

UNIVERSIDADE FEDERAL DE SANTA CATARINA  
CENTRO DE CIÊNCIAS BIOLÓGICAS  
DEPARTAMENTO DE BIOQUÍMICA  
CURSO DE LICENCIATURA EM CIÊNCIAS BIOLÓGICAS

Tâmela Zamboni Madaloz

**Caracterização Estrutural *in silico* dos Receptores Nucleares de *Crassostrea gigas*  
(Thunberg, 1793): Interações com Xenobióticos Estradiol e Tamoxifeno.**

Florianópolis

2019

Tâmela Zamboni Madaloz

**Caracterização Estrutural *in silico* dos Receptores Nucleares de *Crassostrea gigas*  
(Thunberg, 1793): Interações com Xenobióticos Estradiol e Tamoxifeno.**

Trabalho de Conclusão do Curso de Graduação em Ciências Biológicas do Centro de Ciências Biológicas da Universidade Federal de Santa Catarina como requisito para a obtenção do título de Licenciada em Ciências Biológicas.

Orientador: Prof. Dr. Guilherme Razzera Maciel

Coorientadora: Karin dos Santos

Florianópolis

2019

Ficha de identificação da obra elaborada pelo autor,  
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Madaloz, Tâmelá Zamboni

Caracterização Estrutural in silico dos Receptores Nucleares de *Crassostrea gigas* (Thunberg, 1793): Interações com Xenobióticos Estradiol e Tamoxifeno. / Tâmelá Zamboni Madaloz ; orientador, Guilherme Razzera, coorientadora, Karin dos Santos, 2019.

86 p.

Trabalho de Conclusão de Curso (graduação) - Universidade Federal de Santa Catarina, Centro de Ciências Biológicas, Graduação em Ciências Biológicas, Florianópolis, 2019.

Inclui referências.

1. Ciências Biológicas. 2. Bioquímica, Bioinformática, Ecotoxicologia, Receptor nuclear. I. Razzera, Guilherme . II. dos Santos, Karin . III. Universidade Federal de Santa Catarina. Graduação em Ciências Biológicas. IV. Título.

Tâmela Zamboni Madaloz

**Caracterização Estrutural *in silico* dos Receptores Nucleares de *Crassostrea gigas* (Thunberg, 1793): Interações com Xenobióticos Estradiol e Tamoxifeno.**

Este Trabalho de Conclusão de Curso foi julgado adequado para obtenção do Título de Licenciada em Ciências Biológicas e aprovado em sua forma final pelo Curso de Ciências Biológicas.

Florianópolis, 13 de dezembro de 2019.

Prof. Carlos Roberto Zanetti, Dr.

Coordenador do Curso

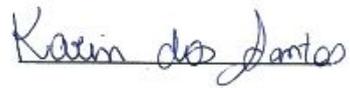
**Banca Examinadora:**



Prof. Guilherme Razzera Maciel, Dr.

Orientador

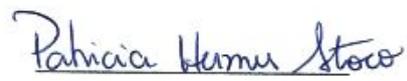
Instituição UFSC



Karin dos Santos

Coorientadora

Instituição UFSC



Profa. Patricia Hermes Stoco, Dra.

Avaliadora

Instituição UFSC



Jacó Joaquim Mattos, Msc

Avaliador

Instituição UFSC

Este trabalho é dedicado ao desenvolvimento e expansão de métodos alternativos ao uso de animais na ciência, para o progresso de uma ciência com inovação, tecnologia e ética.

## **AGRADECIMENTOS**

Agradeço pelo apoio da minha família e dos meus amigos pelas diversas formas de suporte que eu tive o privilégio de obter na realização deste trabalho e ao longo de todo o curso de Ciências Biológicas. Este trabalho não seria possível sem a dedicação, paciência e entrega de meu orientador, Guilherme Razzera e coorientadora Karin dos Santos. Aos membros do Laboratório de Bioinformática Estrutural da Universidade Federal de Santa Catarina, obrigada pelo acolhimento. Os laboratórios parceiros Laboratório de Biomarcadores de Contaminação Aquática e Imunoquímica e Laboratório de Genômica contribuíram com dados, conselhos e espaços de extrema importância para o desenvolvimento do trabalho. Ademais, agradeço a experiência de um curso de excelência em uma instituição de extrema qualidade como a UFSC, que permite transcender barreiras e fornecer uma formação enquanto pessoa, cidadã e profissional.

## RESUMO

Compostos contaminantes, como fármacos e hormônios produzidos por humanos, são encontrados em ambientes aquáticos e possuem potencial de interagir com componentes dos seres vivos que habitam esses locais contaminados. A superfamília dos receptores nucleares, devido a características estruturais de ligação com moléculas que inibem ou induzem a transcrição de genes, são frequentes alvos desses compostos xenobióticos. O presente trabalho busca entender a relação entre os receptores nucleares da ostra do Pacífico (*Crassostrea gigas*), uma espécie filtradora sésil sujeita a mudanças ambientais e exposição a contaminantes, com ligantes xenobióticos que possam interagir com seus sítios de ligação e modular a expressão de genes. Por se tratarem de muitos compostos xenobióticos introduzidos no ambiente e muitos organismos para estudo, análises *in vivo* são dificultadas e a utilização de metodologias *in silico* podem verificar muitas interações possíveis entre xenobióticos e seres vivos. Para entender os mecanismos de interação destes receptores foi feita uma análise da superfamília de receptores nucleares de *C.gigas*. Nossos resultados mostram que a partir de modelagem 3D e caracterização de cavidade do sítio de ligação é possível prever características estruturais para proteínas e selecionar candidatos a interagirem com compostos de interesse em organismos invertebrados não modelo. Através de técnicas *docking* molecular demonstramos interações favoráveis de três receptores nucleares da ostra do Pacífico, NR1P2, NR2E1 e NR2E5 com xenobióticos estradiol e tamoxifeno.

**Palavras-chave:** Bioinformática. Ecotoxicologia. Receptores nucleares.

## ABSTRACT

Contaminant compounds, such as human-produced drugs and hormones, are found in aquatic environments with the potential to interact with living beings components that inhabit these contaminated sites. The nuclear receptor superfamily, due to structural binding characteristics with molecules that inhibit or induce gene transcription, is a frequent target of such xenobiotic compounds. The present work aims to understand the relation between Pacific oyster (*Crassostrea gigas*) nuclear receptors, a protein superfamily from a filter-feeding sessile species submitted to environmental changes and exposure to contaminants, with xenobiotic ligands that can interact with their binding sites and modulate gene expression. Since many xenobiotic compounds are introduced into the environment which also have many organisms to study, *in vivo* analysis may be difficult and the use of *in silico* approaches are able to evaluate the potential interactions between xenobiotics and living beings. In order to understand the mechanisms of these receptors, a *C.gigas* nuclear receptor superfamily analysis was performed. Our data showed that through 3D protein modeling and binding site cavity characterization was possible to predict structural characteristics of the nuclear receptors and also select candidates to interact with target compounds using a non-model invertebrate organism. Through molecular docking technique we demonstrate favorable interactions of three Pacific oyster nuclear receptors, NR1P2, NR2E1 and NR2E5 with xenobiotics estradiol and tamoxifen.

**Keywords:** Bioinformatics. Ecotoxicology. Nuclear receptors.

## LISTA DE FIGURAS

- Figura 1** - Organização dos domínios de um receptor nuclear na proteína. A região A/B, N-terminal, contém a função de ativação AF-1. O domínio de interação com o DNA, região C, inclui dois dedos de zinco. A região D, uma dobradiça (hinge) variável, separa o domínio DBD do Domínio de Interação com o Ligante LBD (região E), o qual possui o fator de ativação em hélice AF-2. A porção C-terminal F é localizada ao final da sequência. ....20
- Figura 2** - Enovelamento tridimensional típico de receptores nucleares. (A) Estrutura quaternária recorrente em NRs, destacando o Domínio de Interação com o Ligante (LBD) e o Domínio de Interação com o DNA (DBD) interagindo com uma molécula de DNA. (B) Estrutura do domínio LBD com  $\alpha$ -hélices bem definidas, um grampo- $\beta$  entre as  $\alpha$ -hélices H6 e H7 e um ligante associado ao sítio de ligação (em amarelo). ....21
- Figura 3** - Estrutura química dos ligantes estradiol e 4-hidroxitamoxifeno. ....24
- Figura 4** - Esquema das etapas metodológicas do trabalho, com os programas utilizados, na sequência de sua realização ao longo do desenvolvimento do estudo. ....29
- Figura 5** - Enovelamento tridimensional correspondente ao modelo controle CgNR3A e às estruturas cristalográficas de referência para este modelo. (A) CgER modelo (verde). (B) CgER estrutura experimental 4N1Y (azul). (C) ER humano 1GWR (roxo). ....34
- Figura 6** - Alinhamento entre as estruturas do CgER modelo (verde) e estrutura experimental CgER 4N1Y (azul). ....34
- Figura 7** - Comparação do volume de superfície da cavidade entre o ER humano, a estrutura cristalográfica de CgER e o modelo CgER, com a presença do estradiol interno ao sítio de ligação. (A) Superfície de volume da cavidade do sítio de ligação da estrutura cristalográfica de ER humano (1GWR), suficientemente grande para comportar o estradiol. (B) Superfície de volume da estrutura experimental de CgER (4N1Y), apresentando cavidade reduzida para comportar o ligante estradiol, assim como o modelo de CgER (C). ....36
- Figura 8** - Estruturas tridimensionais obtidas através da técnica de modelagem. (A) Modelo CgNR1P2, o qual não apresenta formação de alfa-hélice na região que corresponderia a H4 e H6. (B) Modelo CgNR2E1 e (C) Modelo CgNR2E5, ambos pertencentes ao grupo NR2E, não apresentam formação de H2. ....44
- Figura 9** - Enovelamento tridimensional correspondente ao modelo controle e às estruturas cristalográficas de referência para este modelo. Estrutura 3D do modelo CgER (A), da estrutura cristalográfica de CgER 4N1Y (B) e da estrutura cristalográfica de ER humano 1GWR (C). Todos apresentam 11 alfa-hélices bem definidas, com ausência de H2, recorrente nos membros desse grupo. ....46
- Figura 10** - Alinhamento dos esqueletos carbônicos do modelo CgER (verde) e da estrutura resolvida experimentalmente de CgER 4N1Y (azul). ....47

**Figura 11** - Comparação do volume de superfície da cavidade entre o ER humano, a estrutura cristalográfica de CgER e o modelo CgER, com a presença do estradiol interno ao sítio de ligação (A) Superfície de volume da cavidade do sítio de ligação da estrutura cristalográfica de ER humano (1GWR), suficientemente grande para comportar o estradiol. (B) Superfície de volume da cavidade do sítio de ligação da estrutura cristalográfica de CgER (4N1Y), apresentando cavidade reduzida para comportar o ligante estradiol, assim como o modelo de CgER (C). .....49

**Figura 12** - Interações químicas entre resíduos de aminoácidos que compõem o sítio de ligação do domínio LBD da estrutura cristalográfica de ER humano (1GWR) com o estradiol. Linhas tracejadas pretas indicam ligações de hidrogênio, ocorrendo com Glu353, Arg394 e His524. Linhas sólidas verdes mostram interações hidrofóbicas, que ocorrem com os resíduos Ala350, Leu346, Leu384, Met388 e Phe404. Linhas tracejadas verdes mostram interações  $\pi$ - $\pi$ , que ocorre entre o anel aromático do resíduo Phe404 e o anel aromático do estradiol. ....52

**Figura 13** - Alinhamento da estrutura cristalográfica do ER humano (1GWR), com o ligante estradiol (em roxo) interno à cavidade do sítio de ligação e do modelo CgNR2E1 com o estradiol (em verde) interno ao sítio de ligação na posição de menor RMSD em relação à posição que este ligante ocupa no sítio do ER humano. ....54

**Figura 14** - Interações químicas entre o estradiol e os resíduos de aminoácidos que compõem a cavidade do sítio de ligação do modelo CgNR2E1. Nessa posição há formação de uma ponte de hidrogênio entre o resíduo Arg224 e o estradiol, além de interações hidrofóbicas com Phe120, Leu221 e Arg224.....55

**Figura 15** - Interações químicas entre o estradiol e os resíduos de aminoácidos que compõem a cavidade do sítio de ligação do modelo CgNR2E5. Na posição mostrada, muitas interações hidrofóbicas ocorrem entre o estradiol e os resíduos Cys73, Phe76, Met111, Val177, Ala199 e Ile203, além de uma ligação  $\pi$ - $\pi$  entre o anel aromático de Phe76 e o anel aromático do estradiol.....57

**Figura 16** - Interações químicas entre o estradiol e os resíduos de aminoácidos que compõem a cavidade do sítio de ligação do modelo CgNR1P2. Posição 1 (A) apresenta ponte de hidrogênio com o resíduo Gly108, diversas interações hidrofóbicas com resíduos de Glu73, Phe74, Met77, Leu124, Trp141, além de uma ligação  $\pi$ - $\pi$  entre Phe74 e o anel aromático do estradiol.....59

**Figura 17** - (A) Domínio LBD do ER humano com o ligante agonista estradiol presente no sítio de ligação. (B) Domínio LBD do ER humano com o ligante antagonista 4-hidroxitamoxifeno presente no sítio de ligação. É possível notar que os diferentes ligantes causam conformações diferentes no domínio LBD, principalmente no posicionamento da hélice 12, o que influencia na ligação com co-ativadores e correpressores. O ligante estradiol em A e o tamoxifeno em B, estão representados em amarelo. ....61

**Figura 18** - Interações químicas entre resíduos de aminoácidos que compõem o sítio de ligação do domínio LBD da estrutura experimental de um ER humano (3ERT) cristalizado com 4-hidroxitamoxifeno. Linhas tracejadas pretas indicam ligações de hidrogênio, ocorrendo com Arg394 e Glu353. Linhas sólidas verdes mostram interações hidrofóbicas, que ocorrem com os resíduos Met343, Leu346, Ala350, Trp383, Leu387, Leu391, Met421, Ile424 e Leu525.....62

**Figura 19** - Alinhamento da estrutura cristalográfica do ER humano (3ERT), com o ligante 4-hidroxitamoxifeno (em roxo) interno à cavidade do sítio de ligação e do modelo CgNR2E1 com o 4-hidroxitamoxifeno (em verde) interno ao sítio de ligação na posição de menor RMSD em relação à posição que este ligante ocupa no sítio do ER humano. ....64

**Figura 20** - Interações químicas entre o tamoxifeno e os resíduos de aminoácidos que compõem a cavidade do sítio de ligação do modelo CgNR2E1. Esta posição mostra a formação de interações hidrofóbicas com Met44, Ala48, Phe85, Val86, Phe120, Leu221, Arg224, além de uma ligação  $\pi$ - $\pi$  entre Phe85 e um anel aromático de TAM. ....64

**Figura 21** - Interações químicas entre o tamoxifeno e os resíduos de aminoácidos que compõem a cavidade do sítio de ligação do modelo CgNR2E5. Há formação de uma ponte de hidrogênio com Arg206, ligação  $\pi$ - $\pi$  entre Phe76 e dois anéis aromáticos de TAM, além de muitas interações hidrofóbicas envolvendo os resíduos Phe36, Phe76, Ala80, Met111, Ile203 e Leu218. ....66

**Figura 22** - Interações químicas entre o tamoxifeno e os resíduos de aminoácidos que compõem a cavidade do sítio de ligação do modelo CgNR1P2. A posição apresentou diversas interações hidrofóbicas com resíduos de Ala70, Phe74, Met77, Asn110, Glu112, Trp141, além de uma ligação  $\pi$ - $\pi$  entre Phe74 e 2 anéis aromáticos do 4-hidroxitamoxifeno. ....67

**Figura 23** - Análise qualitativa das três posições de menor RMSD mostram as interações químicas entre o estradiol e os resíduos de aminoácidos que compõem a cavidade do sítio de ligação do modelo CgNR2E1. A posição de menor RMSD (A) mostra a formação de uma ponte de hidrogênio entre o resíduo Arg224 e o estradiol, além de interações hidrofóbicas com Phe120, Leu221 e Arg224 na Posição 1. A segunda posição de menor RMSD (B) apresenta interações hidrofóbicas com os resíduos Phe120, Pro41, Cys45, Leu221, além de uma ligação  $\pi$ - $\pi$  entre Phe85 e o anel aromático do estradiol. A terceira posição de menor RMSD (C) também identificou interações hidrofóbicas com Phe120 e Leu221 com formação de ligação  $\pi$ - $\pi$  com Phe120. ....80

**Figura 24** - Análise qualitativa das três posições de menor RMSD mostrando as interações químicas entre o estradiol e os resíduos de aminoácidos que compõem a cavidade do sítio de ligação do modelo CgNR2E5. Na posição de menor RMSD (A) muitas interações hidrofóbicas ocorrem entre o estradiol e os resíduos Cys73, Phe76, Met111, Val177, Ala199 e Ile203, além de uma ligação  $\pi$ - $\pi$  com Phe76. A segunda posição de menor RMSD (B) e a terceira (C) configuraram as mesmas interações químicas fracas no sítio de ligação, compostas por uma ponte de hidrogênio com Asn32 e interações hidrofóbicas com Met111, Ala199 e Ile203. ....81

**Figura 25** - Análise qualitativa das três posições de menor RMSD mostrando as interações químicas entre o estradiol e os resíduos de aminoácidos que compõem a cavidade do sítio de ligação do modelo CgNR1P2. A posição de menor RMSD (A) apresenta ponte de hidrogênio com o resíduo Gly108, diversas interações hidrofóbicas com resíduos de Glu73, Phe74, Met77, Leu124, Trp141, além de uma ligação  $\pi$ - $\pi$  entre Phe74 e o anel aromático do estradiol. Na segunda posição de menor RMSD (B) também ocorreu formação de ponte de hidrogênio entre o estradiol, nesta posição com o resíduo Asn110. Uma ligação  $\pi$ - $\pi$  com Phe74 e outras interações hidrofóbicas foram encontradas com Met77 e Trp141. Interações polares não foram

identificadas para a terceira posição de menor RMSD (C), a qual possui ligação  $\pi$ - $\pi$  com Phe74, além de interações hidrofóbicas com Met77 e Trp141. ....81

**Figura 26** - Análise qualitativa das três posições de menor RMSD mostrando as interações químicas entre o tamoxifeno e os resíduos de aminoácidos que compõem a cavidade do sítio de ligação do modelo CgNR2E1. A posição de menor RMSD (A) mostra a formação de interações hidrofóbicas com Met44, Ala48, Phe85, Val86, Phe120, Leu221, Arg224, além de uma ligação  $\pi$ - $\pi$  entre Phe85 e um anel aromático de TAM. A segunda posição de menor RMSD (B) apresenta uma ligação de hidrogênio com o resíduo Pro113. Interações hidrofóbicas ocorrem com os resíduos Met44, Leu101, Ala103, Val117, Phe120, Ile124, Leu127, Leu221 e Arg224 e o resíduo Phe120 faz ligação  $\pi$ - $\pi$  com um anel aromático do ligante. A terceira posição de menor RMSD (C) também identificou interações hidrofóbicas com Met44, Val47, Ala48, Phe85, Ala89, Phe120, Ile124 e Leu221, além de duas ligações  $\pi$ - $\pi$  entre Phe120 e Phe85 com anéis aromático do tamoxifeno e uma ligação  $\pi$ -cátion entre Phe85 e um nitrogênio de TAM. ....83

**Figura 27** - Análise qualitativa das três posições de menor RMSD mostrando as interações químicas entre o tamoxifeno e os resíduos de aminoácidos que compõem a cavidade do sítio de ligação do modelo CgNR2E5. Na posição de menor RMSD (A) há formação de uma ponte de hidrogênio com Arg206, ligação  $\pi$ - $\pi$  entre Phe76 e dois anéis aromáticos de TAM, além de muitas interações hidrofóbicas envolvendo os resíduos Phe36, Phe76, Ala80, Met111, Ile203 e Leu218. A posição de segundo menor RMSD (B) apresenta interações hidrofóbicas com Phe76, Val77, Leu110, Met111, Ala199, Asp202 e Ile203, além de uma ligação  $\pi$ - $\pi$  entre Phe76 e um anel aromático do ligante. A posição de terceiro menor RMSD (C) configura interações hidrofóbicas ocorrendo entre os resíduos Ala39, Phe76 e Met111 e o ligante, com presença de ligação  $\pi$ - $\pi$  entre um anel aromático do TAM e o resíduo Phe76. ....85

**Figura 28** - Análise qualitativa das três posições de menor RMSD mostrando as interações químicas entre o tamoxifeno e os resíduos de aminoácidos que compõem a cavidade do sítio de ligação do modelo CgNR1P2. A posição de menor RMSD (A) apresentou diversas interações hidrofóbicas com resíduos de Ala70, Phe74, Met77, Asn110, Glu112, Trp141, além de uma ligação  $\pi$ - $\pi$  entre Phe74 e 2 anéis aromáticos do 4-hidroxitamoxifeno. Na posição de segundo menor valor de RMSD (B) também ocorreu formação de ligação  $\pi$ - $\pi$  entre Phe74 e 2 anéis aromáticos do TAM. Outras interações hidrofóbicas foram encontradas com Phe74, Met77, Glu112 e Trp141. Uma interação polar foi formada na terceira posição de menor RMSD (C), uma ponte de hidrogênio com o resíduo His31. Houve formação de ligação  $\pi$ - $\pi$  entre Phe74 com dois anéis aromáticos de TAM e também uma ligação  $\pi$ -cátion entre Phe74 e o nitrogênio do ligante. Interações hidrofóbicas estão presentes com Phe74, Met77, Gly108, Glu112 e Trp141. ....86

## LISTA DE TABELAS

<b>Tabela 1.</b> Parâmetros de qualidade para os modelos gerados segundo Benkert, Biasini e Schwede (2010) e Verli (2014) e número de aminoácidos que compõem a sequência correspondente a cada modelo. ....	<b>31</b>
<b>Tabela 2.</b> Identificação dos valores de área e volume da cavidade do modelo gerado para o CgER, para a estrutura experimental de CgER 4N1Y e para a estrutura experimental de ER humano 1GWR.....	<b>35</b>
<b>Tabela 3.</b> Volumes das cavidades do sítio de ligação dos modelos do domínio LBD dos NRs de <i>C.gigas</i> , com destaque para os modelos com maior probabilidade de comportar o composto xenobiótico de interesse no estudo, o estradiol. ....	<b>37</b>
<b>Tabela 4.</b> Métricas de qualidade dos modelos gerados. O C-score se encontra tipicamente na faixa de -5 a 2, sendo que valores maiores que -1.5 indicam um modelo com topologia global correta. Valores de TM-score maiores que 0,5 representam comparações entre proteínas que assumem geralmente o mesmo enovelamento. O RMSD descreve a distância média por resíduo entre duas estruturas otimamente alinhadas, dessa forma, quanto menor o seu valor, mais próximo de 0, mais similares são as estruturas proteicas comparadas. ....	<b>41</b>
<b>Tabela 5.</b> Métricas de qualidade estrutural e ângulos de conformações permitidos para os modelos obtidos e estruturas do PDB utilizadas como referência para o cálculo. ....	<b>42</b>
<b>Tabela 6.</b> Métricas de qualidade para a estrutura do modelo controle.....	<b>45</b>
<b>Tabela 7.</b> Métricas de qualidade e ângulos de conformações permitidos para o modelo controle e estrutura do PDB utilizada como referência para o cálculo. ....	<b>45</b>
<b>Tabela 8.</b> Volumes das cavidades do sítio de ligação dos modelos do domínio LBD de NR1P2, NR2E1 e NR2E5 de <i>C.gigas</i> . ....	<b>48</b>
<b>Tabela 9.</b> Quantificação do espaço da cavidade do sítio de ligação do modelo controle NR3A de <i>C. gigas</i> .....	<b>48</b>
<b>Tabela 10.</b> Dados de análise quantitativa para posições do teste de <i>docking</i> com estradiol. ....	<b>53</b>
<b>Tabela 11.</b> Dados de análise quantitativa para posições do teste de <i>docking</i> com tamoxifeno.....	<b>63</b>
<b>Tabela 12.</b> Análise quantitativa das três posições do estradiol no teste de <i>docking</i> para cada modelo com menor RMSD em relação ao espaço que este ligante ocupa na estrutura de referência.....	<b>79</b>
<b>Tabela 13.</b> Análise quantitativa das três posições do tamoxifeno no teste de <i>docking</i> para cada modelo com menor RMSD em relação ao espaço que este ligante ocupa na estrutura de referência.....	<b>83</b>

## LISTA DE ABREVIATURAS E SIGLAS

AF-2	Função de ativação 2
Ala	Alanina
Arg	Arginina
Asn	Asparagina
CgER	Receptor de Estrógeno de <i>Crassostrea gigas</i>
Cys	Cisteína
DBD	Domínio de Interação com o DNA
ER	Receptor de estrógeno
Glu	Glutamato
Gly	Glicina
His	Histidina
Ile	Isoleucina
LBD	Domínio de Interação com o Ligante
Leu	Leucina
Met	Metionina
NR	Receptor Nuclear
PDB	Banco de Dados de Proteínas
Phe	Fenilalanina
Pro	Prolina
RMSD	Desvio da raiz média quadrática
SERM	Modulador Seletivo de Receptor de Estrógeno
TAM	Tamoxifeno
TBT	Tributilestanho
Trp	Triptofano
Val	Valina

## SUMÁRIO

<b>1 INTRODUÇÃO</b>	<b>18</b>
1.1 A OSTRAS DO PACÍFICO E O CONTEXTO ECOTOXICOLÓGICO	18
1.2 RECEPTORES NUCLEARES	19
1.3 LIGANTES E XENOBIÓTICOS	22
1.4 ANÁLISES IN SILICO PARA SIMULAÇÃO E ENTENDIMENTO DOS SISTEMAS BIOLÓGICOS	25
<b>2 OBJETIVOS</b>	<b>27</b>
2.1 OBJETIVO GERAL	27
2.2 OBJETIVOS ESPECÍFICOS	27
<b>3 MATERIAL E MÉTODOS</b>	<b>28</b>
3.1 MODELAGEM 3D	28
3.2 CÁLCULO DE CAVIDADES	28
3.3 DOCKING MOLECULAR	29
<b>4 RESULTADOS E DISCUSSÃO</b>	<b>30</b>
4.1 MODELOS TRIDIMENSIONAIS SWISS-MODEL	30
4.2 CARACTERIZAÇÃO DO SÍTIO DE LIGAÇÃO	35
4.3 MODELOS TRIDIMENSIONAIS I-TASSER	40
4.4 CARACTERIZAÇÃO DAS CAVIDADES DO SÍTIO DE LIGAÇÃO	47
4.5 POTENCIAL DE INTERAÇÃO DA CAVIDADE COM ESTRADIOL	49
<b>4.5.1 CgNR2E1</b>	<b>53</b>
<b>4.5.2 CgNR2E5</b>	<b>57</b>
<b>4.5.3 CgNR1P2</b>	<b>58</b>
4.6 POTENCIAL DE INTERAÇÃO DA CAVIDADE COM TAMOXIFENO	60
<b>4.6.1 CgNR2E1</b>	<b>63</b>
<b>4.6.2 CgNR2E5</b>	<b>65</b>
<b>4.6.3 CgNR1P2</b>	<b>66</b>
<b>5 CONCLUSÃO</b>	<b>68</b>

<b>REFERÊNCIAS</b>	<b>69</b>
<b>APÊNDICE A</b>	<b>75</b>
<b>APÊNDICE B</b>	<b>79</b>
<b>APÊNDICE C</b>	<b>83</b>

## 1 INTRODUÇÃO

### 1.1 A OSTRAS DO PACÍFICO E O CONTEXTO ECOTOXICOLÓGICO

A ostra do Pacífico, *Crassostrea gigas*, é uma espécie de invertebrado marinho introduzida em diversos países e que atualmente possui distribuição global. Naturalmente, ocorre em regiões entre-marés e estuarinas, desempenhando importantes funções ecológicas nestes habitats. A ostra do Pacífico é um animal sésil que se fixa a rochas e não consegue se mover para outras regiões (MENG et al., 2013). Devido a esta característica, desenvolveu vários mecanismos de adaptação ao estresse ambiental ao longo da evolução e é conhecida por acumular e tolerar poluentes. O seu mecanismo de alimentação é a filtração, o que a expõe a diversos patógenos microbiológicos e produtos do estilo de vida humano, como xenobióticos (ZHANG et al., 2016). A espécie é sensível a alterações de seu habitat, inclusive ao fenômeno de acidificação dos oceanos, apresentando alterações em vias metabólicas, como as de respostas a antioxidantes, metabolismo de carboidratos, transcrição e tradução (TIMMINS-SCHIFFMAN et al., 2014).

A plasticidade da ostra do pacífico na manutenção de sua sobrevivência em ambientes extremos e poluídos sugere que esta espécie desenvolveu evolutivamente respostas únicas de adaptação ao estresse ambiental (ZHANG et al., 2016). As características e a ecologia da espécie a tornam de extrema importância para o monitoramento de áreas poluídas. Através da utilização de biomarcadores, indicadores de alterações bioquímicas, imunoquímicas e fisiológicas, a nível molecular e celular, é possível caracterizar estas mudanças nos organismos em consequência da exposição ao estresse ambiental. Este monitoramento através de marcadores biológicos possui o potencial de fornecer alertas sobre as condições de áreas contaminadas, permitindo assim o planejamento de estratégias para contornar situações de estresse ambiental antes do desencadeamento de danos irreversíveis (CAJARAVILLE et al., 2000). Alguns xenobióticos como o tributilestanho (TBT), substância extensamente utilizada como anti-incrustante com o intuito de impedir que organismos marinhos aderissem aos cascos de embarcações, causa danos aos organismos marinhos expostos a ele (LARANJEIRO et al., 2018). Este composto teve sua utilização proibida nas tintas que revestem os cascos de

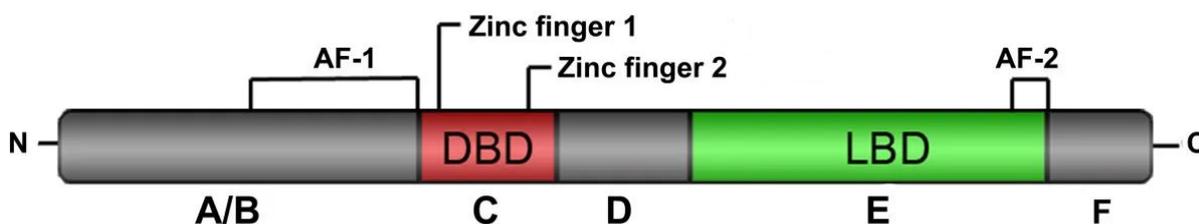
embarcações, pela Organização Marítima Internacional e o efeito mais bem reportado da exposição de organismos marinhos ao TBT é o imposex. Esse fenômeno é caracterizado pelo surgimento de órgãos sexuais masculinos (vaso deferente e pênis) em fêmeas de moluscos, podendo ocasionar a esterilização das fêmeas e declínio em populações de espécies afetadas. O composto causa desregulação endócrina e é bioacumulado pelos organismos (ANTIZAR-LADISLAO, 2008). Apesar de sua utilização ter sido banida, concentrações significativas ainda são encontradas em regiões portuárias de todo o mundo. O imposex vem sendo utilizado como marcador biológico para identificar a contaminação de áreas poluídas por butilestânicos e ressalta a importância da utilização de biomarcadores para monitoramento de regiões poluídas (ROSSATO, 2017). Sabe-se que TBT é capaz de interagir com biomoléculas como receptores nucleares (BAKER et al., 2015), que modulam a transcrição de diversos genes. Portanto, estudar os mecanismos de ligação de xenobióticos passa a ser fundamental para o entendimento dos fenômenos fisiológicos e ecológicos decorrentes destas exposições.

## 1.2 RECEPTORES NUCLEARES

As proteínas são caracterizadas por serem moléculas biológicas muito complexas, uma vez que estão presentes em muitos processos celulares e suas funções dependem de interações específicas com os mais variados elementos (BRIDGHAM et al., 2014; ZHANG et al., 2016). Receptores nucleares (NRs) são proteínas exclusivas de animais (HUANG et al., 2015), pertencentes à superfamília de fatores de transcrição, os quais regulam importantes funções em seres vivos, como processos fisiológicos, desenvolvimento e reprodução, por meio da inibição ou indução da expressão de genes específicos (BRIDGHAM et al., 2014). Receptores Nucleares apresentam uma estrutura característica, que consiste nos domínios A/B, C, D, E e F, como ilustrado na Figura 1. A/B constitui a região inicial N-terminal do receptor e contém a função de ativação AF-1, que é capaz de cooperar com a função de ativação AF-2 (hélice 12), no domínio E, para uma mais estável regulação da modulação gênica. O domínio F configura a região C-terminal, que varia de tamanho em diferentes receptores nucleares. A região C corresponde ao Domínio de Interação com o DNA, *DNA-Binding Domain* (DBD), sendo esta a região que interage com o DNA e é a mais conservada das regiões da proteína. Essa região possui dois “dedos” de zinco, que correspondem a alfa-hélices com funções de

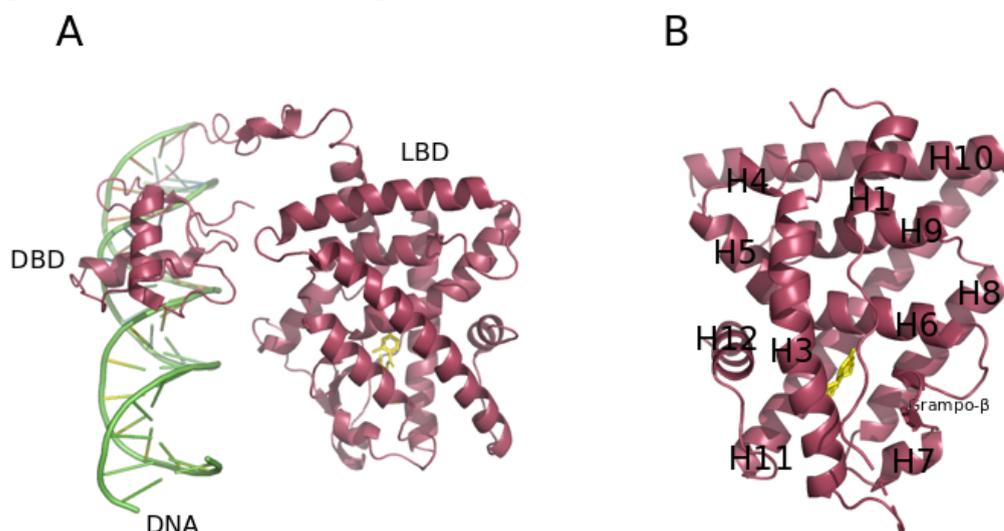
mediar a dimerização do receptor e determinar a especificidade da interação com o DNA. O domínio D é uma “dobradiça” flexível, que conecta o domínio DBD com o domínio E, o segundo domínio mais conservado, chamado de Domínio de Interação com o Ligante, *Ligand-Binding Domain* (LBD). A Figura 2 ilustra a formação de estrutura quaternária típica dos NRs, além da estrutura terciária do domínio LBD. Este domínio, por sua vez, pode interagir com ligantes endógenos, como hormônios, ou com moléculas exógenas, como contaminantes e poluentes (VOGELER et al., 2014). Após a interação com um ligante, em muitos NRs há o recrutamento de proteínas reguladoras positivas, co-ativadoras, ou negativas, correpressoras, que são fundamentais na ativação ou repressão induzida pelo ligante na transcrição gênica (BENOD et al., 2014). Muitos NRs não possuem um ligante natural identificado, estes são classificados como receptores nucleares órfãos.

Figura 1 - Organização dos domínios de um receptor nuclear na proteína. A região A/B, N-terminal, contém a função de ativação AF-1. O domínio de interação com o DNA, região C, inclui dois dedos de zinco. A região D, uma dobradiça (*hinge*) variável, separa o domínio DBD do Domínio de Interação com o Ligante LBD (região E), o qual possui o fator de ativação em hélice AF-2. A porção C-terminal F é localizada ao final da sequência.



Fonte: adaptado de VOGELER et al., 2014.

Figura 2 - Enovelamento tridimensional típico de receptores nucleares. (A) Estrutura quaternária recorrente em NRs, destacando o Domínio de Interação com o Ligante (LBD) e o Domínio de Interação com o DNA (DBD) interagindo com uma molécula de DNA. (B) Estrutura do domínio LBD com  $\alpha$ -hélices bem definidas, um grampo- $\beta$  entre as  $\alpha$ -hélices H6 e H7 e um ligante associado ao sítio de ligação (em amarelo).



Fonte: Própria autora.

Como consequência da característica que muitos NRs possuem de interagir com moléculas através do domínio LBD, eles se tornam suscetíveis a xenobióticos, moléculas ou substâncias encontradas em organismos que não ocorrem naturalmente nos mesmos. Estes xenobióticos podem induzir ou inibir a expressão de genes, por mimetizar a interação do receptor com seu ligante natural (OMIECINSKI et al., 2010). As consequências para o organismo podem incluir expressão anormal de genes e uma desregulação do desenvolvimento e funcionamento endócrino (VOGELER et al., 2014).

A evolução dos receptores nucleares em metazoários sugere que essas proteínas se diversificaram a partir de um único NR de um metazoário ancestral (BRIDGHAM et al., 2010), visto que foram identificados apenas dois receptores nucleares na classe demosponge, do filo Porifera (LARROUX et al., 2006), se diversificando e aumentando a quantidade em outros metazoários: 17 no cnidário *Nematostella vectensis* (REITZEL; TARRANT, 2009), 21 em *Drosophila melanogaster* (KING-JONES; THUMMEL, 2005), 43 em *Crassostrea gigas* (VOGELER et al., 2014), 270 em *Caenorhabditis elegans* (SLUDER; MAINA, 2001) e 48 NRs em humanos. Há uma alta conservação, ao longo da evolução, das regiões DBD e LBD dos NRs em geral e a classificação destes é feita a partir destas regiões da proteína. Dentre os 43 NRs de *C. gigas*, sua maioria possui homologia aos de *Homo sapiens* e de *D.*

*melanogaster*, a mosca da fruta. Os NRs homólogos aos de humanos incluem receptores de hormônios da tireoide e receptores de ácido retinóico, homólogos aos da mosca da fruta são descritos os receptores relacionados ao desenvolvimento, à ecdise e à processos reprodutivos (VOGELER et al., 2014). Durante a evolução dos animais multicelulares, os receptores nucleares duplicaram, divergiram e aumentaram em número consideravelmente (BRIDGHAM et al., 2010).

Dentre a grande diversidade de NRs que ocorrem em animais, encontram-se receptores de hormônios, que possuem representantes como receptores de esteróides adrenais, como glicocorticóides, receptores de esteróides sexuais, como estradiol, testosterona e progesterona, assim como receptores de vitamina D, hormônios da tireóide (T3 e T4) e ecdisona, um importante receptor de esteróide em insetos (BAKER, 1997). A contribuição evolutiva de receptores esteróides em vertebrados é reconhecida, a qual proporcionou melhor regulação de diversas vias fisiológicas de diferenciação, desenvolvimento, reprodução, respostas imunes e resposta ao estresse (BAKER, 2019). Animais invertebrados possuem vários homólogos de receptores esteróides de vertebrados, porém os mecanismos de interação de invertebrados com moléculas esteróides permanece incerto (SCOTT, 2013), demandando estudos que esclareçam essas possíveis interações e relações evolutivas envolvendo esses receptores e seus ligantes.

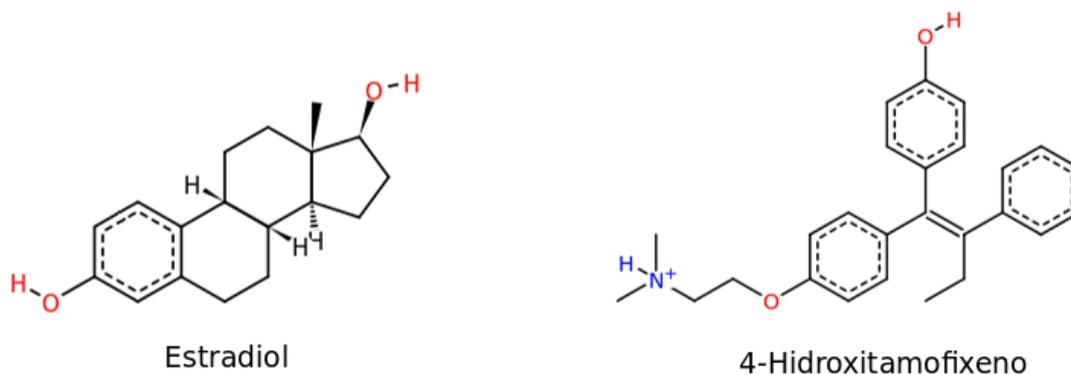
### 1.3 LIGANTES E XENOBIÓTICOS

Hormônios esteróides, os quais incluem moléculas como estrogênio, estradiol e testosterona, são produzidos por animais vertebrados e invertebrados e até mesmo por plantas. Em vertebrados essas moléculas são conhecidas por terem papéis relacionados a diferenciação sexual e reprodução. Níveis significativos destes esteróides foram identificados em moluscos, sugerindo que também podem ter papéis fisiológicos entre organismos invertebrados (JANER; PORTE, 2007). Fontes de esteróides xenobióticos para moluscos, como os produzidos sinteticamente, provenientes de moléculas mimetizadoras de esteróides ou originados do metabolismo de outros animais vertebrados e seu efeito nesses animais vem sendo amplamente investigados. No entanto, correlações destes ligantes com efeitos em moluscos ainda são consideradas ambíguas e não estão estabelecidas cientificamente (SCOTT, 2013). A ostra do Pacífico possui dois NRs análogos aos de hormônios sexuais esteróides (CgNR3A e CgNR3B). Entretanto, os receptores de *C. gigas* identificados como

homólogos aos de esteróides em vertebrados não respondem à estrogênio (MATSUMOTO et al., 2007) e não comportam a molécula de estradiol em sua cavidade do domínio LBD, como o faz o NR de estrogênio homólogo humano. Este domínio de interação com o ligante, em *C. gigas* e outras espécies de moluscos, demonstrou experimentalmente ativar a transcrição sem a adição de nenhum ligante, sugerindo que os NRs de estrogênio da ostra do Pacífico sejam fatores transcricionais independentes de ligante (BRIDGHAM et al., 2014). Não obstante, a ostra do Pacífico possui três homólogos do receptor NR1J, conhecido por responder a hormônios esteróides na espécie *Daphnia pulex* (KARIMULLINA et al., 2012), demonstrando que ligantes esteróides possuem potencial de interagir com outros receptores nucleares que não os anotados como NRs de hormônios sexuais. Este fato incita a possibilidade de que, dentre os 43 NRs identificados no genoma de *C. gigas*, alguns podem interagir com estes ligantes esteróides.

Tamoxifeno (TAM) é um fármaco derivado de trifeniletileno muito utilizado no tratamento de câncer de mama. Uma vez que se estabeleceu o papel central de receptores de estrogênio (ERs) e do estradiol neste tipo de câncer, moléculas antagonistas de estrogênio, como TAM, foram desenvolvidas para ocupar o sítio do ER e suprimir os efeitos que o ligante original desencadeia nos tecidos (LONARD; SMITH, 2002). Posteriormente foi reportado como TAM e outras moléculas desenvolvidas como antagonistas, são moduladores seletivos de receptores de estrogênio (SERMs), pois podem exibir respostas agonistas e antagonistas, conforme o tecido em que se encontram. Essas diferentes respostas também ocorrem pois a ligação de estradiol e de SERMs promovem mudanças conformacionais diferentes no domínio LBD (SHIAU et al., 1998). A Figura 3 ilustra a estrutura química da molécula de estradiol e do SERM 4-hidroxitamoxifeno que ocupam o sítio de ligação do ER humano.

Figura 3 - Estrutura química dos ligantes estradiol e 4-hidroxitamoxifeno.



Fonte: Própria autora.

Atualmente fármacos são compostos considerados ubíquos em ambientes aquáticos. Tamoxifeno, por exemplo, já foi reportado em efluentes hospitalares e urbanos e estudos apontam que é um composto altamente bioacumulável por organismos (ORIAS; SIMON & PERRODIN, 2015), podendo apresentar risco ecotoxicológico significativo para os ecossistemas aquáticos (ORIAS et al., 2015). Sendo uma molécula que possui propriedades químicas e estruturais para ocupar o mesmo sítio de ligação da molécula de estradiol, é possível que os NRs modulados pelo estradiol também sofram modulação por TAM.

Os fármacos produzidos por humanos são caracterizados por uma variedade diversificada de estruturas químicas e alta heterogeneidade de propriedades (ZHANG et al., 2013). Moléculas como cafeína (psicoativo estimulante), ibuprofeno (anti-inflamatório), carbamazepina (anticonvulsivo), novobiocina (antibiótico) e tamoxifeno (quimioterapêutico) são algumas dessas drogas produzidas por humanos encontradas em ambientes aquáticos que já possuem respostas de estresse identificadas por biomarcadores em invertebrados (AGUIRRE-MARTÍNEZ; DELVALLS; MARTÍN-DÍAZ, 2015). A grande quantidade de moléculas continuamente reportadas como contaminantes e o potencial efeito que elas possuem na biota de locais contaminados não é amplamente compreendido, o que demanda estudos na área. Pesquisas *in vivo* são dificultadas, por se tratar de uma variedade muito ampla de compostos e muitas interações possíveis com seres vivos. Estudos *in silico* podem apontar e verificar várias interações potenciais entre essas moléculas e componentes de organismos, direcionando estudos e auxiliando no entendimento do seu efeito em sistemas biológicos.

## 1.4 ANÁLISES *IN SILICO* PARA SIMULAÇÃO E ENTENDIMENTO DOS SISTEMAS BIOLÓGICOS

O desenvolvimento tecnológico e avanço no poder computacional modernizaram a área da biologia, possibilitando a introdução de termos como genômica, proteômica e metabolômica. Hoje possuímos o genoma sequenciado de diversas espécies e muitas informações de variações genéticas em bancos de dados (DOSS et al., 2014). Dados sobre variantes que ocorrem nos organismos, como substituições de um único aminoácido, nos auxiliam a entender diversos padrões que ocorrem nas biomoléculas presentes nos seres vivos, como mutações que levam a substituições de aminoácidos em proteínas associadas com processos de desenvolvimento de câncer (MOSHKOVSKII et al., 2018). Ao estudar proteínas, observamos que suas funções estão diretamente ligadas ao seu enovelamento tridimensional e dinâmica de movimento. Técnicas experimentais para resolução de estruturas proteicas envolvem cristalografia de raios-X e ressonância magnética nuclear (RMN), porém esses métodos para determinação estrutural são limitados pelo custo, consumo de tempo e exigência de instrumentos especializados. Por isso, apesar de muitas sequências de proteínas disponíveis, há um número comparativamente reduzido de estruturas resolvidas no banco de dados de proteínas (PDB) (LEVITT, 2007). Em uma busca nas plataformas de banco de dados de sequências de proteínas, como o NCBI e de banco de dados de estrutura de proteínas, como o PDB, podemos exemplificar essa diferença. Aplicando um filtro para a busca, por exemplo, receptores nucleares de uma espécie muito estudada, como *Homo sapiens*, é possível encontrar 9.628 sequências de proteínas no NCBI. Já no PDB, aplicando o mesmo filtro, encontramos 1.510 estruturas para NRs humanos, o que representa 15% do total de sequências. Nesse sentido, técnicas *in silico* para a obtenção de modelos da estrutura 3D de proteínas auxiliam nesses estudos e são importantes ferramentas para entender e prever a função de proteínas.

As técnicas de modelagem constroem a estrutura 3D de proteínas através de diferentes análises. A modelagem comparativa utiliza *templates* homólogos evolutivamente relacionados, identificados a partir de comparações de sequências entre a proteína de interesse e os bancos de dados. O método de *threading* encontra enovelamentos que correspondem à estrutura da proteína de interesse, mesmo quando os *templates* não possuem a mesma origem evolutiva. Já a modelagem *ab initio* constrói modelos do zero, quando não são encontradas

proteínas estruturalmente relacionadas na biblioteca do PDB (ROY; KUCUKURAL; ZHANG, 2010). Métricas de qualidade são calculadas para garantir a qualidade dos modelos gerados. A fim de prever funções para proteínas, deve-se considerar que estas biomoléculas não são estáticas, possuem certa flexibilidade e dinamismo em sua estrutura. Técnicas como *docking* e dinâmica molecular auxiliam no entendimento das interações de proteínas com proteínas e proteínas com outras moléculas, permitindo uma melhor caracterização estrutural e funcional.

Análises de *docking* buscam investigar o potencial de interação entre duas moléculas, como uma proteína e um ligante. Se a ligação das moléculas é possível, a técnica busca descrever a orientação em que as interações são máximas e onde a energia de ligação está no seu mínimo. Trata-se de uma técnica muito útil no desenvolvimento de novas drogas, por exemplo, que se trata de um processo longo e dispendioso. A aplicação de métodos computacionais auxilia neste desenvolvimento ao direcionar e apontar as interações possíveis entre compostos e os alvos nos organismos (DOSS et al., 2014). A fim de verificar o dinamismo de proteínas em diferentes escalas de tempo, análises de dinâmica molecular são empregadas. Através dessa técnica é possível realizar simulações de sistemas mais realistas, incluindo moléculas de água explícitas, contra-íons e até um ambiente com membrana plasmática. Proteínas assumem diferentes conformações em solução, em análises de dinâmica molecular é incorporada no sistema a flexibilidade do ligante e da proteína, melhorando a simulação de interações e aumentando a complementaridade entre elas, e, portanto, correspondendo ao ajuste induzido que ocorre em proteínas (ALONSO; BLIZNYUK; GREASY, 2006). Essas metodologias computacionais combinadas buscam tornar mais fidedigna possível a descrição de processos biológicos.

Apesar de sua importância ecotoxicológica, a ostra do Pacífico possui apenas uma estrutura cristalográfica de NR resolvida experimentalmente depositada no PDB. A fim de avaliar o potencial de interação e modulação dos NRs de *C. gigas* por compostos xenobióticos, utilizamos métodos computacionais que auxiliaram no entendimento dessas interações e apontaram candidatos possíveis entre os NRs a serem modulados por esses compostos.

## 2 OBJETIVOS

### 2.1 OBJETIVO GERAL

Caracterizar estruturalmente a superfamília de receptores nucleares de *C. gigas* avaliando os mecanismos de interação com xenobióticos.

### 2.2 OBJETIVOS ESPECÍFICOS

- I. Realizar modelagem 3D dos domínios LBD da superfamília de receptores nucleares de *C. gigas*;
- II. Analisar as estruturas modeladas, com enfoque nas cavidades do sítio de ligação;
- III. Identificar sítios de ligação para moléculas xenobióticas como o estradiol e o tamoxifeno.

### 3 MATERIAL E MÉTODOS

#### 3.1 MODELAGEM 3D

A técnica de modelagem para a região LBD dos NRs foi realizada utilizando o programa SWISS-MODEL (WATERHOUSE et al., 2018). Essa plataforma gera modelos 3D de proteínas a partir de uma sequência alvo, extrapolando informações experimentais de uma estrutura de proteína relacionada evolutivamente que serve como *template*. Inicialmente as sequências dos receptores nucleares, obtidas a partir do genoma de *C.gigas* depositado no NCBI, foram submetidas ao BLASTp (CAMACHO et al., 2009), utilizando como referência o *Protein Data Bank* (PDB), e ao programa ClustalW2 2.0 para verificar a similaridade entre a sequência do receptor de *C. gigas* e as sequências depositadas em bancos de dados e para localizar a região correspondente ao domínio LBD. O pacote I-TASSER (ZHANG, 2008) também foi empregado para modelagem para casos em que os modelos gerados pelo SWISS-MODEL não apresentaram enovelamento típico esperado para o domínio LBD de NRs. Sua metodologia divide a sequência *input* em fragmentos, faz alinhamentos com *templates* baseado na sequência de aminoácidos e em recursos estruturais. A topologia dos modelos é construída remontando os fragmentos obtidos do alinhamento com alguns *templates*. No caso de haverem regiões não alinhadas, a estrutura é construída do zero por modelagem *ab initio*. Várias etapas de simulação da topologia são feitas para refinamento da estrutura até a formação final do modelo (ROY; KUCUKURAL; ZHANG, 2010). Ambas as metodologias foram submetidas a etapas de verificação qualidade e confiabilidade dos modelos gerados, descritos nos resultados.

#### 3.2 CÁLCULO DE CAVIDADES

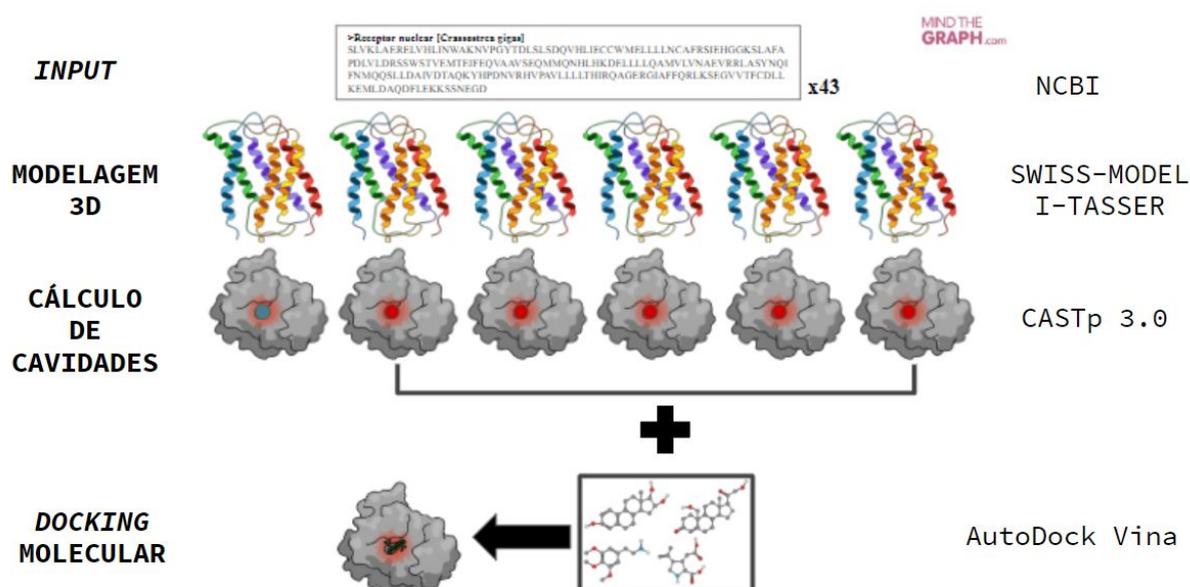
Os modelos 3D obtidos através da modelagem foram submetidos ao programa CASTp (TIAN et al., 2018), com *radius probe* de 1.3 Å como valor padrão para cálculo de área de superfície acessível ao solvente. O programa identificou cavidades, bolsões de superfície, resíduos de aminoácidos que compõem a cavidade, bem como forneceu a área e volume das

mesmas. Os dados obtidos pelo CASTp permitiram a montagem de figuras pelo PyMOL (SCHRÖDINGER, 2017).

### 3.3 DOCKING MOLECULAR

Os alvos de interesse a partir dos resultados das análises anteriores foram submetidos ao teste de *docking* molecular com a molécula de estradiol e tamoxifeno. Para o preparo do teste, caixas de atracamento foram construídas utilizando o programa CHIMERA (PETTERSEN, 2004). Essas caixas foram posicionadas na região da cavidade de ligação dos modelos com os ligantes. Além disso, os resíduos completamente internos às caixas foram considerados flexíveis. Para a realização do *docking* molecular foi utilizado o programa AutoDock Vina (TROTT; OLSON, 2009). Para analisar os resultados do teste, foram utilizados os programas DockRMSD (BELL & ZHANG, 2019), para calcular as posições dos ligantes contra uma estrutura experimental de referência; o *software* PoseView (STIERAND; RAREY, 2010) e LigPlus (WALLACE; LASKOWSKI; THORNTON, 1995), para obter diagramas bidimensionais das interações químicas da cavidade do sítio e a visualização 3D das posições resultantes do *docking* no PyMOL (SCHRÖDINGER, 2017). A Figura 4 demonstra os métodos utilizados para a realização do trabalho na sequência de sua realização.

Figura 4 - Esquema das etapas metodológicas do trabalho, com os programas utilizados, na sequência de sua realização ao longo do desenvolvimento do estudo.



## 4 RESULTADOS E DISCUSSÃO

### 4.1 MODELOS TRIDIMENSIONAIS SWISS-MODEL

Neste trabalho utilizamos duas metodologias para construção de modelos 3D para o domínio LBD das proteínas de interesse. A primeira delas utiliza a plataforma SWISS-MODEL, onde obtivemos dados de modelagem comparativa que serão descritos a seguir.

Inicialmente os 43 NRs identificados no genoma de *C. gigas* (VOGELER et al., 2014) tiveram sua sequência correspondente ao domínio LBD submetidas à técnica de modelagem comparativa pelo programa SWISS-MODEL. A análise dos modelos gerados seguiu critérios para verificar a sua qualidade. Foram considerados valores de identidade de no mínimo 25% entre a sequência utilizada como *template* e a sequência alvo de *C. gigas* (VERLI, 2014). Além disso, o *software* calcula uma métrica de qualidade do modelo chamada QMEAN (*Qualitative Model Energy Analysis*), que avalia as distâncias interatômicas modeladas com informações de estruturas de proteínas resolvidas experimentalmente que possuem homologia com a sequência alvo (WATERHOUSE et al., 2018; BENKERT; BIASINI; SCHWEDE, 2010). Os valores de QMEAN próximos de 0 são considerados de alta qualidade e menores de -4 de baixa qualidade.

Do total de NRs da ostra do Pacífico, 37 foram modelados com qualidade. Para três deles, a técnica empregada não encontrou sequência com identidade suficiente no banco de dados e outros dois NRs não possuem a região LBD identificada (VOGELER et al., 2014). A Tabela 1 apresenta as métricas de qualidade dos 37 modelos obtidos, os quais correspondem a valores de QMEAN maiores que -4 e possuem identidade de sequência maior que 25%, dentro dos limites permitidos para modelos de qualidade.

Tabela 1. Parâmetros de qualidade para os modelos gerados segundo Benkert, Biasini e Schwede (2010) e Verli (2014) e número de aminoácidos que compõem a sequência correspondente a cada modelo.

<b>Modelo</b>	<b>QMEAN</b>	<b>Identidade de sequência (%)</b>	<b>Tamanho da sequência de aminoácidos do modelo</b>
CgNR1A	-3.05	43.20	258
CgNR1B	-3.13	63.18	253
CgNR1C	-1.91	29.31	178
CgNR1CDEFa	-3.78	29.05	220
CgNR1CDEFb	-2.47	30.54	205
CgNR1D	-1.87	29.03	279
CgNR1E	-2.08	32.10	258
CgNR1F	-0.44	39.83	241
CgNR1H	-1.55	37.66	260
CgNR1Ja	-3.27	35.20	213
CgNR1Jb	-2.62	27.07	236
CgNR1Jg	-3.03	35.23	273
CgNR1P1	-3.13	26.37	191
CgNR1P2	-2.94	28.39	159
CgNR1P3	-2.15	28.39	161
CgNR1P4	-3.19	32.93	179
CgNR1P5	-3.09	26.78	188
CgNR1P6	-2.67	27.12	188
CgNR1P7	-2.90	29.14	189
CgNR1P8	-2.84	32.18	181

CgNR1P9	-2.62	29.15	201
CgNR1P10	-2.24	25.77	200
CgNR2A	-2.33	71.17	222
CgNR2B	-0.71	91.93	229
CgNR2CD	-2.04	41.08	232
CgNR2E1	-3.90	52.85	206
CgNR2E2	-2.74	50.00	211
CgNR2E3	-2.27	38.97	218
CgNR2E5	-2.40	36.65	201
CgNR2F	-3.23	38.57	254
CgNR3A	-1.29	40.93	237
CgNR3B	-3.01	42.48	231
CgNR4A	-1.67	54.02	264
CgNR5A	-1.53	48.50	238
CgNR5B	-1.38	33.33	218
CgNR0B	-2.92	30.69	201
CgNHR42	-2.07	30.59	221

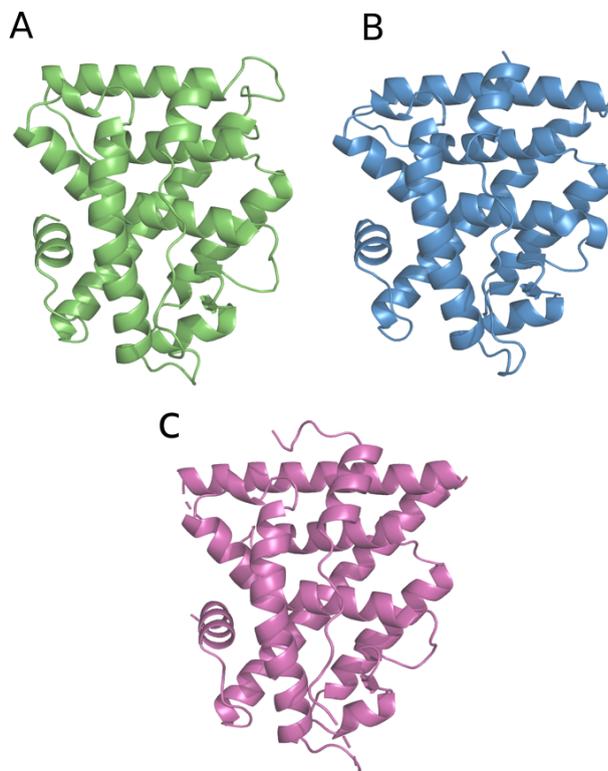
---

Fonte: Própria autora.

Observamos que *templates* que possuem maiores valores de identidade com nossa sequência alvo facilitam a obtenção de modelos com enovelamento esperado para o domínio LBD de NRs, uma vez que a modelagem realizada pelo SWISS-MODEL é comparativa por homologia. No entanto, identidades de sequência no limiar do valor estabelecido por Verli (2014) de 25% também apresentam modelos confiáveis. Além disso, valores de QMEAN próximos de -4 também não limitaram a obtenção de modelos de qualidade.

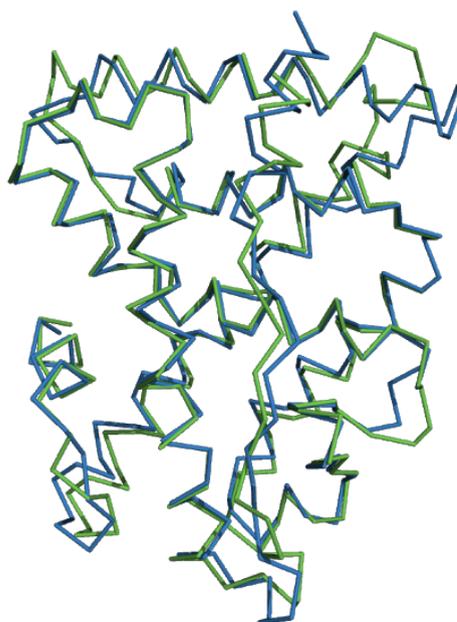
Utilizamos o ER de *C. gigas*, CgNR3A, como modelo controle nesse estudo. Sua escolha para este fim ocorreu devido a ser o único NR de *C. gigas* que possui estrutura cristalográfica resolvida (4N1Y) depositada no PDB. Para inferir se os modelos gerados dos NRs que não possuem estrutura 3D resolvida experimentalmente são representativos e correspondentes ao enovelamento típico de LBD nos organismos, utilizamos a sequência do CgER (Receptor de Estrógeno de *C. gigas*) para gerar um modelo controle utilizando um *template* do banco de dados, que não fosse sua própria estrutura de cristal. Deste modo, foi possível verificar o quão próximo o modelo gerado *in silico* se aproxima da estrutura experimental já descrita (WÄRNMARK et al., 2002). A Figura 5 apresenta a estrutura tridimensional do modelo obtido (A), do *template* utilizado para a modelagem 1GWR (C) e do ER de *C. gigas* resolvido experimentalmente (B). Mesmo utilizando como *template* a estrutura 1GWR, um receptor de estrógeno humano, com 40.93% de identidade com a sequência do ER de *C. gigas*, o modelo resultante possui estrutura semelhante ao cristal do CgER (4N1Y). A Figura 6 demonstra através de um alinhamento como o esqueleto carbônico da estrutura do modelo é muito similar à da estrutura resolvida experimentalmente (4N1Y), com pequenas variações em regiões de *loop*. O valor de RMSD (Desvio Médio da Raiz Quadrática), que descreve a distância média por resíduo entre duas estruturas otimamente alinhadas, para o modelo gerado e o cristal de CgER é de apenas 0.90, sendo que quanto menor for esse valor, mais similares são as estruturas. Esse resultado aponta que utilizando *templates* do PDB é possível construir modelos representativos do que esperamos encontrar em um método experimental que descreva a estrutura tridimensional dos NRs no organismo biológico de interesse.

Figura 5 - Enovelamento tridimensional correspondente ao modelo controle CgNR3A e às estruturas cristalográficas de referência para este modelo. (A) CgER modelo (verde). (B) CgER estrutura experimental 4N1Y (azul). (C) ER humano 1GWR (roxo).



Fonte: Própria autora.

Figura 6 - Alinhamento entre as estruturas do CgER modelo (verde) e estrutura experimental CgER 4N1Y (azul).



Fonte: Própria autora.

## 4.2 CARACTERIZAÇÃO DO SÍTIO DE LIGAÇÃO

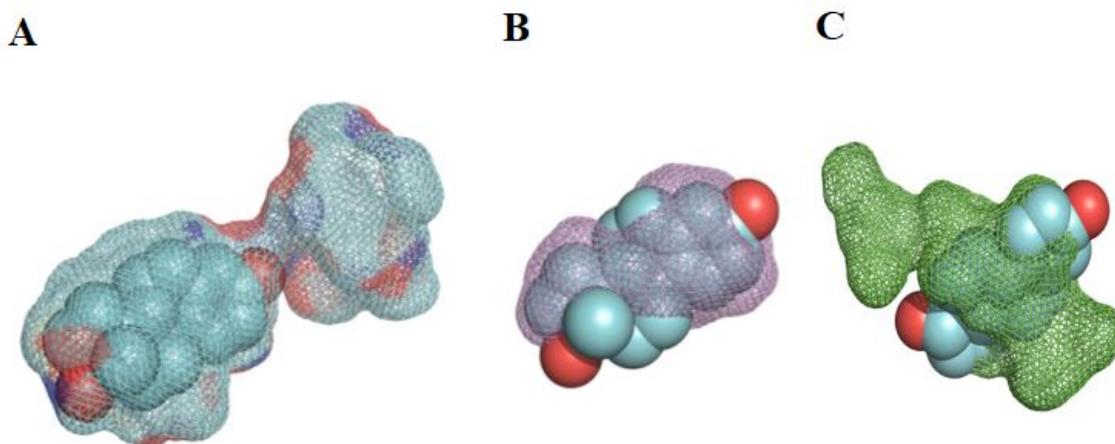
Os modelos gerados, respeitando aos critérios de qualidade citados, passaram para a etapa subsequente de cálculo de cavidades, através do *software* CASTp 3.0. Por ser o NR de *C. gigas* com mais informações de sua estrutura e a fim de padronizar a verificação das etapas metodológicas, o modelo CgER continuou sendo utilizado como controle. As estruturas do cristal CgER (4N1Y) e o cristal de ER humano (1GWR) já possuem o volume de suas cavidades descritos. No entanto, o cálculo de suas cavidades utilizando o programa CASTp 3.0 não apresentou os mesmos valores descritos por Bridgham et al., (2014), como mostra a Tabela 2. Foram realizadas várias tentativas de ajuste da área de superfície acessível ao solvente, *radius probe*, utilizada pelo *software* em seus cálculos, porém os mesmos valores não foram encontrados. A discrepância observada na quantificação de volume das cavidades pode ocorrer devido ao emprego de programas distintos, que se utilizam de metodologias diferentes. Devido aos valores que encontramos para o volume das cavidades de ER humano e CgER serem coincidentemente próximos aos valores de área encontrados por Bridgham et al., (2014) para os mesmos cristais, acreditamos que pode ter ocorrido um equívoco na publicação do artigo, em que os autores confundiram as unidades de medida, quantificando um valor de área como um valor de volume. Como observado por Bridgham e colaboradores, nossos resultados corroboram apresentando um valor de cavidade reduzido do CgER quando comparado à cavidade do ER humano e, assim como descrito por eles, identificamos que a cavidade do CgER possui volume suficientemente reduzido para comportar um ligante como a molécula de estradiol, como mostra a Figura 7.

Tabela 2. Identificação dos valores de área e volume da cavidade do modelo gerado para o CgER, para a estrutura experimental de CgER 4N1Y e para a estrutura experimental de ER humano 1GWR.

	CASTp		Bridgham et al. (2014)
	Área (Å <sup>2</sup> )	Volume (Å <sup>3</sup> )	Volume (Å <sup>3</sup> )
CgNR3A (modelo CgER)	336.711	97.472	
4N1Y (cristal <i>C. gigas</i> )	126.213	63.867	168±8
1GWR (cristal humano)	445.637	265.851	402

Fonte: Própria autora.

Figura 7 - Comparação do volume de superfície da cavidade entre o ER humano, a estrutura cristalográfica de CgER e o modelo CgER, com a presença do estradiol interno ao sítio de ligação. (A) Superfície de volume da cavidade do sítio de ligação da estrutura cristalográfica de ER humano (1GWR), suficientemente grande para comportar o estradiol. (B) Superfície de volume da estrutura experimental de CgER (4N1Y), apresentando cavidade reduzida para comportar o ligante estradiol, assim como o modelo de CgER (C).



Fonte: Adaptado de Santos, 2018.

No decorrer da análise metodológica de cálculo de cavidades identificamos modelos que, mesmo respeitando aos critérios de qualidade estabelecidos, não possuíam um *folding* típico previsto para o domínio LBD de NRs. Uma análise particular dos modelos se demonstrou essencial para garantir que estes se encontrassem dentro do enovelamento possível conservado para o domínio LBD de receptores nucleares. Modelos com hélices faltando ou regiões que supostamente formariam uma alfa-hélice e estavam sem formato de estrutura secundária foram frequentes e, devido a isso, dos 37 modelos gerados pelo SWISS-MODEL, o cálculo de cavidades foi possível, com confiabilidade, para 21 deles. Tentativas de gerar modelos com estrutura tridimensional esperada foram feitas através do SWISS-MODEL, utilizando *templates* diferentes para a modelagem, porém a técnica demonstrou não alcançar modelos com *folding* confiável para muitas sequências do LBD da espécie de invertebrado não-modelo *C. gigas*. Buscamos identificar um padrão que indicasse o porquê da metodologia do SWISS-MODEL não ser capaz de gerar modelos confiáveis para alguns NRs. Observamos se os valores de identidade de sequência ou QMEAN necessitavam de um ajuste mais rígido de seu limite para gerar modelos de qualidade, porém constatamos que, mesmo os modelos com identidade do *template* com a sequência alvo em torno de 25% e

valores de QMEAN próximos de -4 não limitaram a obtenção de modelos com envelhecimento esperado. A Tabela 3 apresenta os volumes de cavidade calculados para os 21 modelos, padronizando o valor de 1,3 como área de superfície acessível ao solvente, *radius probe*, como *input* para o cálculo de quantificação de cavidades pelo programa. A identidade de sequência dos NRs cujo cálculo de cavidade foi possível também compõem a Tabela 3, evidenciando valores variáveis de identidade de sequência para obtenção de modelos de qualidade em que o cálculo de cavidades foi possível.

Através da quantificação do volume da cavidade de 21 NRs de *C. gigas*, foi possível selecionar modelos com maior probabilidade de comportar o composto xenobiótico de interesse no estudo, o estradiol. Como critério para seleção, utilizamos como base o volume da cavidade do ER humano (265.851 Å<sup>3</sup>) calculado pelo CASTp 3.0. O volume do ER humano foi empregado como referência devido a ser um NR cujo ligante natural é o estradiol e o mecanismo de interação entre o receptor e o ligante já ser bem descrito. Desta forma, consideramos valores próximos ou maiores que o calculado para a cavidade de ER humano para selecionar os candidatos possíveis entre os NRs da ostra a serem modulados por estradiol, mostrados também na Tabela 3.

Tabela 3. Volumes das cavidades do sítio de ligação dos modelos do domínio LBD dos NRs de *C.gigas*, com destaque para os modelos com maior probabilidade de comportar o composto xenobiótico de interesse no estudo, o estradiol.

Modelo	Volume (Å <sup>3</sup> )	Identidade de sequência %
CgNR1A	70.543	43.20
CgNR1D	886.311	29.03
CgNR1CDEFb	79.089	30.54
CgNR1F	324.325	39.83
CgNR1H	258.200	37.66
CgNR1Ja	573.292	35.20
CgNR1Jb	292.082	27.07
CgNR1P1	295.562	26.37
CgNR1P6	462.463	27.12
CgNR1P9	214.074	29.15
CgNR1P10	323.575	25.77

CgNR2A	419.179	71.17
CgNR2B	249.815	91.93
CgNR2F	102.615	38.57
CgNR3A	54.966	40.93
CgNR3B	165.674	42.48
CgNR4A	38.722	54.02
CgNR5A	572.062	48.50
CgNR5B	607.353	33.33
CgNR0B	62.174	30.69
CgNHR42	178.433	30.59

Fonte: Própria autora.

Os NRs selecionados com maior probabilidade de sofrer modulação pelo estradiol através da quantificação do volume de suas cavidades incluem receptores pertencentes às subfamílias NR1, NR2 e NR5, muitos deles possuem homólogos identificados em outros invertebrados e vertebrados. A subfamília NR1 inclui CgNR1D, que possui homólogos em *D. melanogaster* e em humanos. Na mosca da fruta o seu homólogo interage com ecdisteróides e está envolvido no desenvolvimento, em processos de muda (DUBROVSKAYA; BERGER; DUBROVSKY, 2004). Homólogos em humanos são silenciadores transcricionais que realizam funções importantes no metabolismo de lipídios, no ritmo circadiano e em inflamação (RAMAKRISHNAN; MUSCAT, 2006). O homólogo de CgNR1H na mosca da fruta é um NR envolvido com desenvolvimento, reprodução e ecdise em insetos (KING-JONES; THUMMEL, 2005). Os homólogos em vertebrados são receptores X do fígado, LXRs, que regulam o metabolismo de lipídios, homeostase do colesterol e de sais biliares (KALAANY; MANGELSDORF, 2006). Além disso, na subfamília NR1 há representantes de um novo grupo monofilético de NRs identificado para *C. gigas* NR1P, que não foi identificado em organismos modelos invertebrados como *D. melanogaster* e *C. elegans* (VOGELER et al., 2014).

A subfamília NR2 possui representantes como CgNR2B, ortólogo de receptores X de retinóide (RXRs), que interagem com ácido retinóide e ácido docosahexaenóico (DHA). O receptor NR2B de *C. gigas* possui similaridade de 90% com este mesmo receptor do molusco

*Biomphalaria glabrata*, que apresentou ligação e ativação por retinóides, sugerindo que vias de sinalização retinóide possuem funcionalidades conservadas evolutivamente através destes NRs (BOUTON, 2005). O representante NR2A da ostra do Pacífico possui um ortólogo em *D. melanogaster* envolvido com beta oxidação e metabolismo de lipídios (PALANKER et al., 2009) e 70% de identidade com domínio LBD deste receptor em humano, envolvido com desenvolvimento de diabetes (LOVE-GREGORY; PERMUTT, 2007).

O membro NR5A possui homologia com o receptor de fígado homólogo-1 (LRH1), um receptor sem ligante identificado envolvido no metabolismo de ácidos biliares e colesterol (FAYARD; AUWERX; SCHOONJANS, 2004). CgNR5B possui homólogo com *D. melanogaster* em receptores envolvidos com a formação do sistema reprodutivo, na reprodução em fêmeas e regulação do citocromo P450 modificador de esteróides (ALLEN; SPRADLING, 2007).

Durante o andamento da análise das cavidades dos modelos, recebemos dados ainda não publicados do Laboratório de Biomarcadores de Contaminação Aquática e Imunoquímica da UFSC (LABCAI), referente a um transcriptoma de *C. gigas* em condições de exposição ao fármaco tamoxifeno (TAM). O tamoxifeno é um fármaco conhecido por modular receptores ER de vertebrados (LONARD; SMITH, 2002). Análises do transcriptoma apontaram para a hiperexpressão de três NRs da ostra, CgNR1P2, CgNR2E1 e CgNR2E5. Devido às características químicas da molécula de TAM de ser um modulador seletivo de ER, e, portanto, ocupar o mesmo sítio de ligação que o estradiol em vertebrados, esses resultados nos apontam esses candidatos, além dos que já havíamos identificado através da caracterização da cavidade do sítio de ligação, como prováveis de interagir com o estradiol. A superexpressão de NRs da ostra do Pacífico na presença de tamoxifeno pode ocorrer, além da interação direta entre ligante e receptor, pela autorregulação que receptores nucleares realizam entre si em diferentes tecidos e condições (BAGAMASBAD; DENVER, 2011).

Os três NRs identificados como candidatos possíveis de interagir com o estradiol incluem um receptor da subfamília NR1, grupo NR1P (CgNR1P2) e dois da subfamília NR2, grupo NR2E (CgNR2E1 e CgNR2E5). O grupo NR1P é composto por 11 NRs em *C. gigas*, sendo um grupo novo identificado na espécie, que não possui homólogo próximo conhecido em espécies modelo de invertebrado ou vertebrado, o que dificulta seu estudo através de similaridade, comparações e homologia, representando um desafio para sua caracterização estrutural e funcional. Vogeler e colaboradores (2014) apontaram que este novo grupo NR1P

possui NRs cujos domínios apresentam homologia com o grupo NR2E, o outro grupo hiperexpressado na presença de TAM no transcriptoma. O grupo NR2E é composto por quatro receptores. CgNR2E1 possui maior identidade com os homólogos de *D. melanogaster tailless* (DmTLL), *dissatisfaction* (DmDSF) e receptores TLX humanos (VOGELER et al., 2014). Todos eles estão envolvidos no desenvolvimento embrionário do eixo ântero-posterior e estão presentes no sistema nervoso, relacionados com comportamento emocional. CgNR2E5 apresenta homologia com DHR83, um receptor de função desconhecida de *D. melanogaster* (KING-JONES; THUMMEL, 2005).

Devido às limitações e artefatos da técnica de modelagem pelo SWISS-MODEL observada durante nosso trabalho, os modelos gerados para os três NRs identificados no transcriptoma demonstraram um enovelamento atípico, com hélices faltando e regiões em que o modelo não adquiriu formação definida de estrutura secundária. Uma segunda etapa de modelagem para o domínio LBD desses três NRs pelo SWISS-MODEL foi realizada, utilizando *templates* diferentes, na tentativa de obter modelos com enovelamentos mais confiáveis. Os modelos de CgNR2E1 e CgNR2E5, apresentaram *folding* mais próximos do esperado para o domínio LBD, porém o modelo para CgNR1P2 continuou apresentando enovelamento com um desvio muito grande do padrão para NRs. A fim de melhoramento e refinamento desses três NRs promissores, optamos por realizar a modelagem através de outro programa, com uma metodologia distinta, utilizando o pacote I-TASSER.

#### 4.3 MODELOS TRIDIMENSIONAIS I-TASSER

A segunda metodologia utilizada neste trabalho foi por modelagem híbrida, comparativa e por fragmentos, disponível na plataforma I-TASSER. Através da construção desta metodologia foi possível obter três modelos de NRs de *C. gigas* identificados como possíveis candidatos de modulação pelo estradiol e tamoxifeno através do transcriptoma. Identificamos que a metodologia do I-TASSER, que utiliza vários *templates* como base para a construção do modelo final demonstrou ser mais promissora para caracterizar estruturalmente o domínio LBD da espécie não modelo de invertebrado *Crassostrea gigas*. Por ser um número menor de modelos para verificação, comparativamente aos 43 que partimos inicialmente, foi possível realizar uma análise de qualidade minuciosa, composta de mais etapas de verificação. A precisão global estimada dos modelos gerados pelo I-TASSER,

ou seja, a confiança de cada modelo, é demonstrada quantitativamente pelo *C-score*, que é calculado com base na significância dos alinhamentos do método de *threading* e nos parâmetros de convergência das simulações de montagem da estrutura. O *C-score* se encontra tipicamente na faixa de -5 a 2 e um *C-score* de maior valor significa um modelo com maior confiança. Valores maiores que -1.5 indicam um modelo com topologia global correta. Com base no *C-score* e comprimento da proteína, valores de *TM-score* e RMSD são estimados para o modelo (ZHANG, SKOLNICK, 2004).

O *TM-score* (*template modeling score*) é uma métrica para calcular a similaridade de duas estruturas de proteínas, para a qual pontuações abaixo de 0,17 correspondem a proteínas não relacionadas escolhidas aleatoriamente, enquanto que estruturas com uma pontuação maior que 0,5 assumem geralmente o mesmo enovelamento (ZHANG, SKOLNICK, 2004). O RMSD (Desvio Médio da Raiz Quadrática) descreve a distância média por resíduo entre duas estruturas otimamente alinhadas. Possui a propriedade de comparar diretamente as coordenadas do espaço entre duas estruturas (BETANCOURT; SKOLNICK, 2001). Na Tabela 4 é possível observar que os modelos gerados apresentam valores de *C-score*, *TM-score* e RMSD permitidos para modelos com qualidade (ROY, KUCUKURAL & ZHANG, 2010).

Tabela 4. Métricas de qualidade dos modelos gerados. O *C-score* se encontra tipicamente na faixa de -5 a 2, sendo que valores maiores que -1.5 indicam um modelo com topologia global correta. Valores de *TM-score* maiores que 0,5 representam comparações entre proteínas que assumem geralmente o mesmo enovelamento. O RMSD descreve a distância média por resíduo entre duas estruturas otimamente alinhadas, dessa forma, quanto menor o seu valor, mais próximo de 0, mais similares são as estruturas proteicas comparadas.

	<b>CgNR1P2</b>	<b>CgNR2E1</b>	<b>CgNR2E5</b>
<i>C-score</i>	-0.26	0.64	0.69
<i>TM-score</i> estimado	0.68±0.12	0.80±0.09	0.81±0.09
RMSD (Å) estimado	6.3±3.9	4.5±2.9	4.2±2.8

Fonte: Própria autora.

A metodologia empregada pelo pacote I-TASSER, diferente do *software* SWISS-MODEL, utiliza vários *templates* como base para a modelagem e as métricas de *C-score*, *TM-score* e RMSD são gerados levando em conta estes diferentes moldes utilizados. Valores de RMSD baixos, entre 1 e 2, indicam modelos com menores desvios em sua

estrutura comparada com o molde utilizado para a construção do mesmo, indicando um modelo com maior qualidade. Os valores de RMSD da Tabela 4 levaram em conta para seu cálculo os diferentes *templates* que compunham a base para a construção do modelo, justificando um desvio maior no seu valor, ainda assim considerado como dentro dos limites de qualidade. O *output* do I-TASSER inclui o alinhamento do modelo gerado com estruturas de proteínas depositadas em bancos de dados com as quais o modelo possui muita similaridade. A Tabela 5 apresenta os valores de *TM-score* e RMSD para os modelos gerados comparativamente às estruturas do banco de dados às quais eles possuem maior identidade estrutural.

O espaço que peptídeos podem ocupar no espaço é restrito, devido a confrontos estéricos entre a cadeia lateral dos resíduos de aminoácidos. Observando o enovelamento refinado de uma estrutura proteica, é possível observar que os ângulos migram para conformações permitidas, indicando que a distribuição dos ângulos em uma proteína pode fornecer um guia para a qualidade de sua estrutura (MORRIS et al., 1992). A Tabela 5 apresenta os ângulos de Ramachandran permitidos para os modelos gerados (LASKOWSKI et al., 1993). As restrições espaciais encontradas para nossos modelos, em geral, são próximas às descritas para estruturas experimentais. Diagramas de Ramachandran exemplificam a distribuição dos ângulos dos resíduos na estrutura proteica (APÊNDICE A).

Tabela 5. Métricas de qualidade estrutural e ângulos de conformações permitidos para os modelos obtidos e estruturas do PDB utilizadas como referência para o cálculo.

Modelo	Estrutura do PDB	TM-score	RMSD(Å)	Ramachandran permitidos
CgNR1P2	2NXX	0.939	1.12	96.9%
CgNR2E1	1QKU	0.950	1.45	99.1%
CgNRE5	1QKU	0.946	1.31	98.5%

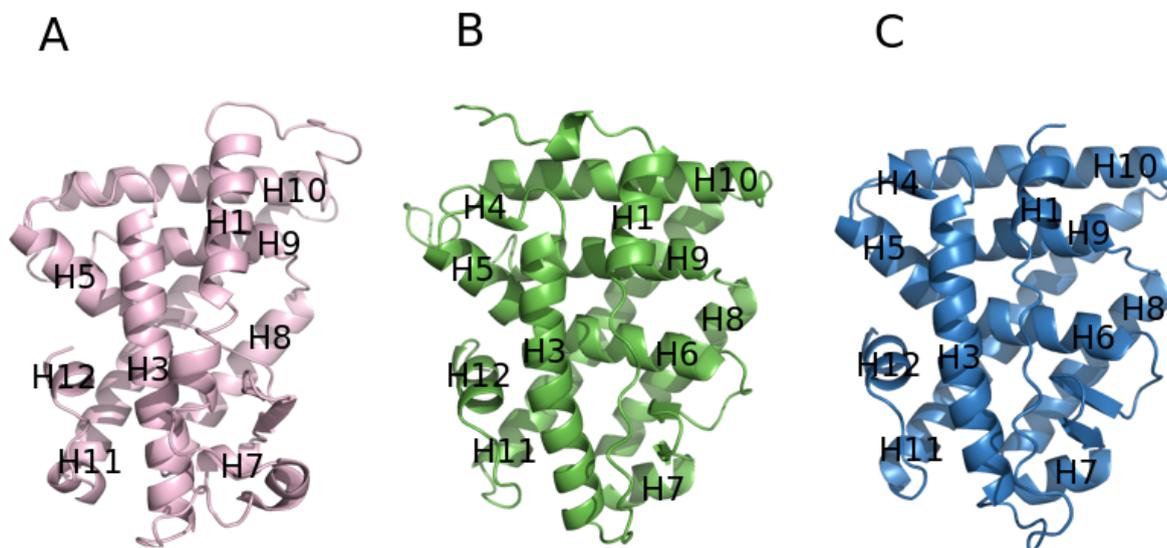
Fonte: Própria autora.

As estruturas resolvidas experimentalmente alinhadas aos modelos de *C. gigas* mostradas na Tabela 5 são todas de receptores que possuem como ligantes moléculas derivadas de colesterol, como o estradiol. O cristal 2NXX, alinhado ao CgNR1P2,

corresponde ao domínio LBD de um receptor de ecdisona de um invertebrado artrópode, *Tribolium castaneum*, cujo ligante é o hormônio ponasterona A. O cristal 1QKU, alinhado por similaridade estrutural ao modelo CgNR2E1 e CgNR2E5, é um ER humano cujo ligante presente no cristal é o estradiol. Podemos notar que os modelos obtidos possuem similaridade com outros NRs cuja estrutura, sítio de ligação e ligantes apresentam características aproximadas aos nossos alvos de interesse, o estradiol e o tamoxifeno.

Além dos parâmetros de qualidade descritos acima, foi realizada uma análise do enovelamento estrutural apresentado por cada modelo, a fim de verificar se estes possuem características típicas do domínio LBD de NRs. Estas incluem a presença de 11 a 13  $\alpha$ -hélices (H0 a H12) em um arranjo espacial de três camadas anti-paralelas que se assemelha a um sanduíche. As hélices H4, H5, H8 e H9 estão presentes na porção do topo do domínio, o que permite a formação de uma cavidade na parte inferior na maioria dos NRs, entre as hélices H3, H7 e H11. Entre as  $\alpha$ -hélices H6 e H7 ocorrem duas folhas- $\beta$  em forma de grampo. Comparações estruturais revelaram que as hélices que compõem o topo do domínio são conservadas, o que mais difere é a região que envolve as hélices H2, H3 e o grampo- $\beta$ , sugerindo que a formação em sanduíche é evolutivamente mantida para que a ligação de pequenas moléculas seja possível na maioria dos NRs (LI; LAMBERT; XU, 2003). A Figura 8 demonstra o enovelamento encontrado para os modelos gerados pelo I-TASSER. CgNR1P2 (A) não forma H4 e H6. Esse NR pertence ao novo grupo descrito para *C. gigas*, sem homólogos próximos, o que dificulta a construção de modelos por homologia e explica a dificuldade de encontrar um enovelamento esperado. Ambos CgNR2E1 (B) e CgNR2E5 (C), membros do mesmo grupo (NR2E) não possuem apenas H2. O homólogo humano de CgNR2E1, receptor TLX é conhecido por não apresentar H1 e H2 (BENOD et al., 2014), sendo possível que representantes desse grupo não apresentem essas hélices.

Figura 8 - Estruturas tridimensionais obtidas através da técnica de modelagem. (A) Modelo CgNR1P2, o qual não apresenta formação de alfa-hélice na região que corresponderia a H4 e H6. (B) Modelo CgNR2E1 e (C) Modelo CgNR2E5, ambos pertencentes ao grupo NR2E, não apresentam formação de H2.



Fonte: Própria autora.

A utilização do CgER como modelo controle na modelagem pelo SWISS-MODEL foi também empregado para os modelos gerados pelo I-TASSER, para as mesmas finalidades de validação e comparação entre parâmetros. Retiramos manualmente na submissão para modelagem no servidor do I-TASSER a possibilidade de utilização do próprio cristal de CgER (4N1Y) como um dos *templates* base para gerar o modelo. A Tabela 6 mostra valores de *C-score*, *TM-score* e *RMSD* que o I-TASSER calcula para o modelo gerado, utilizando como referência os vários *templates* que ele utiliza para a construção do modelo. Os valores das métricas de qualidade encontram-se dentro dos limites para um modelo de qualidade. Na Tabela 7 são demonstrados *TM-score* e *RMSD* referentes ao alinhamento do modelo de CgER com o cristal 4N1Y, o próprio cristal da estrutura de CgER da ostra, que apesar de não ter sido usado para gerar o modelo, possui uma estrutura muito similar ao que a técnica de modelagem gerou e por isso compõe os resultados do I-TASSER da modelagem. Os valores de Ramachandran permitidos também compõe a Tabela 7, demonstrando que os ângulos encontrados para os resíduos de aminoácidos na proteína estão dentro do permitido. O diagrama de Ramachandran para esse modelo se encontra no APÊNDICE A.

Tabela 6. Métricas de qualidade para a estrutura do modelo controle.

	<b>CgNR3A</b>
<i>C-score</i>	0.56
TM-score estimado	0.79±0.09
RMSD estimado	4.6±3.0

Fonte: Própria autora.

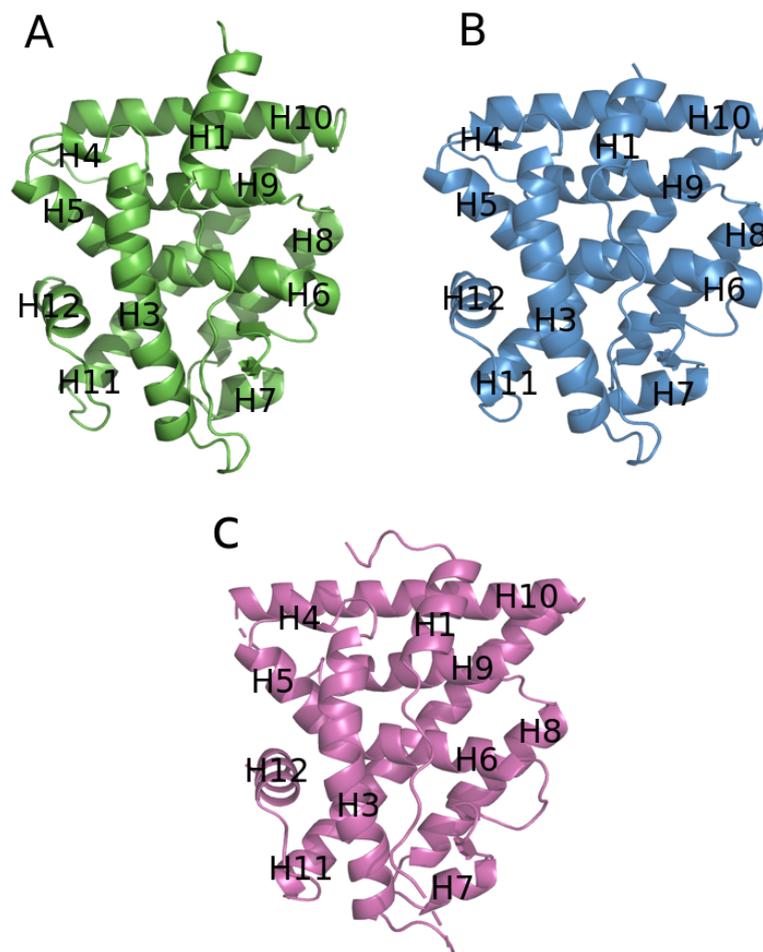
Tabela 7. Métricas de qualidade e ângulos de conformações permitidos para o modelo controle e estrutura do PDB utilizada como referência para o cálculo.

<b>Modelo</b>	<b>Estrutura do PDB</b>	<b>TM-score</b>	<b>RMSD (Å)</b>	<b>Ramachandran permitidos</b>
CgNR3A	4N1Y	0.927	0.32	100%

Fonte: Própria autora.

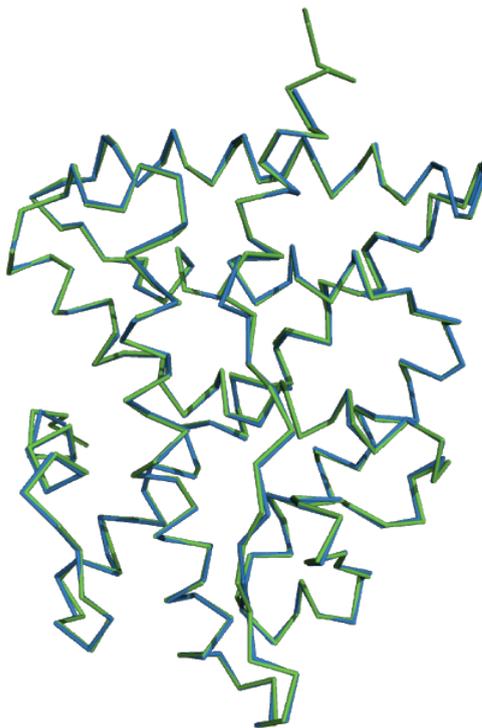
Uma análise do enovelamento da estrutura 3D do modelo CgNR3A demonstrou *fold* típico do domínio LBD. A Figura 9 mostra o modelo de CgER (A) com 11  $\alpha$ -hélices definidas e ausência de H2, conforme o esperado para o domínio LBD de ER (ASCENZI; BOCEDI; MARINO, 2006) e assim como ocorre no cristal de CgER 4N1Y (B) e em seu homólogo humano 1GWR (C). A Figura 10 apresenta o alinhamento entre o modelo do ER de *C. gigas* com sua estrutura de cristal 4N1Y, demonstrando uma grande similaridade (RMSD: 0.32 Å) entre as estruturas alinhadas, reforçando um modelo que se aproxima do observado experimentalmente.

Figura 9 - Enovelamento tridimensional correspondente ao modelo controle e às estruturas cristalográficas de referência para este modelo. Estrutura 3D do modelo CgER (A), da estrutura cristalográfica de CgER 4N1Y (B) e da estrutura cristalográfica de ER humano 1GWR (C). Todos apresentam 11 alfa-hélices bem definidas, com ausência de H2, recorrente nos membros desse grupo.



Fonte: Própria autora.

Figura 10 - Alinhamento dos esqueletos carbônicos do modelo CgER (verde) e da estrutura resolvida experimentalmente de CgER 4N1Y (azul).



Fonte: Própria autora.

#### 4.4 CARACTERIZAÇÃO DAS CAVIDADES DO SÍTIO DE LIGAÇÃO

As cavidades que compõem o sítio de ligação de CgNR1P2, CgNR2E1, CgNR2E5 foram calculadas pelo programa CASTp 3.0, utilizando como padrão o *radius probe* de 1,3 Å como área de superfície acessível ao solvente. A Tabela 8 apresenta o volume da cavidade dos três modelos candidatos possíveis de interagir com o estradiol. O critério de seleção para identificar modelos com cavidades com volume suficiente para comportar estradiol continuou utilizando como base o valor da cavidade de ER humano, que possui volume de 265.851 Å<sup>3</sup>. Os três modelos gerados pelo I-TASSER apresentam volumes de cavidade que se aproximam ao da estrutura de referência 1GWR, confirmando seu *status* de candidatos possíveis de serem modulados por estradiol de acordo com o volume de suas cavidades.

Tabela 8. Volumes das cavidades do sítio de ligação dos modelos do domínio LBD de NR1P2, NR2E1 e NR2E5 de *C.gigas*.

	Volume Å <sup>3</sup>
CgNR1P2	245.742
CgNR2E1	221.582
CgNR2E5	379.810

Fonte: Própria autora

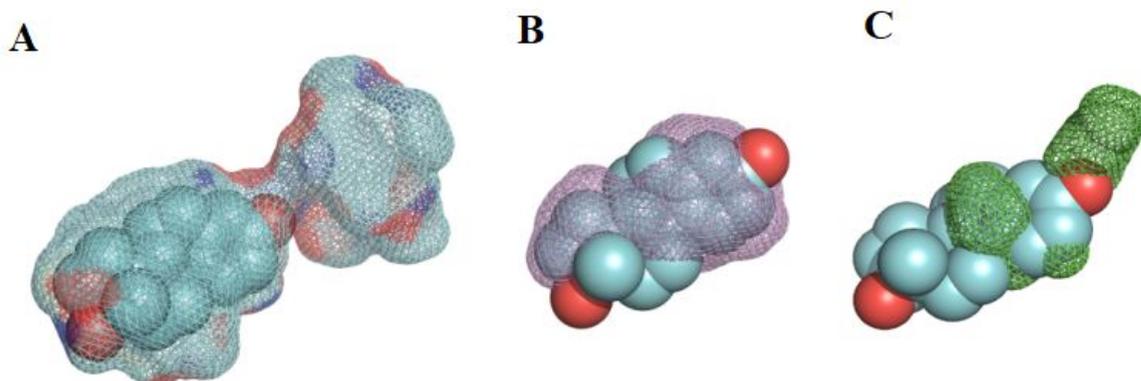
O modelo controle CgER demonstrou uma cavidade reduzida do sítio de ligação, com volume menor que a descrita para a estrutura cristalográfica de CgER. Os dados da quantificação do volume da cavidade demonstram que, assim como a estrutura resolvida experimentalmente do CgER 41NY, o modelo também não comporta o estradiol em seu sítio de ligação, devido a uma cavidade de volume reduzido. A Tabela 9 mostra os valores de cálculo da cavidade para o modelo CgER, a estrutura experimental de CgER (4N1Y) e a estrutura experimental de ER humano 1GWR. É possível notar um volume reduzido da cavidade do modelo tanto na Tabela 9 quanto na Figura 11, que ilustra a superfície das cavidades de (A) estrutura cristalográfica de ER humano (1GWR), (B) estrutura cristalográfica do CgER (4N1Y) e (C) modelo de CgER na presença do ligante estradiol.

Tabela 9. Quantificação do espaço da cavidade do sítio de ligação do modelo controle NR3A de *C. gigas*.

	Área (Å <sup>2</sup> )	Volume (Å <sup>3</sup> )
CgER modelo	98.907	20.425
4N1Y (estrutura experimental ER de <i>C. gigas</i> )	126.213	63.867
1GWR (estrutura experimental ER humano)	445.637	265.851

Fonte: Própria autora.

Figura 11 - Comparação do volume de superfície da cavidade entre o ER humano, a estrutura cristalográfica de CgER e o modelo CgER, com a presença do estradiol interno ao sítio de ligação (A) Superfície de volume da cavidade do sítio de ligação da estrutura cristalográfica de ER humano (1GWR), suficientemente grande para comportar o estradiol. (B) Superfície de volume da cavidade do sítio de ligação da estrutura cristalográfica de CgER (4N1Y), apresentando cavidade reduzida para comportar o ligante estradiol, assim como o modelo de CgER (C).



Fonte: Adaptado de Santos, 2018.

A verificação de que os modelos de NR1P2, NR2E1 e NR2E5 de *C. gigas* possuem cavidades dos sítio de ligação com volume aproximado suficiente para comportar estradiol e que apresentaram hiperexpressão no transcriptoma na presença de tamoxifeno, confirmam a possibilidade de estes três NRs poderem interagir com essas moléculas. Devido a isso, os testes subsequentes de *docking* molecular foram realizados com esses três modelos promissores e os xenobióticos estradiol e tamoxifeno.

#### 4.5 POTENCIAL DE INTERAÇÃO DA CAVIDADE COM ESTRADIOL

Através dos dados que recebemos apontando CgNR1P2, CgNR2E1 e CgNR2E5 como candidatos mais prováveis de comportar o estradiol e padronizando as próximas etapas do estudo com modelos gerados pela mesma metodologia, a verificação da possível interação do estradiol e tamoxifeno foi testada com esses modelos. Primeiramente, o potencial de interação do estradiol com os três modelos foi testado através da técnica de *docking* molecular. O resultado do teste identificou as posições em que a molécula de estradiol possui maior

probabilidade de interagir com o sítio de ligação do domínio LBD a partir da identificação de interações com resíduos de aminoácidos e minimização da energia livre de ativação da reação.

O resultado da técnica identificou várias posições possíveis de interação entre a cavidade do sítio de cada modelo e o ligante. Para selecionar as posições de *docking* que caracterizam as interações mais prováveis de acontecer entre o NR e estradiol, a partir do que conhecemos de estruturas resolvidas experimentalmente, utilizamos três critérios. Primeiramente alinhamos a estrutura experimental de ER humano (1GWR) ao nosso modelo e verificamos as posições do teste de *docking* em que houve sobreposição do estradiol no sítio do modelo de *C. gigas* com a posição que este ligante ocupa na cavidade do sítio de ligação humano. Para as posições sobrepostas foi realizado cálculo do RMSD, a fim de selecionar quais posições do nosso teste apresentam maior identidade com a posição que o ligante ocupa na estrutura experimental de referência. As três posições com menores RMSDs do resultado de *docking* para cada modelo foram selecionadas para seguir com os próximos critérios de análise. Devido a serem muitos dados, a posição de menor RMSD está descrita no texto e as duas posições subsequentes analisadas encontram-se no APÊNDICE B. Utilizamos a ferramenta DockRMSD (BELL & ZHANG, 2019) para realizar os cálculos. A visualização 3D das posições do *docking* e alinhamento com estrutura de referência (1GWR) foram feitas no PyMOL (SCHRÖDINGER, 2017).

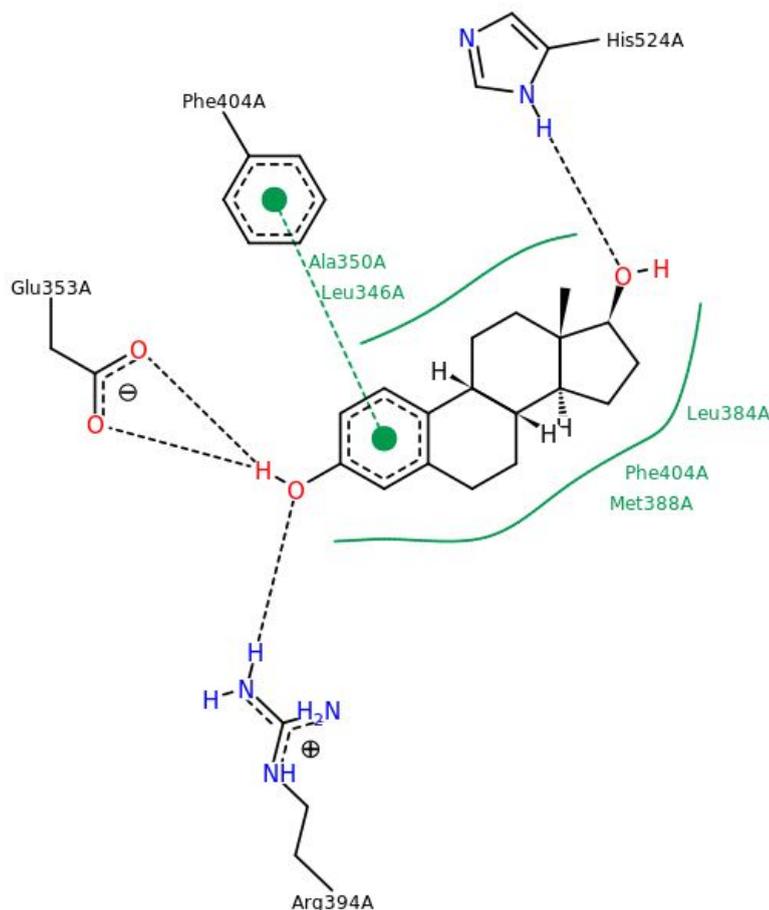
O segundo critério foi composto por uma análise qualitativa das posições do estradiol, observando quais interações químicas fracas foram possíveis no sítio de ligação em cada posição identificada no *docking*. Para esta análise, foi utilizado o programa PoseView (STIERAND; RAREY, 2010) e LigPlus (WALLACE; LASKOWSKI; THORNTON, 1995). Diagramas bidimensionais de complexos da estrutura 3D dos NRs com os ligantes foram geradas através desses *softwares*, de acordo com as convenções de desenho químico. A estrutura experimental do ER humano possui o estradiol ligado em seu sítio de ligação e por isso foi utilizado como referência, a fim de identificar quais interações são cruciais para que o complexo proteína-ligante seja estável.

Os valores de  $\Delta G$  (variação da energia livre) de cada posição do estradiol resultante do teste de *docking* compuseram o terceiro critério de análise. Valores de energia livre correspondem à energia capaz de realizar trabalho durante uma reação à temperatura e pressão constantes. Quando  $\Delta G$  é negativo, os produtos contêm menos energia livre que os reagentes, o que faz com que a reação ocorra espontaneamente em condições padrão. Isso ocorre pois as

reações químicas tendem a ocorrer no sentido de uma diminuição da energia livre do sistema. Isso exemplifica uma reação exergônica (NELSON; COX, 2006). Os valores negativos de  $\Delta G$  observados em nossos resultados caracterizam interações possíveis de ocorrer espontaneamente.

O sítio de ligação da estrutura experimental do ER humano, utilizado como base para o entendimento da interação do domínio LBD de um NR com o estradiol, é composto por resíduos de aminoácidos das hélices 3, 4, 6, 8 e 11. Três resíduos são importantes para a ligação do ER humano com o estradiol: Glu353, Arg394 e His524. Estes são os resíduos de aminoácidos com cadeias laterais capazes de formar pontes de hidrogênio com o ligante, interações importantes para a estabilização deste na cavidade do sítio. Além dessas ligações polares, o sítio de ligação é em sua maioria hidrofóbico (CELIK; LUND; SCHIØTT, 2007). A Figura 12 mostra as interações químicas que ocorrem entre os resíduos de aminoácidos da estrutura cristalográfica de ER humano (1GWR) e o estradiol no sítio de ligação. Por conservar essa rede de interações, foi observado em dinâmicas moleculares que H3 e H11 são mantidas em contato próximo, impedindo que o ligante se dissocie da cavidade do sítio (CELIK; LUND; SCHIØTT, 2007).

Figura 12 - Interações químicas entre resíduos de aminoácidos que compõem o sítio de ligação do domínio LBD da estrutura cristalográfica de ER humano (1GWR) com o estradiol. Linhas tracejadas pretas indicam ligações de hidrogênio, ocorrendo com Glu353, Arg394 e His524. Linhas sólidas verdes mostram interações hidrofóbicas, que ocorrem com os resíduos Ala350, Leu346, Leu384, Met388 e Phe404. Linhas tracejadas verdes mostram interações  $\pi$ - $\pi$ , que ocorre entre o anel aromático do resíduo Phe404 e o anel aromático do estradiol.



Fonte: Representação gráfica gerada pelo programa PoseView.

A posição de interação do estradiol no teste de *docking* com os modelos de *C. gigas* que apresentou menor RMSD em relação a posição que este ligante ocupa no sítio de ligação da estrutura experimental de ER humano são mostradas na Tabela 10, assim como os valores de variação da energia livre para cada uma das posições. As posições de *docking* descritas para cada modelo constituem uma localização espacial do ligante no sítio de ligação aproximado ao que se observa na estrutura resolvida de referência. Em um alinhamento das estruturas (cristal de referência e modelos de *C. gigas*) com os ligantes associados ao sítio de ligação, é possível observar sobreposição dos ligantes na cavidade do sítio, demonstrando que ocupam locais muito aproximados, como indica o RMSD da Tabela 10. Os valores negativos de  $\Delta G$  configuram essas posição de interação como possíveis de acontecer espontaneamente

em condições padrão, uma vez que ocorrem no sentido de minimização da energia livre do sistema. As análises qualitativas da posição de menor RMSD para cada modelo são descritas separadamente a seguir.

Tabela 10. Dados de análise quantitativa para posições do teste de *docking* com estradiol.

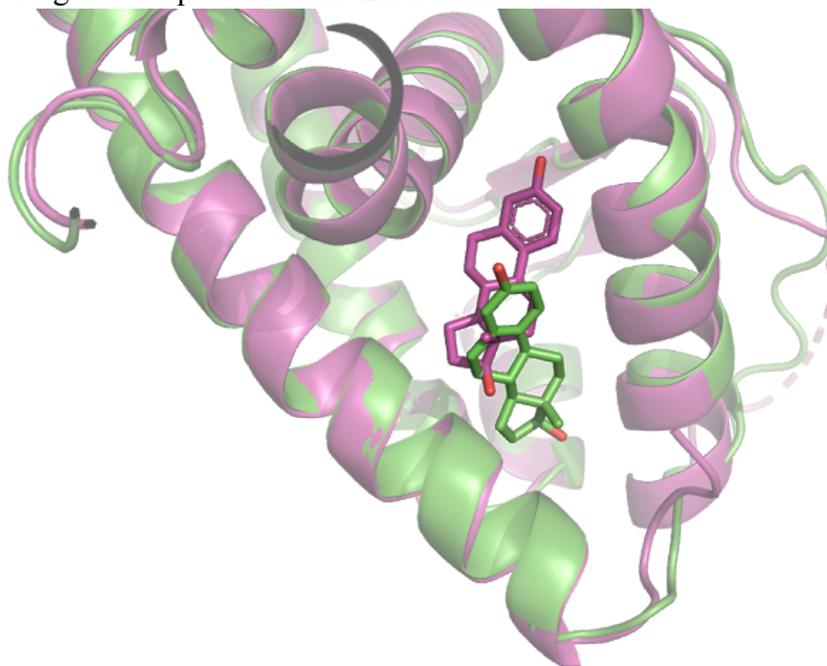
	RMSD (Å)	$\Delta G$
CgNR2E1	4.875	-8.9
CgNR2E5	5.436	-9.5
CgNR1P2	7.981	-9.2

Fonte: Própria autora.

#### 4.5.1 CgNR2E1

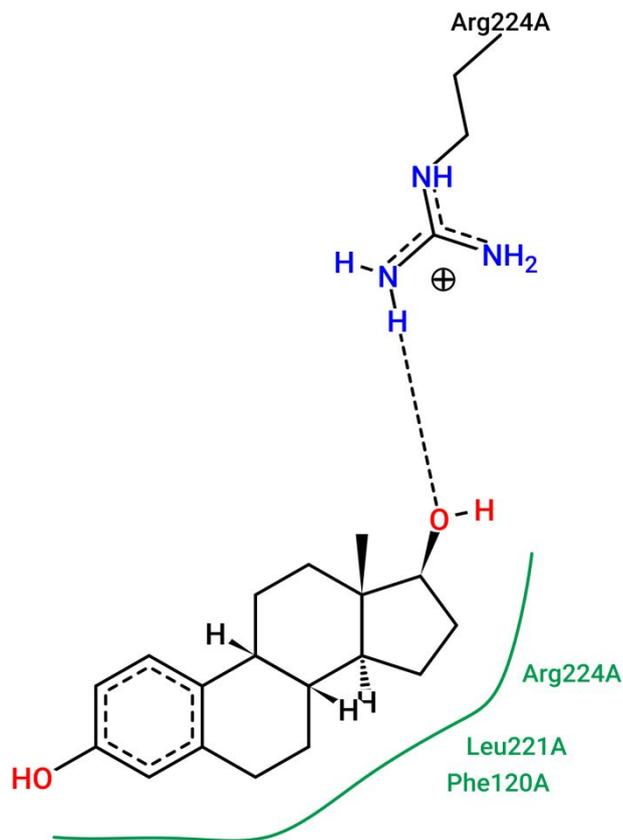
O teste de *docking* com o modelo CgNR2E1 e o estradiol apresentou o menor valor de RMSD entre os três modelos testados, indicando que as posições em que o ligante se acomodou no sítio de ligação foram as mais próximas das que o estradiol ocupa no sítio do ER humano. O valor de RMSD da posição com menor desvio comparado à estrutura de referência encontra-se na Tabela 10, que também apresenta o valor de  $\Delta G$  correspondente à posição descrita. A Figura 13 mostra um alinhamento entre o ER humano e CgNR2E1, demonstrando como o estradiol presente no sítio de ligação de ambos os receptores ocupa espaço similar. Analisando qualitativamente esta posição, identificamos interações químicas que também ocorrem com o estradiol no sítio do ER humano (Figura 12). A Figura 14 mostra a formação de uma ponte de hidrogênio entre o resíduo Arg224 e o estradiol, além de interações hidrofóbicas com Phe120, Leu221 e Arg224 na posição de menor RMSD do teste para este modelo. Os resíduos representados pertencem às hélices 3, 8 e 11, todas com estrutura secundária bem definida. As outras duas posições com menor RMSD identificadas no *docking* para CgNR2E1 possuem seus dados de análise quantitativa e qualitativa presentes no APÊNDICE B.

Figura 13 - Alinhamento da estrutura cristalográfica do ER humano (1GWR), com o ligante estradiol (em roxo) interno à cavidade do sítio de ligação e do modelo CgNR2E1 com o estradiol (em verde) interno ao sítio de ligação na posição de menor RMSD em relação à posição que este ligante ocupa no sítio do ER humano.



Fonte: Própria autora.

Figura 14 - Interações químicas entre o estradiol e os resíduos de aminoácidos que compõem a cavidade do sítio de ligação do modelo CgNR2E1. Nessa posição há formação de uma ponte de hidrogênio entre o resíduo Arg224 e o estradiol, além de interações hidrofóbicas com Phe120, Leu221 e Arg224.



Fonte: Representação gráfica gerada pelo programa PoseView.

O receptor CgNR2E1 pertence a subfamília NR2, grupo NR2E. Este receptor da ostra possui homólogo com os NRs de *D. melanogaster tailless* (DmTLL) e *dissatisfaction* (DmDSF), além de homologia com receptor TLX humano (VOGELER et al., 2014). DmTLL em *Drosophila* está envolvido na embriogênese, é expresso no sistema nervoso em desenvolvimento e também possui funções no desenvolvimento depois de adulto (RUDOLPH et al., 1997). Ele atua principalmente na repressão da transcrição de alguns genes. Seu ortólogo em mamíferos, TLX, é restrito ao prosencéfalo, é muito expresso no cérebro embrionário e nas células-tronco neurais adultas, desempenhando, como em *D. melanogaster*, importantes papéis no sistema nervoso. Semelhante ao seu homólogo da mosca, o TLX funciona como um repressor transcricional. DmDSF, assim como DmTLL e TLX, funciona como repressor de transcrição (KING-JONES; THUMMEL, 2005). Mutações no gene para

DmDSF afetam o comportamento sexual adulto. Os machos mutantes podem ter menor eficiência na cópula, já as fêmeas podem apresentar uma resistência aumentada aos machos na reprodução, além de deixarem de pôr ovos maduros, principalmente porque os músculos uterinos não são inervados pelo conjunto necessário de neurônios motores. As mutações de TLX demonstraram potencial de levar à agressão em mamíferos, indicando que as funções comportamentais dessa classe de NRs foram conservadas através da evolução. King-Jones; Thumel (2005) sugerem que, considerando que os mutantes de TLX também apresentam defeitos graves no sistema límbico, que está envolvido com a libido em humanos, seria possível conjecturar sobre um possível papel de TLX no comportamento sexual de mamíferos.

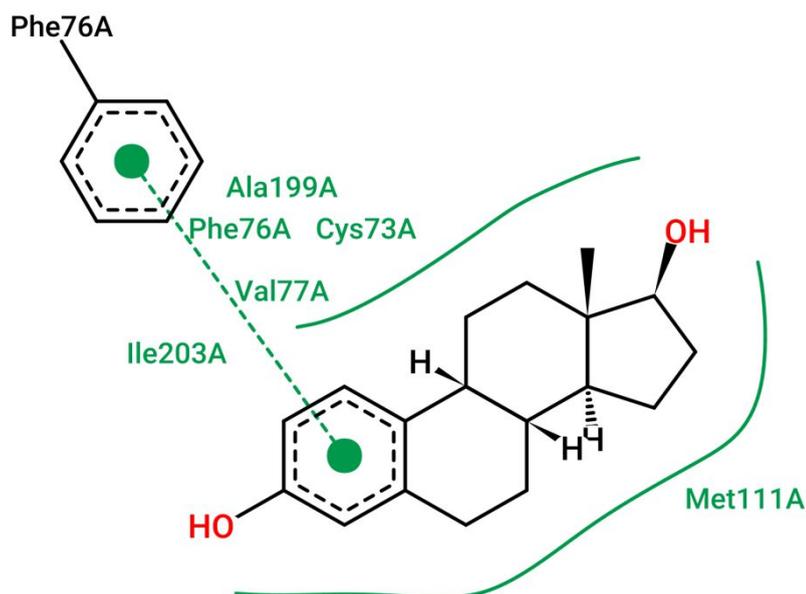
O receptor TLX em humanos é descrito como receptor órfão, sem um ligante identificado. Está presente no sistema nervoso, ligado com células tronco e desenvolvimento do sistema visual (WANG; XIONG, 2016). TLX pertence a uma classe emergente de NRs que não possuem as hélices  $\alpha 1$  e  $\alpha 2$  no LBD e que tem potencial para formar um grande bolsão aberto, Bolsão de Interação do Ligante - *Ligand Binding Pocket* (LBP). Benod et al. (2014) testaram diversos compostos a fim de verificar a possibilidade do LBP deste receptor ser modulado por algum ligante e verificaram a interação com três moléculas. Ccrp1 (famprofazona), um agente anti-inflamatório não esteróide da série pirazolona, ccrp2, que é semelhante ao ccrp1 com grupos funcionais de pirazol e piperazina e ccrp3, drorogesterona, uma progestina esteróide. O estudo aponta que outros esteróides devem ser testados no futuro em TLX, devido às suas estruturas químicas e porque se observou que compostos com estrutura química com esqueleto, *scaffold*, de esteróides se ligam a outros LBPs de NR, sugerindo fortemente que esses compostos se ligam ao LBP de TLX (BENOD et al., 2014).

Esses dados da literatura corroboram com a ligação possível que encontramos do esteróide estradiol com o receptor NR2E1 da ostra do pacífico. É possível discutir sobre funcionalidades deste receptor nuclear a partir das funções conservadas evolutivamente que este desempenha em organismos, tanto invertebrados como vertebrados. Um possível papel como NR repressor da transcrição no sistema nervoso pode ser apontado, sofrendo possível modulação pelo hormônio estradiol.

#### 4.5.2 CgNR2E5

A posição do teste de *docking* para o modelo CgNR2E5 que apresentou o menor RMSD em comparação com a estrutura experimental de referência compõe a Tabela 10, juntamente com o valor de  $\Delta G$  correspondente. Nas hélices 3, 6, 8 e 11 do modelo estão presentes os resíduos de aminoácidos que compõem a cavidade do sítio de ligação e interagem com o estradiol em nossos testes. O diagrama bidimensional para a posição de menor RMSD é mostrado na Figura 15, compondo a análise qualitativa dessa posição do estradiol. Nesta posição muitas interações hidrofóbicas ocorrem entre os resíduos Cys73, Phe76, Met111, Val177, Ala199 e Ile203 e o estradiol. Também ocorre uma ligação  $\pi$ - $\pi$  entre o anel aromático do ligante e de Phe76. As outras duas posições com menor RMSD identificadas no *docking* para CgNR2E5 possuem seus dados de análise quantitativa e qualitativa presentes no APÊNDICE B.

Figura 15 - Interações químicas entre o estradiol e os resíduos de aminoácidos que compõem a cavidade do sítio de ligação do modelo CgNR2E5. Na posição mostrada, muitas interações hidrofóbicas ocorrem entre o estradiol e os resíduos Cys73, Phe76, Met111, Val177, Ala199 e Ile203, além de uma ligação  $\pi$ - $\pi$  entre o anel aromático de Phe76 e o anel aromático do estradiol.



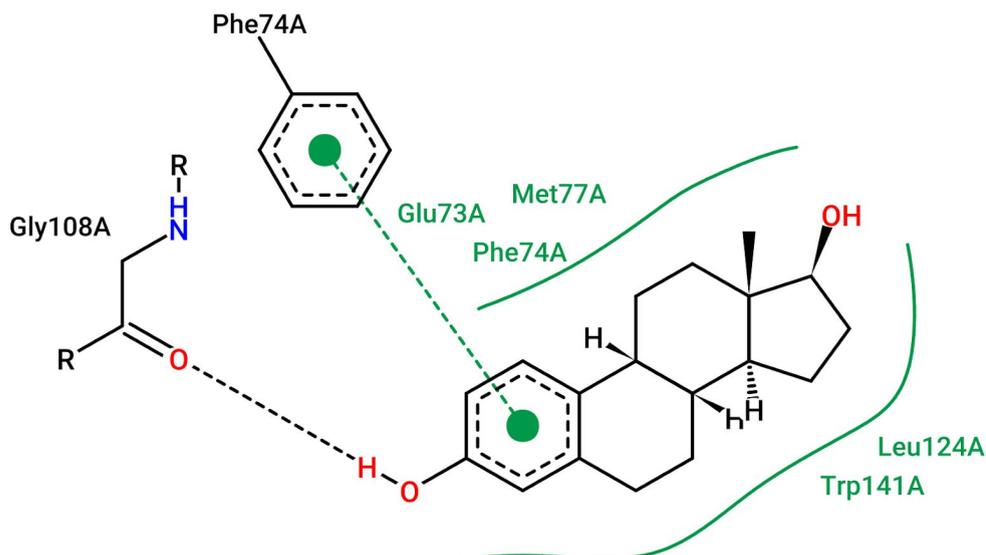
Fonte: Representação gráfica gerada pelo programa PoseView.

O representante NR2E5 de *C. gigas* possui homologia com o NR DHR83 de *D. melanogaster*. Este NR em *Drosophila* é um ortólogo pouco conservado de PNR, receptor nuclear específico de fotorreceptores de vertebrados, que embora não se tenha dados precisos sobre seu mecanismo de funcionamento, parece estar envolvido com desenvolvimento do sistema nervoso, assim como outros membros do grupo NR2E (KING-JONES; THUMMEL, 2005). Através da caracterização estrutural deste membro da subfamília NR2 em *C. gigas* e da quantificação do volume de sua cavidade, é possível propor candidatos de ligantes possíveis que ocupem seu sítio de ligação, além da interação favorável com estradiol verificada em nosso estudo.

#### 4.5.3 CgNR1P2

As posições encontradas no *docking* para o modelo pertencente à subfamília NR1 apresentaram resultados indicando uma interação possível com o estradiol. A posição identificada com base no RMSD é mostrada em um diagrama bidimensional na Figura 16. A posição que representa o menor RMSD encontrado, apresentou ponte de hidrogênio com o resíduo Gly108, diversas interações hidrofóbicas com resíduos de Glu73, Phe74, Met77, Leu124, Trp141, além de uma ligação  $\pi$ - $\pi$  entre Phe74 e o anel aromático do estradiol. As hélices 3, 6, 8, e uma folha-beta do modelo contém os resíduos de aminoácidos que interagem com o estradiol. A interação polar ponte de hidrogênio, importante para estabilização do ligante na cavidade do sítio, foi identificada com o resíduo Gly108, localizado em H6. No entanto, o modelo não possui enovelamento em forma de alfa-hélice nesta região, como tipicamente ocorre na hélice 6 do domínio LBD. Se essa hélice se forma *in vivo* há a possibilidade da orientação desse resíduo mudar de posição e essa interação não ocorrer. As outras duas posições com menor RMSD identificadas no *docking* para CgNR1P2 possuem seus dados de análise quantitativa e qualitativa presentes no APÊNDICE B.

Figura 16 - Interações químicas entre o estradiol e os resíduos de aminoácidos que compõem a cavidade do sítio de ligação do modelo CgNR1P2. Posição 1 (A) apresenta ponte de hidrogênio com o resíduo Gly108A, diversas interações hidrofóbicas com resíduos de Glu73A, Phe74A, Met77A, Leu124A, Trp141A, além de uma ligação  $\pi$ - $\pi$  entre Phe74A e o anel aromático do estradiol.



Fonte: Representação gráfica gerada pelo programa PoseView.

O grupo NR1P possui 11 membros identificados recentemente, dos quais características estruturais e funcionais não podem ser deduzidas com base na filogenia, pois nenhum homólogo próximo foi identificado. No entanto, muitos membros da subfamília NR1 estão envolvidos com crescimento e desenvolvimento em humanos ou nos processos de muda e metamorfose em *D. melanogaster* (VOGELER et al., 2014), sugerindo que este grupo possui funções relacionadas a estas. Caracterizar estruturalmente o domínio LBD de CgNR1P2, quantificar sua cavidade e apontar possíveis interações com moléculas ligantes fornecem informações novas sobre este membro de funcionalidade desconhecida.

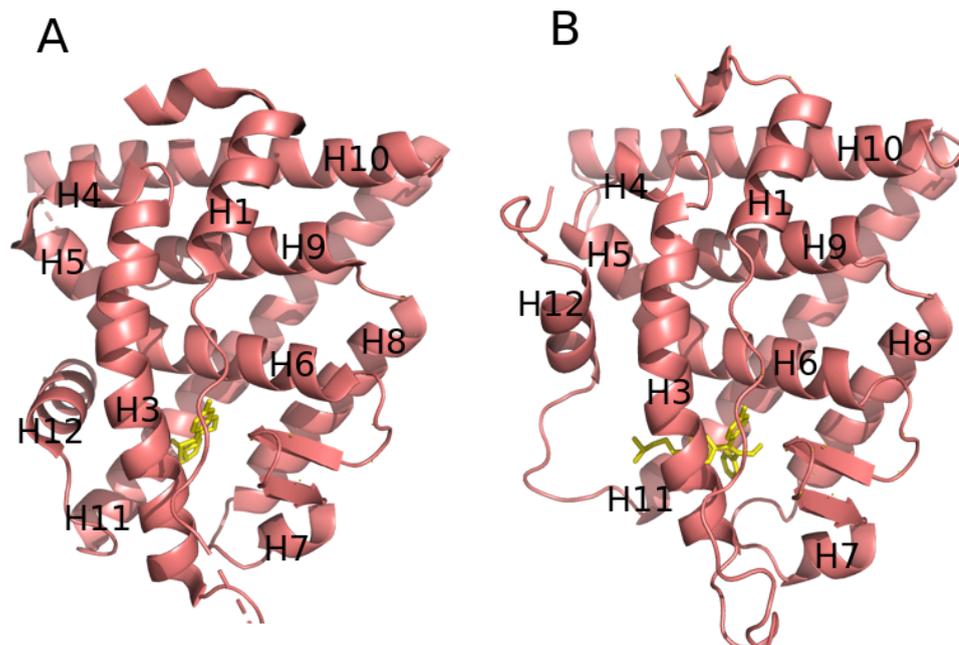
Os resultados do teste de *docking* com o ligante estradiol demonstraram interações favoráveis entre os 3 NRs candidatos possíveis de serem modulados por este ligante. Análises posteriores a essa podem ser realizadas para entender e verificar através de outras metodologias essa possível ligação, como testes de dinâmica molecular e testes *in vitro*.

#### 4.6 POTENCIAL DE INTERAÇÃO DA CAVIDADE COM TAMOXIFENO

Os modelos CgNR1P2, CgNR2E1 e CgNR2E5 tiveram sua possível interação com o fármaco tamoxifeno testada através da técnica de *docking* molecular. Cálculos de RMSD do ligante, análises qualitativas das interações entre resíduos de aminoácidos e o 4-hidroxitamoxifeno e os valores de  $\Delta G$  constituíram as análises dos resultados do teste, as mesmas utilizadas para o *docking* com o estradiol. A estrutura experimental utilizada como referência foi o cristal 3ERT, um ER humano, cristalizado com 4-hidroxitamoxifeno em seu sítio de ligação.

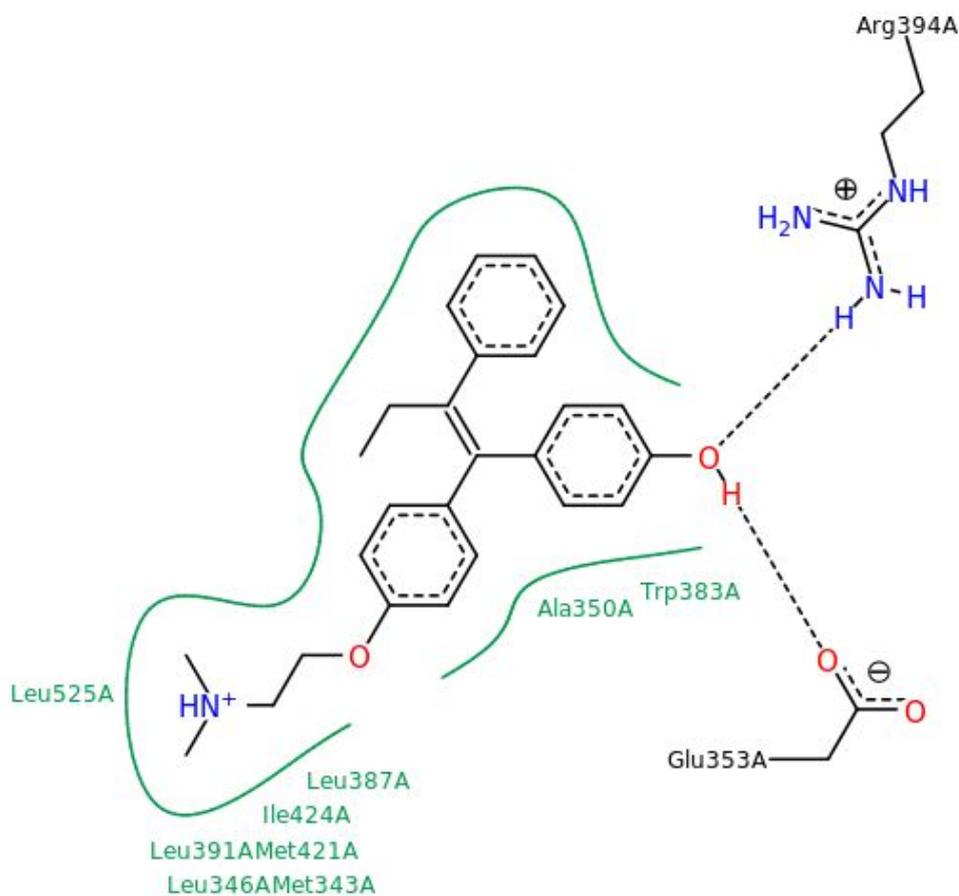
A ligação do fármaco tamoxifeno na cavidade do sítio do ER humano, ligante que possui ação antagonista em condições e tecidos específicos, provoca mudanças conformacionais no domínio LBD diferentes das que ocorrem com o ligante natural de ação agonista, o estradiol (SHIAU et al., 1998). NRs recrutam co-ativadores e correpressores através de sua função de ativação 2 (AF-2), localizada no domínio LBD C-terminal, entre as hélices H3, H5 e H12. A atividade de AF-2 depende da orientação de uma hélice móvel de ativação, H12. A ligação de um ligante agonista, como o ligante natural estradiol no ER humano, estabiliza uma posição ativa de H12, resultando na formação de um local de ligação específico para os co-ativadores. Os ligantes antagonistas de NR, como o tamoxifeno para o ER humano em muitos tecidos, induzem conformações no receptor nas quais H12 assume outra posição, de modo que é completamente afastada do corpo do LBD ou está interagindo com outro local fora da região AF-2. Essa conformação gerada pela ligação de um antagonista pode tornar o NR acessível a interagir com correpressores (HELDRING et al., 2007). A Figura 17 mostra estruturas experimentais do domínio LBD de ER humano ligado ao estradiol (A) e ao tamoxifeno (B), exemplificando as mudanças conformacionais neste domínio. A Figura 18 descreve as interações que ocorrem entre TAM e os resíduos de aminoácidos na cavidade do ER humano (3ERT).

Figura 17 - (A) Domínio LBD do ER humano com o ligante agonista estradiol presente no sítio de ligação. (B) Domínio LBD do ER humano com o ligante antagonista 4-hidroxitamoxifeno presente no sítio de ligação. É possível notar que os diferentes ligantes causam conformações diferentes no domínio LBD, principalmente no posicionamento da hélice 12, o que influencia na ligação com co-ativadores e correpressores. O ligante estradiol em A e o tamoxifeno em B, estão representados em amarelo.



Fonte: Própria autora.

Figura 18 - Interações químicas entre resíduos de aminoácidos que compõem o sítio de ligação do domínio LBD da estrutura experimental de um ER humano (3ERT) cristalizado com 4-hidroxitamoxifeno. Linhas tracejadas pretas indicam ligações de hidrogênio, ocorrendo com Arg394 e Glu353. Linhas sólidas verdes mostram interações hidrofóbicas, que ocorrem com os resíduos Met343, Leu346, Ala350, Trp383, Leu387, Leu391, Met421, Ile424 e Leu525.



Fonte: Representação gráfica gerada pelo programa PoseView.

A posição de interação do 4-hidroxitamoxifeno no teste de *docking* com os modelos de *C. gigas* que apresentou menor RMSD em relação a posição que este ligante ocupa no sítio de ligação do cristal de ER humano são mostradas na Tabela 11, assim como os valores de variação da energia livre para cada uma das posições. Em um alinhamento das estruturas (cristal de referência 3ERT e modelos de *C. gigas* NR1P2, NR2E1 e NR2E5) com os ligantes associados ao sítio de ligação, é possível observar sobreposição dos ligantes na cavidade do sítio, demonstrando que ocupam locais muito aproximados, como indicam os valores de RMSD da Tabela 11. Os valores negativos de  $\Delta G$  configuram essas posições de interação

como possíveis de acontecer espontaneamente em condições padrão, uma vez que ocorrem no sentido de minimização da energia livre do sistema, como ocorreu no teste com estradiol.

Tabela 11. Dados de análise quantitativa para posições do teste de *docking* com tamoxifeno.

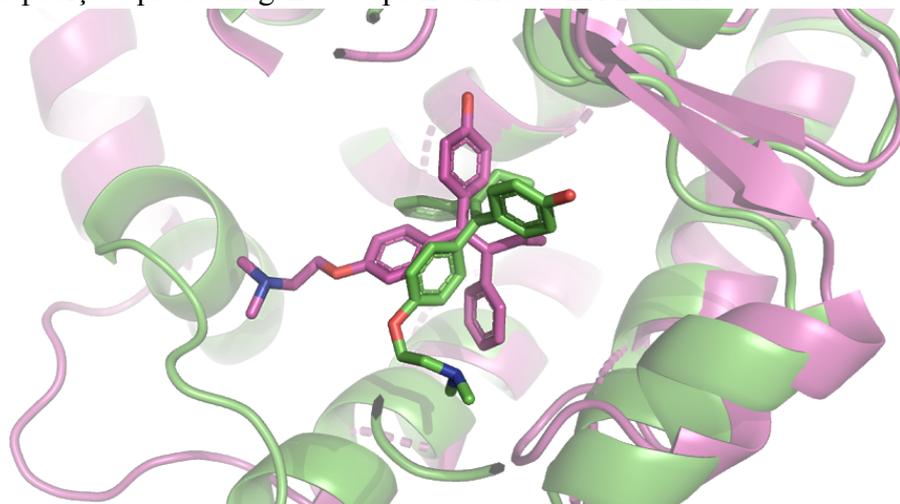
	<b>RMSD</b>	<b><math>\Delta G</math></b>
CgNR2E1	5.313	-9.5
CgNR2E5	5.784	-7.3
CgNR1P2	7.061	-8.5

Fonte: Própria autora.

#### 4.6.1 CgNR2E1

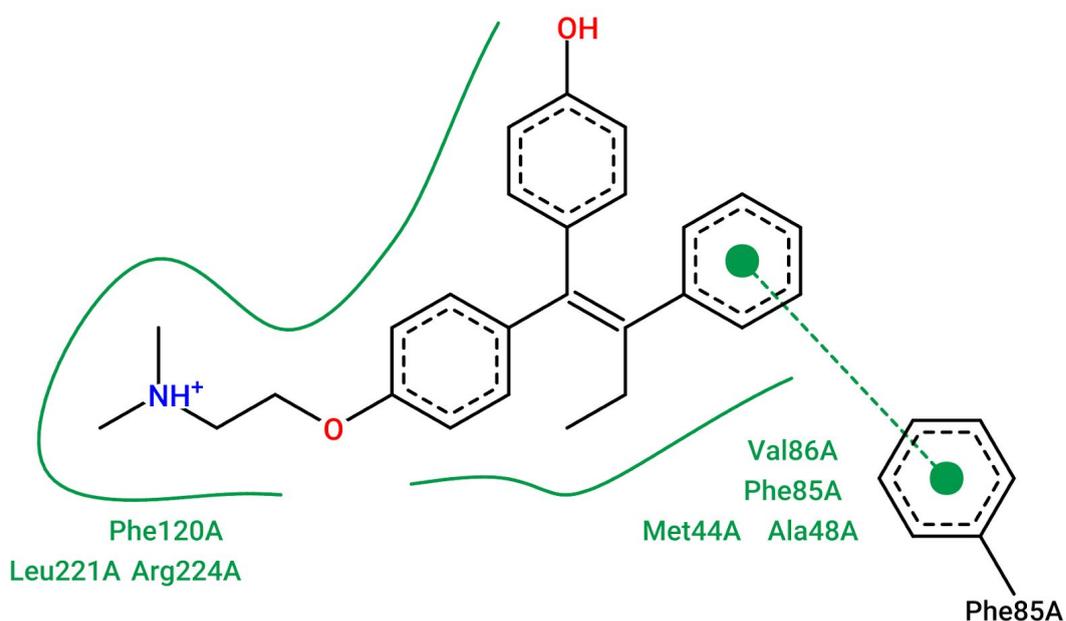
O teste de *docking* com o modelo CgNR2E1 e o ligante 4-hidroxitamoxifeno, assim como ocorreu nos testes com o estradiol, apresentou os menores valores de RMSD entre os três modelos testados, indicando que as posições em que o tamoxifeno se acomodou no sítio de ligação foram as mais próximas das que este ligante ocupa no sítio do ER humano de referência. Consideramos, portanto, este alvo como o candidato mais provável dentre os avaliados, para ser modulado por tamoxifeno. Os valores de RMSD da posição com menor desvio comparado à estrutura experimental 3ERT encontra-se na Tabela 11, que também apresenta o valor de  $\Delta G$  para esta posição. A Figura 19 mostra um alinhamento entre o ER humano (3ERT) e CgNR2E1, demonstrando como o tamoxifeno presente no sítio de ligação de ambos os receptores ocupa espaço similar. A Figura 20 mostra a formação de interações hidrofóbicas com Met44, Ala48, Phe85, Val86, Phe120, Leu221, Arg224, além de uma ligação  $\pi$ -cátion entre Phe85 e um anel aromático de TAM. Os resíduos representados pertencem às hélices 3, 6, 7, 8, 11 e a uma folha-beta, todas com estrutura bem definidas, conforme o esperado para o domínio LBD. As outras duas posições com menor RMSD identificadas no *docking* para CgNR2E1 possuem seus dados de análise quantitativa e qualitativa presentes no APÊNDICE C.

Figura 19. Alinhamento da estrutura cristalográfica do ER humano (3ERT), com o ligante 4-hidroxitamoxifeno (em roxo) interno à cavidade do sítio de ligação e do modelo CgNR2E1 com o 4-hidroxitamoxifeno (em verde) interno ao sítio de ligação na posição de menor RMSD em relação à posição que este ligante ocupa no sítio do ER humano.



Fonte: Própria autora.

Figura 20 - Interações químicas entre o tamoxifeno e os resíduos de aminoácidos que compõem a cavidade do sítio de ligação do modelo CgNR2E1. Esta posição mostra a formação de interações hidrofóbicas com Met44, Ala48, Phe85, Val86, Phe120, Leu221, Arg224, além de uma ligação  $\pi$ - $\pi$  entre Phe85 e um anel aromático de TAM.

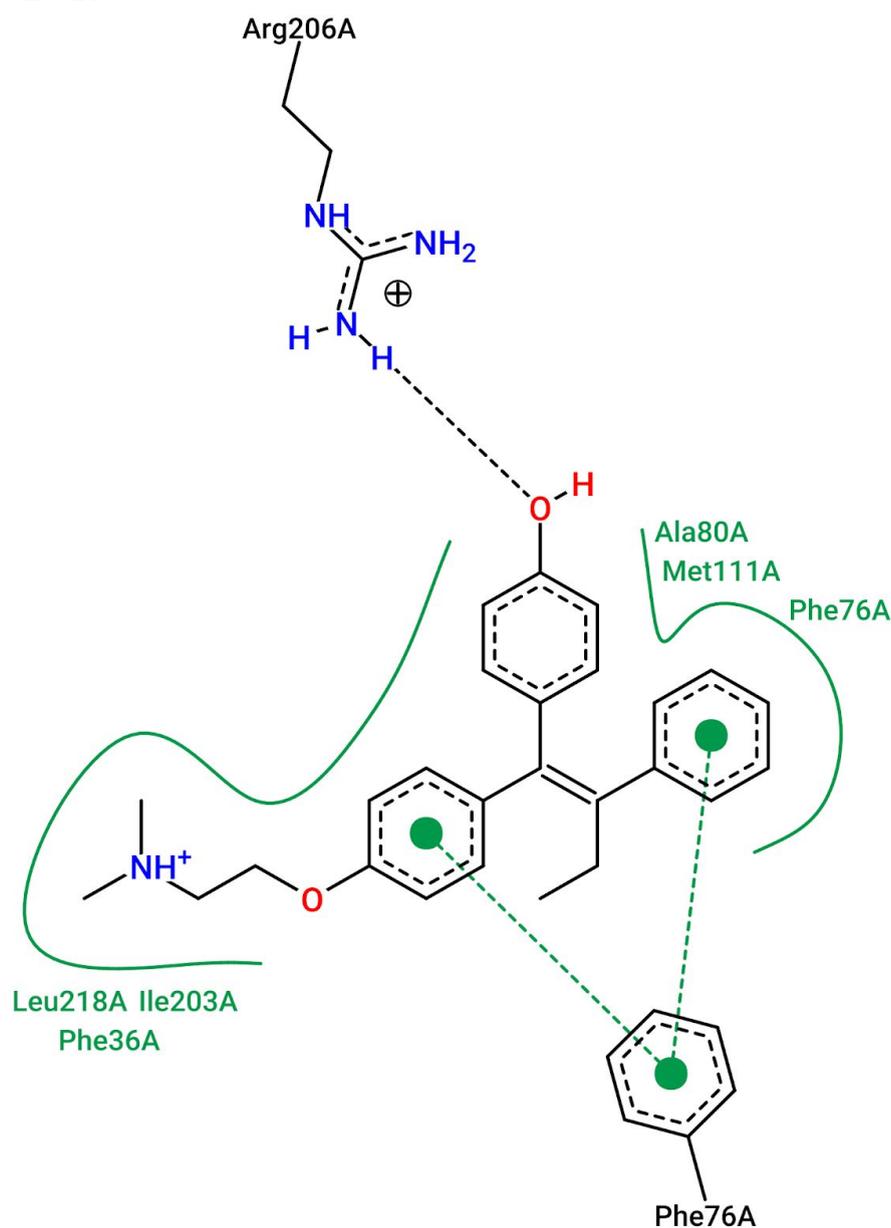


Fonte: Representação gráfica gerada pelo programa PoseView.

#### 4.6.2 CgNR2E5

A posição do teste de *docking* para o modelo CgNR2E5 que apresenta o menor RMSD em relação à estrutura experimental de referência está compoendo a Tabela 11, juntamente com o valor de  $\Delta G$  correspondente. Nas hélices 3, 6, 8, 11 e 12 do modelo estão os resíduos de aminoácidos que compõem a cavidade do sítio e interagem com o tamoxifeno. A Figura 21 apresenta o diagrama bidimensional para a posição do *docking* com menor valor de RMSD, mostrando a análise qualitativa de uma posição do 4-hidroxitamoxifeno com os resíduos de aminoácidos da cavidade do sítio de ligação do modelo. Nesta posição há formação de uma ponte de hidrogênio com Arg206, ligação  $\pi$ - $\pi$  entre Phe76 e dois anéis aromáticos de TAM, além de muitas interações hidrofóbicas envolvendo os resíduos Phe36, Phe76, Ala80, Met111, Ile203 e Leu218. As outras duas posições com menor RMSD identificadas no *docking* para CgNR2E5 possuem seus dados de análise quantitativa e qualitativa presentes no APÊNDICE C.

Figura 21 - Interações químicas entre o tamoxifeno e os resíduos de aminoácidos que compõem a cavidade do sítio de ligação do modelo CgNR2E5. Há formação de uma ponte de hidrogênio com Arg206, ligação  $\pi$ - $\pi$  entre Phe76 e dois anéis aromáticos de TAM, além de muitas interações hidrofóbicas envolvendo os resíduos Phe36, Phe76, Ala80, Met111, Ile203 e Leu218.



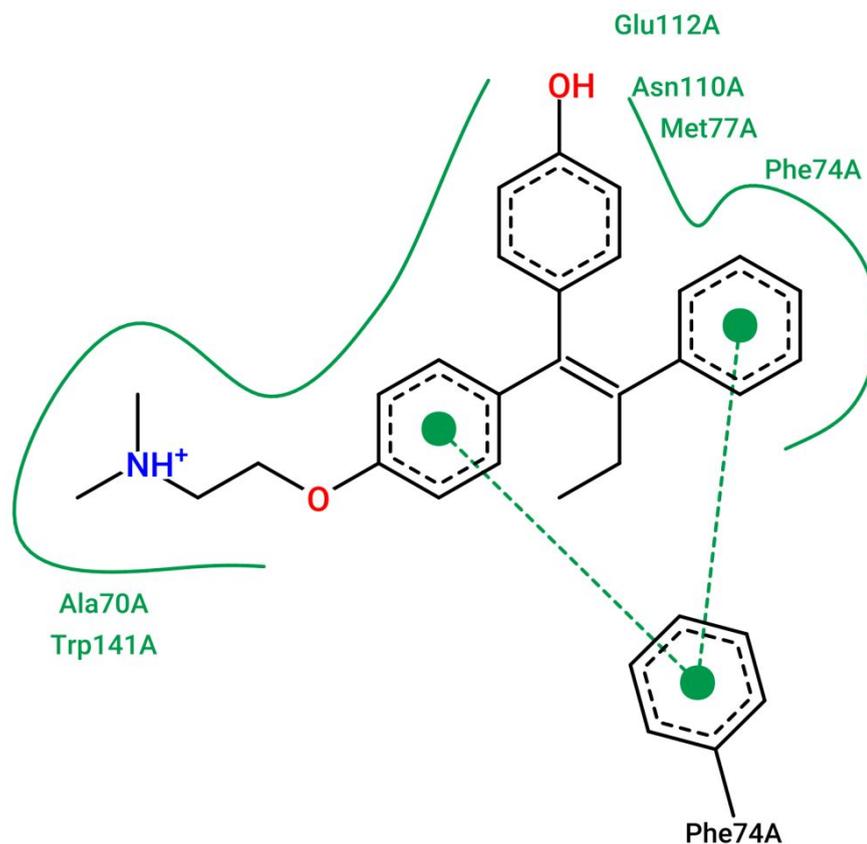
Fonte: Representação gráfica gerada pelo programa PoseView.

#### 4.6.3 CgNR1P2

As posições encontradas no *docking* com TAM para o modelo CgNR1P2 apresentaram resultados indicando uma interação possível entre receptor e ligante. A Tabela 11 apresenta os valores de RMSD da posição do ligante mais próxima à posição que este

ocupa no sítio de ligação do ER humano utilizado como referência (3ERT). O valor de  $\Delta G$  encontrado correspondente a esta posição é mostrado, apresentando valor negativo, o que caracteriza uma reação possível. A posição identificada com base no RMSD é mostrada em ilustração bidimensional na Figura 22. A posição que representa o menor RMSD encontrado, apresentou diversas interações hidrofóbicas com resíduos de Ala70, Phe74, Met77, Asn110, Glu112, Trp141, além de uma ligação  $\pi$ - $\pi$  entre Phe74 e 2 anéis aromáticos do 4-hidroxitamoxifeno. As hélices 2, 3, 6 e 8, contém os resíduos de aminoácidos que interagem com o 4-hidroxitamoxifeno no modelo. As outras duas posições com menor RMSD identificadas no *docking* para CgNR1P2 possuem seus dados de análise quantitativa e qualitativa presentes no APÊNDICE C.

Figura 22 - Interações químicas entre o tamoxifeno e os resíduos de aminoácidos que compõem a cavidade do sítio de ligação do modelo CgNR1P2. A posição apresentou diversas interações hidrofóbicas com resíduos de Ala70, Phe74, Met77, Asn110, Glu112, Trp141, além de uma ligação  $\pi$ - $\pi$  entre Phe74 e 2 anéis aromáticos do 4-hidroxitamoxifeno.



Fonte: Representação gráfica gerada pelo programa PoseView.

Através do teste de *docking* foi possível verificar valores de RMSD,  $\Delta G$  e interações químicas no sítio de ligação dos modelos que sugerem uma ligação possível do 4-hidroxitamoxifeno com CgNR1P2, CgNR2E1 e CgNR2E5. Nossos modelos apresentam conformações do domínio LBD com a posição de H12 diferente da orientação que ela assume na presença do fármaco tamoxifeno no cristal de ER humano. Dinâmicas moleculares de nossos modelos na presença de tamoxifeno, em posições que verificamos interações possíveis entre o fármaco e o NR, podem testar a possibilidade de este ligante provocar conformações diferentes no domínio LBD dos receptores da ostra e apresentar funcionalidades diferentes através dessa possível mudança de conformação.

## 5 CONCLUSÃO

Propomos em nosso trabalho diferentes metodologias para a obtenção de modelos tridimensionais para estudo de uma espécie de invertebrado não-modelo, observando que o pacote I-TASSER demonstrou ser mais adequado para a construção de modelos a partir de sequências de baixa similaridade ou que não conseguiram gerar modelos pela técnicas de modelagem comparativas utilizadas neste trabalho. A caracterização estrutural através da modelagem nos permitiu identificar cavidades do sítio de ligação que possuem potencial de interação com os ligantes alvo de interesse em nosso estudo, o hormônio estradiol e o fármaco tamoxifeno. Testes de ligação com moléculas xenobióticas demonstraram interações favoráveis entre os três receptores nucleares da ostra do Pacífico e os ligantes xenobióticos estradiol e tamoxifeno, com preferência para o receptor CgNR2E1, sendo, portanto, um forte candidato para estudos posteriores.

## REFERÊNCIAS

- AGUIRRE-MARTÍNEZ, Gabriela V.; DELVALLS, Angel T.; MARTÍN-DÍAZ, M. Laura. Yes, caffeine, ibuprofen, carbamazepine, novobiocin and tamoxifen have an effect on *Corbicula fluminea* (Müller, 1774). **Ecotoxicology And Environmental Safety**, v. 120, p.142-154, out. 2015. Elsevier BV.
- ALLEN, A. K.; SPRADLING, A. C.. The Sf1-related nuclear hormone receptor Hr39 regulates *Drosophila* female reproductive tract development and function. **Development**, v. 135, n. 2, p.311-321, dez. 2007. The Company of Biologists.
- ALONSO, Hernán; BLIZNYUK, Andrey A.; GREASY, Jill E.. Combining docking and molecular dynamic simulations in drug design. **Medicinal Research Reviews**, v. 26, n. 5, p.531-568, 2006.
- ANTIZAR-LADISLAO, Blanca. Environmental levels, toxicity and human exposure to tributyltin (TBT)-contaminated marine environment. A review. **Environment International**, v. 34, n. 2, p.292-308, fev. 2008. Elsevier BV.
- ASCENZI, Paolo; BOCEDI, Alessio; MARINO, Maria. Structure–function relationship of estrogen receptor  $\alpha$  and  $\beta$ : Impact on human health. **Molecular Aspects Of Medicine**, v. 27, n. 4, p.299-402, ago. 2006. Elsevier BV.
- BAGAMASBAD, Pia; DENVER, Robert J. Mechanisms and significance of nuclear receptor auto- and cross-regulation. **General And Comparative Endocrinology**, v. 170, n. 1, p.3-17, jan. 2011. Elsevier BV.
- BAKER, Amelia H. et al. Tributyltin Engages Multiple Nuclear Receptor Pathways and Suppresses Osteogenesis in Bone Marrow Multipotent Stromal Cells. **Chemical Research In Toxicology**, v. 28, n. 6, p.1156-1166, maio, 2015. American Chemical Society (ACS).
- BAKER, Michael E. Steroid receptors and vertebrate evolution. **Molecular And Cellular Endocrinology**, v. 496, out. 2019. Elsevier BV.
- BAKER, Michael e. Steroid receptor phylogeny and vertebrate origins. **Molecular And Cellular Endocrinology**, v. 135, n. 2, p.101-107, dez. 1997. Elsevier BV.
- BELL, Eric W.; ZHANG, Yang. DockRMSD: an open-source tool for atom mapping and RMSD calculation of symmetric molecules through graph isomorphism. **Journal Of Cheminformatics**, v. 11, n. 1, jun. 2019. Springer Science and Business Media LLC.
- BENKERT, Pascal; BIASINI, Marco; SCHWEDE, Torsten. Toward the estimation of the absolute quality of individual protein structure models. **Bioinformatics**, v. 27, n. 3, p.343-350, dez. 2010. Oxford University Press (OUP).

BENOD, Cindy et al. The Human Orphan Nuclear Receptor Tailless (TLX, NR2E1) Is Druggable. **Plos One**, v. 9, n. 6, jun. 2014. Public Library of Science (PLoS).

BETANCOURT, Marcos R.; SKOLNICK, Jeffrey. Universal similarity measure for comparing protein structures. **Biopolymers**, v. 59, n. 5, p.305-309, 2001. Wiley.

BOUTON, D. A conserved retinoid X receptor (RXR) from the mollusk *Biomphalaria glabrata* transactivates transcription in the presence of retinoids. **Journal Of Molecular Endocrinology**, v. 34, n. 2, p.567-582, abr. 2005. Bioscientifica.

BRIDGHAM, Jamie T. et al. Protein Evolution by Molecular Tinkering: Diversification of the Nuclear Receptor Superfamily from a Ligand-Dependent Ancestor. **Plos Biology**, v. 8, n. 10, out. 2010. Public Library of Science (PLoS).

BRIDGHAM, Jamie T. et al. Vestigialization of an Allosteric Switch: Genetic and Structural Mechanisms for the Evolution of Constitutive Activity in a Steroid Hormone Receptor. **Plos Genetics**, v. 10, n. 1, jan. 2014. Public Library of Science (PLoS).

CAJARAVILLE, Miren P et al. The use of biomarkers to assess the impact of pollution in coastal environments of the Iberian Peninsula: a practical approach. **Science Of The Total Environment**, v. 247, n. 2-3, p.295-311, mar. 2000. Elsevier BV.

CAMACHO, Christiam et al. BLAST+: architecture and applications. **Bmc Bioinformatics**, v. 10, n. 1, p.10-421, dez. 2009. Springer Nature.

CELIK, Leyla; LUND, Julie Davey Dalsgaard; SCHIØTT, Birgit. Conformational Dynamics of the Estrogen Receptor  $\alpha$ : Molecular Dynamics Simulations of the Influence of Binding Site Structure on Protein Dynamics†. **Biochemistry**, v. 46, n. 7, p.1743-1758, fev. 2007. American Chemical Society (ACS).

DOSS, C. George Priya et al. Computational Approaches and Resources in Single Amino Acid Substitutions Analysis Toward Clinical Research. In: **Advances in Protein Chemistry and Structural Biology**. Ai: Elsevier, 2014. Cap. 10. p. 365-423.

DUBROVSKAYA, Veronica A.; BERGER, Edward M.; DUBROVSKY, Edward B. Juvenile hormone regulation of the E75 nuclear receptor is conserved in Diptera and Lepidoptera. **Gene**, v. 340, n. 2, p.171-177, out. 2004. Elsevier BV.

FAYARD, Elisabeth; AUWERX, Johan; SCHOONJANS, Kristina. LRH-1: an orphan nuclear receptor involved in development, metabolism and steroidogenesis. **Trends In Cell Biology**, v. 14, n. 5, p.250-260, maio 2004. Elsevier BV.

HELDRING, Nina et al. Structural Insights into Corepressor Recognition by Antagonist-bound Estrogen Receptors. **Journal Of Biological Chemistry**, v. 282, n. 14, p.10449-10455, fev. 2007. American Society for Biochemistry & Molecular Biology (ASBMB).

HUANG, Wen et al. Evolution of a novel nuclear receptor subfamily with emphasis on the member from the Pacific oyster *Crassostrea gigas*. **Gene**, v. 567, n. 2, p.164-172, ago. 2015. Elsevier BV.

JANER, Gemma; PORTE, Cinta. Sex steroids and potential mechanisms of non-genomic endocrine disruption in invertebrates. **Ecotoxicology**, v. 16, n. 1, p.145-160, 13 jan. 2007. Springer Science and Business Media LLC.

KALAANY, Nada Y.; MANGELSDORF, David J. LXRS AND FXR: The Yin and Yang of Cholesterol and Fat Metabolism. **Annual Review Of Physiology**, v. 68, n. 1, p.159-191, jan. 2006. Annual Reviews.

KARIMULLINA, Elina et al. *Daphnia* HR96 is a promiscuous xenobiotic and endobiotic nuclear receptor. **Aquatic Toxicology**, v. 116-117, p.69-78, jul. 2012. Elsevier BV.

KING-JONES, Kirst; THUMMEL, Carl S. Nuclear receptors — a perspective from *Drosophila*. **Nature Reviews Genetics**, v. 6, n. 4, p.311-323, abr. 2005. Springer Science and Business Media LLC.

LARANJEIRO, Filipe et al. Fifteen years of imposex and tributyltin pollution monitoring along the Portuguese coast. **Environmental Pollution**, v. 232, p.411-421, jan. 2018. Elsevier BV.

LASKOWSKI, R. A. et al. PROCHECK: a program to check the stereochemical quality of protein structures. **Journal Of Applied Crystallography**, v. 26, n. 2, p.283-291, abr. 1993. International Union of Crystallography (IUCr).

LEVITT, M. Growth of novel protein structural data. **Proceedings Of The National Academy Of Sciences**, v. 104, n. 9, p.3183-3188, fev. 2007. Proceedings of the National Academy of Sciences.

LI, Yong; LAMBERT, Millard H; XU, H.eric. Activation of Nuclear Receptors. **Structure**, v. 11, n. 7, p.741-746, jul. 2003. Elsevier BV.

LONARD, David M.; SMITH, Carolyn L.. Molecular perspectives on selective estrogen receptor modulators (SERMs): progress in understanding their tissue-specific agonist and antagonist actions. **Steroids**, v. 67, n. 1, p.15-24, jan. 2002. Elsevier BV.

LOVE-GREGORY, Latisha; PERMUTT, M Alan. HNF4A genetic variants: role in diabetes. **Current Opinion In Clinical Nutrition And Metabolic Care**, v. 10, n. 4, p.397-402, jul. 2007. Ovid Technologies (Wolters Kluwer Health).

MATSUMOTO, Toshie et al. Oyster estrogen receptor: cDNA cloning and immunolocalization. **General And Comparative Endocrinology**, v. 151, n. 2, p.195-201, abr. 2007. Elsevier BV.

MENG, Jie et al. Genome and Transcriptome Analyses Provide Insight into the Euryhaline Adaptation Mechanism of *Crassostrea gigas*. **Plos One**, v. 8, n. 3, 12 mar. 2013. Public Library of Science (PLoS).

MORRIS, Anne Louise et al. Stereochemical quality of protein structure coordinates. **Proteins: Structure, Function, and Genetics**, v. 12, n. 4, p.345-364, abr. 1992.

MOSHKOVSKII, S. A. et al. Identification of Single Amino Acid Substitutions in Proteogenomics. **Biochemistry**, v. 83, n. 3, p.250-258, mar. 2018. Pleiades Publishing Ltd.

NELSON, David L.; COX, Michael M. **Princípios de Bioquímica de Lehninger**. 4. ed. São Paulo. Si: Sarvier, 2006. 1202 p.

OMIECINSKI, C. J. et al. Xenobiotic Metabolism, Disposition, and Regulation by Receptors: From Biochemical Phenomenon to Predictors of Major Toxicities. **Toxicological Sciences**, v. 120, n. 1, p.49-75, nov. 2010. Oxford University Press (OUP).

ORIAS, Frédéric et al. Tamoxifen ecotoxicity and resulting risks for aquatic ecosystems. **Chemosphere**, v. 128, p.79-84, jun. 2015. Elsevier BV.

ORIAS, Frédéric; SIMON, Laurent; PERRODIN, Yves. Experimental assessment of the bioconcentration of 15 N-tamoxifen in *Pseudokirchneriella subcapitata*. **Chemosphere**, v. 122, p.251-256, mar. 2015. Elsevier BV.

PALANKER, Laura et al. *Drosophila* HNF4 Regulates Lipid Mobilization and  $\beta$ -Oxidation. **Cell Metabolism**, v. 9, n. 3, p.228-239, mar. 2009. Elsevier BV.

PETTERSEN, Eric F. et al. UCSF Chimera - A visualization system for exploratory research and analysis. **Journal Of Computational Chemistry**, v. 25, n. 13, p.1605-1612, out. 2004.

RAMAKRISHNAN, Sathiya N.; MUSCAT, George E.o. The Orphan Rev-Erb Nuclear Receptors: A Link between Metabolism, Circadian Rhythm and Inflammation? **Nuclear Receptor Signaling**, v. 4, n. 1, jan. 2006. SAGE Publications.

REITZEL, Adam M; TARRANT, Ann M. Nuclear receptor complement of the cnidarian *Nematostella vectensis*: phylogenetic relationships and developmental expression patterns. **Bmc Evolutionary Biology**, v. 9, n. 1, p.230-9, 2009. Springer Science and Business Media LLC.

ROSSATO, Martina. **Gastrópodes como bioindicadores da poluição por organoestênicos: possíveis fatores que influenciam na acumulação dos compostos e indução do imposex**. 2017. 129 p. Tese (Doutorado) - Curso de Oceanografia Química, Física e Geológica, Furg, Rio Grande, 2017.

ROY, Ambrish; KUCUKURAL, Alper; ZHANG, Yang. I-TASSER: a unified platform for automated protein structure and function prediction. **Nature Protocols**, v. 5, n. 4, p.725-738, mar. 2010. Springer Science and Business Media LLC.

RUDOLPH, K. et al. Complex regulatory region mediating tailless expression in early embryonic patterning and brain development. **Development**, v. 124. p. 4297-308. 1997.

SANTOS, Karin dos. **Caracterização *in silico* da superfamília de Receptores Nucleares em *Crassostrea gigas* (Thunberg, 1793)**. 2018. 93 f. TCC (Graduação) - Curso de Ciências Biológicas, Departamento de Bioquímica, Universidade Federal de Santa Catarina, Florianópolis, 2018.

SCHRÖDINGER, L. L. C. The PyMOL molecular graphics system, Version 1.8. 2015. Google Scholar, 2017.

SCOTT, Alexander P. Do mollusks use vertebrate sex steroids as reproductive hormones? II. Critical review of the evidence that steroids have biological effects. **Steroids**, v. 78, n. 2, p.268-281, fev. 2013. Elsevier BV.

SHIAU, Andrew K. et al. The Structural Basis of Estrogen Receptor/Coactivator Recognition and the Antagonism of This Interaction by Tamoxifen. **Cell**, v. 95, n. 7, p.927-937, dez. 1998. Elsevier BV.

SLUDER, Ann.; MAINA, Claude V. Nuclear receptors in nematodes: themes and variations. **Trends In Genetics**, v. 17, n. 4, p.206-213, 1 abr. 2001. Elsevier BV.

STIERAND, Katrin; RAREY, Matthias. Drawing the PDB: Protein-Ligand Complexes in Two Dimensions. **Acs Medicinal Chemistry Letters**, v. 1, n. 9, p.540-545, ago. 2010. American Chemical Society (ACS).

TIAN, Wei et al. CASTp 3.0: computed atlas of surface topography of proteins. **Nucleic Acids Research**, v. 46, n. 1, p.363-367, jun. 2018. Oxford University Press (OUP).

TIMMINS-SCHIFFMAN, Emma et al. Shotgun proteomics reveals physiological response to ocean acidification in *Crassostrea gigas*. **Bmc Genomics**, v. 15, n. 1, p.951, 2014. Springer Nature.

TROTT, Oleg; OLSON, Arthur J. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. **Journal Of Computational Chemistry**, p.455-461, jun. 2009.

VERLI, Hugo. **Bioinformática: da Biologia à Flexibilidade Molecular**. São Paulo: Sociedade Brasileira de Bioquímica e Biologia Molecular - Sbbq, 2014. 282 p.

VOGELER, Susanne et al. The nuclear receptor gene family in the Pacific oyster, *Crassostrea gigas*, contains a novel subfamily group. **Bmc Genomics**, v. 15, n. 1, p.15-369, maio 2014. Springer Nature.

WALLACE, Andrew C.; LASKOWSKI, Roman A.; THORNTON, Janet M.. LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. **Protein Engineering**, v. 8, n. 2, p.127-134, 1995. Oxford University Press (OUP).

WANG, Tao; XIONG, Jian-qiong. The Orphan Nuclear Receptor TLX/NR2E1 in Neural Stem Cells and Diseases. **Neuroscience Bulletin**, v. 32, n. 1, p.108-114, jan. 2016. Springer Science and Business Media LLC.

WÄRNMARK, Anette et al. Interaction of Transcriptional Intermediary Factor 2 Nuclear Receptor Box Peptides with the Coactivator Binding Site of Estrogen Receptor  $\alpha$ . **Journal Of Biological Chemistry**, v. 277, n. 24, abr. 2002. American Society for Biochemistry & Molecular Biology (ASBMB).

WATERHOUSE, Andrew et al. SWISS-MODEL: homology modelling of protein structures and complexes. **Nucleic Acids Research**, v. 46, n. 1, p.296-303, maio 2018. Oxford University Press (OUP).

ZHANG, Guofan et al. Molecular Basis for Adaptation of Oysters to Stressful Marine Intertidal Environments. **Annual Review Of Animal Biosciences**, v. 4, n. 1, p.357-381, fev. 2016. Annual Reviews.

ZHANG, Jiefeng et al. Removal of cytostatic drugs from aquatic environment: A review. **Science Of The Total Environment**, v. 445-446, p.281-298, fev. 2013. Elsevier BV.

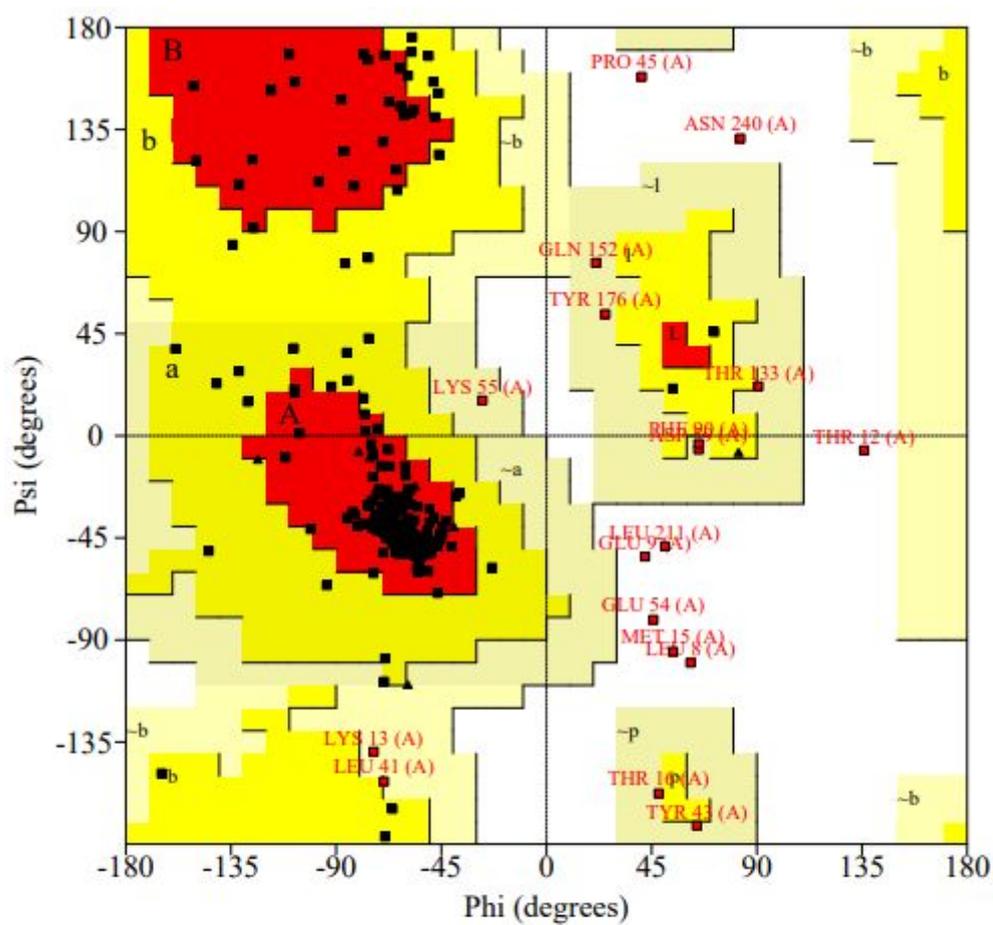
ZHANG, Yang. I-TASSER server for protein 3D structure prediction. **Bmc Bioinformatics**, v. 9, n. 1, p.9-40, 23 jan. 2008. Springer Nature.

ZHANG, Yang; SKOLNICK, Jeffrey. Scoring function for automated assessment of protein structure template quality. **Proteins: Structure, Function, and Bioinformatics**, v. 57, n. 4, p.702-710, dez. 2004.

## APÊNDICE A

### Gráficos de Ramachandran para os modelos do I-TASSER.

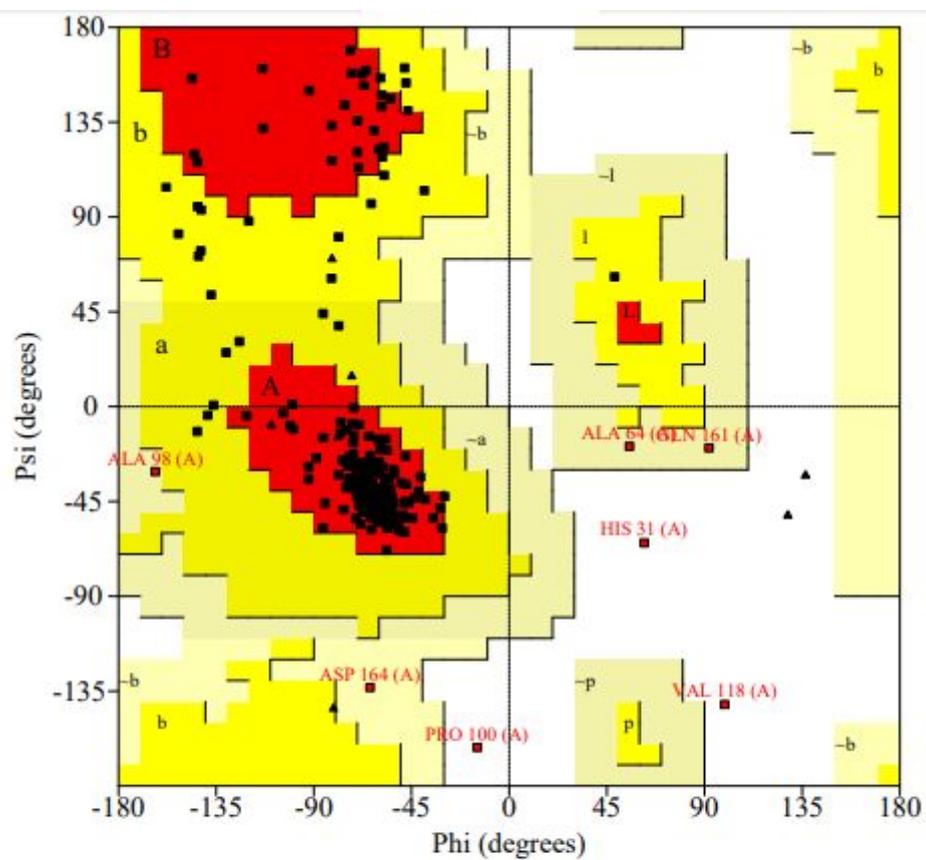
CgNR1P2



#### Plot statistics

Residues in most favoured regions [A,B,L]	176	77.9%
Residues in additional allowed regions [a,b,l,p]	33	14.6%
Residues in generously allowed regions [-a,-b,-l,-p]	10	4.4%
Residues in disallowed regions	7	3.1%
-----		
Number of non-glycine and non-proline residues	226	100.0%
Number of end-residues (excl. Gly and Pro)	2	
Number of glycine residues (shown as triangles)	11	
Number of proline residues	9	
-----		
Total number of residues	248	

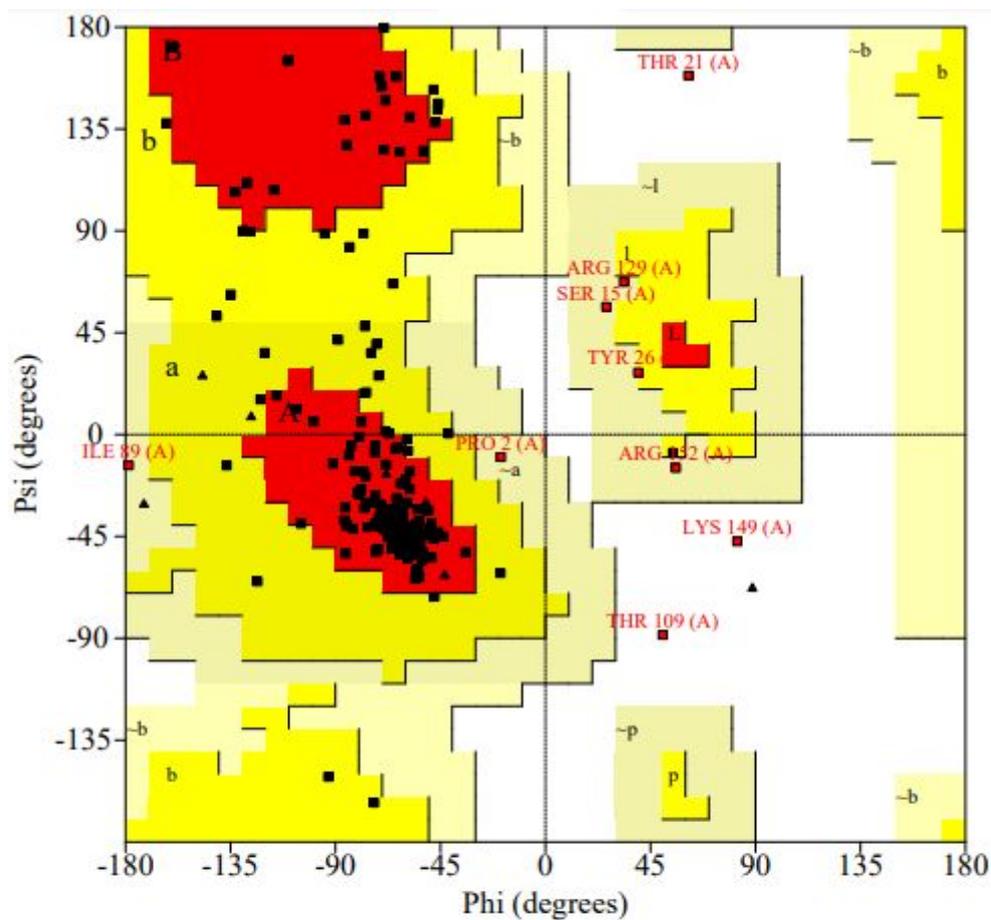
CgNR2E1



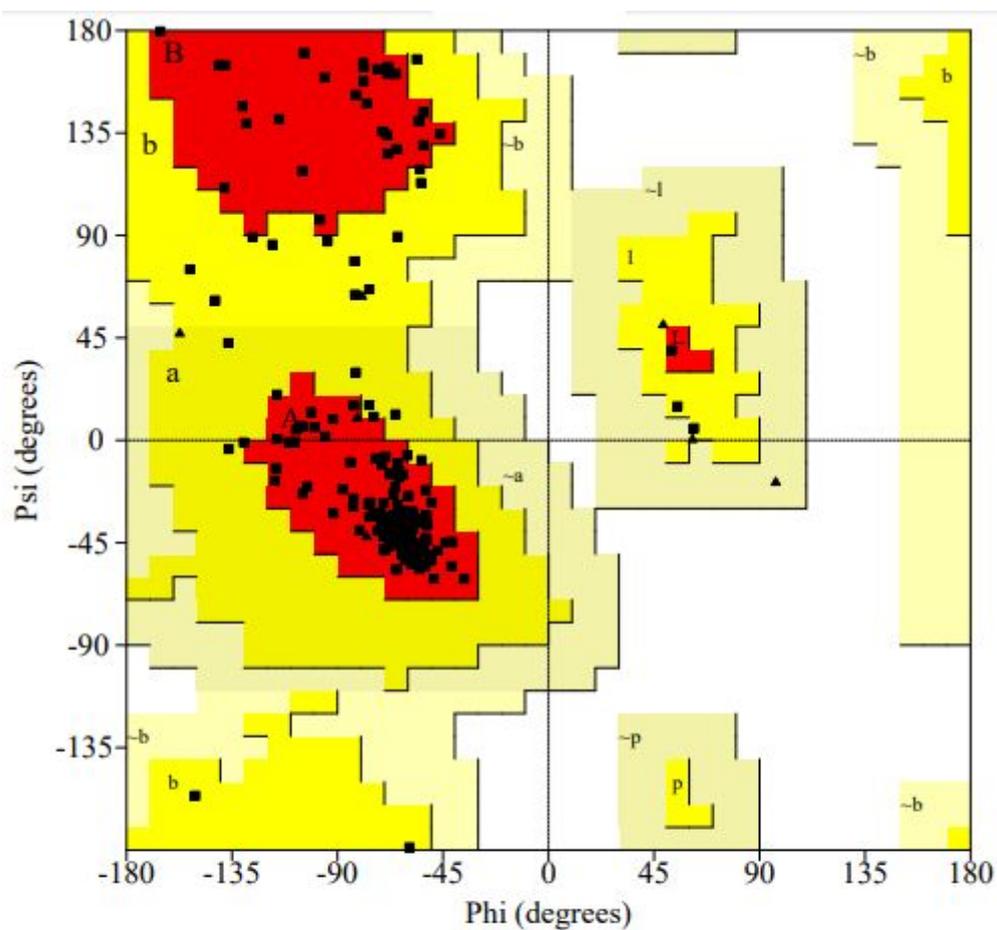
## Plot statistics

Residues in most favoured regions [A,B,L]	182	86.3%
Residues in additional allowed regions [a,b,l,p]	23	10.9%
Residues in generously allowed regions [-a,-b,-l,-p]	4	1.9%
Residues in disallowed regions	2	0.9%
----		
Number of non-glycine and non-proline residues	211	100.0%
Number of end-residues (excl. Gly and Pro)	1	
Number of glycine residues (shown as triangles)	14	
Number of proline residues	19	
----		
Total number of residues	245	

CgNR2E5



## CgNR3A



## Plot statistics

Residues in most favoured regions [A,B,L]	204	90.3%
Residues in additional allowed regions [a,b,l,p]	22	9.7%
Residues in generously allowed regions [-a,-b,-l,-p]	0	0.0%
Residues in disallowed regions	0	0.0%
----		
Number of non-glycine and non-proline residues	226	100.0%
Number of end-residues (excl. Gly and Pro)	2	
Number of glycine residues (shown as triangles)	8	
Number of proline residues	7	
----		
Total number of residues	243	

## APÊNDICE B

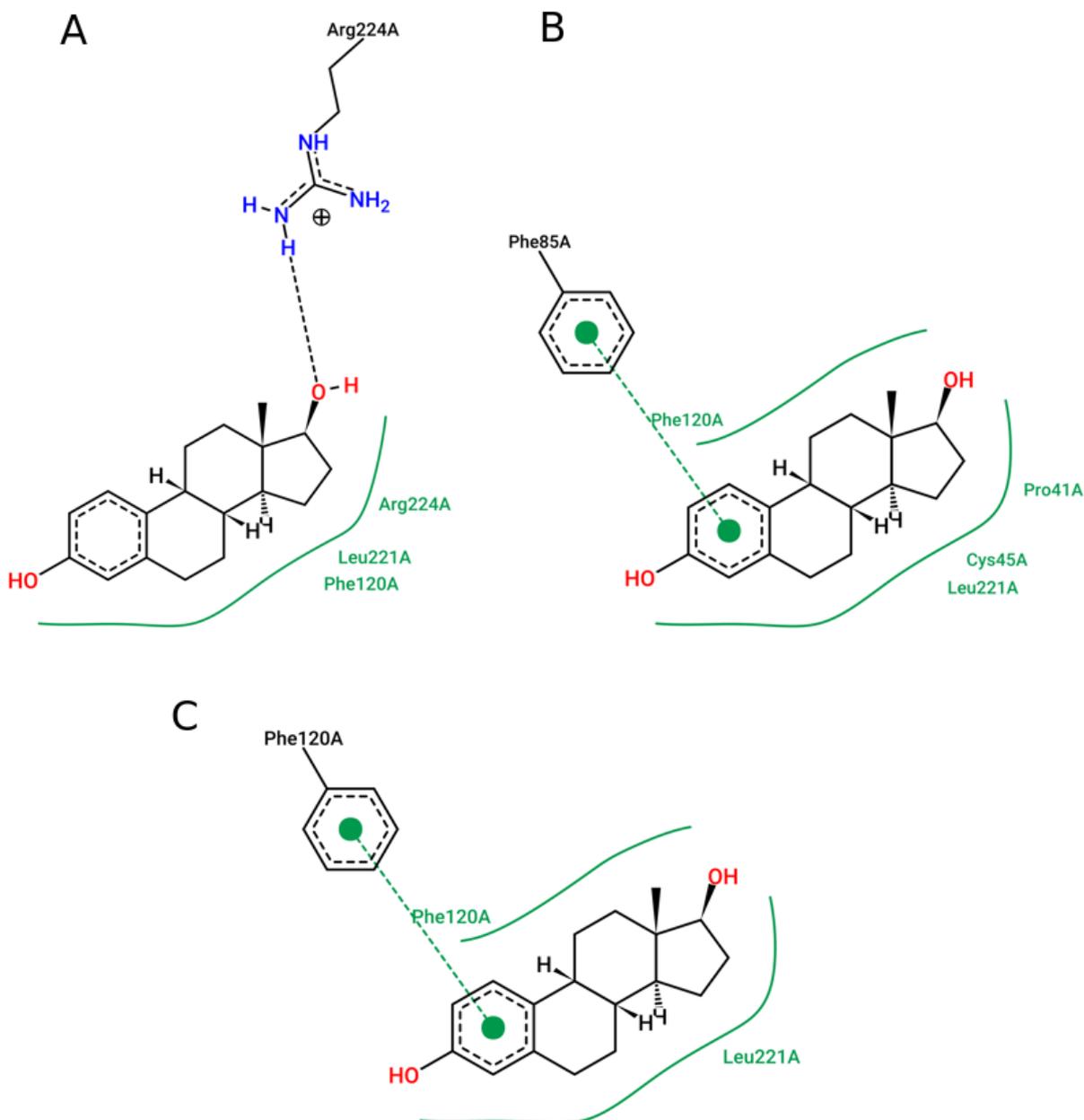
### Potencial de interação da cavidade com Estradiol

Tabela 12. Análise quantitativa das três posições do estradiol no teste de *docking* para cada modelo com menor RMSD em relação ao espaço que este ligante ocupa na estrutura de referência.

	RMSD	$\Delta G$
<b>CgNR2E1</b>	4.875	-8.9
	5.232	-8.5
	6.565	-10.4
<b>CgNR2E5</b>	5.436	-9.5
	6.301	-9.9
	6.314	-8.2
<b>CgNR1P2</b>	7.981	-9.2
	7.893	-10.1
	8.376	-10.2

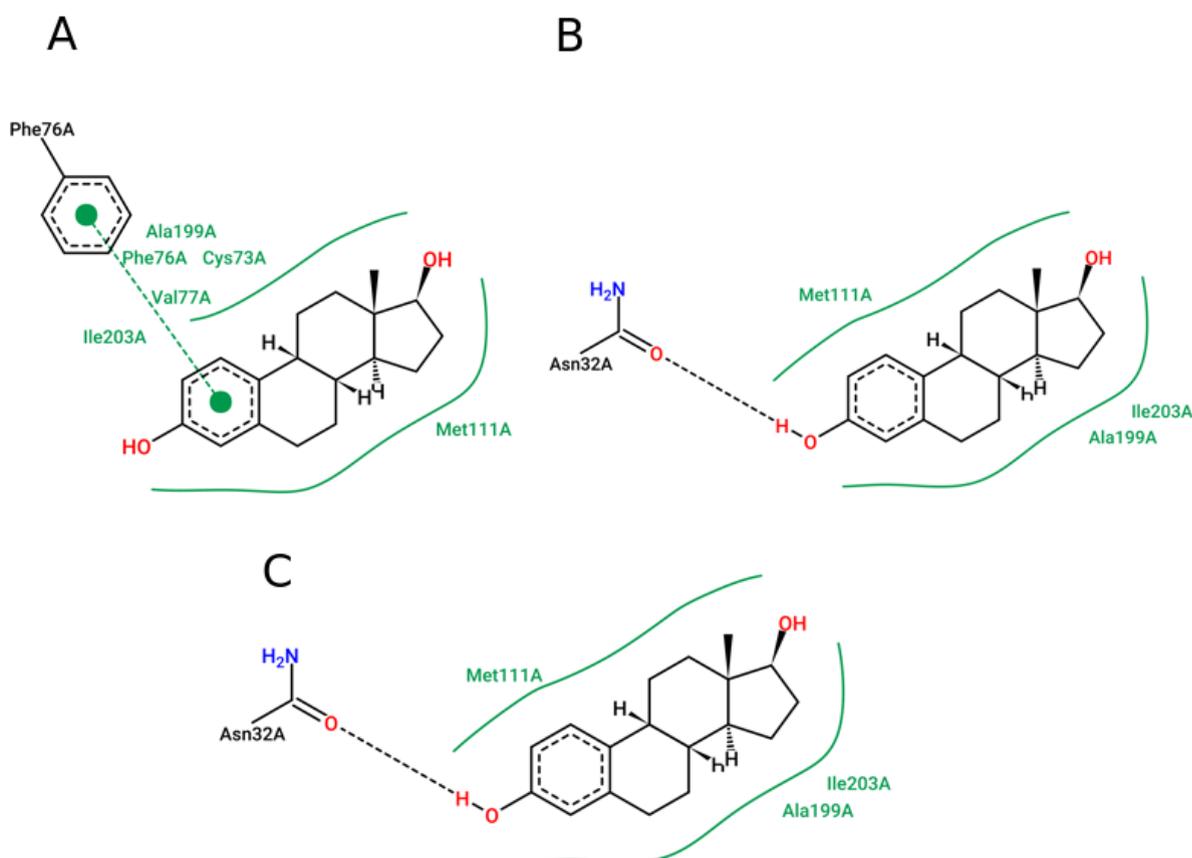
Fonte: Própria autora.

Figura 23 - Análise qualitativa das três posições de menor RMSD mostrano as interações químicas entre o estradiol e os resíduos de aminoácidos que compõem a cavidade do sítio de ligação do modelo **CgNR2E1**. A posição de menor RMSD (A) mostra a formação de uma ponte de hidrogênio entre o resíduo Arg224 e o estradiol, além de interações hidrofóbicas com Phe120, Leu221 e Arg224 na Posição 1. A segunda posição de menor RMSD (B) apresenta interações hidrofóbicas com os resíduos Phe120, Pro41, Cys45, Leu221, além de uma ligação  $\pi$ - $\pi$  entre Phe85 e o anel aromático do estradiol. A terceira posição de menor RMSD (C) também identificou interações hidrofóbicas com Phe120 e Leu221 com formação de ligação  $\pi$ - $\pi$  com Phe120.



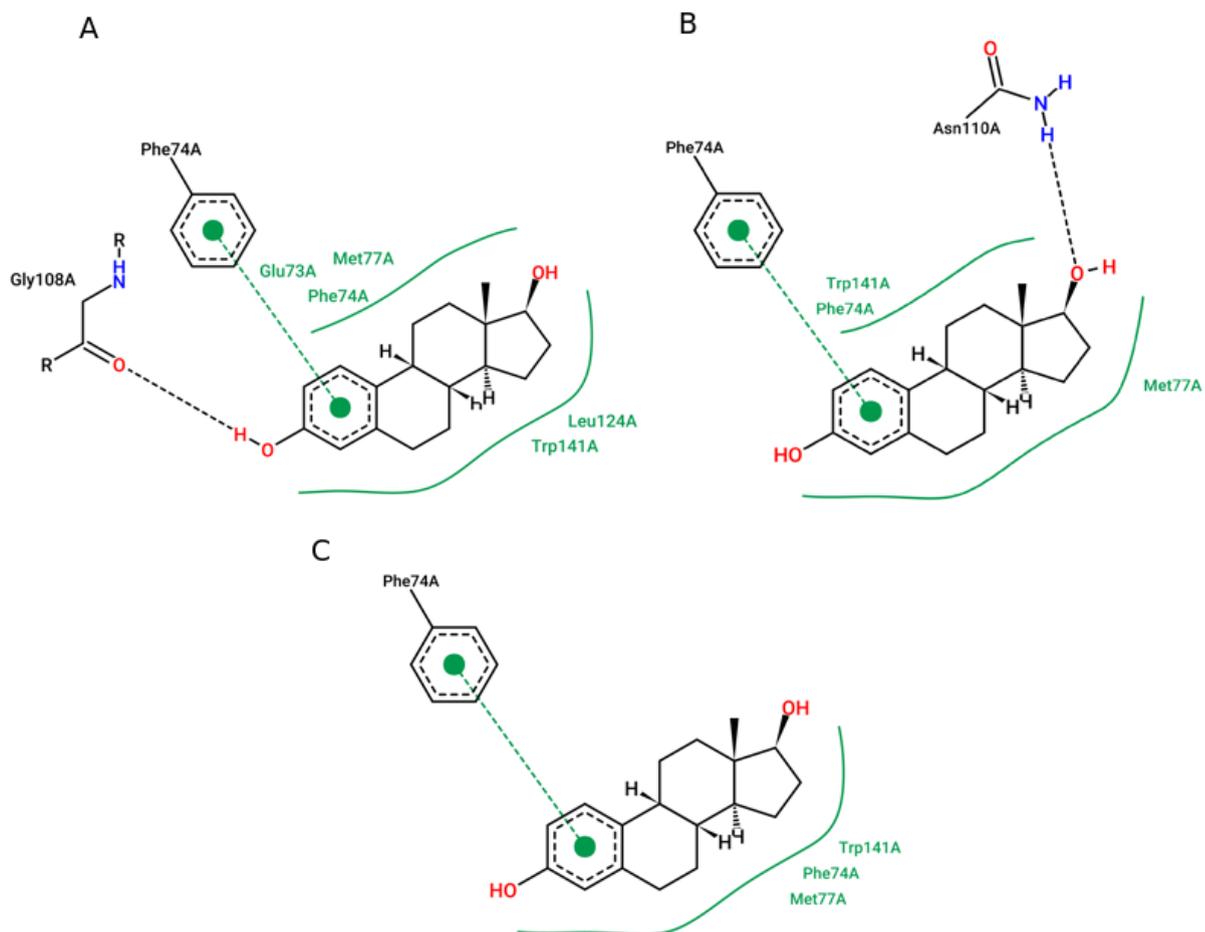
Fone: Representação gráfica gerada pelo programa Poseview.

Figura 24 - Análise qualitativa das três posições de menor RMSD mostrando as interações químicas entre o estradiol e os resíduos de aminoácidos que compõem a cavidade do sítio de ligação do modelo **CgNR2E5**. Na posição de menor RMSD (A) muitas interações hidrofóbicas ocorrem entre o estradiol e os resíduos Cys73, Phe76, Met111, Val177, Ala199 e Ile203, além de uma ligação  $\pi$ - $\pi$  com Phe76. A segunda posição de menor RMSD (B) e a terceira (C) configuraram as mesmas interações químicas fracas no sítio de ligação, compostas por uma ponte de hidrogênio com Asn32 e interações hidrofóbicas com Met111, Ala199 e Ile203.



Fonte: Representação gráfica gerada pelo programa PoseView.

Figura 25 - Análise qualitativa das três posições de menor RMSD mostrando as interações químicas entre o estradiol e os resíduos de aminoácidos que compõem a cavidade do sítio de ligação do modelo **CgNR1P2**. A posição de menor RMSD (A) apresenta ponte de hidrogênio com o resíduo Gly108, diversas interações hidrofóbicas com resíduos de Glu73, Phe74, Met77, Leu124, Trp141, além de uma ligação  $\pi$ - $\pi$  entre Phe74 e o anel aromático do estradiol. Na segunda posição de menor RMSD (B) também ocorreu formação de ponte de hidrogênio entre o estradiol, nesta posição com o resíduo Asn110. Uma ligação  $\pi$ - $\pi$  com Phe74 e outras interações hidrofóbicas foram encontradas com Met77 e Trp141. Interações polares não foram identificadas para a terceira posição de menor RMSD (C), a qual possui ligação  $\pi$ - $\pi$  com Phe74, além de interações hidrofóbicas com Met77 e Trp141.



Fonte: Representação gráfica gerada pelo programa PoseView.

## APÊNDICE C

### Potencial de interação da cavidade com tamoxifeno

Tabela 13. Análise quantitativa das três posições do tamoxifeno no teste de docking para cada modelo com menor RMSD em relação ao espaço que este ligante ocupa na estrutura de referência.

	RMSD	$\Delta G$
<b>CgNR2E1</b>	5.313	-9.5
	6.803	-9.2
	7.034	-9.5
<b>CgNR2E5</b>	5.784	-7.3
	6.684	-9.1
	6.769	-7.6
<b>CgN1P2</b>	7.061	-8.5
	7.608	-7.7
	8.108	-8.1

Fonte: Própria autora.

Figura 26 - Análise qualitativa das três posições de menor RMSD mostrando as interações químicas entre o tamoxifeno e os resíduos de aminoácidos que compõem a cavidade do sítio de ligação do modelo CgNR2E1. A posição de menor RMSD (A) mostra a formação de interações hidrofóbicas com Met44, Ala48, Phe85, Val86, Phe120, Leu221, Arg224, além de uma ligação  $\pi$ - $\pi$  entre Phe85 e um anel aromático de TAM. A segunda posição de menor RMSD (B) apresenta uma ligação de hidrogênio com o resíduo Pro113. Interações hidrofóbicas ocorrem com os resíduos Met44, Leu101, Ala103, Val117, Phe120, Ile124, Leu127, Leu221 e Arg224 e o resíduo Phe120 faz ligação  $\pi$ - $\pi$  com um anel aromático do ligante. A terceira posição de menor RMSD (C) também identificou interações hidrofóbicas com Met44, Val47, Ala48, Phe85, Ala89, Phe120, Ile124 e Leu221, além de duas ligações  $\pi$ - $\pi$  entre Phe120 e Phe85 com anéis aromáticos do tamoxifeno e uma ligação  $\pi$ -cátion entre Phe85 e um nitrogênio de TAM.

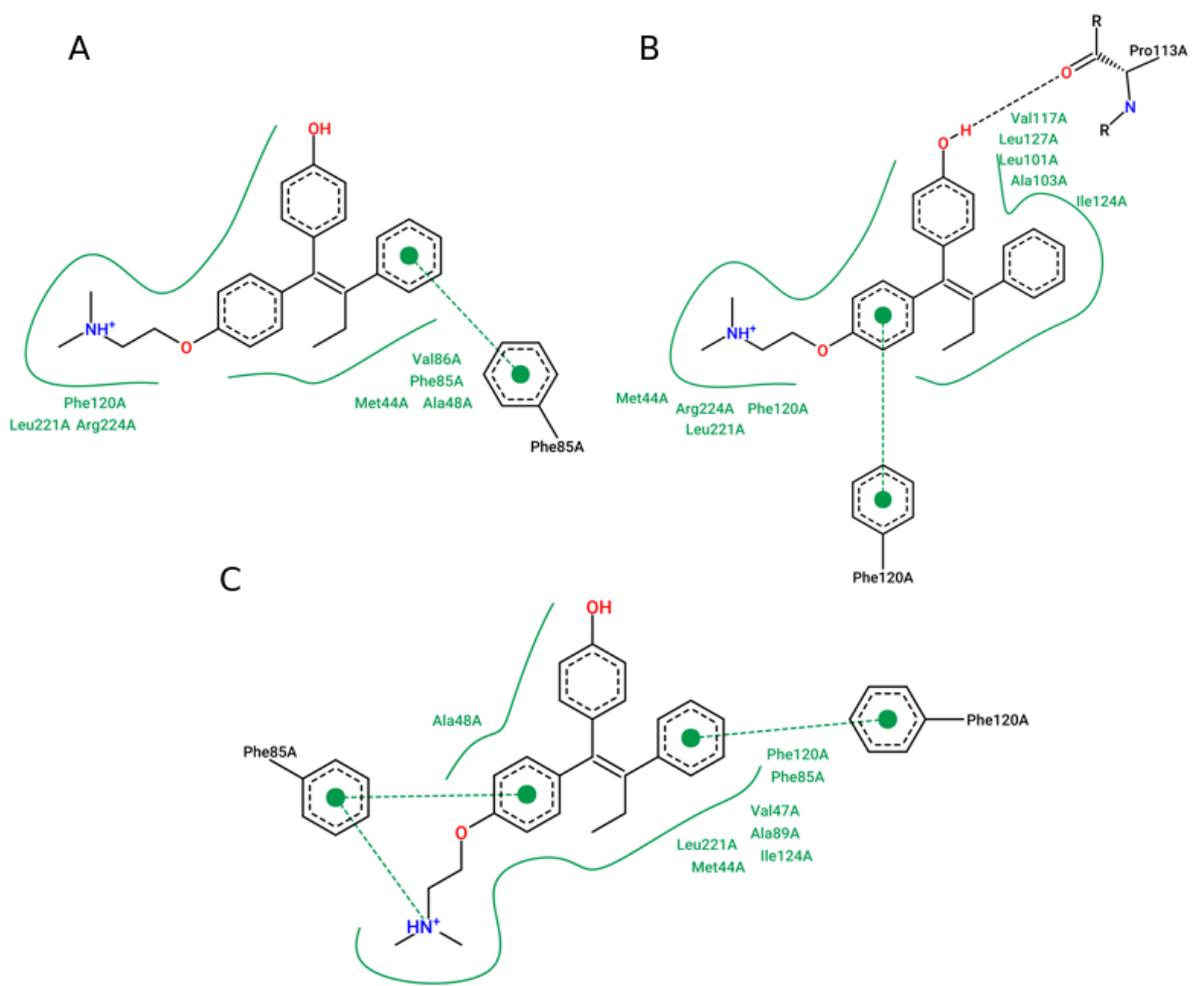
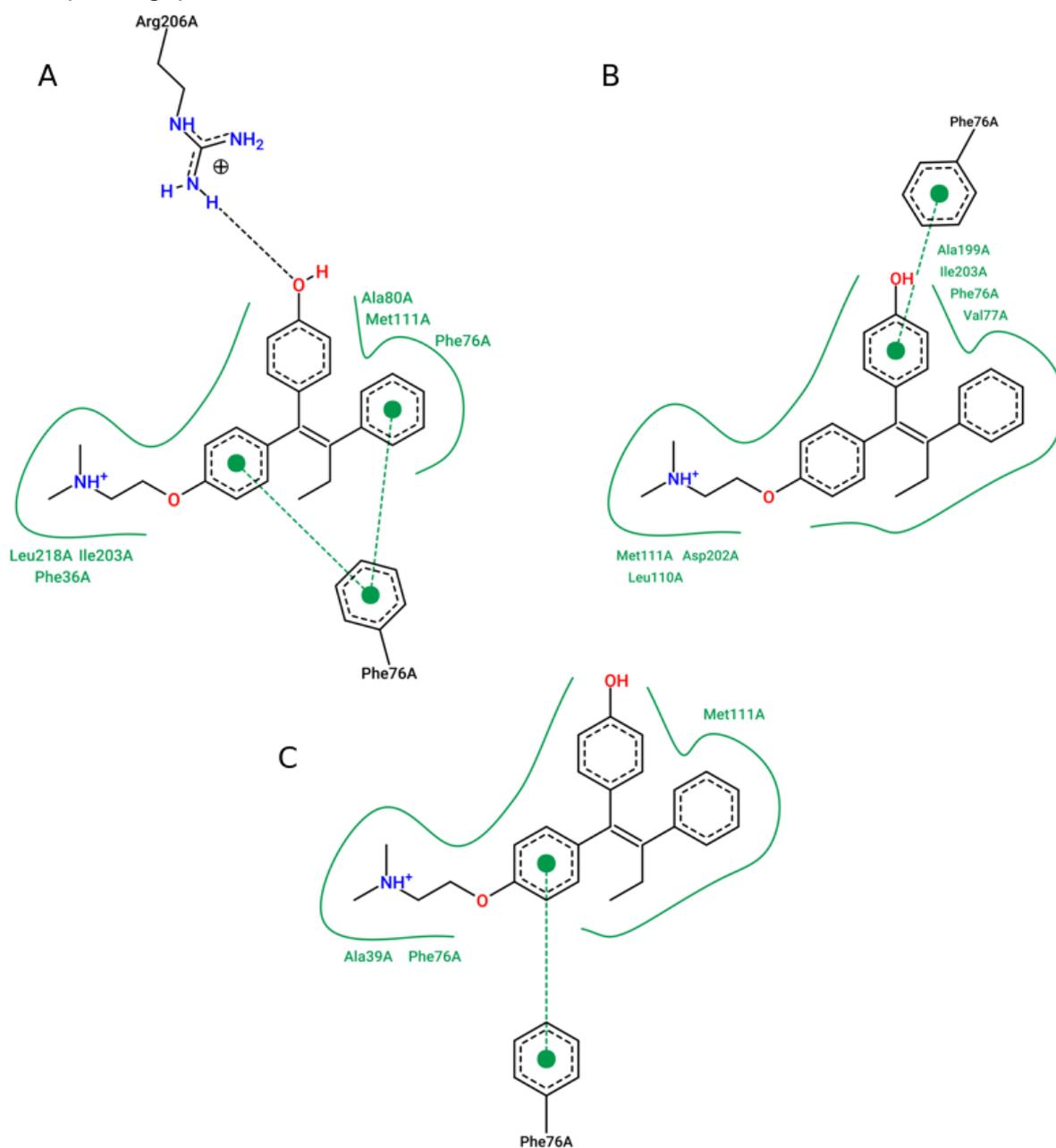
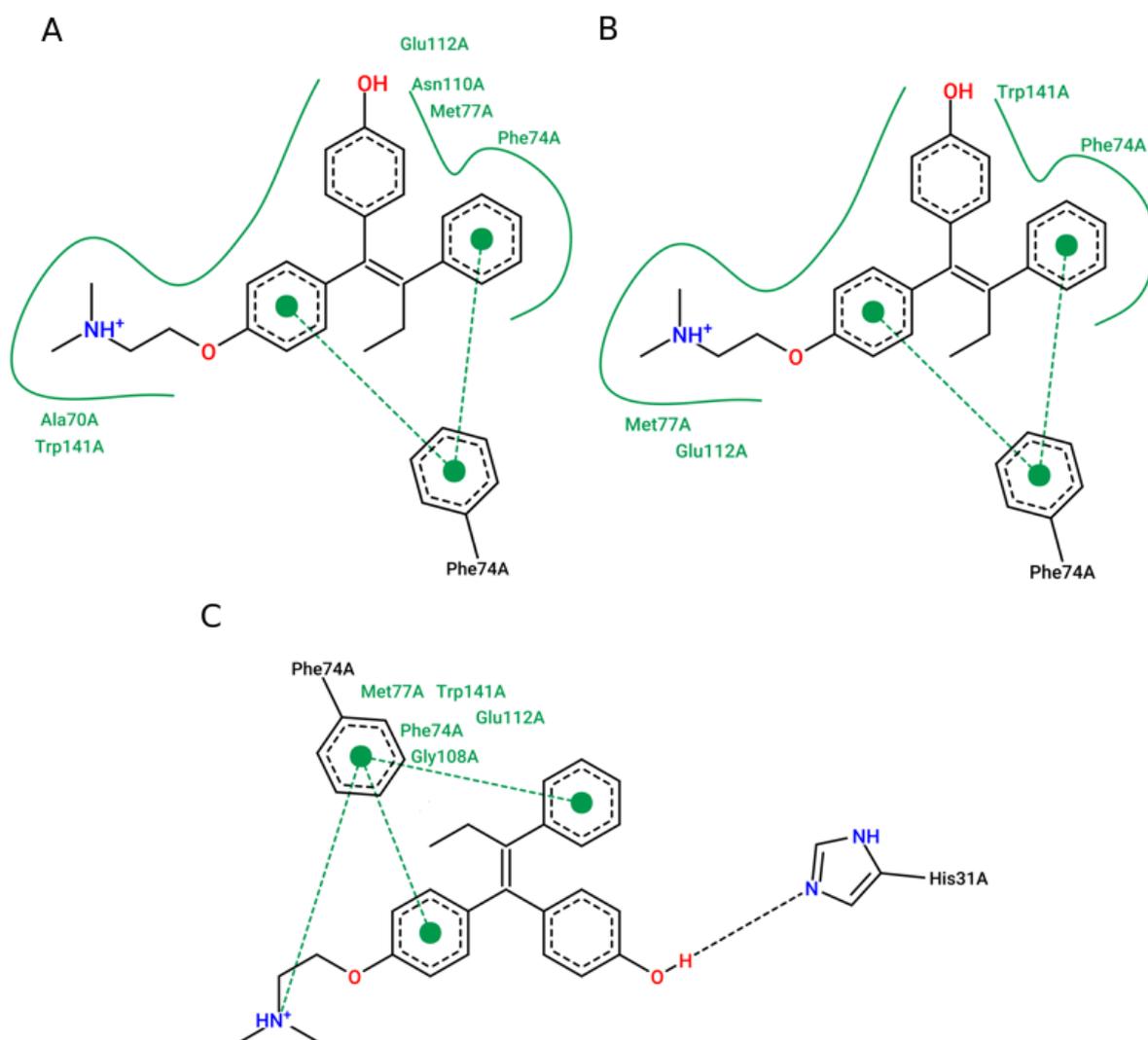


Figura 27 - Análise qualitativa das três posições de menor RMSD mostrando as interações químicas entre o tamoxifeno e os resíduos de aminoácidos que compõem a cavidade do sítio de ligação do modelo CgNR2E5. Na posição de menor RMSD (A) há formação de uma ponte de hidrogênio com Arg206, ligação  $\pi$ - $\pi$  entre Phe76 e dois anéis aromáticos de TAM, além de muitas interações hidrofóbicas envolvendo os resíduos Phe36, Phe76, Ala80, Met111, Ile203 e Leu218. A posição de segundo menor RMSD (B) apresenta interações hidrofóbicas com Phe76, Val77, Leu110, Met111, Ala199, Asp202 e Ile203, além de uma ligação  $\pi$ - $\pi$  entre Phe76 e um anel aromático do ligante. A posição de terceiro menor RMSD (C) configura interações hidrofóbicas ocorrendo entre os resíduos Ala39, Phe76 e Met111 e o ligante, com presença de ligação  $\pi$ - $\pi$  entre um anel aromático do TAM e o resíduo Phe76.



Fonte: Representação gráfica gerada pelo programa PoseView.

Figura 28 - Análise qualitativa das três posições de menor RMSD mostrando as interações químicas entre o tamoxifeno e os resíduos de aminoácidos que compõem a cavidade do sítio de ligação do modelo CgNR1P2. A posição de menor RMSD (A) apresentou diversas interações hidrofóbicas com resíduos de Ala70, Phe74, Met77, Asn110, Glu112, Trp141, além de uma ligação  $\pi$ - $\pi$  entre Phe74 e 2 anéis aromáticos do 4-hidroxitamoxifeno. Na posição de segundo menor valor de RMSD (B) também ocorreu formação de ligação  $\pi$ - $\pi$  entre Phe74 e 2 anéis aromáticos do TAM. Outras interações hidrofóbicas foram encontradas com Phe74, Met77, Glu112 e Trp141. Uma interação polar foi formada na terceira posição de menor RMSD (C), uma ponte de hidrogênio com o resíduo His31. Houve formação de ligação  $\pi$ - $\pi$  entre Phe74 com dois anéis aromáticos de TAM e também uma ligação  $\pi$ -cátion entre Phe74 e o nitrogênio do ligante. Interações hidrofóbicas estão presentes com Phe74, Met77, Gly108, Glu112 e Trp141.



Fonte: Representação gráfica gerada pelo programa PoseView.