptor and on the estrogen formation\(^21\). Although this compound was proven to be a non-selective inhibitor of the 17β-HSD family members\(^22\), study of the EM-139/17β-HSD1 complex structure should help us to better understand the inhibitory mechanism of the dual-site inhibitor, thus facilitating further inhibitor design of the enzyme.

Previously, we have reported the crystallization of the 17β-HSD1/EM-139 complex using both co-crystallization and soaking methods\(^23\). The crystals obtained were isomorphous to the native crystals with a monoclinic space group C\(_2\)\(^2\). After careful analysis of the structures, the inhibitor couldn’t be identified at the binding site of the enzyme due to poor electron density. In the present study, we optimized the co-crystallization procedure and successfully obtained complex crystals with a unique space group never observed in 17β-HSD1 complexes before. The clear electron density at the binding site indicated the presence of the dual-site inhibitor in the enzyme complex.

2.4 Materials and Methods

Protein Preparation and Co-Crystallization

The 17β-HSD1 enzyme was expressed in Sf9 insect cells and purified as described previously\(^24\). After the measurement of specific activity\(^25\), the purified enzyme was concentrated to a final concentration of 15 mg/ml
in the presence of 0.06% β-octyl glucoside (β-OG), and then subjected to a buffer change procedure via centrificon (Emdmillipore, USA) to saturate the enzyme with the inhibitor EM-139. The co-crystallization experiment was carried out using the vapor diffusion method at room temperature. Crystals were obtained under conditions containing 22% - 26% (w/v) polyethylene glycol (PEG) 4000, 0.15 M magnesium chloride, and 0.1 M HEPES buffered to pH 7.5.

Data Collection and Structure Determination

The X-ray diffraction data of the 17β-HSD1-EM-139 crystals were collected at 100 K using synchrotron radiation at Advanced Photon Source (APS) beamline 31-LRL-CAT (Chicago, USA) equipped with a MAR CCD 165 mm detector at a wavelength of 0.9793 Å. The dataset was indexed and intergraded using MOSFLM, and scaled with SCALA from the CCP4 suite. The structure was solved by molecular replacement with Molrep using a reported 17β-HSD1 coordinate (PDB code 1JTV) as search model. The initial model was subjected to multiple rounds of auto-refinement using Refmac and manual rebuild using Coot. Missing portions of the models, inhibitor EM-139, glycerol, polyethylene glycol, and water molecules were progressively added with great caution during the refinement procedure. The final model was verified with PROCHECK. Molecular graphics were presented using the Pymol software (version 2.0 Schrödinger, LLC).

2.5 Results

Crystal utilized in this study belonged to space group I121 and each asymmetric unit contained a dimer, which is known to be the functional unit of the enzyme. The complex structure was refined at 2 Å with good stereochemistries, and the quality of the final model can be assessed in Table 2.1. Similar to most previously reported 17β-HSD1 structures, the highly flexible βFGaG-loop (amino acids Phe192 to Leu197) as well as the C-terminal end of the protein (amino acids 286 to 327) cannot be defined in the electron density (Figure 2.2). This high flexibility of the inhibitor is in accordance with its relatively low affinity for the enzyme.

In the binary complex structure, EM-139 has definable electron density in the A subunit of the dimeric enzyme. However, the ligand density in the B subunit is poorly defined, similar to previously described complex with equilin. Accordingly the ligand was not included in the B subunit of the final model. Even for the A subunit, only the steroid moiety of EM-139 can be defined but with a high average B-factor (97.5 Å²), whereas the 7α-alkyl side chain of the inhibitor cannot be defined in the electron density (Figure 2.3). This high flexibility of the inhibitor is in accordance with its relatively low affinity for the enzyme.
2.6 Discussion

The space group of 17β-HSD1 crystals can be affected by the presence of cations in the crystallization conditions\textsuperscript{41}. The space group of crystals obtained in the presence of Mg\textsuperscript{2+} and Mn\textsuperscript{2+} belong to C2, whereas crystals grown under conditions with Li\textsuperscript{+} and Na\textsuperscript{+} had a space group of P2\textsubscript{1}2\textsubscript{1}2\textsubscript{1}. Despite the presence of Mg\textsuperscript{2+}, the space group of the co-crystallized EM-139 complex crystals has been changed to I\textbar{}21, not observed in any other reported 17β-HSD1 structures. The change in space group may be due to the long alkyl side chain at the C7 of EM-139, which may affect the packing during crystal growth.

When the EM-139 binary and E2 binary (PDB ID 1IOL\textsuperscript{37}) complexes as well as the apo structure of 17β-HSD1 (PDB ID 1BHS\textsuperscript{42}) are superimposed, a similar conformation is observed at the steroid binding site of the enzyme (Figure 2.4A). The root-mean-square deviation (RMSD) for all paired Cα atoms obtained between EM-139 complex and apo structure is 0.456 Å, similar to the value obtained between EM-139 and E2 complexes (0.508 Å). It is worth mentioning that the position observed for the steroidal moiety of EM-139 has roughly 9˚ rotation around the axis at the C-3 atom and perpendicular to its β-face, when compared with the position of E2. This leads to the shifting of the O-17 by 1.4 Å as compared with the position of its counterpart in the E2 complex (Figure 2.4B,C). As a result, the bifurcated hydrogen bonds between the O-17 of EM-139 with Ser\textsuperscript{142} and Tyr\textsuperscript{155} (3.5 and 3.2 Å, respectively) are established, although the bond distances differ from their counterparts observed in the E2 complex (3.1 and 3.5 Å, respectively)\textsuperscript{37}. Moreover, the bifurcated hydrogen bond between the 3-hydroxyl group of EM-139 with His\textsuperscript{221} and Glu\textsuperscript{282} (3.2 and 3.5 Å, respectively) at the recognition end of the steroid binding cleft is conserved. Although much weaker as compared to their counterparts in the E2 complex (3.1 and 2.7 Å, respectively)\textsuperscript{37}, these hydrogen bonds are essential for stabilizing the inhibitor in the steroid binding cavity together with the hydrogen bonds at the O-3 of EM-139.

The 7α-alkyl moiety of EM-139 is facing toward the outside of the steroid binding cavity which is apparently accommodated by the βFαG'-loop. However, both the 7α-alkyl side chain of the inhibitor and the βFαG'-loop of the enzyme are unable to be defined by electron density due to their high degree of flexibility. This bulky 7α-alkyl side chain is essential for the inhibitor to possess anti-estrogenic activity\textsuperscript{43}. It is also safe to conclude that the α conformation of the C-7 is essential for this compound to be able to bind with 17β-HSD1. Similar results can also be observed at the C-16 of the inhibitor where a 16β halogen atom may have steric hindrance with Tyr\textsuperscript{155}. However, no obvious interaction is observed between the 16α-Cl atom and surrounding residues (Figure 2.4B,C).

2.7 Conclusion

The present work was aimed at investigating the molecular basis of the inhibitory mechanism of the dual-site inhibitor EM-139 in 17β-HSD1. We successfully co-crystallized and solved the crystal structure of 17β-HSD1 in
complex with the inhibitor. Through comparative analyses of EM-139 binary complexes and previously reported E2 binary complex as well as the apo structure, we observed a similar binding pattern of the inhibitor to the enzyme. The bifurcated hydrogen bonds between the O-3 of the inhibitor and the recognition end (His\textsuperscript{221} and Glu\textsuperscript{282}) of the binding site as well as the O-17 of the inhibitor and the catalytic end (Ser\textsuperscript{142} and Tyr\textsuperscript{155}) of the binding site are critical in stabilizing the bound inhibitor molecule. However, the introduction of a bulky side chain at the C-7 of the steroid core, which contributes to the anti-estrogenic activity of the dual-site inhibitor, may negatively affect the binding of inhibitor to 17β-HSD1. These results will contribute to the design of more potent and selective inhibitors of 17β-HSD1 for clinical purposes.

Acknowledgements

We thank Dre. M. Zhou for her technical support in enzyme purification and crystal growth. We acknowledge the use of beamline 31-LRL-CAT at the Advanced Photon Source, Argonne National Laboratory, USA.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.
2.8 Reference


**Figure 2.1.** Structure of dual-site inhibitor N-n-Butyl-N-methyl-ll-(16'α-chloro-3',17'β-dihydroxyestra-1',3',5'(10')-trien-7'α-yl)undecanamide (EM-139).
Figure 2.2. Stereo representation of the overall structure of A subunit of 17β-HSD1-EM-139 complex. The protein molecule is shown in cartoon and colored in pink. The bound EM-139 molecule is depicted as stick and colored in blue. The N-terminus and the C-terminus of the protein molecule are indicated. Segment of residues 190-197, which unable to be defined in the electron density, is represented as dash line.
Figure 2.3. Front and side views of the electron density of EM-139 in the 17β-HSD1-EM-139 complex structure. EM-139 (ligand ID EM9) was shown in the omit Fo-Fc and 2Fo-Fc electron density. 2Fo-Fc map draw in gray and contoured at 0.8σ; Fo-Fc map draw in green and contoured at 1.5σ. The occupancy of the inhibitor was refined to 1. No significant negative density features were present in the region of binding site.
Figure 2.4. Superposition of A subunit of EM-139 (EM9) binary complex (pink) and E2 binary complex (cyan) along with 17β-HSD1 apo structure (orange), showing the steroid ligand binding sites. (a) General view of the active sites within the A subunit of EM-139 and E2 complex structures as well as the apo structure; (b) Top and (c) side view of the steroid binding sites in the superposed structures. Residues Ser^{142}, Leu^{149}, Tyr^{155}, His^{221}, and Glu^{282} are labeled and shown in sticks. Hydrogen bonds between EM-139 and surrounding residues are drawn in green dash lines and labeled. Chloride atom is colored in green.
### Table 2.1. Data collection and refinement statistics

<table>
<thead>
<tr>
<th>Parameter</th>
<th>17β-HSD1-EM-139</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data Collection</strong></td>
<td></td>
</tr>
<tr>
<td>Space group</td>
<td>I121</td>
</tr>
<tr>
<td>Unit cell</td>
<td></td>
</tr>
<tr>
<td>a,b,c (Å)</td>
<td>120.76, 42.19, 122.67</td>
</tr>
<tr>
<td>α,β,γ (°)</td>
<td>90, 102.07, 90</td>
</tr>
<tr>
<td>Resolution range (Å)</td>
<td>35.67-2.00 (2.11-2.00)(^a)</td>
</tr>
<tr>
<td>Number of reflections</td>
<td>138222 (20278)</td>
</tr>
<tr>
<td>Unique reflections</td>
<td>38108 (5536)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>92.7 (92.5)</td>
</tr>
<tr>
<td>I/σ(I)</td>
<td>8.5 (2.8)</td>
</tr>
<tr>
<td>Rmerge(^b)</td>
<td>0.086 (0.315)</td>
</tr>
<tr>
<td>Multiplicity</td>
<td>3.6 (3.7)</td>
</tr>
<tr>
<td>Wilson B-factor (Å(^2))</td>
<td>30.5</td>
</tr>
<tr>
<td><strong>Refinement</strong></td>
<td></td>
</tr>
<tr>
<td>R-work(^c)</td>
<td>0.20</td>
</tr>
<tr>
<td>R-free(^d)</td>
<td>0.24</td>
</tr>
<tr>
<td>r.m.s.d</td>
<td></td>
</tr>
<tr>
<td>Bond lengths (Å)</td>
<td>0.010</td>
</tr>
<tr>
<td>Bond angles (°)</td>
<td>1.483</td>
</tr>
<tr>
<td>Ramachandran plot(^e) (%)</td>
<td></td>
</tr>
<tr>
<td>Most favored regions</td>
<td>94.4</td>
</tr>
<tr>
<td>Additional allowed regions</td>
<td>5.6</td>
</tr>
<tr>
<td>Generously allowed regions</td>
<td>0.0</td>
</tr>
<tr>
<td>Disallowed regions</td>
<td>0.0</td>
</tr>
<tr>
<td>Average B, all atoms (Å(^2))</td>
<td>42.0</td>
</tr>
<tr>
<td>PDB ID</td>
<td>6DTP</td>
</tr>
</tbody>
</table>

\(^a\) Data statistics for the outer shell are given in parentheses.

\(^b\) The redundancy-independent \(R_{merge}/R_{sym}\), \(R_{means} = \sum_{hkl} \left( \frac{n}{n-1} \sum_{i=1}^{n} |I_{hkl,i} - \langle I_{hkl} \rangle| \right) / \sum_{hkl} \sum_{i} I_{hkl,i} \)

\(^c\) \(R_{work} = \sum_{hkl} \left| F_{obs}(hkl) \right| - \left| F_{calc}(hkl) \right| / \sum_{hkl} \left| F_{obs}(hkl) \right| \)

\(^d\) \(R_{free} = \) the cross-validation \(R\) factor for 5% of reflections against which the model was not refined.

\(^e\) Calculated with PROCHECK.
Chapitre 3 Crystal structures of human 17β-hydroxysteroid dehydrogenase type 1 complexed with estrone and cofactor reveal the mechanism of substrate inhibition

3.1 Résumé
La 17β-HSD1 catalyse la dernière étape de la bioactivation de l’estradiol, l’œstrogène le plus puissant et est également capable de convertir la dihydrotestostérone en 3β, 17β-androstanediol par le biais de son activité 3β-hydroxystéroïde déshydrogénase. À la différence des autres membres des 17β-HSDs, la 17β-HSD1 subit une inhibition induite par le substrat que nous avons récemment rapportée. Afin d’élucider les bases moléculaires de l’inhibition du substrat, on a résolu les structures cristallines binares et ternaires de la 17β-HSD1 en complexe avec l’estrone et l’analoque du cofacteur, le NADP+, qui fournissent une image complète des interactions enzyme-substrat-cofacteur. Ces structures complexes ont révélé un mode de liaison inversé de l’œstrone dans la 17β-HSD1 jamais trouvé dans les complexes d’estradiol. Cela conduit à la formation d’un complexe sans issue, similaire au mécanisme d’inhibition du substrat décrit dans la 5β-réductase, l’aldéhyde déshydrogénase et la déhydroépiandrostérone sulfotransférase.

3.2 Abstract
Human type 1 17β-hydroxysteroid dehydrogenase (17β-HSD1) catalyzes the last step in the bioactivation of the most potent estrogen estradiol, and is also able to convert dihydrotestosterone into 3β,17β-androstanediol through its 3β-hydroxysteroid dehydrogenase activity. Unlike in other member of 17β-HSDs, 17β-HSD1 undergoes a substrate induced inhibition that we have recently reported. In order to elucidate the molecular basis of the substrate inhibition, here we solved the binary and ternary crystal structures of 17β-HSD1 in complex with estrone and cofactor analog NADP+ that provide a complete picture of enzyme-substrate-cofactor interactions. These complex structures revealed a reversely orientated binding mode of estrone in 17β-HSD1, never found in estradiol complexes. This leads to the formation of a dead-end complex, similar as the substrate inhibition mechanism described in 5β-reductase, aldehyde dehydrogenase, and dehydroepiandrosterone sulfotransferase. Structural comparison with 17β-HSD1-estradiol/testosterone binary complexes confirmed that residue His²²¹ is responsible for the recognizing and stabilizing the reversely bound estrone, leading to the formation of dead-end complex. Thus, the overall catalytic activity of 17β-HSD1 is modulated through its substrate inhibition, indicating a simple mechanism for regulation of enzyme activity in physiological background, which may be used more widely across this family of enzymes.
3.3 Introduction

Estradiol (E2) is well known to play an important role in promoting the genesis and development of estrogen-dependent diseases such as breast cancer, endometrial cancer, endometriosis and ovarian cancer \(^1\)\(^-\)\(^4\). Under normal circumstances E2 is acquired from the circulating plasma. However, for postmenopausal women, ovarian-derived estrogens are withdrawn and replaced by estrogens synthesized in an intracrine manner\(^5\). The E2 concentration is significantly higher in malignant breast tissues than in plasma levels in postmenopausal women\(^6\). Human 17β-hydroxysteroid dehydrogenase type 1 (17β-HSD1, EC. 1.1.1.62) catalyzes the conversion of an inactive estrogen, estrone (E1), into the biologically active estrogen, E2, in living cells \(^7\)\(^-\)\(^8\). It is also involved in the reduction of dehydroepiandrosterone (DHEA) into 5β-androstane-3β,17β-diol (A-diol), and dihydrotestosterone (DHT) into 5α-androstane-3β,17β-diol (3β-diol) \(^9\)\(^-\)\(^10\). A-diol has been proposed to be the major estrogens present after menopause \(^11\)\(^-\)\(^12\), whereas 3β-diol was able to induce estrogen receptor (ER) \(\alpha\) activation and proliferation \(^13\). Therefore, inhibiting the 17β-HSD1 activity is a promising approach for the treatment of estrogen-dependent diseases.

17β-HSD1 is a membrane-associated protein whose first structure was solved in this laboratory \(^14\)\(^-\)\(^15\). It requires the presence of a dinucleotide cofactor (NADP\(^+\)/NADPH or NAD+/NADH) during the conversion of estrogens. It uses NAD(H) and NADP(H) as cofactors \(in \textit{vitro}\)\(^16\); however, only NADPH was subsequently confirmed to be used by the enzyme as a cofactor during the reduction of E1 in cells and \(in \textit{vivo}\) \(^17\). The cofactor binding site of the enzyme molecule involves βA to βF and forms a typical Rossmann fold, while the substrate binding site involving βD to βG only partially belongs to the Rossmann fold \(^18\)\(^-\)\(^19\). The mechanism for estrogen recognition as well as androgen discrimination by 17β-HSD1 was previously studied using the crystallographic structure of the enzyme in complex with E2, DHT and testosterone (T) \(^20\)\(^-\)\(^21\). The E2 complex showed that the hydrogen bonds and van der Waals interactions were major contributors to the binding energy, indicating that these non-bond interactions help to orientate and stabilize E1 in such a way that the carbonyl group is close to the catalytic triad (Ser142, Tyr155 and Lys159) of the enzyme and undergoes reduction to a 17β-hydroxyl group \(^20\). The 17β-HSD1-C19-steroid complexes further demonstrated the role of residue Leu\(^149\) in discriminating the binding of C19-steroids to the enzyme \(^10\)\(^,\)\(^21\).

Substrate inhibition has previously been reported in many enzymes, such as 5β-reductase (AKR1D1) \(^22\)\(^-\)\(^23\), aldehyde dehydrogenase (ALDH) \(^24\), dehydroepiandrosterone sulfotransferase (SULT2A1) \(^25\), indoleamine 2,3-dioxygenase (IDO) \(^26\), lactate dehydrogenase \(^27\), trimethylamine dehydrogenase \(^28\), etc. Their enzymatic activities can be inhibited by their own substrate, which causes the reaction velocity curve to rise to the maximum as substrate concentration increases and then decreases to zero or to a non-zero asymptote \(^29\). The substrate inhibition mechanism of an enzyme which has a coenzyme can generally be divided into three major groups: one is the formation of a dead-end complex resulted from the nonproductive binding of the substrate
molecule; another is a reversed binding order of substrates which leads to a slowing down of the reaction; the third is a limited dissociation rate of cofactor. Substrate inhibition resulting in a dead-end complex of 17β-HSD1-E1 was first proposed by our group, illustrating its important role in enzyme activity regulation in cells. In order to better understand the peculiarities of 17β-HSD1 in terms of substrate binding and stabilization as well as the molecular basis of substrate inhibition, we crystallized this enzyme in complex with E1 and NADP+. The dead-end complex of reversely bound E1 inside the substrate-binding site of the 17β-HSD1 revealed the crucial information on the mechanism by which steroid substrate can influence the activity of this enzyme. Moreover, these complex structures confirmed the role of His221 in the substrate inhibition mechanism of the enzyme.

3.4 Materials and Methods

Protein preparation and co-crystallization.

The 17β-HSD1 enzyme was expressed in Sf9 insect cells and purified by a procedure comprising three chromatographic steps: Q-Sepharose anion exchange, Blue-Sepharose affinity, and phenyl-Superose hydrophobic interaction columns, as described by Zhu et al. The purified enzyme was then subjected to a buffer change procedure via centrifugation (Emdmillipore, USA) to saturate the enzyme with E1 for binary complex, and NADP+ was added to a final concentration of 1 mM to generate ternary complexes. These binary and ternary complexes were then concentrated to 18–20 mg/ml and used for crystal growth. Crystals were obtained at 27°C using the hanging-drop vapor diffusion method. The 17β-HSD1-E1 binary crystals were grown under conditions containing 26% (w/v) PEG 3350, 150 mM magnesium chloride, 20% glycerol and 100 mM HEPES; while the ternary complex crystals were grown under conditions of 24–29% (w/v) PEG 8K, 100 mM Tris buffer (pH 7.5–7.8) and 50 mM KH₂PO₄.

Data collection and structure determination.

Diffraction data of the 17β-HSD1-E1 binary crystal were collected using an R-AXIS llc image plate area detector and Rigaku RU300 rotating anode generator at 298 K with a wavelength of 1.5418 Å. The 17β-HSD1-E1-NADP+ ternary crystal diffraction data were collected using synchrotron radiation at Canadian Light Source (CLS) beamline 08B1-1(Saskatoon, Canada) equipped with a RAYONIX MX300HE CCD detector at 100 K using a wavelength of 0.9795 Å. Mineral oil was used as the cryoprotectant for all crystals. The datasets were intergraded using MOSFLM and scaled with SCALA from the CCP4 suite. The structures were solved by molecular replacement with Molrep using the coordinates of 17β-HSD1, with the highest resolution (PDB ID: 1JTV) as search model. The initial models issued from rigid body refinements were subjected to multiple rounds of refinement using Refmac and manual rebuild using Coot. After the E1 and NADP+
being added, models were further refined by isotropic B-factor refinement (restrained, individual B-factor refinement) and corrected by manual rebuilding. Missing portions of the models, glycerol, polyethylene glycol, and water molecules were progressively added during the refinement procedure. The final model was verified with \textit{PROCHECK} \cite{41}. Final statistics for all the refined structures are summarized in \textbf{Table 3.1}. Molecular graphics were derived using the \textit{Pymol} (version 2.0 Schrödinger, LLC). A plot showing the interaction between NADP$^+$ and surrounding residues was prepared using the \textit{LigPlot+} version 1.4 program \cite{42}.

\textit{In silico} studies.

The manual compound edition was performed using \textit{SeeSAR} \cite{43} software. The pseudo E1 complex was build based on E2 in E2 complex (PDB ID 1IOL\cite{20}). Whereas the new compound designed in light of substrate inhibition mechanism was built from reversely bound E1 in the E1 ternary complex (PDB ID 6BBC). Ten binding poses of steroid ligands were generated with \textit{SeeSAR}, whose geometry were optimized by the Hydrogen bond and Dehydration (HYDE) \cite{44} as implemented in \textit{SeeSAR}, and ranked according to their estimated affinity according to the HYDE affinity assessment.

The docking studies were carried out using \textit{Gold} software \cite{45}. The 3D structure of proteins (6BBC for 17βHSD1 and 1ERE for estrogen receptor α ligand binding domain) were taken from PDB. The ligands for docking studies were prepared in \textit{OpenBabel} (http://openbabel.org/wiki/Main_Page) and the energy minimization was carried out in \textit{Avogadro} \cite{46}. Genetic algorithm \textit{Gold} \cite{47} was used in docking studies and the ChemPLP scoring was used for ranking the binding poses.

\section*{3.5 Results}

Previously reported substrate inhibition in 5β-reductase, aldehyde dehydrogenase and dehydroepiandrosterone sulfotransferase revealed an alternative binding mode of the substrate, which resulted in the formation of dead-end complex \cite{22,24-25}. Particularly, a possibility for E1 to adopt an alternative conformation in the binding site was observed in several 17β-HSD1-C-19 steroid complexes \cite{10,21,48}. Thus, we have employed molecular docking to investigate the different binding modes of E1 in 17β-HSD1. We manually built the pseudo E1 complex structure from the previous reported E2 complex (PDB ID 1IOL \cite{20}) using the \textit{SeeSAR} \cite{43} software. The top two poses of E1 having similar calculated binding affinities in the Hydrogen bond and Dehydration (HYDE) \cite{44} assessment are presented in \textbf{Figure 3.1}. Interestingly, the two poses have very different steroid orientations. The first pose of E1 has a normal oriented conformation with its 17-ketone group close to the catalytic triad. The hydrogen bonds between the 17-ketone group and residues Tyr$^{155}$ and Ser$^{142}$ were maintained (\textbf{Figure 3.1A}). The second pose of E1 is almost reversely oriented with the 17-ketone group close to His$^{221}$ whereas the O-3 hydroxyl group facing toward residues Ser$^{142}$ and Cys$^{185}$, which were stabilized
by the hydrogen bonds between them (Figure 3.1B). This \textit{in silico} analysis urged us to engage co-crystallization experiment for 17β-HSD1-E1 to further clarify the mechanism.

**Overall Structure and Model Quality**

Crystals utilized in this study belonged to the space group P2\textsubscript{1}2\textsubscript{1}2\textsubscript{1}, and contained a dimer per asymmetric unit, similar to previously described ternary complexes with equilin \textsuperscript{49}. The 17β-HSD1-E1-NADP\textsuperscript{+} ternary complex was refined at 1.86 Å, whereas the 17β-HSD1-E1 binary complex was refined at 2.4 Å. The two models show good stereochemistry \textsuperscript{50} and the quality of the final models is demonstrated in Table 3.1. Similar to the most previously reported 17β-HSD1 complex structures, no clear electron density was present either for the highly flexible βFoG'-loop (amino acids Ala\textsuperscript{191} to Gly\textsuperscript{198}) or for the C-terminal end of the protein (amino acids 286 to 327) \textsuperscript{10, 18, 20, 48}.

**Crystal structure of 17β-HSD1 in complex with E1**

In order to understand the interactions between 17β-HSD1 and E1, and also the mechanism of the substrate inhibition, we first solved the crystal structure of 17β-HSD1 in complex with E1. In this binary complex, E1 has a well-defined electron density in the B subunit of the dimer (Figure 3.2A), whereas the ligand density in the A subunit is poorly defined. The presence of only a few disconnected density peaks in this region led to the conclusion that the ligand is disordered in this subunit in the crystal. Thus, the ligand was not included in the A subunit. The protein portions of the two subunits are almost identical, with a root mean-square deviation (RMSD) of 0.59 Å for the Cα of 276 amino acids. However, it is worth mentioning that residues Phe\textsuperscript{226} and Phe\textsuperscript{259} in the ligand-binding pocket of the two subunits exhibit significant differences (Figure 3.3A). The side chains of residues Phe\textsuperscript{226} and Phe\textsuperscript{259} in B subunit face toward E1, forming a “closed” conformation to favor the van der Waals contacts with the ligand. In contrast, Phe\textsuperscript{226} and Phe\textsuperscript{259} in the A subunit adopt an “opened” conformation with their side chains rotating about 60º and 100º respectively.

Interestingly, we found that E1 is bound in a very different manner to 17β-HSD1 compared with the binding mode previously described for E2 and other steroids \textsuperscript{18, 20, 49, 51}, but is similar to the mode described for C19-steroid complexes \textsuperscript{10, 21, 45}. It is reverse-orientated in the substrate binding site with its A-ring facing toward the catalytic site while its D-ring faces the recognition end (His\textsuperscript{221}, Glu\textsuperscript{282}) of the binding site (Figure 3.3A). This is similar to the second pose of E1 in the pseudo E1 complex described above with a roughly 32º rotation around the axis perpendicular to its β-face. Obviously, the reverse-orientated E1 cannot be catalyzed by 17β-HSD1, suggesting that a potential dead-end complex likely accounts for its substrate inhibition. Moreover, the residue Leu\textsuperscript{149}, which is responsible for the reverse binding of C19-steroids \textsuperscript{21, 48}, is not likely to be involved in the reverse binding of E1 since the closest distance between the Cδ of Leu\textsuperscript{149} and C18 of E1 is more than 4 Å.
However, the presence of NADP(H) may have significant influence on the binding mode of E1 considering that NADP+ can significantly increase the affinity of 17β-HSD1 to E1 with a $K_D$ of 1.6 ± 0.2 μM. Thus we further co-crystallized E1 ternary complex with NADP+, the product of cofactor NADPH, used as an analogue of the cofactor.

**Crystal structure of 17β-HSD1 in the presence of NADP+ and E1**

The electron density in the substrate binding pocket of 17β-HSD1-E1-NADP+ ternary complex clearly indicates the presence of E1 in both subunit of the dimeric protein (Figure 3.2B). However, similar to previously reported 17β-HSD1-A-dione-NADP and 17β-HSD1-4-dione-NADP complex structures, only the ADP moiety and the 2'-phosphate group of the adenine ribose can be unambiguously defined for the bound NADP+ molecule (Figure 3.2C). The nicotinamide and the attached ribose of the NADP+ molecules are poorly defined in the electron densities in both subunits and thus omitted from the final model. It indicates that the major interactions between NADP+ and the enzyme happen at the ADP part in agreement with previous structure-function study.

The overall conformations of ADP moiety and surrounding residues are almost identical in the two subunits of the ternary complex and similar to those in the previously described 17β-HSD1 complexes. The adenine ring adopts an anti conformation, stabilized through the hydrogen bonds established with Asp65 and Val66. The important hydrogen bond interactions between the ribose and pyrophosphate groups and the residues Ser11, Ser12, and Ile14 are conserved (Figure 3.4). However, it is worth mentioning that noticeable difference exists at the binding of the 2'-phosphate group of NADP+ in the A and B subunits. In the B subunit, the 2'-phosphate of NADP+ is stabilized by a salt bridge formed with side chain of Arg37 (Figure 3.4B), similar to what was seen in several NADP+ complexes reported previously. Despite an identical 2'-phosphate in the A and B subunits, there is no direct interaction between the Arg37 and the 2'-phosphate group (Figure 3.4A) in subunit A, a phenomenon observed in the complexes of 17β-HSD1 with C19-steroids such as 5α-Androstan-3,17-dione and 4-Androstene-3,17-dione.

As for the bound substrate, E1 in both subunits adopts the same reverse binding mode similar as described in the E1 binary complex (Figure 3.3B). Although E1 in the A subunit of the ternary complex can be unambiguously defined in the electron density, it shows higher mobility (average B-factor 61.9 Å$^2$) compared to its counterpart (average B-factor 41.2 Å$^2$) in the B subunit. Some residues in the substrate binding site are differently oriented in these two subunits (Figure 3.3B). Phe226 in the B subunit adopts a “closed” conformation, whereas it has an “opened” conformation in the A subunit, similar to what is seen in the E1 binary complex. As a result, the space around the A-ring of E1 is less compact in the A subunit, and E1 shifts away from Tyr155. Thus, no hydrogen bonds can form with surrounding residues at the O3 end of E1 (Figure 3.3B).
**17β-HSD1 inhibitor design based on substrate inhibition mechanism**

An inhibitor design, in the light of reversible binding of E1, was conducted by manual editing using SeeSAR. The complex structures of 17β-HSD1 with inhibitor CC-156, 2-MeO-CC-156 and PBRM have showed a space in the active site which was not occupied by the native substrates and suitable to accommodate an extra benzylamide ring. Thus we added a benzylamide ring moiety at the O-3 of E1 to form a novel compound 3-(((8R,9S,13S,14S)-13-methyl-17-oxo-7,8,9,11,12,13,14,15,16,17-decahydro-6H-cyclopenta[a]phenanthren-3-yl)oxy) benzamide (SX7) (Figure 3.5). After structure optimization and poses generation, the pose with the highest estimated binding affinity (43nM) by HYDE is presented below (Figure 3.6). As a comparison, the estimated affinity of the best pose of CC-156 in CC-156 ternary complex calculated by HYDE is 31 nM, whereas its IC₅₀ obtained by experiments is 44 nM. The added benzylamide ring moiety established a hydrogen bond network with residues Leu and Gln, similar to that observed in the CC-156 ternary complex. The energetically favorable edge-to-face π-π interaction between the benzylamide ring of the inhibitor and the phenol ring of Tyr was also formed with a distance of 4.4 Å, similar to that in the CC-156 binary complex (4.3 Å) but longer than in the CC-156 ternary complex (3.8 Å). Moreover, a π-donor hydrogen bond contributing to the stabilization of the local 3D structures was also observed between the benzylamide ring of the inhibitor and the OH group of Tyr with a distance of 3.5 Å. As for the recognition end of the binding site, a hydrogen bond was formed between the inhibitor and His, similar to that observed in the E1 binary and ternary complexes.

**3.6 Discussion**

The present work was carried out to investigate the molecular basis of the substrate inhibition observed in 17β-HSD1. We solved the crystal structures of 17β-HSD1 in complex with E1 as well as the ternary complexes with NADP⁺. Interestingly, E1 in both binary and ternary complexes adopt the same reverse binding mode. This binding orientation is judged to be nonproductive. The reversely oriented E1 acts as a competitive inhibitor and the presence of a phenolic hydroxyl group of E1 at the catalytic triad of the enzyme can be the key for the formation of a non-catalytic dead-end complex. The substrate binding pocket in 17β-HSD1 is narrow and deep (Figure 3.7), which is similar to that of dehydroepiandrosterone sulfotransferase but differs from 5β-reductase which has a relatively large steroid binding pocket. Besides, in both the 17β-HSD1-E1 binary and ternary complexes, the bound E1 adopted a reverse orientation. The well-defined electron density for the reversely bound E1 and the resulted B factors after refinement indicate almost no normally bound E1 existed in the crystal structures. This may be due to the relatively high concentration of E1 (>600μM) with multiple cycles of buffer exchange via centricon to saturate the enzyme with E1 used in co-crystallization experiments. Moreover, the Kᵢ (1.3 μM) of E1 to 17β-HSD1 is similar to the apparent Kᵢ (1.6 μM) of the steroid to the enzyme, suggested that the affinity between 17β-HSD1 and normally oriented E1 is lower than that of the
reversely oriented one. The reverse binding mode is relatively energy favourable compared to the normal oriented one in the steroid binding pocket, which is strengthened by the presence of NADP⁺.

Further superposing E1, E2 ²⁰ and T ²¹ in the 17β-HSD1 complex crystal structures, varying conformations are observed (Figure 3.8). In general, the position observed for E1 is roughly in 180° rotation around the axis perpendicular to its β-face as compared with the position of E2. The orientation of T is similar to that of E1 with an approximate 26° rotation around the axis perpendicular to its β-face and a rotation of 20° around its long axis (O3–O17). At the catalytic end of the substrate binding pocket in 17β-HSD1, the phenolic hydroxyl of E1 establishes a hydrogen bond with the OH group of Tyr¹⁵⁵ (3.3 Å), similar to the E2 complex structure (Figure 3.8B). However, the hydrogen bond between E1 and Tyr¹⁵⁵ is not presented in the chain A of the E1 ternary complex (Figure 3.3B), indicating that this residue does not play a critical role in the binding mode of E1. Meanwhile, at the recognition end of the substrate binding pocket, the 17-carbonyl group of E1 faces toward His²²¹, forming a strong hydrogen bond (2.9 Å) with the side chain of His²²¹ (Figure 3.8). As for the nearby Glu²⁸² residue, it does not form hydrogen bond with E1, indicating that it does not significantly contribute to the binding of the reversely oriented steroid. In contrast, the residues His²²¹ and Glu²⁸² are both involved in hydrogen bond formation with the O17 of T and O3 of E2 in their complexes (Figure 3.8). Furthermore, the mutation of His²²¹ indeed diminished the substrate inhibition of 17β-HSD1 in intact cells ⁵³. Thus it further substantiates that the His²²¹ is a key residue, responsible for the substrate inhibition of 17β-HSD1 through its binding to the reversely oriented E1 and the formation of a dead-end complex.

Moreover, the residue at position 36 plays an essential role in the discrimination of cofactor NADP(H)/NAD(H) in the SDR family ⁵³. A negatively charged residue at this position will serve to repel the 2'-phosphate of NADP(H) and accept hydrogen bonds from the 2' and 3' ribose hydroxyls ⁶¹, typically found in NAD(H)-preferring enzyme. Mutagenesis study demonstrated that the sole mutation of the Leu³⁶ into aspartic acid residue indeed changed the cofactor preference of 17β-HSD1 from NADP(H) to NAD(H) ⁵³, and eliminated the substrate inhibition of the enzyme in the presence of NADPH⁶¹. Although the (phosphor-)adenosine moiety of NADP is distal from the catalytic site of the enzyme, subtle perturbations to the (phosphor-) adenosine binding pocket was proved to have a dramatic effect on activity ⁶² and mutations at the 2'-phosphate binding site was demonstrated to affect substrate specificity ⁶³. Thus the stabilization of the 2'-phosphate group of NADPH is essential for maintaining the substrate inhibition in 17β-HSD1.

3.7 Conclusion
Taking together, the stabilization of the reversely oriented E1 requires the presence of His²²¹ at the recognition end of the substrate binding site, and the mutation of His²²¹ is sufficient to destabilize the reversely bound E1, and prevent the formation of a dead-end complex. Meanwhile, the presence of NADP⁺ may strengthen the
reverse binding mode of E1. Besides, this mechanism may play a protective role under physiological background by limiting the E2 levels upon an increase in intracellular E1 levels. Moreover, 17β-HSD1 is primarily expressed in the placenta and ovarian granulose cells 64, and the physiological E1 level in human placenta was measured to reach 1.5 µM 65, while the threshold concentrations required to exhibit substrate inhibition in both molecular level (0.2 µM) 31 and cell level (0.65 µM) 32. Therefore, it is likely that the substrate inhibition of 17β-HSD1 takes place in living cells. Furthermore, based on this dead-end complex, we employed the *in silico* method to design 17β-HSD1 inhibitor, yielding a novel compound with a high estimated binding affinity. This substrate inhibition mechanism described in 17β-HSD1 may widely exist in NADP(H)-preferred enzymes for regulation of their enzymatic activity. These results will contribute to advance the knowledge of enzyme inhibition and encourage the development of inhibitors for clinical purposes.

Notes

The authors declare no competing financial interests.

The PDB ID of 17β-HSD1-E1 and 17β-HSD1-E1-NADP⁺ are 6MNC and 6MNE, respectively.

Acknowledgements

We thank Dr. D. W. Zhu and Dre. M. Zhou for their technical support in enzyme purification. We would also like to thank Dr. Muriel Steel for her editing of the manuscript. We acknowledge the use of beamline 08B1-1 at the Canadian Light Source, Brookhaven National Laboratory.
3.8 Reference


43. SeeSAR version 7.3; BioSolveIT GmbH, Sankt Augustin, Germany, 2018.


Figures and Legends

Figure 3.1. Results from the in silico building of E1 at the binding site of 17β-HSD1. The normal (A) and reverse (B) binding poses of E1, derived in SeeSAR analysis are shown as sticks and colored in blue. The binding site conformation of E2 complex structure (PDB ID 1IOL) is represented (magenta). Residues Tyr^{155}, Ser^{142}, Cys^{185} and His^{221} are labeled and shown in sticks. Hydrogen bonds between E1 and surrounding residues are drawn in green dash lines.
Figure 3.2. Front and side views of the electron density in E1 and NADP+ of the B subunit of E1 binary (A, green) and ternary (B and C, blue) complexes. E1 and the ADP moiety of NADP+ are shown in the omit Fo-Fc electron density contoured at 2.5σ level. The positive and negative densities are drawn in gray and red, respectively.
Figure 3. Superimposition of the steroid binding site in chain-A (pink) and chain-B (blue) in 17β-HSD1-E1 (A) and 17β-HSD1-E1-NADP+ (B) complexes. Residues Val^{143}, Tyr^{155}, His^{221}, Phe^{226}, Phe^{259}, and Glu^{282} are labeled and shown in sticks. Hydrogen bonds between E1 and surrounding residues are drawn in green dash lines.
The 2’-phosphate of NADP in A subunit is stabilized through water (W533) bridged hydrogen bond with residues Thr$^{41}$ and Asp$^{38}$, whereas that in B subunit is stabilized by salt bridge with Arg$^{37}$ and water bridged hydrogen bond with residues Thr$^{41}$ and Asp$^{38}$. The NADP$^+$ and protein side chains are shown in ball-and-stick representation, with the NADP$^+$ bonds colored in purple. Hydrogen bonds are shown as green dotted lines, while the spoked arcs represent protein residues making nonbonded contacts with the NADP$^+$. Figure is prepared using the LigPlot+ version 1.4 program.
Figure 3.5. The 2D structure of modeled 17β-HSD1 inhibitor 3-(((8R,9S,13S,14S)-13-methyl-17-oxo-7,8,9,11,12,13,14,15,16,17-decahydro-6H-cyclopenta[α]phenanthren-3-yl)oxy)benzamide (SX7).
Figure 3.6. Top (A) and side (B) view of binding residues (blue sticks) and the best pose of the SX7 (pink stick) in 17β-HSD1. Hydrogen bonds between inhibitor and surrounding residues are drawn in green dash lines. Several important distances are labeled and shown in black dash line.
Figure 3.7. Surface representation of the substrate binding pocket of 17β-HSD1. The chain B of E1 binary (green) and ternary (blue) complexes are superimposed and shown in cartoon. The surface of the steroid binding pocket in E1 ternary complex is presented in side view (A) and top view (B) and colored by elements. Residues Tyr$^{155}$, His$^{221}$, and Glu$^{282}$ are labeled and shown in sticks.
Figure 3.8. Superimposition of estrone, estradiol and testosterone binding in 17β-HSD1. Side view (A) and top view (B) of the active site residues of the 17β-HSD1-E2 (magenta, PDB ID 1IOL), 17β-HSD1-T (orange, PDB ID 1JTV), and the B subunit of 17β-HSD1-E1-NADP+ (blue) complexes. The steroid molecules are colored the same as their binding residues. Hydrogen bonding interactions between steroid molecules and the enzyme residues are represented by green dash lines. Water molecule (W647 from 17β-HSD1-T) is shown as red spheres.
Table 3.1. Data collection and refinement statistics

<table>
<thead>
<tr>
<th>Parameter</th>
<th>17β-HSD1-E1</th>
<th>17β-HSD1-E1-NADP+</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data Collection</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Space group</td>
<td>P2₁2₁2₁</td>
<td>P2₁2₁2₁</td>
</tr>
<tr>
<td>Unit cell (Å)</td>
<td>43.56, 110.02, 117.14</td>
<td>43.78, 108.24, 117.67</td>
</tr>
<tr>
<td>a,b,c (Å)</td>
<td>90, 90, 90</td>
<td>90, 90, 90</td>
</tr>
<tr>
<td>Resolution range (Å)</td>
<td>50-2.4 (2.53-2.4)</td>
<td>50-1.86 (1.96-1.86)</td>
</tr>
<tr>
<td>Number of reflections</td>
<td>77771 (9146)</td>
<td>353641 (51177)</td>
</tr>
<tr>
<td>Unique reflections</td>
<td>21966 (2955)</td>
<td>47986 (6911)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>96.6 (91.7)</td>
<td>100.0 (100.0)</td>
</tr>
<tr>
<td>I/σ(I)</td>
<td>12.2 (2.3)</td>
<td>13.7 (4.5)</td>
</tr>
<tr>
<td>R&lt;sub&gt;merge&lt;/sub&gt;&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.065 (0.543)</td>
<td>0.091 (0.516)</td>
</tr>
<tr>
<td>Multiplicity</td>
<td>3.5 (3.1)</td>
<td>7.4 (7.4)</td>
</tr>
<tr>
<td>Wilson B-factor (Å²)</td>
<td>50.0</td>
<td>25.5</td>
</tr>
<tr>
<td><strong>Refinement</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-work&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.21</td>
<td>0.19</td>
</tr>
<tr>
<td>R-free&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.28</td>
<td>0.23</td>
</tr>
<tr>
<td>r.m.s.d</td>
<td>0.013</td>
<td>0.012</td>
</tr>
<tr>
<td>Bond lengths (Å)</td>
<td>1.676</td>
<td>1.722</td>
</tr>
<tr>
<td>Bond angles (°)</td>
<td>0.013</td>
<td>0.012</td>
</tr>
<tr>
<td>Ramachandran plot&lt;sup&gt;e&lt;/sup&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Most favored regions</td>
<td>91.8</td>
<td>96.7</td>
</tr>
<tr>
<td>Additional allowed regions</td>
<td>7.6</td>
<td>3.3</td>
</tr>
<tr>
<td>Generously allowed regions</td>
<td>0.4</td>
<td>0.0</td>
</tr>
<tr>
<td>Disallowed regions</td>
<td>0.2</td>
<td>0.0</td>
</tr>
<tr>
<td>Average B, all atoms (Å²)</td>
<td>57.0</td>
<td>31.0</td>
</tr>
<tr>
<td>PDB ID</td>
<td>6MNC</td>
<td>6MNE</td>
</tr>
</tbody>
</table>

<sup>a</sup> Data statistics for the outer shell are given in parentheses.

<sup>b</sup> The redundancy-independent R<sub>merge</sub>/R<sub>sym</sub> = ∑ᵢ₌₁ⁿ ||I<sub>hkl</sub;i|| - (1/∑ᵢ₌₁ⁿ I<sub>hkl</sub>) / ∑ᵢ₌₁ⁿ I<sub>hkl</sub>i

<sup>c</sup> R<sub>work</sub> = ∑ᵢ₌₁ⁿ ||F<sub>obs</sub>(hkl)| - |F<sub>calc</sub>(hkl)|| / ∑ᵢ₌₁ⁿ F<sub>obs</sub>(hkl)

<sup>d</sup> R<sub>free</sub> = the cross-validation R factor for 5% of reflections against which the model was not refined.

<sup>e</sup> Calculated with PROCHECK.
Chapitre 4 Remarkable steroid-converting enzyme and receptor regulations in large number breast tumor samples: molecular correlation and combined therapies

4.1 Résumé

4.2 Abstract
Endocrine therapy is a cornerstone against hormone-dependent breast-cancer (BC), represented by aromatase inhibitors (AIs). Despite the effectiveness of AI-treatment, resistance often occurs. Recently, the joint estradiol accumulation and dihydrotestosterone degradation stimulating BC growth has been demonstrated by reductive 17beta-hydroxysteroid dehydrogenases in vitro and in vivo, indicating an aromatase-independent pathway for estrogen-synthesis. A systematic study of the expression and correlation of steroid enzymes in clinical samples becomes critical. In the present study, the RNA sequencing dataset of The Cancer Genome Atlas Breast Invasive Carcinoma (TCGA-BRCA) cohort (N=1079) was retrieved through Genomic Data Commons (GDC) data portal, which included post-menopausal normal breast tissues (N=56) and estrogen receptor positive breast tumors (N=526). Differential expressions and gene expression correlation were analyzed by Mann–Whitney U test and Spearman’s rho test. Differential expression analysis showed significant up-regulation of reductive 17β-HSD7 (2.61-fold, p=5.57E-26) in BC, supporting its sex-hormone effect. Besides, suppression of 11β-HSD1 expression (-8.33-fold, p=1.51E-23) and elevation of 11β-HSD2 expression (2.30-fold, p=2.17E-09), provide a low glucocorticoid level environment diminishing BC anti-proliferation effects. Furthermore, 3α-HSDs were significantly down-regulated (-1.51-fold, p=0.002; -8.18-fold, p=1.63E-28; -35.07-fold, p=2.56E-29; -30.38-fold, p=5.08E-30 for type 1-4 respectively), while 5α-reductases
significantly up-regulated (1.35-fold, \(p=3.42\times10^{-5}\); 3.11-fold, \(p=1.33\times10^{-11}\); 1.68-fold, \(p=1.56\times10^{-15}\) for type 1-3 respectively) in BC compared with normal tissues, reducing cell proliferation suppressers 4-pregnenes, increasing cell proliferation stimulators 5α-pregnanes. Expression correlation analysis indicates significant correlations between 11β-HSD1 with 3α-HSD4 (\(r_s=0.55\), \(p=7.42\times10^{-41}\)). Significant expression correlations between 3α-HSDs were also observed. A 3D schema vividly presents the regulation of steroid enzymes, extensively demonstrating their roles in BC. Our strategy can also contribute to other cancers. Our results support novel therapy targeting the reductive 17β-HSD7 and the combined therapy targeting 11β-HSD2 and 17β-HSD7.

### 4.3 Introduction

Breast cancer (BC) is the most commonly diagnosed cancer in women in North America, and the second leading cause of cancer death in women\(^1\). Molecular therapies for BC have developed rapidly during the recent decades and two milestone treatments for hormone-receptor-positive BC have been achieved: the selective estrogen receptor modulators (SERMs) represented by tamoxifen and aromatase inhibitors (AIs) such as letrozole and anastrozole\(^2\). However, significant side effects have occurred in response to AI treatment and resistance was evident in approximately 37% of patients during AI therapy\(^3\). Several hypotheses have been proposed to explain the mechanism of AI resistance, including constitutive estrogen receptor (ER) α activation caused by growth factor receptor pathways\(^4\); activation of growth-signaling pathways independent of estrogens and ERα\(^5\); and aromatase-independent estrogen biosynthesis pathway such as sulfatase pathway involving the generation of dehydroepiandrosterone (DHEA) and estradiol (E2) from dehydroepiandrosterone sulfate (DHEAS) and estrone sulfate (E1S) through steroid sulfatase (STS), and androst-5-ene-3β,17β-diol (5-diol) from DHEA, 5α-androstane-3β,17β-diol (3β-diol) from dihydrotestosterone (DHT) through 17β-hydroxysteroid dehydrogenase type1 and type 7 (17β-HSD1,7)\(^6,7\). Moreover, recent studies have demonstrated the important role of glucocorticoids (GCs, predominantly cortisol in humans and corticosterone in rodents) in human BC development. GCs, primarily involved in the regulation of glucose metabolism, inflammation inhibition and immune suppression\(^8\), not only exert important effects on the development and functions of the mammary gland\(^9\), but also act as inhibitors of human BC cell proliferation\(^10\). And the expression modulation of 11β-hydroxysteroid dehydrogenases type 1 and type 2 (11β-HSD1,2) in human BC resulted in a low intratumoral GC environment, direct contribute to AI resistance. Furthermore, progesterone metabolites 4-pregnenes and 5α-pregnanes possess important effects on the control of BC development\(^11\). The maintaining of a high 5α-pregnanes/4-pregnenes ratio through down-regulation of 3α-hydroxysteroid dehydrogenases (3α-HSDs) and up-regulation of 5α-reductases (5αRs) expression provide a favorable environment for cancer cell growth, contributing to AI resistance.
In the present study, with The Cancer Genome Atlas Breast Invasive Carcinoma (TCGA-BRCA) RNA sequencing dataset from clinical samples, we analyzed the differential expression and expression correlation of key steroid-converting enzymes directly involved in the modulation of estrogen and androgen, cortisol and cortisone, 4-pregnene and 5α-pregnane, together with their related receptors. The in depth understanding of the joint control of breast cancer by related steroid-hormones will lay down the base for more efficient combined endocrine therapies.

4.4 Materials and Methods

Ethics statement

The usage of RNA sequencing data from TCGA in this study meets the data use policies set by TCGA (https://cancergenome.nih.gov/abouttcga/policies/ethicslawpolicy).

RNA sequencing dataset

TCGA is a community resource project and the TCGA database is currently the largest database of cancer genetic information of over 30 kinds of human tumours. TCGA database contains a large number of RNA-seq data from clinical samples and provides most complete clinical information of each patient, thus is widely used in many studies. To avoid introducing errors when merging RNA-seq data from different cohorts, here we choose to use the RNA-seq data from the TCGA database. In this study, we focused on the transcriptome profiling of primary tumor in post-menopausal ER+ female BC patients, and cases that did not meet this criterion were excluded from the analysis. RNA sequencing dataset of TCGA-BRCA cohort (n=1097) was downloaded through the Genomic Data Commons (GDC) data portal service. The gene level expression values in the dataset were generated through the GDC mRNA quantification analysis pipeline by first aligning reads to the GRCh38 reference genome and then by quantifying the mapped reads, which finally normalized to fragments per kilobase of transcript per million mapped reads (FPKM). Due to the highly skewed nature of RNA-seq data, the FPKM values were then log2-transformed to bring them closer to normal distribution. Since we were focused on the transcriptome profiling of primary tumor in post-menopausal ER+ female BC patients, totally 526 tumor samples and 56 normal breast samples were used in following analysis. Moreover, in the analysis of differential expression of key steroid converting enzyme genes between pre- and post-menopausal ER+ BC, totally 163 pre-menopausal cases were included.

Statistical Analysis

Samples were separated into different groups (such as tumor and normal, pre-menopause and post-menopause) according to the variables used in following analysis, which then displayed in Boxplot to show the
distribution of data among groups. Case with a value larger than 1.5 times of interquartile range (IQR) have been considered as an outlier and excluded from following statistical analysis.

Student's t-test and the Mann-Whitney U test (also called Wilcoxon rank-sum test) are commonly used in identification of differentially expressed genes in two user-defined groups in statistic analysis\textsuperscript{15}. Both tests assume that the data distributions of the two groups have the same shape, and the student's t-test additionally assuming normal distributions. However, similar to DNA microarray gene expression data, the assumption of a normal distribution of intensities of every gene in RNA-seq may not be valid even after log transformation\textsuperscript{16}. Thus to be conservative and robust, the differential gene expression analysis was evaluated by the Mann–Whitney U test (2-tailed). The fold change (FC) was defined as the ratio of means of the two compared groups. Positive FC value indicates up-regulation and negative one indicates down-regulation. For the gene expression correlation coefficient test, Spearman's rank correlation coefficient (2-tailed) was employed. For all statistical analysis, the Benjamini-Hochberg Procedure was performed to decrease the false discovery rate (FDR)\textsuperscript{17}, as a correction of significance; and a p<0.05 was considered statistically significant and represented by*, a p<0.001 was represented by**.

4.5 Results

17β-HSD7 over-expressed in post-menopausal ER positive (ER+) BC compared to adjacent normal breast tissues

We first examined the differential expression of 17β-HSD1 (gene HSD17B1) and 17β-HSD7 (gene HSD17B7) based on cancer and normal tissues. A Boxplot of the TCGA-BRCA data showed the distribution of values of both genes in normal and cancer groups (Figure 4.1A and Table 4.1). The expression levels of 17β-HSD1 remained controversial in literatures, with some reports indicating an up-regulation\textsuperscript{18-19} whereas others showing a down-regulation\textsuperscript{20-21}, but both are modest. These different results may due to their limited sample size. With large number of clinical samples, the expression level of 17β-HSD1 exhibited no significant difference between ER+ BC and normal adjacent breast tissues in post-menopausal women (p=0.073) (Figure 4.1A and Table 4.1). Although the expression of 17β-HSD1 was not changed during BC development, considering its high specific activity in E1 to E2 conversion, 96 ± 10 s\textsuperscript{-1}(µM)\textsuperscript{-1} at the molecular level\textsuperscript{22}, but a very significant substrate inhibition\textsuperscript{23}, its enzyme role in maintaining a high intratumoral concentration of E2 is still worth consideration\textsuperscript{24}. Moreover, the enzyme may also contribute to significant 17β-HSD7 regulation (see below). Further analysis showed that the expression level of the enzyme in pre- and post-menopausal groups were similar (Table 4.2).
For 17β-HSD7, the immunohistological study conducted by Shehu et al. showed the enzyme’s high expression in both invasive and in situ breast carcinoma. The immunoreactivity of 17β-HSD7 was detected in 20 of 41 cases (49%) in BC and 24 of 41 cases (58%) in non-malignant adjacent tissues. The results from present study indicated that its expression in ER+ BC was significantly up-regulated in post-menopausal women (2.61-fold, p=6.08E-26) (Figure 4.1A and Table 4.1). However, there was no significant difference between the expression of 17β-HSD7 in pre- and post-menopausal groups (Table 4.2).

**11β-HSD1 under-expressed while 11β-HSD2 over-expressed in post-menopausal ER+ BC compared to adjacent normal breast tissues**

11β-HSD1 (gene HSD11B1) has been detected in most BC tissues and normal adjacent tissues by Immunohistochemical studies, and its expression was significantly down-regulated in BC specimens compared with normal adjacent tissues. Whereas the expression of 11β-HSD2 (gene HSD11B2) has been detected in 8 out of 12 breast tumor specimens (66%) by western blot. We examined their expression in both normal breast tissues and ER+ BCs in post-menopausal women, in the TCGA-BRCA cohort. A Boxplot showed a clear different data distributions of the two genes in normal and tumor groups (Figure 4.1B and Table 4.1). Mann-Whitney U tests and FC calculation demonstrated the significant down-regulation of 11β-HSD1 (~8.33-fold, p=1.64E-23) and the significant up-regulation of 11β-HSD2 (2.30-fold, p=2.17E-09) in ER+ BCs compared with normal breast tissues (Table 4.1). No significant difference in expression level of these two genes was observed between pre- and post-menopausal groups (Table 4.2).

**3α-HSDs under-expressed in post-menopausal ER+ BC compared to adjacent normal breast tissues**

The down-regulation of 3α-HSD4 (also known as 3α(20α)HSD, gene AKR1C1), 3α-HSD3 (gene AKR1C2) and 3α-HSD2 (gene AKR1C3) in human breast tumors as compared to normal breast tissues has been demonstrated by qRT-PCR with a large number of clinical samples. These expression modifications were also observed in BC cells (such as MCF7, T-47D and MDA-MB-231) as compared to normal breast cell MCF-10A. Here we examined 3α-HSDs expression in ER+ BCs in comparison with normal breast tissues. A Boxplot showed the obvious different data distributions of these genes in normal and cancer groups (Figure 4.1C and Table 4.1). Mann-Whitney U tests and FC calculation showed significantly down-regulation of all four isoforms of 3α-HSDs by ~1.51-fold (p=0.002), ~8.18-fold (p=1.63E-28), ~35.07-fold (p=2.56E-29) and ~30.38-fold (p=5.08E-30) respectively in breast cancerous tissues as compared with normal breast tissues (Table 4.1). Their expression levels in pre- and post-menopausal patients were comparable (Table 4.2).

**5α-reductases over-expressed in post-menopausal ER+ BC compared to adjacent normal breast tissues**
The expression modulation of 5α-reductases (5αRs) in BCs is remaining controversial. In vitro experiments with breast cell lines and BC cell lines indicated a significant up-regulation of 5αR1 in cancer cells than normal cells\(^28\). On the contrary, Zhao et al. reported a significant down-regulation of 5αR1 in breast carcinoma compared to adjacent normal tissues\(^29\). In the present study, the expression status of 5αRs (gene SRD5As) in ER+ BCs and normal breast tissues was displayed with Boxplot, showing obvious different data distributions in normal and cancer groups (Figure 4.1D and Table 4.1). Mann-Whitney U tests and FC calculation indicated significant up-regulation of all three isoforms of SRD5As by 1.35-fold (\(p=3.42\times 10^{-05}\)), 3.11-fold (\(p=1.33\times 10^{-11}\)) and 1.68-fold (\(p=1.56\times 10^{-15}\)) respectively in breast cancerous tissues as compared with normal breast tissues (Table 4.1). All three isoforms showed similar expression level in pre- and post-menopausal patients (Table 4.2).

**Differential expressions of steroid hormone receptors in post-menopausal ER+ BC and adjacent normal breast tissues**

To clearly understand the effects of the expression modification of these key steroid enzymes to breast cancer development, we also examined the expression of their related receptors. The data distribution of receptor genes in normal and cancer groups were displayed by a Boxplot (Figure 4.1E and Table 4.1). Since all cancer samples used in this study were ER+, ERα (gene ESR1) in those samples were over-expressed (4.01-fold, \(p=6.74\times 10^{-18}\)) compared to normal breast tissues. The expression of ERα in post-menopausal cases were significantly higher than premenopausal cases (2.21-fold, \(p=4.43\times 10^{-21}\)), which significantly increased the estrogen sensitivity of cancer cells.

Androgen receptor (AR) expression was found to be a favorable prognostic indicator of disease outcomes\(^30\). It can be detected in 61% of BCs and in 75% of ER+ cases, and it is the most commonly expressed hormone receptor in “in situ”, invasive and metastatic BC\(^30\). We examined AR differential expression with clinical samples, results showed a significant up-regulation (1.50-fold, \(p=3.36\times 10^{-08}\)) in ER+ BCs compared to normal breast tissues (Table 4.1). Its expression levels in pre- and post-menopausal patients were comparable (Table 4.2).

Progesterone receptor (PR, gene PGR) expression is driven by estrogen-bound ER\(^31\), and its role in BC remains controversial\(^32\). According to the present study, no statistical difference of expression has been detected between ER+ BCs and normal breast tissues (1.11-fold, \(p=0.340\)) (Table 4.1). However, in contrast with the ERα, the expression level of PR in post-menopausal women was significantly lower as compared with pre-menopausal ones (-1.57-fold, \(p=0.048\)) (Table 4.2).
Glucocorticoid receptor (GR, gene NR3C1) appeared in approximately 50-70% of human invasive BC samples through ligand-binding assays, and its levels decrease significantly during cancer progression\textsuperscript{33}. In the present study, GR expression in ER+ BCs was significantly down-regulated compared with normal breast tissues (-3.36-fold, p=2.14E-28) (Table 4.1). Its expression levels in pre- and post-menopausal patients were comparable (Table 4.2).

**Expression correlation of key steroid-converting enzymes and related receptors in post-menopausal ER+ BC**

To better understand the expression correlation between these key steroid enzymes and related receptors, we further performed a Spearman’s rank correlation coefficient test. Totally 526 ER+ cases from post-menopausal women were involved in this study. Results showed that AKR1C1, AKR1C2 and AKR1C3 expression was strongly positively correlated with each other (r=0.886, p=1.29E-173; r=0.698, p=4.80E-76 and r=0.682, p=3.03E-71 respectively) (Table 4.3 and Figure 4.2). This may be related to their location in chromosomes in the same region and may be subjected to similar regulation mechanisms. Besides, the expression of AKR1C1 and AKR1C2 were also positively correlated with HSD11B1 (r=0.548, p=7.42E-41 and r=0.491, p=1.06E-31 respectively) (Table 4.3 and Figure 4.2). Interestingly, ESR1 has some expression correlations with several other genes. It positively correlated with HSD17B7 (r=0.239, p=1.38E-07), AR (r=0.476, p=9.89E-30) and PGR (r=0.382, p=1E-18), whereas negatively correlated with HSD11B1 (r=-0.237, p=1.89E-07), AKR1C1 (r=-0.268, p=2.26E-09), AKR1C2 (r=-0.227, p=7.23E-07), AKR1C3 (r=-0.154, p=0.001), SRD5A1 (r=-0.35, p=1.05E-15) (Table 4.3). AR and PGR also positively correlated with each other with an r value of 0.33 (p=5.67E-14). As expected, the expression of HSD11B1 was positively correlated with NR3C1 (r=0.299, p=1.4E-11).

**4.6 Discussion**

ER activation by estrogens synthesized through multiple aromatase-independent pathways is still one of the major mechanisms of AI-resistance. Besides E1 and E2, androgen metabolites, such as 5-diol and 3β-diol, were also reported to have estrogenic activities. Both of them possess dual and opposite effect on BC growth: they act as stimulators on their own through ER, but counteract the growth-stimulatory effect of E2 through the AR under the physiological concentrations, contributing to AI-resistance\textsuperscript{34}. 5-diol was synthesized from DHEA by 17β-HSD1, and 3β-diol could be converted from DHT by both 17β-HSD1 and 17β-HSD7\textsuperscript{35}. With mRNA-sequencing data from a large number of clinical samples, we observed a significant up-regulation of 17β-HSD7 in ER+ BCs in post-menopausal women compared with normal breast tissues, while no significant change has been observed for 17β-HSD1. However, 17β-HSD1 may still contribute to the maintaining of E2 level due to the high enzyme activity. Moreover, the remarkable regulation of 17β-HSD7 by 17β-HSD1 via E2 modulations in BC cells has recently been demonstrated \textsuperscript{36}. Inhibition of 17β-HSD7 in breast cancer cells led to E2
decrease and DHT accumulation, resulting in a cell cycle arrest and feedback down-regulation of the enzyme. Thus the significant over-expression of 17β-HSD7 in ER+ BC directly contributes to the high levels of intratumoral estrogens and low levels of intratumoral androgens (Figure 4.3). This was consistent with the report by Stanczyk et al. that androgen levels were generally lower in cancerous tissue than in benign tissue. The possible use of 17β-HSD7 as target for ER+ BC treatment awaits the study of the enzyme role in cholesterol biosynthesis. Furthermore, we also observed a significant higher expression of ER in post-menopausal BC patients than in premenopausal BC patients, which may remarkably increase the estrogen sensitivity of cancer cells.

The stimulating effect of estrogens on BC proliferation is modulated by GCs. It has been reported that GCs inhibited estrogen responses, and the activation of GR by DEX can attenuate estrogen responses through the induction of the expression of estrogen sulfotransferase (SULT1E1). GCs inhibited the proliferative activity of MCF-7 cells in the presence of GR, and also have the ability to block the stimulatory effect of E2 on MCF-7 cell proliferation. In T47D BC cells, GCs inhibited cell migration by disrupting the cytoskeletal dynamic organization. In peripheral tissues, the concentrations of intracellular GCs were modulated by the 11β-HSD enzymes. It has been reported that GR-rich normal tissues express 11β-HSD1, while cancerous tissues express 11β-HSD2, and high GR expression has been reported to be associated with better prognosis than low or no GR expression. The significant down-regulation of 11β-HSD1 and GR whereas the up-regulation of 11β-HSD2 in breast cancerous tissues was demonstrated by clinical samples in the present study, in which we also observed that the expression of 11β-HSD1 was negatively correlated with ERα but positively correlated with GR. The up-regulation of 11β-HSD2 may be due to the stimulation of E2, since there is an estrogen response element (ERE) located in the promoter of the gene according to the human ERE databases reported by Bourdeau et al. The down-regulation of 11β-HSD1 and GR together with the up-regulation of 11β-HSD2 consequentially led to a lower intratumoral cortisol level as well as a decreased GC signal in cells, diminishing the anti-inflammation effect and anti-proliferative effect of GCs. This contributed to the favorable tumor growth environment, and relieved the estrogen deprivation stress coursed by aromatase inhibition (Figure 4.3). 11β-HSD2 acts as an enzymatic shield maintaining and facilitating BC cell growth, and the inhibition of 11β-HSD2 activity elevates the anti-proliferative effect of GCs on BC cells.

More and more evidence indicate that the metabolites rather than progesterone itself played important roles in AI-resistant of BC. Progesterone was metabolized to 5α-pregnane-3,20-dione (5αP) by 5α-reductase or to 3α-hydroxy-4-pregnen-20-one (3αHP) and 4-pregnen-20α-ol-3-one (20αDHP) by 3α-HSDs in breast tumors. 5αP has been demonstrated to be able to promote BC cell proliferation and detachment in vitro and tumor formation in vivo regardless of the presence or absence of ER or PR, whereas 3αHP and 20αDHP suppress proliferation and detachments of MCF-7 cells, and those effects were mediated through their
receptors. Through the down-regulation of 5αP receptor, 3αHP and 20αDHP suppress mitogenic and metastatic activity in BC cells. However, the genes encoded these receptors have not been reported yet. Experiments also indicated that 3αHP and 20αDHP decreased ER levels or block the stimulation of E2 and 5αP on ER expression in MCF-7 cells in a dose-dependent manner. 3α-HSDs are responsible for the conversion of progesterone to 4-pregnenes, while 5αRs metabolize progesterone and 4-pregnenes to 5α-pregnanes. In the present study, we observed significant down-regulation of 3α-HSDs and up-regulation of 5αRs in ER+ BC patients. This selective expression loss of AKR1Cs in breast tumors may augment progesterone signaling by its nuclear receptors, or more importantly, may suppress the formation of mitogen/metastasis inhibitors 3αHP and promote the formation of cancer stimulator 5αP. The expression of AKR1C1 and AKR1C2 in human breast carcinoma cells was positively correlated with disease-free and overall survival; and the expression status of AKR1C1 in tumor cells was proposed as an independent prognostic marker. Moreover, these significant expression modifications of 3α-HSDs and 5αRs lead to lower levels of 3αHP and 20αDHP, and higher level of 5αP (Figure 4.3), consequentially stimulating cancer cell proliferation, providing an escape pathway for AI-resistance.

Beside all the steroid-converting enzymes analysed in the present work, aromatase is still one of the most important enzyme associated with estrogen-dependent BC development. With large number of clinical samples, our study showed a down-regulation of the enzyme in ER+ BC with statistical significance (-1.70-fold, p=0.002).

4.7 Conclusion
The dual role on E2 and DHT by 17β-HSD7 was recently reported in detail and the enzyme inhibition yields successful reduction of cell proliferation and xenograft tumor shrinkage of the estrogen-dependent cancer. The significant up-regulation of the enzyme in ER+ BC strongly suggests it a novel target for endocrine treatment. Furthermore, different combinatory use of inhibitors targeting dual steroid hormones may yield novel endocrine therapeutic approaches. The inhibition of 17β-HSD7 will not only decrease the E2 level and restore the DHT level, but will also arrest the cell cycle in the G0/G1 phase and trigger apoptosis. The decreased E2 level will relieve its suppression to 11β-HSD1 expression. Combined with the use of an 11β-HSD2 inhibitor will lead to the restoration of cortisol levels, that may subsequently elevate endogenous anti-inflammatory and anti-proliferative effects. We are confident that the understanding of expression and regulation of steroid enzymes and their receptors, as well as their correlation, will facilitate BC mechanism study and novel therapy design.
Conflict of Interest

The authors declare that they have no conflict of interest.

Acknowledgements

The work has been supported by the Canadian Institute of Health Research (CIHR) grant to Lin SX et al. (MOP97017). We acknowledge Dr. Muriel Steel for her editing of the manuscript.
4.8 Reference


32. Kuhl, H.; Schneider, H. P., Progesterone--promoter or inhibitor of breast cancer. *Climacteric* 2013, 16 Suppl 1, 54-68.


48. Wiebe, J. P.; Rivas, M. A.; Mercogliano, M. F.; Elizalde, P. V.; Schillaci, R., Progesterone-induced stimulation of mammary tumorigenesis is due to the progesterone metabolite, 5alpha-dihydroprogesterone (5alphaP) and can be suppressed by the 5alpha-reductase inhibitor, finasteride. *J Steroid Biochem Mol Biol* 2015, 149, 27-34.


Figure 4.1. Boxplot display gene expression distribution of several key steroid-converting enzymes and related receptors in normal breast and ER+ BC in post-menopausal women. N, normal adjacent breast tissue. T, primary breast tumor. IQR, Interquartile Range. *, p<0.05 (2-tailed). **, p<0.001 (2-tailed).
Figure 4.2. Scatter plot of gene expression correlation between AKR1C1, AKR1C2, AKR1C3 and HSD11B1. Relationships between genes were examined using Spearman’s rank correlation coefficient test, and the correlation coefficient ($r_s$), $p$ values and case numbers were indicated.
Figure 4.3. Schematic representation of important regulation of steroid-converting enzymes in BC based on a large number of clinical samples from TCGA cohort. The Red arrows indicate up-regulation; the green arrows indicate down-regulation; the red squares indicate cancer stimulators; the green squares indicate cancer suppressers. *, fold change was significant at the 0.05 level; **, fold change was significant at the 0.001 level with Mann–Whitney U test (2-tailed). FC, fold change; DHEA, dehydroepiandrosterone; 4-Dione, androstenedione; A-Dione, 5α-androstanedione; ADT, androsterone; E1, estrone; E2, estradiol; T, testosterone; 5-Diol, androst-5-ene-3β,17β-diol; DHT, dihydrotestosterone; 3β-diol, 5α-androstane-3β,17β-diol; 3αHP, 3α-hydroxy-4-pregnen-20-one; 20αDHP, 4-pregnen-20α-ol-3-one.
Table 4.1. Differential expression of several key steroid-converting enzymes in post-menopausal ER+ BC vs. normal breast tissue.

<table>
<thead>
<tr>
<th>Gene (Protein)</th>
<th>Case Number</th>
<th>FPKMmean</th>
<th>p</th>
<th>FC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>T</td>
<td>N</td>
<td>T</td>
</tr>
<tr>
<td>HSD17B1 (17β-HSD1)</td>
<td>56</td>
<td>526</td>
<td>0.71</td>
<td>0.67</td>
</tr>
<tr>
<td>HSD17B7 (17β-HSD7)</td>
<td>56</td>
<td>526</td>
<td>1.37</td>
<td>3.58</td>
</tr>
<tr>
<td>HSD11B1 (11β-HSD1)</td>
<td>56</td>
<td>526</td>
<td>8.33</td>
<td>1.00</td>
</tr>
<tr>
<td>HSD11B2 (11β-HSD2)</td>
<td>56</td>
<td>526</td>
<td>1.62</td>
<td>3.73</td>
</tr>
<tr>
<td>AKR1C1 (3α-HSD4)</td>
<td>56</td>
<td>525</td>
<td>9.52</td>
<td>0.31</td>
</tr>
<tr>
<td>AKR1C2 (3α-HSD3)</td>
<td>56</td>
<td>524</td>
<td>13.65</td>
<td>0.39</td>
</tr>
<tr>
<td>AKR1C3 (3α-HSD2)</td>
<td>56</td>
<td>526</td>
<td>17.82</td>
<td>2.18</td>
</tr>
<tr>
<td>AKR1C4 (3α-HSD1)</td>
<td>44</td>
<td>326</td>
<td>0.03</td>
<td>0.02</td>
</tr>
<tr>
<td>SRD5A1 (5αR1)</td>
<td>56</td>
<td>526</td>
<td>1.53</td>
<td>2.06</td>
</tr>
<tr>
<td>SRD5A2 (5αR2)</td>
<td>53</td>
<td>510</td>
<td>0.02</td>
<td>0.06</td>
</tr>
<tr>
<td>SRD5A3 (5αR3)</td>
<td>56</td>
<td>526</td>
<td>5.92</td>
<td>9.95</td>
</tr>
<tr>
<td>ESR1(ERα)</td>
<td>56</td>
<td>526</td>
<td>9.94</td>
<td>39.87</td>
</tr>
<tr>
<td>AR</td>
<td>56</td>
<td>526</td>
<td>9.15</td>
<td>13.77</td>
</tr>
<tr>
<td>PGR (PR)</td>
<td>56</td>
<td>526</td>
<td>2.73</td>
<td>3.03</td>
</tr>
<tr>
<td>NR3C1(GR)</td>
<td>56</td>
<td>526</td>
<td>18.25</td>
<td>5.43</td>
</tr>
</tbody>
</table>

N\textsubscript{N}, number of normal breast tissue samples. N\textsubscript{T}, number of breast tumor tissue samples. FC, fold change, positive value indicates up-regulation in tumor tissues and negative indicates down-regulation. *, p<0.05; **, p<0.001 with Mann–Whitney U test (2-tailed). 95% CI, 95% confidence intervals.
Table 4.2. Differential expression of several key steroid-converting enzymes between post- and pre-menopausal ER+ BC.

<table>
<thead>
<tr>
<th>Gene (Protein)</th>
<th>N&lt;sub&gt;pre&lt;/sub&gt;</th>
<th>N&lt;sub&gt;post&lt;/sub&gt;</th>
<th>p</th>
<th>FC</th>
</tr>
</thead>
<tbody>
<tr>
<td>HSD17B1 (17β-HSD1)</td>
<td>163 526</td>
<td>0.222</td>
<td>1.16</td>
<td></td>
</tr>
<tr>
<td>HSD17B7 (17β-HSD7)</td>
<td>163 526</td>
<td>0.707</td>
<td>1.03</td>
<td></td>
</tr>
<tr>
<td>HSD11B1 (11β-HSD1)</td>
<td>163 526</td>
<td>0.689</td>
<td>-1.04</td>
<td></td>
</tr>
<tr>
<td>HSD11B2 (11β-HSD2)</td>
<td>163 526</td>
<td>0.095</td>
<td>1.13</td>
<td></td>
</tr>
<tr>
<td>AKR1C1 (3α-HSD4)</td>
<td>163 525</td>
<td>0.576</td>
<td>1.15</td>
<td></td>
</tr>
<tr>
<td>AKR1C2 (3α-HSD3)</td>
<td>163 524</td>
<td>0.558</td>
<td>1.10</td>
<td></td>
</tr>
<tr>
<td>AKR1C3 (3α-HSD2)</td>
<td>163 526</td>
<td>0.683</td>
<td>-1.04</td>
<td></td>
</tr>
<tr>
<td>AKR1C4 (3α-HSD1)</td>
<td>101 326</td>
<td>0.448</td>
<td>1.13</td>
<td></td>
</tr>
<tr>
<td>SRD5A1 (5αR1)</td>
<td>163 526</td>
<td>0.071</td>
<td>-1.14</td>
<td></td>
</tr>
<tr>
<td>SRD5A2 (5αR2)</td>
<td>160 510</td>
<td>0.754</td>
<td>-1.06</td>
<td></td>
</tr>
<tr>
<td>SRD5A3 (5αR3)</td>
<td>163 526</td>
<td>0.819</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>ESR1(ERα)</td>
<td>163 526</td>
<td>4.43E-21</td>
<td>2.21**</td>
<td></td>
</tr>
<tr>
<td>AR</td>
<td>163 526</td>
<td>0.362</td>
<td>1.10</td>
<td></td>
</tr>
<tr>
<td>PGR (PR)</td>
<td>163 526</td>
<td>0.048</td>
<td>-1.57*</td>
<td></td>
</tr>
<tr>
<td>NR3C1(GR)</td>
<td>163 526</td>
<td>0.332</td>
<td>-1.09</td>
<td></td>
</tr>
</tbody>
</table>

N<sub>pre</sub>, number of pre-menopausal ER+ BC cases. N<sub>post</sub>, number of post-menopausal ER+ BC cases. FC, fold change, positive value indicates up-regulation in post-menopausal ER+ BC and negative indicates down-regulation. *, p<0.05; **, p<0.001 with Mann–Whitney U test (2-tailed). 95% CI, 95% confidence intervals.
Table 4.3. Spearman’s rank correlation coefficient test of several key steroid-converting enzymes in ER+ BC.

<table>
<thead>
<tr>
<th>Genes</th>
<th>HSD17B 1</th>
<th>HSD11B 1</th>
<th>HSD11B 2</th>
<th>AKR1C1</th>
<th>AKR1C2</th>
<th>AKR1C3</th>
<th>AKR1C4</th>
<th>SRD5A1</th>
<th>SRD5A2</th>
<th>SRD5A3</th>
<th>AR</th>
<th>ESR1</th>
<th>PGR</th>
<th>NR3C1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>rs</td>
<td>p</td>
<td>N</td>
<td>rs</td>
<td>p</td>
<td>N</td>
<td>rs</td>
<td>p</td>
<td>N</td>
<td>rs</td>
<td>p</td>
<td>N</td>
<td>p</td>
<td>N</td>
</tr>
<tr>
<td>HSD17B 1</td>
<td>0.04</td>
<td>0.02</td>
<td>0.00</td>
<td>0.01</td>
<td>0.01</td>
<td>0.03</td>
<td>-0.10</td>
<td>-0.05</td>
<td>-0.03</td>
<td>-0.13</td>
<td>-0.04</td>
<td>-0.09</td>
<td>0.22</td>
<td>0.04</td>
</tr>
<tr>
<td>HSD11B 1</td>
<td>0.58</td>
<td>0.82</td>
<td>0.99</td>
<td>0.85</td>
<td>0.91</td>
<td>0.68</td>
<td>0.12</td>
<td>0.45</td>
<td>0.70</td>
<td>0.01</td>
<td>0.53</td>
<td>0.09</td>
<td>0.54</td>
<td>0.54</td>
</tr>
<tr>
<td>HSD11B 2</td>
<td>-0.157</td>
<td>0.151</td>
<td>-0.137</td>
<td>-0.122</td>
<td>-0.01</td>
<td>0.03</td>
<td>-0.190</td>
<td>0.122</td>
<td>0.03</td>
<td>0.190</td>
<td>0.239</td>
<td>0.351</td>
<td>-0.095</td>
<td>0.04</td>
</tr>
<tr>
<td>HSD11B 1</td>
<td>0.001</td>
<td>0.002</td>
<td>0.005</td>
<td>0.01</td>
<td>0.86</td>
<td>0.71</td>
<td>0.05</td>
<td>0.01</td>
<td>0.68</td>
<td>0.05</td>
<td>0.17</td>
<td>0.17</td>
<td>0.07</td>
<td>0.07</td>
</tr>
<tr>
<td>HSD11B 2</td>
<td>-0.01</td>
<td>-0.02</td>
<td>0.02</td>
<td>0.03</td>
<td>-0.06</td>
<td>-0.03</td>
<td>0.05</td>
<td>0.00</td>
<td>0.093</td>
<td>0.00</td>
<td>-0.188</td>
<td>-0.188</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>HSD11B 1</td>
<td>0.94</td>
<td>0.81</td>
<td>0.75</td>
<td>0.72</td>
<td>0.27</td>
<td>0.71</td>
<td>0.46</td>
<td>0.98</td>
<td>0.07</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td>HSD11B 1</td>
<td>0.526</td>
<td>0.526</td>
<td>0.526</td>
<td>0.526</td>
<td>0.526</td>
<td>0.526</td>
<td>0.526</td>
<td>0.526</td>
<td>0.526</td>
<td>0.526</td>
<td>0.526</td>
<td>0.526</td>
<td>0.526</td>
<td>0.526</td>
</tr>
<tr>
<td>AKR1C1</td>
<td>0.866</td>
<td>1.29E-4</td>
<td>7.42E-6</td>
<td>0.548</td>
<td>0.491</td>
<td>0.373</td>
<td>-0.05</td>
<td>0.151</td>
<td>-0.04</td>
<td>-0.06</td>
<td>-0.08</td>
<td>-0.237</td>
<td>0.00</td>
<td>0.299</td>
</tr>
<tr>
<td>AKR1C1</td>
<td>0.76</td>
<td>0.75</td>
<td>0.75</td>
<td>0.75</td>
<td>0.47</td>
<td>0.98</td>
<td>0.68</td>
<td>0.01</td>
<td>0.09</td>
<td>0.68</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
</tr>
<tr>
<td>AKR1C2</td>
<td>0.682</td>
<td>3.03E-4</td>
<td>7.17</td>
<td>0.698</td>
<td>4.80E-6</td>
<td>0.03</td>
<td>0.05</td>
<td>0.00</td>
<td>0.03</td>
<td>-0.129</td>
<td>-0.268</td>
<td>-0.03</td>
<td>0.03</td>
<td>0.172</td>
</tr>
<tr>
<td>AKR1C2</td>
<td>0.75</td>
<td>0.53</td>
<td>0.53</td>
<td>0.53</td>
<td>0.53</td>
<td>0.53</td>
<td>0.53</td>
<td>0.53</td>
<td>0.53</td>
<td>0.53</td>
<td>0.53</td>
<td>0.53</td>
<td>0.53</td>
<td>0.53</td>
</tr>
<tr>
<td>AKR1C3</td>
<td>0.129</td>
<td>-0.101</td>
<td>-0.01</td>
<td>0.00</td>
<td>0.02</td>
<td>-0.01</td>
<td>-0.15</td>
<td>0.02</td>
<td>-0.209</td>
<td>6.12E-6</td>
<td>0.02</td>
<td>0.06</td>
<td>0.06</td>
<td>0.06</td>
</tr>
<tr>
<td>AKR1C3</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td>AKR1C4</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>AKR1C4</td>
<td>0.89</td>
<td>0.89</td>
<td>0.89</td>
<td>0.89</td>
<td>0.89</td>
<td>0.89</td>
<td>0.89</td>
<td>0.89</td>
<td>0.89</td>
<td>0.89</td>
<td>0.89</td>
<td>0.89</td>
<td>0.89</td>
<td>0.89</td>
</tr>
</tbody>
</table>
| SRD5A1 | 0.03 | -0.05 | -0.246 | -0.350 | -0.146 | 0.08 | 5.80E-6 | 0.02 | 0.17
<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>510</th>
<th>526</th>
<th>526</th>
<th>526</th>
<th>526</th>
<th>526</th>
</tr>
</thead>
<tbody>
<tr>
<td>SRD5A2</td>
<td>rs</td>
<td>0.04</td>
<td>-0.02</td>
<td>-0.07</td>
<td>0.02</td>
<td>-0.03</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.62</td>
<td>0.82</td>
<td>0.26</td>
<td>0.82</td>
<td>0.71</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>510</td>
<td>510</td>
<td>510</td>
<td>510</td>
<td>510</td>
<td></td>
</tr>
<tr>
<td>SRD5A3</td>
<td>rs</td>
<td>-0.02</td>
<td>0.01</td>
<td>-0.181**</td>
<td>-0.155*</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.75</td>
<td>0.91</td>
<td>1.07E-04</td>
<td>0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>526</td>
<td>526</td>
<td>526</td>
<td>526</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AR</td>
<td>rs</td>
<td>0.476**</td>
<td>0.330**</td>
<td>0.08</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>9.89E-30</td>
<td>5.67E-14</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>526</td>
<td>526</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ESR1</td>
<td>rs</td>
<td>0.382**</td>
<td>0.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.00E-18</td>
<td>0.84</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>526</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PGR</td>
<td>rs</td>
<td>0.07</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>526</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$r_s$, Spearman's rank correlation coefficient. N, number of samples. *, $p<0.05$ (2-tailed). **, $p<0.001$ (2-tailed).
Conclusion

The results reported in this thesis have been discussed in chapter I to IV. In this chapter, we would like to highlight the major points of the previous discussions, and also try to highlight the links between the different results that facilitate the EDDs treatment. Besides, the prospects of the study are indicated.

The new generation 17β-HSD1 inhibitor PBRM forms a covalent bond with the enzyme.

The 17β-HSD1 has a well established role in estrogen-dependent cancer especially in breast cancer, however no candidate inhibitor has eventually reached the clinical trials\(^8\). The major obstacle in the development of an inhibitor for the 17β-HSD1, which is generally associated with previous series of inhibitors, is the presence of undesirable estrogenic activity. This may be largely due to the fact that 17β-HSD1 has a high affinity for estrogens\(^67,132\). Thus the potent inhibitors usually contain an estrogen core making it difficult to eliminate their estrogenic activity\(^8\). To overcome this obstacle, decades of research accompanied by trial and error as well as structure based rational design were devoted and finally lead to the development of PBRM, which has shown promising efficacy in both breast cancer cells and human tumor xenografts in nude mice\(^98-99\). It was derived from the most potent 17β-HSD1 inhibitor CC-156, with a substitution of the C3-end hydroxyl group with a bromoethyl group\(^98\). This modification slightly decreases the inhibitor activity of PBRM to 17β-HSD1 compared to CC-156 with an IC\(_{50}\) value of 68nM for the E1 to E2 conversion\(^99\). However, the presence of a bromide instead of a hydroxyl group at the C3 end of the inhibitor significantly eliminates the binding of PBRM to the estrogen receptor alpha. Moreover, PBRM was further demonstrated to be an irreversible inhibitor of 17β-HSD1\(^101\), which was further proved by the 17β-HSD1-PBRM-NADP+ complex structure reported by Li et al\(^133\). This ternary complex structure reveals a covalent bond between the C-31 of PBRM and the Ne of residue His\(^221\), and is by far the first example of N-alkylation between a human enzyme and a low-reactivity alkyl halide derivative. The successful design of this highly specific irreversible inhibitor opens the door to a new design of alkyl halide-based specific covalent inhibitors and ligands as potential therapeutic agents.

Residues His\(^221\) is responsible for the substrate inhibition of 17β-HSD1.

As an important enzyme in the biosynthesis of estradiol, 17β-HSD1 has been studied since the late 1950s\(^134\). Although the major function of the enzyme is the reversible 17β oxido-reduction of steroids\(^135\), it can also, to a much lower extent, catalyse the 3β oxido-reduction of steroids\(^79\), suggesting the existence of a different substrate recognition mechanism than previously proposed. Indeed, crystal structures of 17β-HSD1 in complexes with androgens such as testosterone, demonstrated the existence of a normal and a reversely orientated binding mode in the substrate binding cavity\(^63,77,136\). In the reverse binding mode, the A-ring of the steroid facing toward the catalytic triad while the D-ring binds to the recognition end of the cavity. No reverse

References

\(^{101}\)
orientation has been observed in the E2 complexes so far. Besides, no 17β-HSD1-E1 complex has been reported prior to us. The previous reported alternative binding mode of steroid in 17β-HSD1 lead us to assess the possible binding mode of E1 in the enzyme as well as its impact on the observed substrate inhibition of the enzyme. Thus we co-crystallized the 17β-HSD1 in complex with E1. Moreover, previous experiments showed that cofactor NADPH has a significant role in the binding affinity of 17β-HSD1 to E1, while NADPH and NADH possess profound different effect in the substrate inhibition of the enzyme. To illustrate the role of NADPH on substrate binding of 17β-HSD1, we also solved the ternary complex structures containing the cofactor analog NADP+. From the binary and ternary complex structures, we indeed observed the reversely oriented E1 in all complexes, and the dead-end complex 17β-HSD1-E1-NADP+ can be responsible for the observed substrate inhibition of the enzyme.}

**Rational design of 17β-HSD1 inhibitor based on substrate inhibition mechanism**

Interestingly, the E1 molecule in both E1 binary and ternary complex crystal structures reported here were observed in a reversed binding mode, indicating the energy favoring of the reverse binding mode of the steroid. Thus it prompts us to design novel inhibitor by using the SeeSAR. The O-3 of estradiol is essential for its binding to ERα. On the basis of substrate inhibition mechanism, we conducted a structural modification at O-3 of E1 in an attempt to modulate interaction with residues at the catalytic site of 17β-HSD1, especially the Tyr as observed in CC complex, and to reduce the undesired residual estrogenic activity. An extra benzylamide ring was added to the O-3 of E1 resulting in the formation of a novel compound 3-(((8R,9S,13S,14S)-13-methyl-17-oxo-7,8,9,11,12,13,14,15,16,17-decahydro-6H-cyclopenta [α]phenanthren-3-yl)oxy)benzamide (SX7). Binding analysis using SeeSAR showed that the E1 moiety of SX7 adopts a reverse binding mode, whereas the benzylamide moiety of the inhibitor interacts with 17β-HSD1 in a similar pattern as CC did. The binding affinity of SX7 to the enzyme calculated by the Hydrogen bond and Dehydration (HYDE) shown a high estimated affinity (43nM). A docking study against ERα ligand binding
domain showed the unfavourable binding of the new compound. Further laboratory experiments need to be performed to investigate its inhibitory properties as well as the estrogenic activity.

**Significant modification in gene expression and correlation analysis of gene expression suggests novel therapy for breast cancer treatment.**

The common goal of endocrine therapy for EDDs treatment is to reduce the production of estrogens, especially the most potent one estradiol, or to block the stimulation of estrogens through binding with the estrogen receptor. The two concepts yield two milestone represented by aromatase inhibitor and tamoxifen. However, significant side effects have occurred in response to AI treatment and resistance was evident in approximately 37% of patients during AI therapy. It was demonstrated that dynamic changes in the genome usually accompanied with tumorigenesis. The modulation of the expression of genes determined the availability of steroid-converting enzymes which consequentially affect the concentration of related steroid hormones. Thus, it is reasonable to identify potential target through the statistically analysis of differentially expressed genes with RNA-seq dataset. The results from the TCGA-BRCA cohort analysis showed significant down-regulation of 3α-HSDs and up-regulation of 5α-reductases, resulting in the decreasing of cell proliferation suppressor 4-pregnenes and increasing of cell proliferation stimulators 5α-pregnanes. Besides, a significant up-regulation of 17β-HSD7 and 11β-HSD2, accompanied by a significant down-regulation of 11β-HSD1 were observed. This resulted in the accumulation of E2 and reduction of cortisol, favoring an environment for BC proliferation. Thus we propose a novel therapy targeting the reductive 17β-HSD7 and the new combined therapy targeting 11β-HSD2 and 17β-HSD7.

In this thesis, we have investigated the interactions of 17β-HSD1 at the atomic level with three inhibitors (EM-139, 2-MeO-CC-156 and PBRM) through crystallographic methods. We demonstrated that the steroid core of the reversible inhibitor EM-139 is responsible for the major interactions with 17β-HSD1, whereas the bulky 7α-alkyl moiety of the inhibitor, which is essential for its anti-estrogenic activity, compromises its inhibitory effect on the enzyme. The other reversible inhibitor 2-MeO-CC-156, which is derived from CC-156 with a reduced intrinsic estrogenic activity but also a decrease inhibitory potency, compromised its potential for further development. The addition of a bromoethyl side chain at position C-3 of CC-156 produced a potent and non-estrogenic covalent inhibitor PBRM, which interacts similarly to CC-156 with 17β-HSD1. The structural analysis of 17β-HSD1-PBRM-NADP⁺ complex clearly shows the formation of a covalent bond between His²²¹ and the bromoethyl side chain of the inhibitor, providing insight into molecular interactions that favor the binding and subsequent N-alkylation event in the enzyme catalytic site. Also, the bromoethyl group at position C-3 of the PBRM warrants its non-estrogenic profile, slows down its metabolism, and secures its specific action of 17β-HSD1 through the formation of a covalent bond with Nε of residue His²²¹. Furthermore, structural
analysis of E1 binary and ternary complexes demonstrates the reverse binding mode of E1, which is stabilized by residue His\textsuperscript{221} and led to the formation of dead-end complex. Based on this substrate inhibition mechanism, we employed the \textit{in silico} method to design a 17β-HSD1 inhibitor, yielding a novel compound SX7 with a high estimated binding affinity to the enzyme. Our present studies provide profound details in the structure-function and inhibitor-enzyme relations of 17β-HSD1, facilitating further development of inhibitors of the enzyme for clinical purposes. Besides, with large number of clinical samples, RNA sequencing data analysis demonstrates the significant up-regulation of 17β-HSD7 and 11β-HSD2. We thus propose a novel combined therapy targeting 11β-HSD2 and 17β-HSD7.
Bibliographie


82. Huang, Y. W.; Pineau, I.; Chang, H. J.; Azzi, A.; Bellemare, V.; Laberge, S.; Lin, S. X., Critical Residues for the Specificity of Cofactors and Substrates in Human Estrogenic 17β-Hydroxysteroid


140. SeeSAR version 7.3; BioSolveIT GmbH, Sankt Augustin, Germany, 2018.


Annexe A Cold-active extracellular lipase: expression in Sf9 insect cells, homogenization, and catalysis

2.1 Résumé
Les lipases actives à froid font l’objet d’une attention particulière de nos jours, car elles sont de plus en plus utilisées dans diverses industries, telles que la synthèse chimique fine, la transformation des aliments et les détergents à lessive. Dans cette étude, un gène de lipase extracellulaire provenant de la Yarrowia lipolytica (LIPY8) a été cloné et exprimé par le système d’expression baculovirus. La lipase recombinante (LipY8p) a été purifiée en chromatographies, donnant un facteur de purification de 25,7 fois avec une activité spécifique de 1102,9 U/mg pour l’huile d’olive. L’enzyme était la plus active à un pH 7,5 et à 17ºC. Son activité maximale est en vers des esters à chaîne moyenne (C10). L’activité de la lipase était affectée par les métaux de transition, les détergents et les solvants organiques. Ces propriétés enzymatiques confèrent à cette lipase un potentiel considérable pour les applications biotechnologiques.

2.2 Abstract
Cold-active lipases are gaining special attention nowadays as they are increasingly used in various industries such as fine chemical synthesis, food processing, and washer detergent. In the present study, an extracellular lipase gene from Yarrowia lipolytica (LIPY8) was cloned and expressed by baculovirus expression system. The recombinant lipase (LipY8p) was purified using chromatographic techniques, resulting in a purification factor of 25.7-fold with a specific activity of 1102.9U/mg toward olive oil. The apparent molecular mass of purified LipY8p was 40kDa. The enzyme was most active at pH 7.5 and 17ºC. It exhibited maximum activity toward medium chain (C10) esters. The presence of transition metals such as Zn²⁺, Cu²⁺, and Ni²⁺ strongly inhibited the enzyme activity, whereas it was enhanced by EDTA. The lipase activity was affected by detergents and was elevated by various organic solvents at 10% (v/v). These enzymatic properties make this lipase of considerable potential for biotechnological applications.
2.3 Introduction

Lipase (EC 3.1.1.3) enzymes are able to hydrolyze triacylglycerol to glycerol and long-chain fatty acids, in addition to the reverse reaction of ester synthesis using a broad range of unnatural substrates. The amount of water in the reaction medium can influence lipase behavior. As a consequence of their useful features, such as independence from cofactors, broad range of substrate specificity, chemoselectivity, regioselectivity, stereoselectivity and stability in organic solvents, they have been used in various biotechnological applications, including organic synthesis, detergent manufacturing, food processing, biodiesel production, the chemical industry and biomedical sciences.

Lipases from different sources have been characterized and commercialized for industrial utilities. However, with intensification of global warming and the energy crisis, the development of cold-active lipases has attracted increased attention. Cold-adapted lipases possess relatively high catalytic activities at a low temperature range between 0 and 30°C whereas normal lipases exhibit dramatically reduced or no catalytic activities. Thus, cold-active lipases are desirable in many areas for their lower energy costs, reduced microbial contamination in industrial processes, reduced chemical side-reactions and product stabilization. Cold-active lipases primarily originate from psychrophilic and psychrotrophic microorganisms, which exist in low temperature environments such as deep seawater and Antarctic/polar regions.

In a previous study, we isolated and characterized the LipY lipase from a psychrotrophic Yarrowia lipolytica (Bohaisea-9145), which exhibited high catalytic activity at low temperatures. We also cloned the LIPY8 lipase gene from this strain, which was previously reported by Song et al. Preliminary experiments indicated that the cold-active feature of the encoded extracellular lipase LipY8p has not been fully characterized. In this paper, we heterologously overexpressed the LIPY8 gene in a baculovirus expression system, followed by purification and careful characterization of the recombinant lipase, with the aim of facilitating the industrial utility of this cold-active lipase.

2.4 Materials and Methods

Materials

Plasmid pUC57-LipY8 containing the LIPY8 gene (GenBank accession number DQ200800) without the N-terminal signal peptide coding sequence was obtained from Dr. Sun’s laboratory. Enzymes used for manipulating DNA, such as Pfu polymerase, T4 DNA ligase, EcoRI and NotI were purchased from NEB (Canada). All primers were synthesized by IDT-DNA (Canada). The Bac-to-Bac Baculovirus Expression System kit, which includes the pFastBac1 vector, the E. coli competent cell DH10Bac and Cellfectin II reagent was from Invitrogen (Canada). Spodoptera frugiperda insect cell line Sf9 and Sf-900 III SFM serum-free media...
were purchased from ThermoFisher Scientific (Canada). i-MAX serum-free media was from Wisent (Canada). Ni-NTA agarose resin was from ThermoFisher Scientific (Canada). Mono Q HR 5/5 columns were obtained from GE Healthcare (USA). The different lipase substrates were purchased from Sigma and Alfa Aesar. All reagents were of analytical grade. All curve fitting were performed using GraphPad Prism version 7 for Windows, GraphPad Software, La Jolla California USA, www.graphpad.com. This protein sequence alignment figure was generated with MEGA 7 software.

Construction of pFastBacSP6His Vector

The original pFastBac1 vector from Invitrogen does not have a signal peptide and is unsuitable for secreted protein expression. Based on pFastBac1, the signal peptide coding sequence (MGGLLLAFLALVSVPRAQA) from human lipocalin-6 (NCBI code NM_198946) was added downstream of the polyhedron promoter (P<sub>HH</sub>), followed with a 6His purification tag. This reconstructed vector was named pFastBacSP6His.

Construction of Recombinant Transfer Vector

The <i>LIPY8</i> gene was amplified using a primer pair designed for the pFastBacSP6His vector. The signal peptide coding sequence of the <i>LIPY8</i> gene was deleted from this construct. The sequence of the forward primer (F) was 5′- GCGCGAATTCCGGCGGCTGAGCCAGGGT -3′, the added EcoRI restriction site is underlined. The reverse primer (R) was 5′- GCGCTCTGAAGTTATGCAGGCCCGCGTTTTTC -3′ bearing an Xhol restriction site (underlined). The PCR was performed using 32 cycles of: denaturation at 94°C for 30s, an annealing step at 63°C for 30s, extension at 72°C for 1.5 min followed by a 5-min final extension at 72°C. The amplified product separated on a 1% agarose gel, purified by gel-extraction kit (Qiagen, Canada) and digested with EcoRI and Xhol, was ligated into the EcoRI-Xhol sites of the pFastBacSP6His vector. The recombinant vector pFastBacSP6His-LipY8 was transformed into competent <i>E.coli</i> DH5α cells. The integrity of the recovered plasmid was confirmed by restriction endonuclease digestion with EcoRI and Xhol, and sequencing (service provided by the genome sequencing and genotyping platform of the research center of University Laval) using the primers described above. The recombinant pFastBacSP6His-LipY8 plasmid was extracted from DH5α cells and transformed into competent <i>E.coli</i> DH10Bac cells. The cells were spread on blue/white selective LB agar plates containing 50µg/ml kanamycin, 7 µg/ml gentamicin, 10 µg/ml tetracycline, 100 µg/ml Bluo-gal and 40 µg/ml IPTG, and incubated overnight at 37°C. Recombinant Bacmid-LipY8 DNA was isolated and integration of the target gene into the Bacmid DNA was detected by PCR using the pUC/M13 forward and pUC/M13 reverse primers as described by the Bac-to-Bac Baculovirus Expression System kit user manual.

Cell Culture and Virus Preparation
The Sf9 cells were grown as monolayers at 27°C in Sf-900 III SFM or I-MAX serum-free media. Purified recombinant Bacmid DNA was used to transfect monolayers of Sf9 cells with Cellfectin II reagent to produce the low-titer P1 viral stock, which was then used to generate a high-titer P2 viral stock through a second infection of Sf9 cells. The titer of the baculoviral stocks was determined by plaque assay. Two percent (v/v) fetal bovine serum was added to all viral stocks, which were stored at 4°C and protected from light. The wild-type Bacmid DNA was subjected to the same procedures and served as a negative control for lipase expression.

Lipase Overexpression and Purification

Sf9 cells were infected with recombinant or wild-type virus at a multiplicity of infection (MOI) > 10. One-milliliter aliquots of the expression culture were collected every 24 h for 7 days for determination of the optimal expression period using the activity tests described below.

All purification performances were carried out at 4°C unless otherwise stated. The cells and debris were precipitated by centrifugation at 500 g for 10 min. The supernatant was collected and Tris buffer pH 8.0 was added to a final concentration of 50 mM. Ammonium sulfate powder was added gradually with constant agitation to 75% saturation over a 2-h period. Protein pellets were collected by centrifugation at 3,200 g for 30 min, and dialyzed overnight against 20 mM Tris-HCl, pH 8.0, with constant agitation.

For Ni-NTA affinity chromatography, 50 ml of concentrated lipase solution was loaded onto a Ni-NTA column (10 ml, 1.6 × 5 cm) equilibrated with buffer A (20 mM Tris-HCl pH 8.0). The lipase was eluted by a stepwise imidazole gradient with increasing concentration in buffer A. The eluted fraction was collected and the solution buffer was changed to buffer B (50 mM Tris pH 7.5, 20% (v/v) glycerol, 50 mM NaCl, 1 mM ethylenediaminetetraacetic acid (EDTA) by repeated concentration and dilution with Centricon filtration units (EMD Millipore, Canada). Ion-exchange chromatography was performed on an AKTA Explorer FPLC system (GE, USA) with a Mono Q HR 5/5 column. Lipase solution was loaded onto the column equilibrated with buffer B, and was eluted with a linear salt gradient using 1 M NaCl (pH 7.5).

SDS-PAGE

Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) was performed on a 12% polyacrylamide gel on a vertical mini gel apparatus (Bio-Rad, Canada). Molecular mass markers were obtained from Bio-Rad. Proteins were stained with Coomassie Brilliant Blue R250 (Bio-Rad, Canada).
Lipase Deglycosylation

The pOPH6 plasmid containing the PNGase F gene was purchased from addgene.org. The PNGase F protein was purified as described by Loo et al.\textsuperscript{21}. Enzymatic deglycosylation was performed at 30°C for 30 h using 1.2 mg of 1 mg/ml purified LipY8 with 125 ug of purified PNGase F. PNGase F was then removed by Mono Q HR 5/5 column.

Lipase Assay

Lipase activity was measured spectrophotometrically (410 nm) using p-nitro phenyl dodecanoate (pNPL) as substrate by the method described by Winkler and Stuckmann\textsuperscript{22}. In brief, 100 μl of substrate stock solution (0.3% (w/v) pNPL) was added to 1 ml standard reaction buffer (50 mM Na\textsubscript{2}HPO\textsubscript{4} pH 7.5, 0.2% (w/v) Na deoxycholate, 0.1% (w/v) gum arabic) and incubated in a water bath with constant shaking at 200 rpm at 22°C for 5 min. The reaction was initiated by the addition of 2 μl of enzymes and terminated by the addition of 1.2 ml acetone-ethanol (1:1) solution. The reaction duration was 2 min, and the release of pNP was recorded at 410 nm using a UV/Vis spectrophotometer (UV70, Beckman Coulter, USA). Enzyme activity was calculated by constructing a standard curve with pNP under the same buffer conditions as the reaction. One unit (U) of lipase activity was defined as the amount of enzyme that liberated 1 μmol of pNP per min under standard assay conditions.

Lipase activity was also measured by the fluorescence-based rhodamine B (RhB) assay using olive oil emulsion\textsuperscript{23} with some modifications. An RhB-olive oil emulsion mixture (RhB-OOe) containing 50 mM Na\textsubscript{2}HPO\textsubscript{4} (pH 7.5), 1% (w/v) gum arabic, 0.001% (w/v) RhB, and 2% (v/v) olive oil was emulsified with a DrinkMaster for 5 min, and then the pH was adjusted. The enzymatic assays were performed in a 45 mm × 12.5 mm quartz cuvette with magnetic stirring at pH 7.5 and 22°C using a fluorescence spectrofluorometer (HORIBA Fluorolog, USA). The enzymatic reactions were initiated by the addition of 2 μl of enzyme solution to 1 ml of emulsion. The liberated fatty acids were calculated from the fluorescence emitted at 580 nm (excitation wavelength is 350 nm). The reaction emulsion with heat-denatured enzyme solution was measured in the same way and used as a blank control. A standard curve for oleic acid in the presence of RhB and gum arabic was prepared, and a linear regression was performed allowing the calculation of lipase activity. The fluorescence emission changes were converted into the hydrolysis rate using polynomial equations. One unit (U) of lipase activity was defined as the amount of enzyme releasing 1 μmol of fatty acid per min under the assay conditions.

Effect of Temperature and pH on Lipase Activity and Stability
The activity of the lipase at different temperatures and pH was determined by a pNP release assay using pNPL as substrate. To investigate temperature stability, the lipase solution was incubated for 1 h at different temperatures ranging from 0 to 45ºC. For pH stability the lipase solution was incubated for 1 h at different pH at 22ºC. Buffers used for different pH values included 50 mM sodium phosphate buffer (pH 6–8), 50 mM Tris-HCl (pH 8.5, 9), 50 mM CHES (pH 9.5), and 50 mM CAPS (pH 10). Residual activity was measured by pNP release assay using pNPL as substrate.

**Substrate Specificity**

For the determination of substrate specificity, several p-Nitro phenyl esters including pNP-acetate (pNPA, C2), pNP-butyrate (pNPB, C4), pNP-decanoate (pNPD, C10), pNP-dodecanoate (pNPL, C12), pNP-myristate (pNPM, C14), and pNP-palmitate (pNPP, C16) were used as substrates.

**Effect of Metal Ions and Inhibitors on Lipase Activity**

The pNP release assay was used to determine the effect of metal ions and inhibitors on lipase activity. The reaction buffer was preloaded with different chemicals at the desired final concentrations.

**Effect of Detergents on Lipase Activity and Stability**

The effects of detergents on enzyme activity and stability were evaluated by pNP release assay using pNPL as substrate. For the effect on lipase activity, different detergents were pre-loaded into the reaction buffer. To determine lipase stability, the purified enzyme was pre-incubated with various detergents for 2–72 h at 22ºC and the residual activity was determined by standard assay. Several detergents were used in this study including SDS, Triton-X100, Tween 20, NP40, n-Dodecyl-β-D-Maltoside (β-DDM), n-octyl-β-D-glucoside (β-OG) and octaethylene glycol monododecyl ether (C_{12}E_{8}).

**Effect of Organic Solvents on Lipase Activity and Stability**

Seven different organic solvents including methanol, ethanol, isopropanol, acetone, Dimethyl sulfoxide (DMSO), Dimethylformamide (DMF) and ethyl ether were used to determine their effects on lipase activity and stability. The lipase residual activity was measured by pNP release assay using pNPL as substrate. To determine the effects on enzyme activity, the standard reaction buffer was prepared with the addition of different solvents to yield the desired final solvent concentrations (10 or 20% v/v). For the lipase stability test the enzymes were incubated with different solvents (20% v/v) for 2 h at 22ºC, and the residual activity was measured.
2.5 Results and Discussion

Protein Sequence Analysis

LipY8p contains 371 amino acids (AAs) with a 28-AAs signal sequence, resulting in a 343-AAs mature protein. The lipase engineering database search indicated that the lipase belongs to the abH23 superfamily with a highly conserved GX pattern in the amino acid sequence \(^{24}\). Sequence alignment between LipY8 and the closely related \emph{Y. lipolytica} lipase genes exhibited 99.2\% identity with LipY (Uniprot: E0Z5H2), 99.2\% with Lip8 (Uniprot: Q872L3), 78.1\% with Lip7 (Uniprot: Q872L4) and 40.9\% with Lip2 (Uniprot: Q9P8F7). Blast analysis with the Uniprot database revealed homology of LipY8 to several yeast lipases such as those from \emph{Candida galli} (CgLIP8, 91.3\%; CgLIP7, 77.5\%), \emph{Candida deformans} (CdLIP3, 90.7\%; CdLIP2, 71\%), and \emph{Candida alimentaria} (CaLIP7, 66\%) (Figure 1). The conserved GHSLG(G/A)A motif characteristic of the triacylglycerol hydrolases, shared by the filamentous fungi lipase family \(^{25}\), was found at position \(\sim 190\). The lipase catalytic triad containing the serine, aspartic acid, and histidine residues were located at conserved positions. Eight highly conserved Cys residues were also found at conserved positions in all of these lipases and are hypothesized to form disulfide bridges (Figure 1).

Cloning and Recombinant Baculovirus Preparation

\emph{Pichia Pastoris} has historically been the first choice for over expression of yeast genes \(^ {26}\). However, here we secretly expressed the \emph{LIPY8} gene in baculovirus-infected insect cells, which also providing sufficient post-translational modification. The 1038-bp \emph{LIPY8} gene fragment was successfully amplified by PCR from plasmid pUC57-LipY8 using a primer pair designed for the pFastBacSP6His vector. The target gene was subcloned downstream of the \(P_{PH}\) promoter of the pFastBacSP6His vector in-frame with the N-terminal signal sequence and 6His tag. After amplification in \emph{E. coli} strain DH5\(\alpha\), the pFastBacSP6His-LipY8 recombinant plasmid was then introduced into the \emph{E. coli} host strain DH10Bac. Integration of the target gene was confirmed by PCR and further confirmed by sequencing using the primers described above. The recombinant Bacmid DNA was extracted and transfected into Sf9 cells with Cellfectin reagent. After 4 or 5 days of incubation, the P1 viral stock was prepared and further amplified to generate the P2 stock. The viral plaque assay indicated the titer of P2 viral stock reached approximately \(3.7 \times 10^8\) pfu/ml.

Expression of Recombinant LipY8p Lipase

The Sf9 cells were infected with recombinant virus from the P2 viral stock. The time course of recombinant extracellular lipase production was monitored by analyzing the activity of the culture medium every 24 h for up to 7 days (Figure 2). The initial lipase activity resulted from the introduction of the enzyme from the viral stock. Maximum lipase activity was attained 3 days post infection when cell viability decreased to around 75\%.
Thereafter, lipase activity stabilized until at least 7 days post infection, indicating strong resistance to protein degradation. The maximal value of lipase specific activity in culture medium reached 17.37 U/mg by \( p \)-Nitrophenyl (\( p \)NP) release assay at 3 days post infection. No activity was detected in the wild-type virus-infected cell group.

**Purification of Recombinant LipY8p Lipase**

Lipase homogenization was achieved using ammonium sulfate precipitation followed by Ni-NTA affinity and Mono Q anion exchange chromatography. In brief, the lipase solution obtained from dialysis after ammonium sulfate precipitation was applied to a Ni-NTA column. Stepwise elution with increasing concentrations of imidazole in buffer A was carried out. Peak 3 with the highest lipase activity was collected (Figure 3A). The active fractions were pooled and applied to a Mono Q HR 5/5 column. A linear gradient of increasing NaCl concentration from zero to 1 M was performed over 100 min and LipY8p was eluted at about 10 mS/cm conductivity, resulting in a homogenous preparation as evaluated by SDS-PAGE (Figure 3B and C). The purification process resulted in an approximate 25.7-fold purification factor and a final recovery of 23.2% of the enzyme protein with a molecular mass of 40 kDa and specific activity of 446.85 U/mg by \( p \)NP release assay (Table 1).

The LipY8p lipase is a glycoprotein and endoglycosidase treatment of the heterologously expressed lipase in *Pichia pastoris* showed a 2 kDa decrease in molecular mass. A similar result was observed for the lipase expressed by insect cells: the molecular mass of the heterologously expressed lipase was reduced by approximately 2 kDa after treatment with PNGase F (Figure 3C). Glycosylation is essential for the activity of a secretory expressed glycoprotein, and deglycosylation was reported to have a significant effect on enzyme activity. However, the residual activity of the LipY8p lipase following deglycosylation retained 90.1 ± 1.3% activity of the untreated protein.

**Effect of Temperature and pH on Lipase Activity and Stability**

The insect cells expressing LipY8p lipase exhibited an extraordinary cold-active property that was not observed in previous report. Cold-active lipases show optimal reaction temperatures at lower than 30°C. LipY8p had optimal activity at a temperature of 17°C and retained 70.6% of the highest activity at 8°C, which is similar to the reported cold-active lipases from *P. lynferdii* NRRL Y-7723, *Geotrichum* sp. SYBC WU-3 and *Candida albicans*. The optimal temperature of LipY8p is lower than many reported cold-active lipases, but higher than the lipase from *Microbacterium luteolum*. Moreover, similar to these reported cold-active lipases, the activity of LipY8p drastically declined as the temperature rose above 25°C and approached inactivity at temperatures above 45°C (Figure 4A). However, LipY8p showed less thermo stability.
than the cold-active lipases from *P. lynferdii* NRRL Y-7723, *Geotrichum* sp. SYBC WU-3 and *Candida albicans* 
7, 31-32. Its activity was essentially maintained from 0 to 30°C temperature, whereas a sharp decrease in stability 
was observed as temperatures rose above 35°C (Figure 4A).

The majority of cold-active microbial lipases exhibit optimal activity at near neutral or alkaline conditions 36. 
LipY8p showed considerable stability over the pH ranges 5–9 with optimal activity at pH 7.5 (Figure 4B), 
which is similar to the lipases from *Rhizomucor endophyticus* 36 and *Candida zeylanoides* 33. The wide range 
of stability of the lipase indicated its potential use in both acidic and alkaline conditions.

**Substrate Specificity of Lipase**

To investigate the substrate specificity of LipY8p, various lengths of *p*-Nitro phenyl esters were used as the 
substrates. The lipase showed the highest specific activity toward *p*-nitro phenyl decanoate (*p*NPD) (C10) 
(relative activity of 155.0%) at 791.3 ± 9.5 U/mg. *p*-nitro phenyl palmitate (*p*NPP) (C16), *p*-nitro phenyl myristate (*p*NPM) (C14) and *p*-nitro phenyl butyrate (*p*NPB) (C4) were equally utilized as substrates. The 
shorter carbon chain ester (C2) was poorly hydrolyzed (Figure 5). This indicated that LipY8p preferred 
medium chain esters 18, which is a typical property of the GX class lipase 24. Similar results were reported for 
cold-active lipase from *Pseudomonas proteolytica* (GBPL_Hb61)10 and *Pseudomonas* sp. strain KB700A40. 
However LipY8p lipase exhibited much higher hydrolysis activity toward olive oil with a specific activity of 
1102.9 U/mg (Figure 6), which was much higher than the AMS8 lipase (394.43U/mg) from Antarctic 
*Pseudomonas* sp.37.

**Effect of Metal Ions and Inhibitors on Lipase Activity**

Lipase activity was assayed in the presence of various metal ions at 1 mM concentrations (Table 2). 
Remarkable inhibition of the enzyme activity was observed in the presence of various transition metals such as 
Zn2+, Cu2+, as well as Ni2+. Similarly, cold-active lipases from *Psychrobacter cryohalolentis* K5T 38 were 
reported to be inhibited by these three metals, and lipase from Antarctic *Pseudomonas* (AMS8 lipase)37 and 
*Pseudomonas* sp. Strain B11-112 were inhibited by Zn2+, Cu2+ and Fe2+. The lipase activities were fairly stable 
in the presence of Mg2+ and Ca2+, and activated by K+ (118.4 ± 5.8%). In contrast, the presence of EDTA (1 
mM) resulted in a considerable stimulation of lipase activity (136.1 ± 4.5%), and the inhibitory effect of Ni2+ 
was eliminated by the addition of EDTA, indicating that the lipase was not a metalloenzyme. Similar results 
were reported for the YILip2 lipase from *Yarrowia lipolytica* 41 and the lipase from *Psychrobacter cryohalolentis* 
K5T 38. In contrast, certain cold-active lipases require metal ions as the enzyme cofactor40. Of interest are the 
failed attempts to inhibit lipase activity through reduction of disulfide bonds in the protein despite sequence 
analysis revealing that the protein may contain several conserved disulfide bonds (Figure 1). The addition of
different concentrations of β-mercaptoethanol (β-ME) to the reaction buffer gave rise to significant activation of lipase activity. However, the simultaneous addition of 1 mM β-ME and 0.1% (w/v) n-octyl-β-D-glucoside (β-OG) led to a marked inactivation of the enzyme (77.2 ± 2.1%). This observation indicated that additional destabilizing factors, such as a detergent, were necessary for the reductant to gain access to the disulfide bond. The activation effect of β-ME on enzyme activity was also reported with lipases from the *P. aeruginosa* mutant and *S. bambergiensis* OC 25-4 where lipase activity was enhanced by 19.6% and 8%, respectively, after treatment with a concentration of 0.1% (v/v) β-ME. This can be explained by the requirement for sulfhydryl groups for lipase activity. As a serine protease inhibitor, phenylmethanesulfonyl fluoride (PMSF) (1 mM and 4 mM) showed significant inhibitory effects with 80.1% and 47.6% residual activities, respectively, demonstrating that the lipase is of the serine hydrolase type.

**Effect of Detergents on Lipase Activity and Stability**

Detergents such as Tween-20 and Triton-X100 are commonly used as emulsifying agents to improve the emulsion of substrates, thereby making the substrate more accessible. However, the present of detergents in the reaction system may affect the catalytic activity of lipase depending on the concentration used. At a concentration of 0.1% (v/v or w/v), the detergents SDS, Triton-X100, NP40, Tween-20 and β-DDM strongly inhibited LipY8p activity. β-OG had a mild positive effect at 0.1% w/v (105.5 ± 2.2%), but significantly inhibited lipase activity as the concentration increased to 0.3% w/v (2.2 ± 0.1%). The inhibitory effect was also observed with C₁₂E₈ at 0.001% w/v (63.1 ± 5.2%) and 0.002% w/v (7.9 ± 1.4%) (Table 3). These results can likely be attributed to the hydrophobic property of the long chains of these detergents making them act as substrates and therefore competitive inhibitors of the enzyme.

Although most of the tested detergents have inhibitory effect on LipY8p activity when present in the reaction buffer system, almost all of them exhibited activating effect on the enzyme activity when be added into the enzyme stock for pre-incubation. When LipY8p was pre-incubated with 0.1% (v/v or w/v) Triton-X100 (128.1 ± 4.1%), β-OG (147.1 ± 0.3%), or C₁₂E₈ (144.2 ± 5.5%) for 2 h at 22°C, we observed a strong activation of lipase activity (Table 4). This positive effect on lipase activity was retained for up to 72 h for Triton-X100 and longer for β-OG and C₁₂E₈ (Table 4). Thus, this indicates that the detergents were able to weaken the hydrophobic interaction within the lipase protein, resulting in disaggregation and stabilization of the enzyme. However, as the incubation time increased, the denaturation effects of these detergents became dominant and the enzyme activity decreased. Pre-incubating of lipase with 0.1% (w/v) β-DDM produced a sharp decrease in lipase activity (39.9 ± 3.3%), and the destabilization effect was more pronounced with SDS, NP40 and Tween-20 at the same concentration (Table 4). This suggested that the lipase showed greater sensitivity to these detergents, which may have induced conformational changes and denaturation of the protein.
Effect of Organic Solvents on Lipase Activity and Stability

Enzymes could be used to perform reactions in organic solvents that are not possible in aqueous systems. However, activity and stability of enzymes in organic solvents show a strong dependence on the nature of the enzymes\textsuperscript{49}. As proteins, enzymes tend to lose their activity in solutions containing higher than 10–20% organic co-solvents\textsuperscript{50}. Thus, reaction buffers containing 10% or 20% (v/v) various organic solvents were used to examine their effects on lipase activity and stability. LipY8p activity was dramatically increased by the presence of 20% (v/v) DMSO (416.9 ± 22.7%) during the reaction. A similar phenomenon was also observed for P. fluorescens lipase whose activity increased up to 4.0-fold in the presence of 50% (v/v) DMSO\textsuperscript{51}. This significant activation of lipase activity may be attributed to a conformational change and increased flexibility of the protein caused by the solvent. Activation of lipase activity was also observed with 20% (v/v) methanol (180.3 ± 1.7%). Similar effects were recorded for ethanol (266.7 ± 5.4%), acetone (361.5 ± 11.3%) as well as isopropanol (174.4 ± 5.3%) at concentrations of 10% (v/v). These became inhibitory as the concentration increased to 20% (v/v) (Table 5). The effects of organic solvents on lipase stability are recorded in Table 6. The enzyme lost almost 90% activity after exposure to 20% (v/v) ethyl ether or Dimethylformamide (DMF) at 22°C for 2 h, and activity was virtually eliminated following addition of acetone, ethanol or isopropanol. However, LipY8p exhibited relatively higher stability in methanol and DMSO retaining 72.8 ± 1.4% and 88.1 ± 4.6% residual activity, respectively, after treatment. These results suggest that longer chain length alcohols have a stronger inhibitory effect. Binding of a thin layer of water molecules to the surface is essential for the enzyme protein to maintain its native conformation\textsuperscript{41}. Water is a particular solvent type that shows lower affinity toward the protein surface in comparison to water-miscible organic solvents \textsuperscript{52}. Water patches on the protein surface are formed by a limited number of directly-bound water molecules and also by water–water interactions. Thus, the presence of water-miscible organic solvents deprives the enzyme of bound water leading to enzyme inactivation. Lipases show diversity in their tolerance to water-miscible organic solvents \textsuperscript{46}. The cold-active lipase from Pseudomonas proteolytica (GBPI_Hb61) showed decreased stability after a 30-min exposure to various water-miscible organic solvents, with the exception of methanol (103.5%) \textsuperscript{10}.

2.6 Conclusion

In the present work, we report the cloning and expression of the LIPY8 gene by baculovirus expression system, as well as purification and characterization of the enzyme. The results from this study revealed that the purified recombinant enzyme was highly active in cold temperatures ranging from 8 to 21°C with maximal activity at 17°C. The lipase showed high stability over a wide range of pH values from 5 to 9 with optimal activity at 7.5. The enzyme also exhibited stability in the presence of a selection of inhibitors, metal ions, detergents and organic solvents. It is particularly interesting that the LipY8p expressed by insect cells showed a marked difference in enzymatic characterization with regard to optimal pH values and temperatures to that
expressed by *Pichia Pastoris* reported by Song et al. These differences also exist between LipY8p and LipY despite both originating from marine *Y. lipolytica* and sharing high sequence identity. To the best of our knowledge, with regard to closely related lipases of LipY8p, only the 3D structure of *Y. lipolytica* Lip2 lipase (40.9% identity) has been solved. Thus, solving the 3D structure of LipY8p will shed light on the enzyme structure and function, and also contribute to the understanding of enzymatic activities at low temperatures as well as their optimization for biotechnological applications.

**Acknowledgments**

We thank Dr. Muriel Steel for her editing of the manuscript.

**Funding**

This work was with the International collaboration program (PSR-SIIRI) supported by and the Ministry of Economy, Science and Innovation of Quebec, Canada (MESI) and the Ministry of Science and Technology of China.

**Conflict of interest statement**

The authors declare that they have no conflict of interest.
2.7 Reference


35. Yan, Q.; Duan, X.; Liu, Y.; Jiang, Z.; Yang, S., Expression and characterization of a novel 1,3-regioselective cold-adapted lipase from Rhizomucor endophyticus suitable for biodiesel synthesis. Biotechnol Biofuels 2016, 9, 86.


Figure 1. Multiple sequence alignments between LipY8p and highly homologous lipases from Y. lipolytica (LIPY, E0Z5H2; LIP8, Q872L3; LIP7, Q872L4; LIP2, Q9P8F7), Candida galli (CgLIP8, A0A078BRV6; CgLIP7, A0A078BNS3), Candida deformans (CdLIP3, Q875G8; CdLIP2, Q875G9) and Candida alimentaria (CaLIP7, A0A078BMP3). Cysteine residues are marked in gray and conserved residues of the active site including serine, aspartic acid and histidine are marked in black.
Figure 2. Time course of LipY8p expression and Sf9 cell viability. Lipase activity reached a plateau three days after infection with virus.
Figure 3. Purification of the recombinant LipY8p lipase. (A) Ni-NTA affinity chromatography. Peak1, peak2 and peak3 were eluted by buffer A containing 5 mM, 20 mM and 150 mM imidazole, respectively; (B) Mono Q anion exchange. LipY8p was eluted at around 10 mS/cm conductivity. (C) SDS-PAGE analysis of purified LipY8p. M, protein marker; 1, purified LipY8p; 2, LipY8p after deglycosylation.
Figure 4. Optimum temperature (A) and optimum pH (B) on activity and stability of LipY8p lipase.
Figure 5. Substrate specificity of LipY8p lipase against different chain length pNP esters. Activity of pNP dodecanoate (pNPL) was considered as 100%. pNPP, pNP palmitate; pNPM, pNP myristate; pNPD, pNP decanoate; pNPB, pNP butyrate; pNPA, pNP acetate.
Figure 6. Quantification of fatty acid released by LipY8p hydrolysis of olive oil. (A) Hydrolysis of olive oil in RhB-OOe leads to fluorescence emission. (B) Standard curve prepared with RhB-OOe using 3–18 mM oleic acid. (C) Quantification of fatty acid released by LipY8p hydrolysis of olive oil. The excitation wavelength was set to 350 nm, and fluorescence emission was recorded at 580 nm. Each measurement was performed three times, and standard deviations were indicated.
## Tables

### Table 1. A summary of LipY8p lipase purification

<table>
<thead>
<tr>
<th>Purification steps</th>
<th>Protein (mg)</th>
<th>Lipase activity (kU(^a))</th>
<th>Specific activity (U/mg)</th>
<th>Yield (%)</th>
<th>Purification (fold)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Culture medium</td>
<td>104.37</td>
<td>1.81</td>
<td>17.37</td>
<td>100</td>
<td>1</td>
</tr>
<tr>
<td>Ammonium sulfate</td>
<td>61.93</td>
<td>1.57</td>
<td>25.36</td>
<td>86.63</td>
<td>1.46</td>
</tr>
<tr>
<td>Ni-NTA column</td>
<td>3.61</td>
<td>0.64</td>
<td>177.08</td>
<td>35.30</td>
<td>10.19</td>
</tr>
<tr>
<td>Mono Q column</td>
<td>0.94</td>
<td>0.42</td>
<td>446.85</td>
<td>23.19</td>
<td>25.72</td>
</tr>
</tbody>
</table>

\(^a\)Activity test was carried out by spectrophotometer in phosphate buffer pH 7.5 at 17ºC, using pNPL as substrate. One unit (U) of enzyme activity was defined as the amount of enzyme required for the liberation of 1.0 μmol p-nitrophenol per min under the assay conditions.
Table 2. Effect of metal ions and inhibitors on lipase activity

<table>
<thead>
<tr>
<th>Compounds</th>
<th>Concentration (mM)</th>
<th>Residual activity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>none</td>
<td>$100 \pm 1.18$</td>
</tr>
<tr>
<td>β-ME</td>
<td>1</td>
<td>$117.6 \pm 1.22$</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>$129.8 \pm 8.32$</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>$125.8 \pm 1.69$</td>
</tr>
<tr>
<td>PMSF</td>
<td>1</td>
<td>$80.1 \pm 2.22$</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>$47.6 \pm 3.92$</td>
</tr>
<tr>
<td>KCl</td>
<td>1</td>
<td>$118.4 \pm 5.83$</td>
</tr>
<tr>
<td>CaCl₂</td>
<td>1</td>
<td>$96.3 \pm 3.99$</td>
</tr>
<tr>
<td>MgCl₂</td>
<td>1</td>
<td>$96.8 \pm 1.50$</td>
</tr>
<tr>
<td>ZnSO₄</td>
<td>1</td>
<td>$1.2 \pm 0.10$</td>
</tr>
<tr>
<td>CuCl₂</td>
<td>1</td>
<td>$12.5 \pm 0.02$</td>
</tr>
<tr>
<td>NiSO₄</td>
<td>1</td>
<td>$51.8 \pm 2.11$</td>
</tr>
<tr>
<td>EDTA</td>
<td>1</td>
<td>$136.1 \pm 4.49$</td>
</tr>
</tbody>
</table>
Table 3. Effect of detergents on lipase activity when present in reaction buffer

<table>
<thead>
<tr>
<th>Detergents</th>
<th>Concentration (v/v or w/v)</th>
<th>Residual activity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>none</td>
<td>100 ± 1.18</td>
</tr>
<tr>
<td>SDS</td>
<td>0.1%</td>
<td>0.89 ± 0.16</td>
</tr>
<tr>
<td>Triton-X100</td>
<td>0.1%</td>
<td>-</td>
</tr>
<tr>
<td>NP40</td>
<td>0.1%</td>
<td>-</td>
</tr>
<tr>
<td>Tween-20</td>
<td>0.1%</td>
<td>0.14 ± 0.02</td>
</tr>
<tr>
<td>β-OG</td>
<td>0.1%</td>
<td>105.5 ± 2.18</td>
</tr>
<tr>
<td></td>
<td>0.3%</td>
<td>2.2 ± 0.10</td>
</tr>
<tr>
<td>β-DDM</td>
<td>0.1%</td>
<td>-</td>
</tr>
<tr>
<td>C₁₂E₈</td>
<td>0.001%</td>
<td>63.1 ± 5.16</td>
</tr>
<tr>
<td></td>
<td>0.002%</td>
<td>7.9 ± 1.38</td>
</tr>
</tbody>
</table>

-, Activity undetectable.
<table>
<thead>
<tr>
<th>Detergents</th>
<th>Concentration (v/v or w/v)</th>
<th>Residual activity (%)</th>
<th>Incubation time (h)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>none</td>
<td>100 ± 1.85</td>
<td>2</td>
</tr>
<tr>
<td>SDS</td>
<td>0.1%</td>
<td>-</td>
<td>2</td>
</tr>
<tr>
<td>Triton-X100</td>
<td>0.1%</td>
<td>128.1 ± 4.05</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>0.1%</td>
<td>140.2 ± 4.15</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>0.1%</td>
<td>110.4 ± 2.75</td>
<td>48</td>
</tr>
<tr>
<td></td>
<td>0.1%</td>
<td>96.4 ± 4.44</td>
<td>72</td>
</tr>
<tr>
<td>NP40</td>
<td>0.1%</td>
<td>-</td>
<td>2</td>
</tr>
<tr>
<td>Tween-20</td>
<td>0.1%</td>
<td>-</td>
<td>2</td>
</tr>
<tr>
<td>β-OG</td>
<td>0.1%</td>
<td>147.1 ± 0.25</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>0.1%</td>
<td>124.7 ± 4.19</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>0.1%</td>
<td>119.2 ± 2.83</td>
<td>48</td>
</tr>
<tr>
<td></td>
<td>0.5%</td>
<td>143.9 ± 4.12</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>0.5%</td>
<td>127.7 ± 1.58</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>0.5%</td>
<td>121.6 ± 2.49</td>
<td>48</td>
</tr>
<tr>
<td>β-DDM</td>
<td>0.1%</td>
<td>39.9 ± 3.29</td>
<td>2</td>
</tr>
<tr>
<td>C_{12}E_8</td>
<td>0.1%</td>
<td>144.2 ± 5.50</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>0.01%</td>
<td>140.0 ± 2.85</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>0.1%</td>
<td>140.8 ± 3.56</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>0.1%</td>
<td>130.0 ± 1.88</td>
<td>48</td>
</tr>
</tbody>
</table>

-, Activity undetectable.
Table 5. Effect of organic solvents on lipase activity

<table>
<thead>
<tr>
<th>Solvents</th>
<th>Concentration (v/v)</th>
<th>Residual activity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>none</td>
<td>100 ± 7.15</td>
</tr>
<tr>
<td>Iso-propanol</td>
<td>10%</td>
<td>174.4 ± 5.27</td>
</tr>
<tr>
<td></td>
<td>20%</td>
<td>5.2 ± 0.29</td>
</tr>
<tr>
<td>Methanol</td>
<td>20%</td>
<td>180.3 ± 1.71</td>
</tr>
<tr>
<td>Ethanol</td>
<td>10%</td>
<td>266.7 ± 5.4</td>
</tr>
<tr>
<td></td>
<td>20%</td>
<td>50.1 ± 3.51</td>
</tr>
<tr>
<td>Acetone</td>
<td>10%</td>
<td>361.5 ± 11.3</td>
</tr>
<tr>
<td></td>
<td>20%</td>
<td>38.7 ± 2.09</td>
</tr>
<tr>
<td>DMSO</td>
<td>20%</td>
<td>416.9 ± 22.7</td>
</tr>
</tbody>
</table>
Table 6. Stability of lipase in different organic solvents

<table>
<thead>
<tr>
<th>Solvents</th>
<th>Concentration (v/v)</th>
<th>Residual activity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>none</td>
<td>100 ± 2.24</td>
</tr>
<tr>
<td>Methanol</td>
<td>20%</td>
<td>72.81 ± 1.37</td>
</tr>
<tr>
<td>Ethanol</td>
<td>20%</td>
<td>0.90 ± 0.03</td>
</tr>
<tr>
<td>Acetone</td>
<td>20%</td>
<td>1.38 ± 0.04</td>
</tr>
<tr>
<td>Iso-propanol</td>
<td>20%</td>
<td>-</td>
</tr>
<tr>
<td>DMSO</td>
<td>20%</td>
<td>88.05 ± 4.64</td>
</tr>
<tr>
<td>DMF</td>
<td>20%</td>
<td>13.21 ± 0.13</td>
</tr>
<tr>
<td>Ethyl Ether</td>
<td>20%</td>
<td>10.86 ± 0.16</td>
</tr>
</tbody>
</table>

-, Activity undetectable.