UNIVERSIDADE ESTADUAL DE CAMPINAS

INSTITUTO DE QUÍMICA

Julia Nogueira Bezerra

Estudos de interação *in silico* da tirosinase humana e componentes do extrato de *Schinus terebinthifolius*

CAMPINAS
2022

Julia Nogueira Bezerra

Estudos de interação *in silico* da tirosinase humana e componentes do extrato de *Schinus terebinthifolius*

Dissertação de Mestrado apresentada ao Instituto de Química da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do título de Mestra em Química na área de Química Orgânica.

Orientador: Prof. Dr. Rodrigo Cormanich

Coorientadora: Profa. Dra. Wanda Pereira

O arquivo digital corresponde à versão final da Dissertação defendida pela aluna Julia Nogueira Bezerra e orientada pelo Prof. Dr. Rodrigo Cormanich.

CAMPINAS
2022

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca do Instituto de Química
Simone Luiz Alves - CRB 8/9094

B469e

Bezerra, Julia Nogueira, 1996-
Estudos de interação *in silico* da tirosinase humana e componentes do extrato de *Schinus terebinthifolius* / Julia Nogueira Bezerra. – Campinas, SP : [s.n.], 2022.

Orientador: Rodrigo Antonio Cormanich.
Dissertação (mestrado) – Universidade Estadual de Campinas, Instituto de Química.

1. Produtos naturais. 2. Tirosinase. 3. Melanina. 4. Docking molecular. 5. Dinâmica molecular. I. Cormanich, Rodrigo Antonio, 1986-. II. Universidade Estadual de Campinas. Instituto de Química. III. Título.

Informações para Biblioteca Digital

Título em outro idioma: *In silico* interactions of the components from the *Schinus terebinthifolius* extract with human tyrosinase
Palavras-chave em inglês:
Natural products
Tyrosinase
Melanin
Molecular docking
Molecular dynamics
Área de concentração: Química Orgânica
Titulação: Mestra em Química na área de Química Orgânica
Banca examinadora:
Rodrigo Antonio Cormanich [Orientador]
Daniel Fabio Kawano
Elaine Fontes Ferreira da Cunha
Data de defesa: 11-05-2022
Programa de Pós-Graduação: Química

Identificação e informações acadêmicas do(a) aluno(a)
- ORCID do autor: https://orcid.org/0000-0003-0724-318X
- Currículo Lattes do autor: http://lattes.cnpq.br/9253995806366487

BANCA EXAMINADORA

Prof. Dr. Rodrigo Cormanich (Orientador)

Prof. Dr. Daniel Fábio Kawano (Universidade Estadual de Campinas)

Profa. Dra. Elaine Fontes Ferreira da Cunha (Universidade Federal de Lavras)

A Ata da defesa, assinada pelos membros da Comissão Examinadora, consta no SIGA/Sistema de Fluxo de Dissertação/Tese e na Secretaria do Programa da Unidade.

Este exemplar corresponde à redação final da Dissertação de Mestrado defendida pela aluna **Julia Nogueira Bezerra**, aprovada pela Comissão Julgadora em 11 de maio de 2022.

Agradecimentos

Em primeiro lugar agradeço às forças que regem o universo e a Deus, por me dar força e coragem para realizar esse trabalho.

Agradeço também aos meus pais Beatriz e Jose Luis por sempre me incentivarem a me desenvolver intelectualmente. Às minhas irmãs por sempre acreditarem que eu sou capaz.

Agradeço aos amigos que me acompanharam durante a trajetória, me incentivando e também me lembrando o quanto momentos de descanso são necessários.

Ao meu orientador Prof. Rodrigo e coorientadora Prof. Wanda por todo seu trabalho de me guiar ao longo do mestrado.

A Joyce e Matheus pela orientação nos procedimentos de MIA-QSAR.

Ao João Paulo, membro do grupo de pesquisa, que me deu suporte nos procedimentos de dinâmica e foi parceiro de discussão de resultados. Ao Wilton pelo apoio e amizade ao longo desses dois anos.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

Agradecimento ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPQ) - Código de Financiamento 001.

Agradecimento à Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) - Código de Financiamento 2018-03910-1

Resumo

A tirosinase é uma enzima chave na melanogênese, uma via biossintética para a formação do pigmento melanina na pele humana, cuja inibição é importante como base de formulações utilizadas para o tratamento de manchas e hiperpigmentação da pele. Fontes botânicas de inibidor para uso cosmético são altamente desejáveis por sua compatibilidade com a pele e diminuição dos efeitos colaterais. Nesse sentido, estudos anteriores relataram a atividade anti-tirosinase do extrato de *Schinus terebinthifolius* e, no presente contexto, avaliamos a interação dos componentes desse extrato com a tirosinase humana utilizando ferramentas computacionais. Como a estrutura cristalina da tirosinase humana não está disponível no *Protein Data Bank*, realizou-se a modelagem de homologia em três softwares (SWISS MODELL, Schrödinger e Modeller) para prever a estrutura da tirosinase humana. A qualidade estereoquímica, por sua vez, foi avaliada usando gráficos de Ramachandran e Z-score; e, além disso, várias abordagens *in silico*, como *docking* molecular, MM-GBSA dinâmica molecular (MD), MM-PBSA, MIA-QSAR e predição da permeabilidade da pele, foram empregadas para rastrear os compostos responsáveis pela regulação negativa da atividade enzimática a fim de propor moléculas similares e mais potentes. Após realizados os cálculos e análises dos resultados, temos que o composto que apresentou maior potencial para aplicação cosmética foi a luteolina **17**. Na proposta de moléculas similares mais potentes, pode ser interessante estudar os derivados de 1,2,3,4,6-penta-O-galoil-glucose **22**, que apresentaram menor energia livre de ligação (**49**, **50**, **47** e **29**) para aplicação na indústria alimentícia com o intuito de evitar perdas decorrentes do escurecimento de frutas. Em ambos os casos seriam necessários testes *in vitro* e *in vivo* para confirmar o potencial de ação das moléculas.

Abstract

Tyrosinase is a key enzyme in melanogenesis, a biosynthetic pathway for the formation of the pigment melanin in human skin, the inhibition of this enzyme is important as a base of formulations used for the treatment of skin blemishes and hyperpigmentation. Botanical sources of inhibitors for cosmetic use are highly desirable for their skin compatibility and lessening of side effects. Previous studies had reported the anti-tyrosinase activity of extract of *Schinus terebinthifolius*; in this context, we evaluated the interaction of the components with human tyrosinase using computational tools. Since the crystal structure of human tyrosinase is not available in Protein Data Bank, homology modeling in three software (SWISS MODELL, Schrödinger and Modeller) was used to predict the structure of human tyrosinase. The stereochemical quality was evaluated by using Ramachandran plots and Z-score. Multiple in silico approaches such as molecular *docking*, MM-GBSA, molecular dynamic (MD), MM-PBSA, MIA-QSAR and skin permeability prediction were employed to screen the compounds responsible for the down-regulation of enzyme activity and to propose related and more potent molecules. After performing the calculations and analyzing the results, the compound that presented the most significant potential for cosmetic application was luteolin **17**. Searching for more potent molecular analogues, it may be interesting to study the derivatives of 1,2,3,4,6-penta-O-galloyl-glucose **22**, which showed lower binding free energy (**49**, **50**, **47** and **29**) for application in the food industry, in order to avoid losses, resulting from fruit browning. In both cases, *in vitro* and *in vivo* tests would be necessary to confirm the potential of the molecules.

LISTA DE FIGURAS

Figura 1: Rota biossintética da melanina 18

Figura 2: Folhas e frutos da árvore de *Schinus terebinthifolius* 20

Figura 3: Compostos do extrato das folhas de *Schinus terebinthifolius* em metanol, acetona e etanol[18–21] 22

Figura 4: Estrutura da proteína relacionada à tirosinase humana (TYRP1 - PDB 5M8L.A) 28

Figura 5: Modelo da proteína tirosinase construído no MODELLER antes (laranja) e depois (rosa) do *script* de enovelamento. 31

Figura 6: Detalhe do sítio ativo da proteína *template* co-cristalizada com a) ácido kójico, onde a molécula de água desempenha um papel de ponte entre ligante e metais e b) Tropolona, onde a mesma molécula de água se mostra irrelevante. Legenda: fitas brancas – proteínas, tubos verdes – átomos de carbono, tubos vermelhos – átomos de oxigênio, tubos brancos – átomos de hidrogênio, esferas azuis – Zn^{2+}. 33

Figura 7: Estrutura dos ligantes modificados, utilizados no modelo de MIA-QSAR 37

Figura 8: Estruturas do GaussView dos ligantes utilizados no modelo sobrepostas entre si. 38

Figura 9: Gráfico do Verify 3D do modelo de homologia knowledge-based da tirosinase humana modelada usando 5M8L como modelo, onde 97,71% dos resíduos têm pontuação média 3D-1D >= 0,2 42

Figura 10: Gráfico de Ramachandran do modelo de homologia *knowledge based*. O percentual de resíduos na área mais favorável é de 83,20% e na área proibida é de 1,30%. 42

Figura 11: Diagrama de interação ligante-proteína (a, c, e) e vista 3D da pose de *docking* molecular Δ*G* MM-GBSA mais negativa (b, d, f) do 1,2,3,4,6-penta-O-galoil-glucose **22** (a, b), ácido gálico **15** (c, d) e ácido cafeico **10** (e, f). Legenda: Fitas brancas – proteína, tubos verdes – carbono, tubos vermelhos – oxigênio, tubos brancos – hidrogênio, esferas de cobre – Cu^{2+}, rosa – distância entre as interações ligante e cobre. 45

Figura 12: Gráfico com a média dos valores de MM-PBSA de interação ligante proteína, juntamente com desvio padrão ao longo da dinâmica de 100ns para os ligantes 1-22. 48

Figura 13: Gráfico com a média dos valores de MM-PBSA de interação ligante com os íons Cu^{2+}, juntamente com desvio padrão ao longo da dinâmica de 100 ns para os ligantes 1-22. 48

Figura 14: Gráfico com a média dos valores de MM-PBSA de interação ligante proteína somado a interação ligante com íons Cu^{2+}, juntamente com desvio padrão ao longo da dinâmica de 100 ns para os ligantes 1-22. 49

Figura 15: Gráfico de William baseado em *studentized residues* e *sample leverages* para detectar dados discrepantes. 51

Figura 16: Gráfico MIA que mostra a importância da variável em diferentes regiões das estruturas sobrepostas (a) e os coeficientes de regressão PLS que mostram o impacto da variável (positivo ou negativo) nas estruturas sobrepostas (b). 52

Figura 17: Compostos que apresentaram melhor Δ*G* MM-GBSA em relação ao ligante original **22**. 58

Figura 18: Modelo *energy based* - Maestro 67

Figura 19: Modelo *energy based* – maestro após minimização. 67

Figura 20: Modelo *knowledge based* - Maestro 67

Figura 21: Modelo *knowledge based* – Maestro, após minimização 67

Figura 22: Modelo L8B082 (repositório) do Swiss-Model 68

Figura 23: Modelo L8B082 (repositório) do Swiss-Model após minimização. 68

Figura 24: Modelo Swiss-Model 68

Figura 25: Modelo Swiss-Model após minimização 68

Figura 26: Modelo Modeller 68

Figura 27: Modelo Modeller após minimização 68

Figura 28: modelo Modeller após aplicado o *script* loop (100x) 69

Figura 29: Modelo Modeller após aplicado *script* loop (100x) e minimizado 69

Figura 30: Modelo P14679 (repositório do Swiss-Model) 69

Figura 31: Modelo P14679 (repositório do Swiss-Model) após minimização 69

Figura 32: Vista 3D da pose de *docking* Δ*G* MM-GBSA mais negativa do ácido kójico (1) obtida por cada programa (Glide -azul, Gold -amarelo e Autodock- rosa) – a) sem água complexada com Cu^{2+} e b) com água complexada com Cu^{2+}. c) Diagrama de interação ligante da melhor pose - Glide com água Δ*G* = -13,76 kcal/mol 73

Figura 33: Vista 3D da pose de *docking* Δ*G* MM-GBSA mais negativa do metil galato (2) obtida por cada programa (Glide -azul, Gold -amarelo e Autodock- rosa) – a) sem água complexada com Cu^{2+} e b) com água complexada com Cu^{2+}. c) Diagrama de interação ligante da melhor pose - Autodock com água Δ*G* = -11,42 kcal/mol 74

Figura 34: Vista 3D da pose de *docking* Δ*G* MM-GBSA mais negativa do miricetina (3) obtida por cada programa (Glide -azul, Gold -amarelo e Autodock- rosa) – a) sem água complexada com Cu^{2+} e b) com água complexada com Cu^{2+}. c) Diagrama de interação ligante da melhor pose - Glide com água ΔG = -33,49 kcal/mol 75

Figura 35: Vista 3D da pose de *docking* Δ*G* MM-GBSA mais negativa do etil galato (4) obtida por cada programa (Glide -azul, Gold -amarelo e Autodock- rosa) – a) sem água complexada com Cu^{2+} e b) com água complexada com Cu^{2+}. c) Diagrama de interação ligante da melhor pose - Glide sem água ΔG = -17,26 kcal/mol 76

Figura 36: Vista 3D da pose de *docking* Δ*G* MM-GBSA mais negativa do catequina (5) obtida por cada programa (Glide -azul, Gold -amarelo e Autodock- rosa) – a) sem água complexada com Cu^{2+} e b) com água complexada com Cu^{2+}. c) Diagrama de interação ligante da melhor pose - Glide com água ΔG = -17,99 kcal/mol 77

Figura 37: Vista 3D da pose de *docking* Δ*G* MM-GBSA mais negativa do quercetina ramnosídeo (6) obtida por cada programa (Glide -azul, Gold -amarelo e Autodock- rosa) – a) sem água complexada com Cu^{2+} e b) com água complexada com Cu^{2+}. c) Diagrama de interação ligante da melhor pose – Gold sem água Δ*G* = -31,16 kcal/mol 78

Figura 38: Vista 3D da pose de *docking* Δ*G* MM-GBSA mais negativa do miricetina ramnosídeo (7) obtida por cada programa (Glide -azul, Gold -amarelo e Autodock- rosa) – a) sem água complexada com Cu^{2+} e b) com água complexada com Cu^{2+}. c) Diagrama de interação ligante da melhor pose – Glide sem água Δ*G* = -27,71 kcal/mol 79

Figura 39: Vista 3D da pose de *docking* Δ*G* MM-GBSA mais negativa do ácido elágico (8) obtida por cada programa (Glide -azul, Gold -amarelo e Autodock- rosa) – a) sem água complexada com Cu^{2+} e b) com água complexada com Cu^{2+}. c) Diagrama de interação ligante da melhor pose – Glide com água Δ*G* = -26,05 kcal/mol 80

Figura 40: Vista 3D da pose de *docking* Δ*G* MM-GBSA mais negativa do quercetina (9) obtida por cada programa (Glide -azul, Gold -amarelo e Autodock- rosa) – a) sem água complexada com Cu^{2+} e b) com água complexada com Cu^{2+}. c) Diagrama de interação ligante da melhor pose – Glide com água Δ*G* = -31,28 kcal/mol 81

Figura 41: Vista 3D da pose de *docking* Δ*G* MM-GBSA mais negativa do ácido cafeico (10) obtida por cada programa (Glide -azul, Gold -amarelo e Autodock- rosa) – a) sem água complexada com Cu^{2+} e b) com água complexada com Cu^{2+}. c) Diagrama de interação ligante da melhor pose – Glide com água Δ*G* = -37,67 kcal/mol 82

Figura 42: Vista 3D da pose de *docking* Δ*G* MM-GBSA mais negativa do ácido siríngico (11) obtida por cada programa (Glide -azul, Gold -amarelo e Autodock- rosa) – a) sem água complexada com Cu^{2+} e b) com água complexada com Cu^{2+}. c) Diagrama de interação ligante da melhor pose – Autodock sem água Δ*G* = -26,63 kcal/mol 83

Figura 43: Vista 3D da pose de *docking* Δ*G* MM-GBSA mais negativa do ácido coumárico (12) obtida por cada programa (Glide -azul, Gold -amarelo e Autodock- rosa) – a) sem água complexada com Cu^{2+} e b) com água complexada com Cu^{2+}. c) Diagrama de interação ligante da melhor pose – Gold sem água Δ*G* = -31,73 kcal/mol 84

Figura 44: Vista 3D da pose de *docking* Δ*G* MM-GBSA mais negativa do 1,6-Digaloil-glucose (13) obtida por cada programa (Glide -azul, Gold -amarelo e Autodock- rosa) – a) sem água complexada com Cu^{2+} e b) com água complexada com Cu^{2+}. c) Diagrama de interação ligante da melhor pose – Glide com água Δ*G* = -30,48 kcal/mol 85

Figura 45: Vista 3D da pose de *docking* Δ*G* MM-GBSA mais negativa do ácido clorogênico (14) obtida por cada programa (Glide -azul, Gold -amarelo e Autodock- rosa) – a) sem água complexada com Cu^{2+} e b) com água complexada com Cu^{2+}. c) Diagrama de interação ligante da melhor pose – Gold sem água Δ*G* = -35,32 kcal/mol 86

Figura 46: Vista 3D da pose de *docking* Δ*G* MM-GBSA mais negativa do ácido gálico (15) obtida por cada programa (Glide -azul, Gold -amarelo e Autodock- rosa) – a) sem água complexada com Cu^{2+} e b) com água complexada com Cu^{2+}. c) Diagrama de interação ligante da melhor pose – Gold sem água Δ*G* = -39,65 kcal/mol 87

Figura 47: Vista 3D da pose de *docking* Δ*G* MM-GBSA mais negativa do ácido 5-p-coumaroilquinico (16) obtida por cada programa (Glide -azul, Gold -amarelo e Autodock- rosa) – a) sem água complexada com Cu^{2+} e b) com água complexada com Cu^{2+}. c) Diagrama de interação ligante da melhor pose – Glide com água Δ*G* = -15,57 kcal/mol 88

Figura 48: Vista 3D da pose de *docking* Δ*G* MM-GBSA mais negativa da luteolina (17) obtida por cada programa (Glide -azul, Gold -amarelo e Autodock- rosa) – a) sem água complexada com Cu^{2+} e b) com água complexada com Cu^{2+}. c) Diagrama de interação ligante da melhor pose – Glide sem água Δ*G* = -30,91 kcal/mol 89

Figura 49: Vista 3D da pose de *docking* Δ*G* MM-GBSA mais negativa do miricetina ramnosídeo galactosídeo (18) obtida por cada programa (Glide -azul, Gold -amarelo e Autodock- rosa) – a) sem água complexada com Cu^{2+} e b) com água complexada com Cu^{2+}. c) Diagrama de interação ligante da melhor pose – Glide com água Δ*G* = -35,01 kcal/mol 90

Figura 50: Vista 3D da pose de *docking* Δ*G* MM-GBSA mais negativa do miricetina glucoronídeo (19) obtida por cada programa (Glide -azul, Gold -amarelo e Autodock- rosa) – a) sem água complexada com Cu^{2+} e b) com água complexada com Cu^{2+}. c) Diagrama de interação ligante da melhor pose – Gold com água Δ*G* = -25,02 kcal/mol 91

Figura 51: Vista 3D da pose de *docking* Δ*G* MM-GBSA mais negativa do miricetina galactosídeo (20) obtida por cada programa (Glide -azul, Gold -amarelo e Autodock- rosa) – a) sem água complexada com Cu^{2+} e b) com água complexada com Cu^{2+}. c) Diagrama de interação ligante da melhor pose – Glide sem água Δ*G* = -34,77 kcal/mol 92

Figura 52: Vista 3D da pose de *docking* Δ*G* MM-GBSA mais negativa do robustaflavona (21) obtida por cada programa (Glide -azul, Gold -amarelo e Autodock- rosa) – a) sem água complexada com Cu^{2+} e b) com água complexada com Cu^{2+}. c) Diagrama de interação ligante da melhor pose – Glide com água Δ*G* = -23,38 kcal/mol 93

Figura 53: Vista 3D da pose de *docking* Δ*G* MM-GBSA mais negativa do 1,2,3,4,6-penta-O-galoil-glucopiranosideo (22) obtida por cada programa (Glide -azul, Gold -amarelo e Autodock- rosa) – a) sem água complexada com Cu^{2+} e b) com água complexada com Cu^{2+}. c) Diagrama de interação ligante da melhor pose – Glide com água Δ*G* = -61,11 kcal/mol 94

Figura 54: Distância entre proteína e ligantes, catequina (5-magenta), quercetina-3-O-α-ramnosídeo (6-violeta), miricetina-3-O-α-ramnosídeo (7-ciano), ácido elágico (8-laranja) . 95

Figura 55: Distância entre proteína e ligantes, ácido kójico (1-magenta), metil galato (2-violeta), miricetina (3-ciano), etil galato (4-laranja). 95

Figura 56: Distância entre proteína e ligantes, 1,6-digalloil-ß-D-glicose (13-magenta), ácido clorogênico (14-violeta), ácido gálico (15-ciano), ácido 5-p-cumaroilquínico (16-laranja) . 95

Figura 57: Distância entre proteína e ligantes, 1,6-digalloil-ß-D-glicose (13-magenta), ácido clorogênico (14-violeta), ácido gálico (15-ciano), ácido 5-p-cumaroilquínico (16-laranja) . 95

Figura 58: Distância entre proteína e ligantes, flavona robusta (21-magenta), 1,2,3,4,6-penta-O-galloil-glucopiranosídeo (22-violeta). 96

Figura 59: Distância entre proteínas e ligantes, luteolina (17-magenta), miricetina- ramnosídeo 1"'->6" galactosídeo (18-violeta), miricetina glucoronídeo (19-ciano), miricetina galactosídeo (20-laranja). 96

Figura 60: RMSD (nm) do sítio de ligação, a partir da proteína simulada com ácido kójico (1-magenta), metil galato (2-violeta), miricetina (3-ciano), etil galato (4-laranja). 96

Figura 61: RMSD (nm) do sítio de ligação, da proteína simulada com catequina (5-magenta), quercetina-3-O-α-ramnosídeo (6-violeta), miricetina-3-O-α-ramnosídeo (7-ciano), ácido elágico (8-laranja). 96

Figura 62: RMSD (nm) do sítio de ligação, a partir da proteína simulada com quercetina (9-magenta), ácido cafeico (10-violeta), ácido siríngico (11-ciano), ácido coumárico (12-laranja). 97

Figura 63: RMSD (nm) do sítio de ligação, da proteína simulada com 1,6-digalloil-ß-D-glicose (13-magenta), ácido clorogênico (14-violeta), ácido gálico (15-ciano), 5- ácido p-coumaroilquínico (16-laranja). 97

Figura 64: RMSD (nm) do sítio de ligação, da proteína simulada com luteolina (17-magenta), miricetina- ramnosídeo 1"'->6" galactosídeo (18-violeta), miricetina glucoronídeo (19-ciano), miricetina galactosídeo (20-laranja). 97

Figura 65: RMSD (nm) do sítio de ligação, da proteína simulada com robustaflavona (21-magenta), 1,2,3,4,6-penta-O-galloil-glucopiranosídeo (22-violeta). 97

Figura 66: RMSD (nm) do C-alfa, da proteína simulada com catequina (5-magenta), quercetina-3-O-α-ramnosídeo (6-violeta), miricetina-3-O-α-ramnosídeo (7 -ciano), ácido elágico (8-laranja). 98

Figura 67: RMSD (nm) do C-alfa, a partir da proteína simulada com ácido kójico (1-magenta), metil galato (2-violeta), miricetina (3-ciano), etil galato (4-laranja). 98

Figura 68: RMSD (nm) do C-alfa, a partir da proteína simulada com quercetina (ácido cafeico 9-magenta (10-violeta), ácido siríngico (11-ciano), ácido coumárico (12-laranja). 98

Figura 69: RMSD (nm) do C-alfa, da proteína simulada com 1,6-digalloil-ß-D-glicose (13-magenta), ácido clorogênico (14-violeta), ácido gálico (15-ciano), 5 ácido -p-coumaroilquínico (16-laranja). 98

Figura 70: RMSD (nm) do C-alfa, da proteína simulada com luteolina (17-magenta), miricetina- ramnosídeo 1"'->6" galactosídeo (18-violeta), miricetina glucoronídeo (19-ciano), miricetina galactosídeo (20-laranja). 99

Figura 71: RMSD (nm) do C-alfa, da proteína simulada com robustaflavona (21-magenta), 1,2,3,4,6-penta-O-galloil-glucopiranosídeo (22-violeta). 99

Figura 72: RMSD (nm) de ligantes, catequina (5-magenta), quercetina-3-O-α-ramnosídeo (6-violeta), miricetina-3-O-α-ramnosídeo (7-ciano), ácido elágico (8-laranja). 99

Figura 73: RMSD (nm) dos ligantes, ácido kójico (1-magenta), metil galato (2-violeta), miricetina (3-ciano), etil galato (4-laranja). 99

Figura 74: RMSD (nm) dos ligantes, quercetina (ácido cafeico 9-magenta (10-violeta), ácido siríngico (11-ciano), ácido coumárico (12-laranja). 100

Figura 75: RMSD (nm) dos ligantes, da proteína simulada com 1,6-digalloil-ß-D-glucose (13-magenta), ácido clorogênico (14-violeta), ácido gálico (15-ciano), 5-p -ácido coumaroilquínico (16-laranja). 100

Figura 76: RMSD (nm) dos ligantes, luteolina (17-magenta), miricetina- ramnosídeo 1"'->6" galactosídeo (18-violeta), miricetina glucoronídeo (19-ciano), miricetina galactosídeo (20-laranja). 100

Figura 77: RMSD (nm) dos ligantes robustaflavona (21-magenta), 1,2,3,4,6-penta-O-galloil-glucopiranosídeo (22-violeta). 100

Figura 78: RMSF de backbone, proteína e cadeia lateral da simulação de dinâmica molecular com metil galato (2). 101

Figura 79: RMSF de backbone, proteína e cadeia lateral da simulação de dinâmica molecular com ácido kójico (1). 101

Figura 80: RMSF de backbone, proteína e cadeia lateral da simulação de dinâmica molecular com etil galato (4). 101

Figura 81: RMSF de backbone, proteína e cadeia lateral da simulação de dinâmica molecular com miricetina (3). 101

Figura 82: RMSF de backbone, proteína e cadeia lateral da simulação de dinâmica molecular com quercetina ramnosídeo (6) 102

Figura 83: RMSF de backbone, proteína e cadeia lateral da simulação de dinâmica molecular com catequina (5). 102

Figura 84: RMSF de backbone, proteína e cadeia lateral da simulação de dinâmica molecular com miricetina-3-O-α-ramnosídeo (7) 102

Figura 85: RMSF de backbone, proteína e cadeia lateral da simulação de dinâmica molecular com ácido elágico (8) 102

Figura 86: RMSF de backbone, proteína e cadeia lateral da simulação de dinâmica molecular com ácido cafeico (10) 103

Figura 87: RMSF de backbone, proteína e cadeia lateral da simulação de dinâmica molecular com quercetina (9) 103

Figura 88: RMSF de backbone, proteína e cadeia lateral da simulação de dinâmica molecular com ácido coumárico (12) 103

Figura 89: RMSF de backbone, proteína e cadeia lateral da simulação de dinâmica molecular com ácido siríngico (11) 103

Figura 90: RMSF de backbone, proteína e cadeia lateral da simulação de dinâmica molecular com ácido clorogênico (14). 104

Figura 91: RMSF de backbone, proteína e cadeia lateral da simulação de dinâmica molecular com 1,6-digalloil-ß-D-glucose (13). 104

Figura 92: RMSF de backbone, proteína e cadeia lateral da simulação de dinâmica molecular com ácido 5-p-cumaroilquínico (16). 104

Figura 93: RMSF de backbone, proteína e cadeia lateral da simulação de dinâmica molecular com ácido gálico (15) 104

Figura 94: RMSF de backbone, proteína e cadeia lateral da simulação de dinâmica molecular com luteolina (17). 105

Figura 95: RMSF de backbone, proteína e cadeia lateral da simulação de dinâmica molecular com miricetina- ramnosídeo 1'''->6'' galactosídeo (18). 105

Figura 96: RMSF de backbone, proteína e cadeia lateral da simulação de dinâmica molecular com miricetina glucoronídeo (19). 105

Figura 97: RMSF de backbone, proteína e cadeia lateral da simulação de dinâmica molecular com miricetina galactosídeo (20). 105

Figura 98: RMSF de backbone, proteína e cadeia lateral da simulação de dinâmica molecular com robustaflavona (21) 106

Figura 99: RMSF de backbone, proteína e cadeia lateral da simulação de dinâmica molecular com 1,2,3,4,6-penta-O-galloil-glucopiranosídeo (22-violeta). 106

Figura 100: Raio de giro do ácido kójico (1-magenta), metil galato (2-violeta), miricetina (3-ciano), etil galato (4-laranja). 106

Figura 101: Raio de giro de catequina (5-magenta), quercetina-3-O-α-ramnosídeo (6-violeta), miricetina-3-O-α-ramnosídeo (7-ciano), ácido elágico (8-laranja). 106

Figura 102: Raio de giro da quercetina (ácido cafeico 9-magenta (10-violeta), ácido siríngico (11-ciano), ácido coumárico (12-laranja). 110

Figura 103: Raio de giro de 1,6-digalloil-ß-D-glicose (13-magenta), ácido clorogênico (14-violeta), ácido gálico (15-ciano), ácido 5-p-cumaroilquínico (16-laranja). 107

Figura 104: Raio de giro da luteolina (17-magenta), miricetina- ramnosídeo 1'''->6'' galactosídeo (18-violeta), miricetina glucoronídeo (19-ciano), miricetina galactosídeo (20-laranja). 107

Figura 105: Raio de giro da robustaflavona (21-magenta) e com 1,2,3,4,6-penta-O-galloil-glucopiranosídeo (22-violeta). 107

LISTA DE TABELAS

Tabela 1: Resultados dos testes de validação estrutural dos modelos de homologia construídos. 39

Tabela 2: Resultados de R^2 de cada um dos modelos e seus respectivos ligantes 41

Tabela 3: Valor mais negativo de Δ*G* MM-GBSA, comparando os programas Glide, AutodockVina e Gold para a proteína com e sem água coordenada aos átomos de Cu^{2+} dos ligantes 1-22. 43

Tabela 4: Menor valor de Δ*G* MM-GBSA para os ligantes modificados utilizados na construção do modelo de MIA-QSAR. 50

Tabela 5: Valores de Log Kps para os componentes do extrato de Schinus terebinthifolius e ácido kójico. 53

Tabela 6: O país da planta, o solvente utilizado na extração e a quantidade de cada componente do extrato de *Schinus terebinthifolius* de acordo com sua respectiva referência. 56

Tabela 7: O país da planta, o solvente utilizado na extração e a quantidade de cada componente do extrato de *Schinus terebinthifolius* de acordo com sua respectiva referência.70

Sumário

1. Introdução 18

1.1 Melanogênese 18

1.2 Tirosinase 18

1.3 *Schinus terebinthifolius* 20

1.4 Estudos *in silico* de interações proteína-ligante 23

1.4.1 Modelagem por homologia 23

1.4.2 *Docking* 24

1.4.3 MM-GBSA 25

1.4.4 Dinâmica molecular 26

1.4.5 MM-PBSA 26

1.4.6 MIA-QSAR 27

1.4.7 Permeabilidade cutânea teórica 27

2 Objetivos 28

2.1 Objetivo geral 28

2.2 Objetivos específicos 28

2.2.1 Modelagem por homologia 28

2.2.2 Validação estrutural do modelo 28

2.2.3 Cálculos de *docking* molecular 28

2.2.4 Cálculos de ΔG (MM-GBSA) 29

2.2.5 Dinâmica molecular 29

2.2.6 MM-PBSA 29

2.2.7 MIA-QSAR 29

2.2.8 Permeabilidade cutânea teórica 29

3 Metodologia 30

3.1 Modelagem por homologia 30

3.1.1 Modeller[25] 30

3.1.2 SWISS-MODEL[26] 31

3.1.3 Prime 31

3.1.4 Protein preparation wizard[57] 31

3.2 Verificação da estrutura 32

3.2.1 Gráfico de Ramachandran gerado no software PROCHECK [50] 32

3.2.2 Verify 3D [51] 32

3.3 Validação do modelo 32
3.3.1 *Re-docking* e *Root Mean Square Deviation* (RMSD) 32
3.3.2 Cálculo de ΔG (MM-GBSA) 33
3.4 *Docking* Molecular 34
3.4.1 Preparação dos ligantes 34
3.4.2 *Docking* molecular 34
3.4.2.1 Glide 34
3.4.2.2 Gold 34
3.4.2.3 Autodock 34
3.4.3 MM-GBSA 35
3.4.4 Dinâmica Molecular 35
3.4.5 MM-PBSA 35
3.4.6 MIA-QSAR 36
3.4.7 Permeabilidade cutânea teórica 38
4 Resultados 39
4.1 Modelagem por homologia 39
4.2 Verificação da estrutura 39
4.3 *RMSD* e *re-docking* 40
4.4 *Docking* molecular 42
4.5 Dinâmica molecular 46
4.6 MM-PBSA 47
4.7 MIA-QSAR 49
4.8 Permeabilidade cutânea teórica 52
5 Discussão 54
6 Conclusões 58
7 Bibliografia 60

1. Introdução

1.1 Melanogênese

A pele é o maior órgão do corpo humano e a primeira barreira de proteção contra microrganismos e raios UV[1] e é classificada histologicamente em três camadas: epiderme, hipoderme e derme, tendo, cada uma delas, células com funções específicas para a proteção do organismo. Na epiderme se encontram os melanócitos, células que fornecem pigmentação à pele através da produção de melanina, a qual é produzida através da oxidação do aminoácido tirosina pela enzima tirosinase.[2] O processo de pigmentação começa com a oxidação da L-tirosina para L-dopaquinona, seguida da formação do dopacromo, o qual posteriormente passa por um processo de ciclização e oxidação intramolecular que acaba por produz o substrato para a produção da melanina (Figura 1).[3] A etapa limitante para todo o processo é a hidroxilação da L-tirosina para L-3,4- dihidroxifenilalanina (L-DOPA) catalisada pela tirosinase.[4]

1.2 Tirosinase

A tirosinase se trata de uma metaloenzima de cobre com crucial importância na rota de síntese da melanina. Em sua estrutura há um sítio ativo composto por dois átomos de cobre, cada um coordenado com três resíduos de histidina; os átomos de cobre ligam-se ao oxigênio e atuam tanto na hidroxilação de um monofenol (tirosina), quanto na conversão do ortodifenol (L-DOPA) resultante em sua ortoquinona correspondente (Figura 1). Após algumas reações bioquímicas, essa ortoquinona se converterá em melanina.[5]

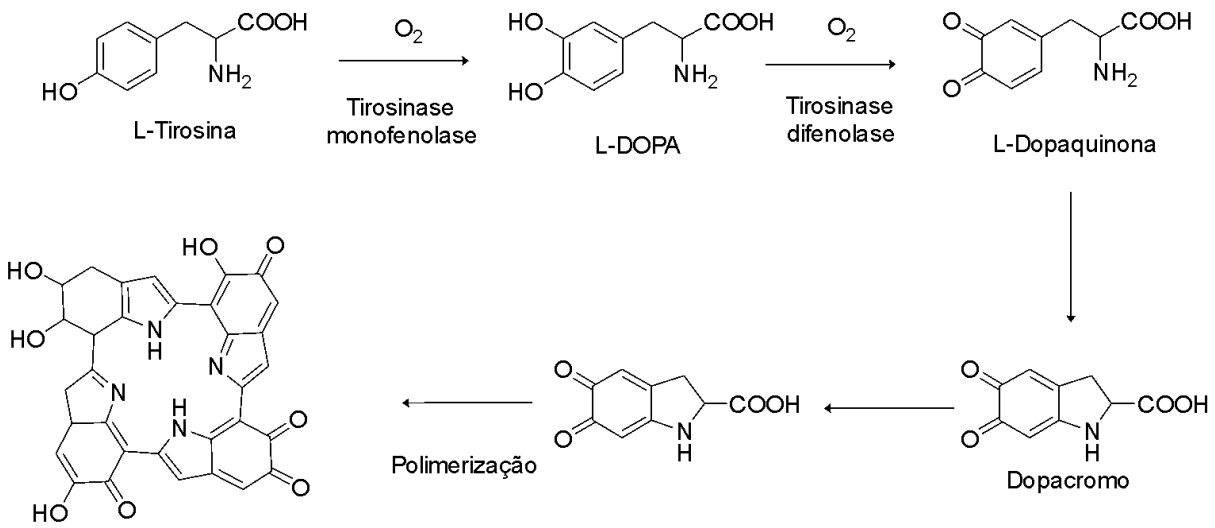


Figura 1: Rota biossintética da melanina

O envolvimento da tirosinase na produção de melanina justifica sua escolha como alvo para inibidores com aplicações na indústria cosmética, assim como na alimentícia,

especialmente considerando quea melanina é a principal proteção química do corpo à radiação ultravioleta. Existem dois tipos principais de melanina, a eumelanina e a feomelanina: a primeira apresenta uma cor preta/marrom, enquanto a segunda apresenta tons mais amarelados e avermelhados. São esses dois pigmentos que, a depender da proporção em que ocorrem, definem a tonalidade de pele e cabelo de um indivíduo[6]; além de serem os responsáveis, sob o ponto de vista fisiológico, por proteger a pele contra raios UV. Entretanto, distúrbios na produção de melanina podem desencadear problemas de ordem estética, como hiperpigmentação e melasma, que podem ser agravados por exposição ao sol e desequilíbrios hormonais.[7]

Ademais, a tirosinase também é responsável pelo escurecimento de frutas e vegetais como resposta ao corte e exposição ao oxigênio, sendo alguns desses alimentos mais sujeitos a esse tipo de reação, como: batatas, maçãs, cogumelos, bananas, pêssegos, sucos de frutas e vinhos. Como se pode depreender a partir do funcionamento da reação, o escurecimento é mais grave quando o alimento foi submetido a danos na superfície, enquanto que nas frutas e hortaliças não cortadas ou não danificadas, os substratos fenólicos naturais são separados da enzima por compartimentação e o escurecimento não ocorre.[8] Nesse sentido, a possibilidade de inibição da tirosinase revela-se relevante para a indústria de alimentos, uma vez que o escurecimento não só prejudica a aparência de frutas e vegetais, como também causa perda de valor nutricional, o que se atribui à destruição de aminoácidos essenciais e diminuição na digestibilidade.[9]

Apesar de um grande número de inibidores para a tirosinase já terem sido identificados, poucos são de fato utilizados devido a questões de segurança e performance *in vivo*. Alguns estudos apontam que o ácido kójico e em especial a hidroquinona, ainda que muito usados na indústria cosmética, podem causar reações adversas, de tal modo que hoje a hidroquinona é um ingrediente proibido na União Europeia.[10] Já no caso da indústria alimentícia, os inibidores usados são restritos pelo sabor, segurança e viabilidade econômica. Tradicionalmente eram usados sulfitos, mas como podem causar reações alérgicas e alteração do sabor, atualmente é usada uma mistura de ácido cítrico com ascórbico, mas que também é uma alternativa pouco eficiente, uma vez que o ácido ascórbico é rapidamente consumido pelo processo de redução da tirosinase.[11]

Dessa maneira, incentiva-se a constante busca por novos inibidores que não causem efeitos adversos. Tendo isso em vista que metabólitos e derivados de plantas têm se apresentado como alternativas promissoras tanto para a indústria de alimentos, uma vez que a percepção de ‘natural’ é um fator importante para a melhor aceitação de tecnologias alimentícias; assim como

para a indústria cosmética, pois são uma boa direção e fonte de inspiração, pois são mais propensos a ter maior compatibilidade biológica com a pele e melhor aceitação pela demanda cosmética do consumidor.[12] Um exemplo é a aroeira vermelha (*Schinus terebinthifolius*), cujo extrato apresenta, dentre muitas propriedades, atividade anti-tirosinase.[13]

1.3 *Schinus terebinthifolius*

Também conhecida como aroeira vermelha, pimenta rosa e aroeira da praia, a *Schinus terebinthifolius* é uma planta utilizada na medicina popular para tratamentos no sistema urogenital e, segundo a literatura, possui atividade antimicrobiana e anti-inflamatória. Trata-se de uma espécie nativa brasileira que ocorre em biomas brasileiros como mata atlântica, cerrado e pampas[14], além de também estar presente em outros países da América do Sul, partes da América Central, Flórida, Europa Mediterrânea, Norte da África, Sul da Ásia e África do Sul.[15] Na Figura 2[16] é mostrada uma imagem dos frutos e folhas da árvore.

Figura 2: Folhas e frutos da árvore de *Schinus terebinthifolius*

Foi reportado na literatura que os extratos em acetona e metanol de aroeira vermelha apresentaram alta atividade inibitória da enzima tirosinase, abrindo espaço para estudos complementares de identificação dos compostos responsáveis por essa atividade. Os extratos que apresentaram melhor desempenho na inibição enzimática foram os das folhas em acetona e metanol,[17] e os componentes desses extratos foram identificados através de técnicas como RMN e HPLC e reportados na literatura.[6] Aqueles provenientes das folhas dos extratos em metanol, acetona e etanol foram selecionados para o presente trabalho (Figura 3). Os métodos de extração e quantidades encontradas de cada composto são apresentados no Anexo 2.

Até o momento não é descrito na literatura qual desses compostos possuiria atividade anti-tirosinase. Para realizar tal determinação é necessário avaliar a interação proteína-ligante de cada um dos compostos individualmente. Como se trata de um grande número de compostos, cuja separação e estudo experimental *in vitro* não seria simples, o presente trabalho propõe a aplicação de uma seleção preliminar, utilizando-se uma metodologia de menor custo: estudos *in silico* de interação proteína-ligante.

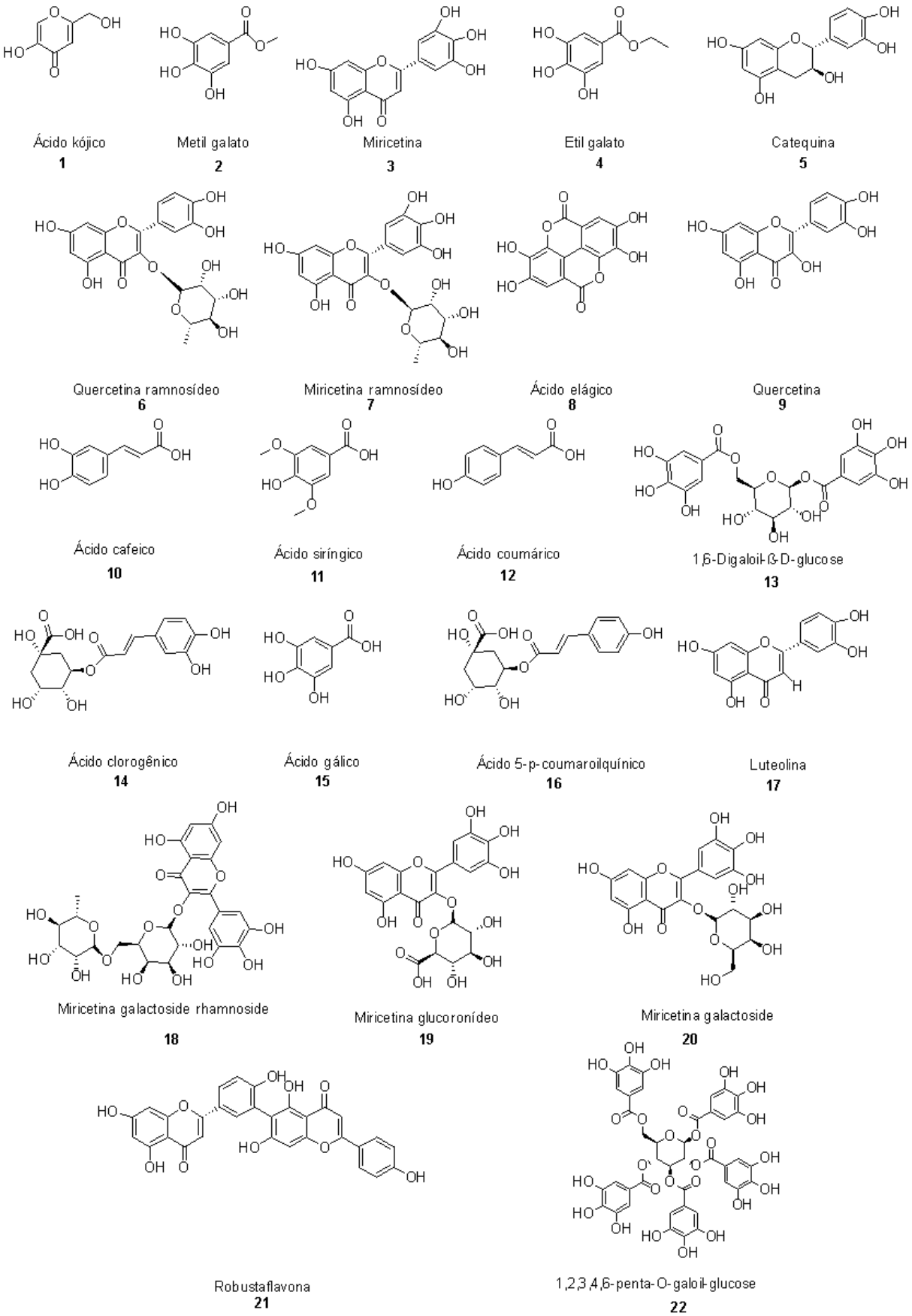


Figura 3: Compostos do extrato das folhas de *Schinus terebinthifolius* em metanol, acetona e etanol[18–21]

1.4 Estudos *in silico* de interações proteína-ligante

Estudos computacionais de interação proteína-ligante vêm sendo bastante utilizados principalmente nas etapas iniciais do desenvolvimento de fármacos, uma vez que permitem o estudo preliminar de um grande número de moléculas com tempo e custo consideravelmente menores do que estudos *in vitro*.[22] Vale ressaltar que, apesar de esses resultados serem apenas um estudo preliminar, sendo ainda imprescindível o estudo *in vitro*, por outro lado permitem selecionar de maneira mais eficaz as moléculas que serão direcionadas para os estudos experimentais. Com o desenvolvimento e maior acessibilidade dos recursos computacionais, técnicas e cálculos para a análise de grandes sistemas químicos têm também se desenvolvido.

Hoje há disponível uma razoável gama de ferramentas para a realização de estudos *in silico* de interações proteína-ligante que possibilita obter resultados acurados para a identificação de potenciais inibidores.[23] O processo de realização desses estudos envolve várias etapas, desde a seleção dos potenciais ligantes e obtenção da estrutura da proteína, passando por cálculos de *docking* e dinâmica molecular e, em alguns casos, estudos de energia de interação por MM-GBSA (*Molecular mechanics with generalised Born and surface area solvation*), MM-PBSA (*Molecular Mechanics Poisson-Boltzmann Surface Area*) e FEP (*Free Energy Perturbation*). O trabalho inicia-se com a obtenção da estrutura da proteína de interesse, neste caso através da modelagem por homologia.

1.4.1 Modelagem por homologia

A modelagem por homologia é realizada através de softwares que usam diferentes algoritmos para construir o modelo 3D da proteína. Cálculos de homologia se baseiam em uma estrutura *template* cuja sequência de aminoácidos (estrutura primária) é similar a uma estrutura cristalográfica análoga e com estrutura tridimensional conhecida e reportada na literatura. Um dos softwares disponíveis para cálculos de modelagem por homologia é o SWISS-MODEL,[24] que usa métodos de montagem de corpo rígido em que um modelo é montado a partir de um pequeno número de corpos rígidos obtidos do centro das regiões alinhadas. A construção envolve encaixar os corpos rígidos na estrutura e reconstruir as partes não conservadas, isto é, ligações e cadeias laterais. Já o software MODELLER[25] executa a modelagem pela satisfação das restrições espaciais, que são derivadas do alinhamento, de forma que o modelo é obtido minimizando as violações a essas restrições.[26]

Além desses, o software Prime[27], dentro do pacote Schroedinger 2020-3, calcula o alinhamento usando uma combinação de sequência e informações da estrutura secundária da proteína *template*. A estrutura é construída levando em consideração solvente, ligante, campo

de força e outras contribuições por meio de uma série de algoritmos. São usadas bibliotecas de rotâmeros de cadeia lateral e diédricas de cadeia principal, derivadas de estruturas do PDB (*Protein Data Bank*)[28] para a construção de cadeia principal e cadeias laterais. Partes da sequência que não se alinham ao modelo, como *loops*, são construídas usando um procedimento *ab initio* que incorpora solvatação.[29]

Uma vez construído o modelo, são realizados os cálculos de interação entre proteína e ligante, o *docking* molecular propriamente dito. Nos cálculos de *docking* molecular, um algoritmo gera uma série de poses para os ligantes e atribui a elas uma pontuação de acordo com as interações que o ligante nessa pose apresenta com a proteína[30]; semelhante ao problema "chave-fechadura", na qual o cálculo de *docking* procura a posição correta para a "chave". Entretanto, devido à flexibilidade tanto da proteína quanto do ligante, alguns autores consideram mais apropriado a analogia "mão-luva".[31]

1.4.2 *Docking*

O *docking* molecular é uma ferramenta importante que faz parte de uma estratégia utilizada na descoberta de substâncias bioativas denominada *structure-based drug design* (SBDD), isto é, se baseia em uma estrutura 3D da proteína alvo, que pode ser obtida tanto através de técnicas de cristalografia de Raios-X e RMN, como através da modelagem por homologia. Trata-se do estudo do comportamento de pequenas moléculas dentro do sítio ativo de proteínas. O método consiste em determinar a melhor posição entre ligante e proteína e prever a afinidade entre ambos.[32] Essa afinidade é calculada usando diferentes parâmetros dependendo do programa utilizado. Neste trabalho foram explorados vários programas com diferentes pontuações, de forma a obter uma diversidade maior de poses a serem analisadas.[33]

No programa Glide,[34] a forma e as propriedades do receptor são representadas em um *grid* e usando diversos conjuntos de campos, que por sua vez fornecem uma pontuação progressivamente mais acurada da pose do ligante, conforme ela é gerada. Esses campos são calculados previamente, logo necessitam ser calculados apenas uma vez para cada receptor. A partir de um conjunto de conformações do ligante inicial, as quais são selecionadas a partir de uma enumeração exaustiva dos mínimos no ligante, são criadas poses em toda a área da caixa pré-definida . Essa pré-seleção reduz drasticamente a região do espaço sobre o qual são realizadas avaliações de energia computacionalmente caras e também evita o uso de métodos estocásticos.[34]

No caso do Autodock[35] e do Gold,[36] ambos utilizam o algoritmo genético para gerar as poses. Cada grau de liberdade do ligante é um gene e o mesmo pode sofrer mutação ou *crossing*

over. No caso da mutação, uma alteração aleatória é gerada e no *crossing over* ocorre uma troca de genes. Essas poses geradas são ranqueadas de acordo com a pontuação de cada programa. Apesar de a mutação ser um método estocástico, com a realização de vários ciclos de cálculos feitos de forma independente, o resultado será estatisticamente o mesmo.

Essas pontuações, tanto no caso do Gold e Autodock, quanto no caso do Glide, não se correlacionam diretamente à força de ligação ou ΔG de interação proteína-ligante. Uma vez que cálculos de *docking* molecular são utilizados para escanear vastas bibliotecas de potenciais ligantes, e para tal, um baixo custo computacional é necessário, uma série de aproximações são feitas. Para uma melhor predição de energias de interação, são utilizados cálculos mais acurados e dispendiosos, como o MM-GBSA.[37]

1.4.3 MM-GBSA

MM-GBSA é uma sigla vem de *Molecular Mechanics-Generalized Born Surface Area*, técnica em que é calculado o ΔG de interação entre pequenas moléculas e proteínas através de uma combinação entre energia da fase gasosa (MM – *molecular mechanics*), energia de solvatação eletrostática (GB – *generalised Born*) e contribuição da energia de solvatação não-eletrostática (SA – *surface area*).[38] Essa energia pode ser calculada em figuras estáticas provenientes de estruturas criptográficas, tanto em poses de *docking* , como também a partir da média das poses geradas por uma dinâmica molecular.

O MM-GBSA possui maior nível de complexidade e custo computacional do que a pontuação gerada pelo *docking* molecular, sendo mais eficaz quando aplicado após uma dinâmica molecular. Além disso, existem ainda técnicas de melhor correlação com dados experimentais, por exemplo o FEP (*Free Energy Perturbation*),[39] que calcula a diferença de energia livre entre estados. Desse modo, com o FEP é possível traçar um mapa de energia livre ao longo do tempo da interação proteína-ligante, levando-se em conta efeitos como solvatação e flexibilidade. Porém, como a abordagem do FEP possui um alto custo computacional, o método MM-GBSA torna-se mais conveniente, pois permite uma rápida e acurada estimativa da energia de ligação. Um estudo que comparou FEP e MM-GBSA com dados experimentais para uma série de inibidores de PLK1 mostrou que enquanto o FEP apresentou um $r^2 = 0{,}854$, o MM-GBSA teve seu $r^2 = 0{,}767$, enquanto requer apenas um oitavo do tempo das simulações de FEP.[40]

O cálculo de MM-GBSA é indicado para filtrar modos de ligação irrelevantes, facilitar a escolha de poses e reduzir o número de cálculos de maior acurácia e custo computacional,

assim como uma melhor seleção de ligantes que podem ser levados para uma fase de estudo *in vivo*.[41] Neste trabalho, as melhores poses geradas pelo *docking* molecular foram definidas utilizando-se o MM-GBSA, uma vez que não seria possível comparar as pontuações de *docking* entre os programas Autodock, Gold e Glide. Essas poses foram também o ponto de partida para os cálculos de dinâmica molecular, a fim de avaliar como essa interação se comporta ao longo do tempo.

1.4.4 Dinâmica molecular

Cálculos de dinâmica molecular (MD – *Molecular Dynamics*) vêm emergindo como uma técnica eficaz no entendimento de sistemas biológicos, providenciando melhor compreensão do comportamento dinâmico de sistemas macromoleculares usando conhecimentos de nível atômico.[42] Trata-se de uma ferramenta muito poderosa, uma vez que permite observar como grandes sistemas se comportam ao longo do tempo, conjuntamente com o fato de que é possível controlar uma série de variáveis, como a exata posição da proteína, concentração de eletrólitos, posição do ligante, solvente e tempo de duração.

Cálculos de MD utilizam o método numérico baseado na equação de Newton do movimento de um sistema contendo N-corpos, que atualiza iterativamente as forças das partículas e as energias potenciais do sistema compostas por potenciais inter e intra-atômicos ou campos de força. Durante as simulações de MD, coordenadas, velocidades, forças e energias potenciais são calculadas a cada passo, cujo intervalo temporal é denominado “passo de tempo”. Quanto menor o intervalo de tempo, mais acurada a simulação se torna, porém com maior custo computacional.[43] No presente estudo, o cálculo de dinâmica molecular será utilizado para avaliar a interação proteína-ligante ao longo do tempo, utilizando como ponto de partida as poses advindas dos cálculos de *docking* e que apresentaram valor ΔG teórico mais negativo (também nos cálculos de *docking*).

1.4.5 MM-PBSA

O MM-PBSA se trata de uma técnica muito semelhante ao MM-GBSA, descrito anteriormente. A sigla MM-PBSA vem de *Molecular Mechanics Poisson-Boltzmann Surface Area* e a principal diferença entre as técnicas é que no caso do MM-PBSA a energia de solvatação eletrostática é calculada utilizando-se a aproximação de Poisson-Boltzmann[44], enquanto que no caso do MM-GBSA a aproximação utilizada é *Generalized Born.*[45] Apesar de o MM-PBSA apresentar algumas limitações no cálculo de entropia, há uma boa integração com o software usado para calcular as dinâmicas moleculares, o GROMACS 2019.[46]

1.4.6 MIA-QSAR

O MIA-QSAR (*Multivariate Image Analysis* aplicada ao QSAR) é uma técnica de análise multivariada de imagem aplicada à relação quantitativa estrutura-atividade na qual os descritores são pixels de imagens bidimensionais de estruturas químicas 2D, que se correlacionam à propriedades bioativas. A aplicação dessa técnica tem como resultado ferramentas gráficas que usam coeficientes de regressão PLS (do inglês – *Partial Least Squares*) e as pontuações de importância variável na projeção (VIP – *Variable Importance in Projection*) para fornecer informações sobre a mudanças estruturais responsáveis por atividades biológicas aumentadas ou atenuadas em uma série de compostos.[47] Foi realizado estudo com o MIA-QSAR para propor melhorias sintéticas e entender melhor as interações entre a proteína e o ligante, com o objetivo de aumentar o percentual de inibição.

1.4.7 Permeabilidade cutânea teórica

Pensando em aplicações cosméticas, a absorção de ativos através da pele é um fator de crucial importância. Por isso, e a fim de enriquecer o presente trabalho, foram realizados estudos teóricos de permeação cutânea utilizando-se o software *web-based* SWISS-ADME.[48]

2 Objetivos

2.1 Objetivo geral

Analisar a interação proteína-ligante entre a tirosinase humana e os componentes dos extratos em metanol, acetona e etanol das folhas de aroeira vermelha (*Schinus terebinthifolius*) usando as técnicas de *docking* molecular, MM-GBSA, MIA-QSAR e dinâmica molecular. Busca-se, principalmente, avaliar o potencial de inibição dos componentes dos extratos.

2.2 Objetivos específicos

2.2.1 Modelagem por homologia

Levando-se em conta que a estrutura cristalográfica da tirosinase humana não está disponível no PDB (*Protein Data Bank*), tornou-se necessário construir um modelo por homologia. Nesse sentido, foram construídos modelos em três programas diferentes: SWISS-MODEL, Modeller e Prime, usando como *template* a proteína relacionada à tirosinase humana (Figura 4, TYRP1 - PDB 5M8L.A).[49]

Figura 4: Estrutura da proteína relacionada à tirosinase humana (TYRP1 - PDB 5M8L.A)

2.2.2 Validação estrutural do modelo

Uma vez construídos os modelos da tirosinase humana, suas estruturas foram validadas através de softwares *web based* como o PROCHECK [50] e o Verify 3D,[51] de forma a escolher o melhor modelo.

2.2.3 Cálculos de *docking* molecular

Realizaram-se cálculos de *docking* molecular entre o melhor modelo escolhido e os compostos encontrados nos extratos de aroeira vermelha (*Schinus terebinthifolius*) usando os softwares GOLD, Autodock Vina e Glide.

2.2.4 Cálculos de Δ*G* (MM-GBSA)

Foi calculado o Δ*G* teórico utilizando o programa Prime a fim de comparar as poses geradas pelos diferentes programas.

2.2.5 Dinâmica molecular

A partir das poses de *docking* molecular com Δ*G* MM-GBSA mais negativo, foram realizados cálculos de dinâmica molecular buscando avaliar a interação ao longo do tempo do sistema ligante-enzima-solvente.

2.2.6 MM-PBSA

As trajetórias de dinâmica molecular foram submetidas ao cálculo de MM-PBSA a fim de avaliar a força de interação entre ligante e proteína.

2.2.7 MIA-QSAR

Foram propostas alterações nos grupos ligantes, e os novos ligantes submetidos a um estudo de correlação imagem-propriedade, a fim de propor melhorias sintéticas e melhor compreender as interações proteína-ligante. Para tanto, a propriedade avaliada foi o Δ*G* MM-GBSA.

2.2.8 Permeabilidade cutânea teórica

Com o auxílio do software Swiss-ADME, foram avaliados os valores de K_{ps} dos componentes do extrato de *Schinus Terebinthifolius*, os quais foram comparados ao composto referência da literatura: o ácido kójico.

3 Metodologia

3.1 Modelagem por homologia

Uma vez que a estrutura cristalográfica da tirosinase humana não se encontra disponível no *Protein Data Bank*, tornou-se necessário construir um modelo por homologia, usando como *template* a proteína com maior índice de identidade e resolução.[52] Para isso, foram construídos modelos utilizando-se os programas SWISS-MODEL, Schroedinger 2020 e Modeller. A sequência utilizada no formato FASTA foi obtida na base UniProt[53] (Código L8B082) e a partir disso foi realizada uma busca usando a plataforma BLAST,[54] que analisa, em toda a base do PDB, as proteínas com sequência semelhante, ou seja, com maior identidade. Por fim, a proteína *template* (modelo) escolhida foi a proteína relacionada à tirosinase humana (TYRP1 - PDB 5M8L.A) por ser a de maior homologia com a proteína de interesse (ID 43.32%) e de maior resolução (2.35 Å).

Em todos os softwares utilizados, após gerado o modelo, foi necessário realizar uma alteração estrutural, pois o metal de coordenação do *template* é o zinco, enquanto que o da proteína de estudo é o cobre. Como método de seleção do modelo, usou-se como base os resultados da etapa de verificação da estrutura; além disso, vale mencionar que o protocolo de geração de modelo em cada um dos programas é ligeiramente diferente, sendo mais automatizado em alguns casos do que em outros. Contudo, as etapas gerais foram as seguintes: o alinhamento entre a estrutura *template* e a sequência da proteína que se deseja modelar e, em seguida, a construção do modelo em si.

3.1.1 Modeller[25]

O Modeller é, dentre os softwares analisados, o menos automatizado. Com ele, foi realizado o alinhamento da sequência alvo e *template* utilizando-se o software *web-based* Clustal W.[55] Nessa etapa, foram gerados 100 modelos com sua pontuação, sendo o de maior pontuação escolhido para as etapas seguintes. A fim de refinar o modelo, foi aplicado o *script* fornecido pelo próprio software para correção de enovelamento: loop (100x) do aminoácido 453 ao 529. A Figura 5 mostra a diferença que o *script* fez no enovelamento (apesar de uma eficácia limite, não gerando mais melhorias significativas acima de 100x).

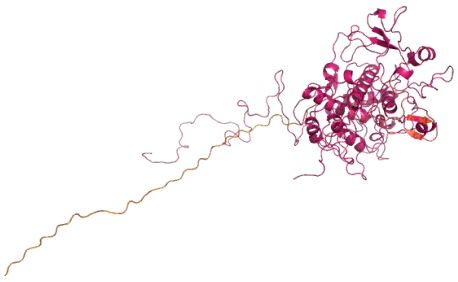


Figura 5: Modelo da proteína tirosinase construído no MODELLER antes (laranja) e depois (rosa) do *script* de enovelamento.

3.1.2 SWISS-MODEL[26]

A plataforma SWISS-MODEL viabiliza todo o processo de alinhamento e geração de modelo automatizado, sendo necessário apenas fornecer a sequência de aminoácidos. A partir disso, a plataforma sugeriu a proteína *template* utilizada neste trabalho e a partir da qual se obteve um modelo; ademais, também foram sugeridos modelos já construídos anteriormente, disponíveis em seu repositório[56] (P14679 e L8B082), os quais foram baixados e submetidos à mesma análise que os demais.

3.1.3 Prime

Para a modelagem no programa maestro foi utilizado o *Prime - Structure Prediction Wizard*, que disponibiliza dois algoritmos de alinhamento: o Clustal W e o Prime STA. Entre eles, o primeiro foi o escolhido, uma vez que gerou melhores resultados de identidade (41%), enquanto o segundo gerou apenas 28%. Na etapa seguinte foram gerados dois modelos incluindo apenas os metais da proteína *template*, utilizando dois métodos disponíveis no software: um que usa como base segmentos de estruturas conhecidas para a construção das lacunas ou gaps do *template* (*knowledge-based*), e outro que usa como base a energia para o preenchimento das mesmas (*energy-based*).

3.1.4 Protein preparation wizard[57]

As estruturas geradas foram otimizadas utilizando o programa *Protein Preparation Wizard* do pacote Schroedinger 2020-3, como qual é possível corrigir pequenos erros na estrutura, como hidrogênios faltantes, cadeias laterais e loops incompletos, estados de protonação ambíguos e resíduos invertidos. As correções foram feitas em três etapas, sendo a primeira um pré-processamento, no qual foram atribuídas ordens de ligação, adicionados hidrogênios, definida como zero a ordem de ligação de metais e protonação de acordo com *p*H

fisiológico (7±2). A etapa seguinte consistiu em otimizar a rede de ligações de hidrogênio, sendo escolhida a opção de otimização padrão. Já na terceira e última etapa, a proteína foi minimizada utilizando um campo de força, no caso o OPLS3e.[58]

3.2 Verificação da estrutura

A fim de definir qual modelo seria usado nas etapas seguintes, foram realizados os seguintes testes de verificação de estrutura.

3.2.1 Gráfico de Ramachandran gerado no software PROCHECK [50]

O software PROCHECK verifica a qualidade estereoquímica de uma estrutura de proteína através de análise da geometria resíduo a resíduo e da geometria geral da estrutura. Desse modo, com esse servidor foi possível observar o gráfico de Ramachandran para cada um dos modelos construídos. Os dados coletados para análise foram as porcentagens de resíduos em regiões não permitidas e em regiões mais favoráveis.

3.2.2 Verify 3D [51]

A plataforma Verify 3D determina a compatibilidade de um modelo atômico (3D) com sua própria sequência de aminoácidos (1D) atribuindo uma classe estrutural com base em sua localização e ambiente (alfa, beta, loop, polar, apolar etc.) e comparando os resultados com boas estruturas. O requisito mínimo é 80% os aminoácidos em regiões permitidas.Foi utilizado como ferramenta de avaliação da qualidade estrutural do modelo de homologia gerado.

3.3 Validação do modelo

Os testes estruturais não foram suficientes para a escolha de um único modelo. Por isso, a fim de verificar se o modelo gerado se comportava de forma similar à proteína *template*, foram realizados cálculos semelhantes a técnica de *re-docking*.[59]

3.3.1 *Re-docking* e *Root Mean Square Deviation* (RMSD)

Foi realizada comparação entre: a estrutura do ligante co-cristalizado com a proteína *template* que são alguns inibidores já conhecidos (PDB 5M8M - ácido kójico; PBD 5M8T – tropolona; PDB 5M8N - mimosina) e a pose gerada através do *docking* molecular com os modelos construídos, com esses mesmos inibidores. Para facilitar a análise calculou-se, com o software Schroedinger 2020-3, o RMSD do ligante, usando como referência o ligante co-cristalizado.

No PDB tem-se a mesma proteína *template* co-cristalizada com três diferentes ligantes: o ácido kójico, a tropolona e a mimosina. Desse modo, foram geradas 10 poses de *docking* para

cada um dos ligantes com os modelos do SWISS-MODELLER e do Maestro provenientes da etapa anterior. Para definição do grid, as proteínas referência do PDB foram alinhadas com os modelos, e a coordenada do sítio ativo foi definida através dos ligantes co-cristalizados, sendo elas: ácido kójico (X= -31.14 Y=-4.37 Z= -25.21), mimosina (X = -30.45 Y= -5.07 Z= -24.8) e tropolona (X=-31.64 Y= -3.07 Z= -25.2).

Durante a análise de sítio ligante, verificou-se a importância de uma molécula de água coordenada com os metais no caso do ácido kójico (5M8M) e da mimosina (5M8R). Já, no caso da tropolona (5M8O), a água não apresentava relevância na coordenação do ligante aos átomos de Cu^{2+} (como pode ser observado na Figura 6). Essa molécula de água foi adicionada aos modelos antes de ser realizado o *docking* molecular, quando necessário, a fim de obter resultados de maior acurácia possível. Feito isso, calculou-se o RMSD entre ligante co-cristalizado e pose de *docking* molecular para as 10 poses. A fim de analisar a correlação entre o desvio das poses de *docking* para a posição do ligante co-cristalizado (RMSD) e a força de ligação, foi plotado um gráfico com RMSD *vs* ΔG para cada um dos modelos e ligantes. A partir do gráfico, foi traçada uma linha de tendência e seu R^2.

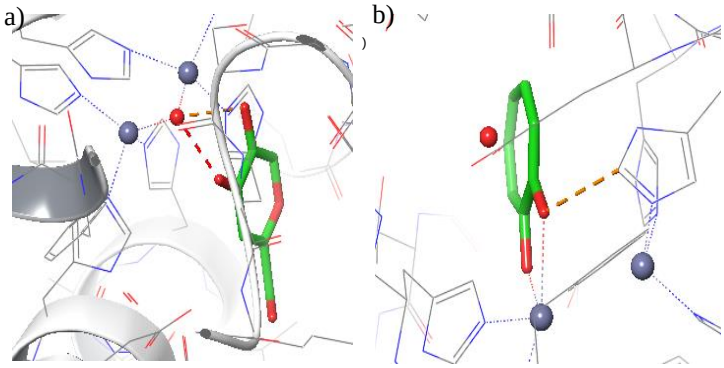


Figura 6: Detalhe do sítio ativo da proteína *template* co-cristalizada com a) ácido kójico, onde a molécula de água desempenha um papel de ponte entre ligante e metais e b) Tropolona, onde a mesma molécula de água se mostra irrelevante. Legenda: fitas brancas – proteínas, tubos verdes – átomos de carbono, tubos vermelhos – átomos de oxigênio, tubos brancos – átomos de hidrogênio, esferas azuis – Zn^{2+}.

3.3.2 Cálculo de ΔG (MM-GBSA)

Posteriormente foi realizado o cálculo de ΔG (MM-GBSA) a fim de analisar a correlação entre o desvio das poses de *docking* para a posição do ligante co-cristalizado (RMSD) e a força de ligação para cada um dos ligantes considerados: mimosina, tropolona e ácido kójico. Com esses resultados foi definido o melhor modelo que seguiria para as etapas seguintes e com o qual seria realizado o *docking* dos ligantes que esse trabalho se propõe analisar.

3.4 *Docking* Molecular

3.4.1 Preparação dos ligantes

Todos os ligantes foram construídos no software ChemDraw[60] e suas estruturas processadas pelo software EpiK.[61] Uma vez definido o melhor modelo, foi realizado o *docking* do mesmo com os ligantes da *Schinus terebinthifolius*, gerando dez poses em cada um dos três programas - Glide, Gold e Autodock Vina - usando como base o modelo com e sem a água coordenada aos íons cobre. Devido ao fato de as pontuações serem atribuídas de forma diferentes em cada um desses programas, temos uma maior diversidade de poses, o que também aumenta as chances de obter resultados experimentalmente coerentes.

3.4.2 *Docking* molecular

Foram realizados cálculos de *docking* em três programas: Glide, Gold e Autodock Vina. Para todos os cálculos de *docking* molecular, o grid gerado foi uma caixa de 20 Å centralizada no ligante ácido kójico, estando co-cristalizado à proteína *template*, alinhada ao modelo. Sendo suas coordenadas na proteína: X = - 31; Y = -4, Z = -25. A proteína *template* possuía em sua estrutura uma molécula de água coordenada aos dois átomos de Cu^{2+}. Como a molécula de água foi identificada como importante na coordenação de ligantes dessa enzima durante a análise de *re-docking*,foi realizado o *docking* dos componentes do extrato de *Schinus Terebinthifolius* com e sem a presença dessa molécula de água.

3.4.2.1 Glide

Inicialmente um grid de *docking* molecular foi gerado, centralizando-o no ligante ácido kójico co-cristalizado com a proteína *template* (PDB - 5M8M), e definido um raio de 20 Ă. A partir desse grid foi definido o sítio ativo no cálculo de *docking* dos demais programas. Todos os cálculos foram feitos utilizando protocolo de *flexible docking* e pontuação *extra precision* (Glide XP).

3.4.2.2 Gold

Foram realizados os cálculos de *docking* molecular no programa Gold, utilizando como sítio ativo a mesma região escolhida para o Glide. A função de pontuação escolhida foi o CHEM PLP, especialmente porque vem apresentado melhor performance em estudos recentes.[62] A configuração padrão de pesquisa, o algoritmo genético, foi utilizada.

3.4.2.3 Autodock

Os arquivos de coordenadas de proteínas e ligantes foram convertidos em um formato PDB estendido denominado pdbqt. A grade foi definida através das coordenadas obtidas na

etapa de *docking* molecular com Glide. Arquivo de parâmetro de grade (gpf) e arquivo de parâmetro de encaixe (dpf) foram gerados. O algoritmo genético Lamarckiano foi usado como mecanismo de busca, e foram obtidas 10 poses por ligante.

3.4.3 MM-GBSA

Com o objetivo de comparar as poses geradas nos diferentes programas, foi executado o cálculo de Δ*G* (MM-GBSA). Nesse processo, contudo, observou-se certa dificuldade em transferir as poses geradas pelo Autodock Vina para o Schroedinger 2020-3. Isso ocorreu porque as poses são geradas em formato pdbqt pelo Autodock Viana, não lido pelo Maestro. Por conta disso, as imagens foram convertidas para arquivos mol2 utilizando-se o software Open Babel[63]. Em seguida a proteína foi preparada no *Preparation Wizard* do pacote Schroedinger 2020-3.[57]

3.4.4 Dinâmica Molecular

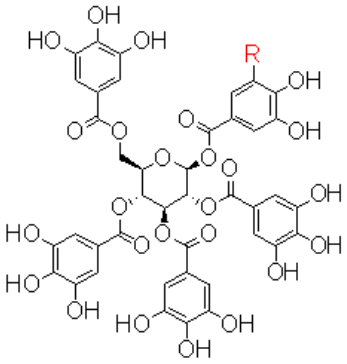
Na etapa seguinte, os ligantes **1-22** foram submetidos a simulações de dinâmica molecular utilizando-se o software Gromacs 2019[64] e aplicando-se o campo de força OPLS-AA[65], de forma a prever o comportamento dos ligantes no sítio ativo da enzima ao longo do tempo. As poses com energia de ligação mais negativa (Δ*G* MM-GBSA) calculadas pelo Prime foram parametrizadas usando ACPYPE[66] e MKTOP.[67] Na sequência, as estruturas foram inseridas em uma caixa cúbica de solvente para simular o ambiente fisiológico, enquanto íons Na^+ foram adicionados para neutralizar a carga do sistema. O sistema foi submetido, respectivamente, a: (a) minimização *steepest descent* com restrição de posição por 20.000 passos; (b) minimização *steepest descent* sem restrição de posição por 20.000 passos; (c) minimização do gradiente conjugado com restrição de posição para 20.000 passos; (d) realização de 500ps de dinâmica molecular em condições isotérmicas-isobáricas (NPT) com restrição de posição, 310 K e 1 bar; (e) Simulações de 100 ns em condições NPT, com 310 K e 1 bar, utilizando o termostato Berendsen *speed-rescaling*[68] com período de 0,1 ps o baróstato [68,69] para 1 ps e registrado a cada 50 ps, totalizando 2000 frames.

3.4.5 MM-PBSA

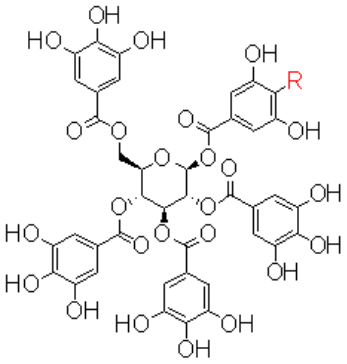
As trajetórias de dinâmica molecular (100 ns) dos ligantes **1-22** foram analisadas utilizando o *script* MM-PBSA.py[70], o qual retira as moléculas de solvente e sal de cada *frame* da dinâmica e calcula uma estimativa de energia de ligação. Esse procedimento acontece através de uma combinação entre energia da fase gasosa (MM), energia de solvatação eletrostática (Poisson–Boltzmann) e contribuição da energia de solvatação não-eletrostática (SA).[71]

3.4.6 MIA-QSAR

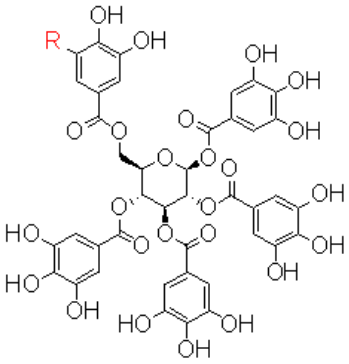
Para o modelo QSAR, foram propostas modificações para os ligantes de energia de ligação (Δ*G* MM-GBSA) mais negativa: 1,2,3,4,6-penta-O-galloil-glicose **22**, ácido gálico **15** e ácido cafeico **10**. As modificações foram realizadas substituindo-se o grupo hidroxila no anel aromático por F, Cl, Br, I, S-metil e para os grupos metila e etila. A partir disso, as novas estruturas análogas (Figura 7, **23-78**) foram submetidas ao mesmo procedimento de *docking* molecular descrito anteriormente. O modelo QSAR foi então construído utilizando-se descritores MIA como variáveis preditoras (x) e o Δ*G* MM-GBSA calculado por Prime como variáveis respostas (y). Todo o procedimento de modelagem MIA-QSAR foi realizado usando-se o programa Chemoface.[72]



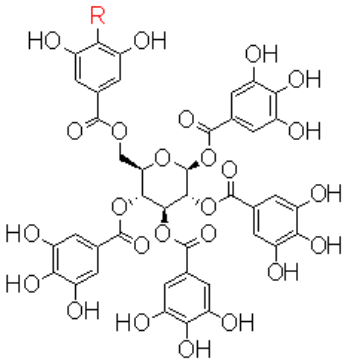
23 R = F
24 R = Cl
25 R = Br
26 R = I
27 R = CH_3
28 R = S-CH_3
29 R = $CHCH_2$



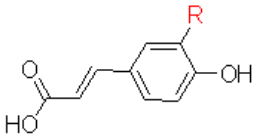
30 R = F
31 R = Cl
32 R = Br
33 R = I
34 R = CH_3
35 R = S-CH_3
36 R = $CHCH_2$



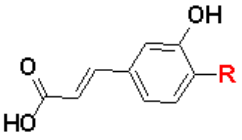
37 R = F
38 R = Cl
39 R = Br
40 R = I
41 R = CH_3
42 R = S-CH_3
43 R = $CHCH_2$



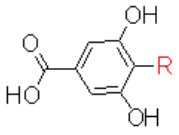
44 R = F
45 R = Cl
46 R = Br
47 R = I
48 R = CH_3
49 R = S-CH_3
50 R = $CHCH_2$



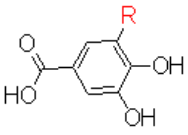
51 R = F
52 R = Cl
53 R = Br
54 R = I
55 R = CH_3
56 R = S-CH_3
57 R = $CHCH_2$



58 R = F
59 R = Cl
60 R = Br
61 R = I
62 R = CH_3
63 R = S-CH_3
64 R = $CHCH_2$



65 R = F
66 R = Cl
67 R = Br
68 R = I
69 R = CH_3
70 R = S-CH_3
71 R = $CHCH_2$



72 R = F
73 R = Cl
74 R = Br
75 R = I
76 R = CH_3
77 R = S-CH_3
78 R = $CHCH_2$

Figura 7: Estrutura dos ligantes modificados, utilizados no modelo de MIA-QSAR

Primeiramente, vale destacar que as estruturas foram desenhadas no software GaussView 5.0,[73] de forma a ficarem sobrepostas (Figura 8), e a partir do que obteve-se um *snapshot* de cada figura. Em seguida, essas figuras foram convertidas em matrizes a partir dos pixels de cada imagem, onde cada cor apresenta um valor específico. Em seguida, foi feito um tratamento de dados no qual partes irrelevantes da figura são removidas. O pixel de fundo tem seu valor substituído por 0, e o valor de cada cor é substituído pelo valor de uma propriedade específica que se deseja correlacionar com a variável de resposta. No caso desse estudo, cada pixel (descritor MIA) foi numerado de acordo com a eletronegatividade atômica de Pauling (ε), raios de van der Waals (rvdW) e a razão entre esses dois parâmetros (rvdW/ε); já como variável de resposta foram utilizados os valores ΔG MM-GBSA. O modelo mais confiável foi obtido usando a eletronegatividade atômica de Pauling (ε), o que implica que interações eletrostáticas-dependentes podem ser mais relevantes para as energias de interação do que interações relacionadas ao tamanho atômico.

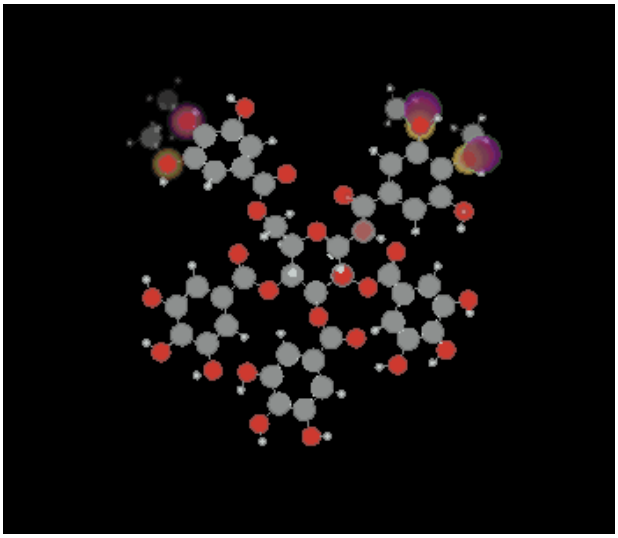


Figura 8: Estruturas do GaussView dos ligantes utilizados no modelo sobrepostas entre si.

3.4.7 Permeabilidade cutânea teórica

O software SwissADME[74] do Instituto Suíço de Bioinformática foi acessado em um servidor web, a partir do que as estruturas 2D dos compostos **1-22** foram desenhadas no próprio servidor da Chem Axons Marvin JS. Em seguida, as estruturas foram transferidas como uma lista de códigos SMILES para o lado direito da página de submissão, que é a entrada para cálculo, e o arquivo de saída SwissADME, por sua vez, foi composto por um painel por molécula que contém suas propriedades calculadas, como o LogKp.[75]

4 Resultados

4.1 Modelagem por homologia

As Figura 18- 31 referentes a cada um dos modelos construídos por homologia se encontram no Anexo 2.

4.2 Verificação da estrutura

A Tabela 1 mostra os resultados obtidos pelos modelos gerados nos testes estruturais propostos. Usando como parâmetro de corte a proteína *template* (5M8L_A) minimizada no *Preparation Wizard*, é possível observar uma melhor performance dos modelos construídos usando o SWISS-MODELLER e Schrödinger 2020-3, principalmente de acordo com os valores do Verify 3D, ainda que não tenha havido uma diferença significativa entre eles. Assim, foi proposta mais uma etapa de validação do modelo, através do *re-docking* e cálculos de RMSD com relação ao ligante do *template*. Os modelos do MODELLER, nesse caso, foram excluídos da etapa seguinte, já que apresentaram porcentagem menor ou muito próxima a 80%.

Tabela 1: Resultados dos testes de validação estrutural dos modelos de homologia construídos.

Modelo	Descrição	Verify 3D	Gráfico de Ramachandran	
			Região mais favorável	**Região proibida**
5M8L_A	*Template*	94.39%	86.50%	0.30%
5M8L_A *Preparation Wizard*	***Template* minimizado no PrepWizard**	**94.39%**	**86.50%**	**0.30%**
Modeller	Melhor modelo de 100 construídos	80.53%	88.50%	0.00%
Modeller *Preparation Wizard*	Melhor modelo de 100 construídos e minimizado no PrepWizard	76.75%	87.60%	0.00%
Modeller L- 100	Melhor modelo de 100 construídos, aplicado o *script* loop 100x	81.66%	84.80%	0.20%
Modeller L-100 *Preparation Wizard*	Melhor modelo de 100 construídos, aplicado o *script* loop 100x e minimizado no *PrepWizard*	78.07%	84.10%	0.20%
Swiss P14679	Modelo do repositório do Swiss-Model	94.71%	86.90%	1.30%
Swiss P14679 *Preparation Wizard*	Modelo do repositório do Swiss-Model e minimizado no *PrepWizard*	95.86%	83.90%	1.10%

Swiss L8B082	Modelo do repositório do Swiss-Model	94.71%	86.90%	1.30%
Swiss L8B082 *Preparation Wizard*	Modelo do repositório do Swiss-Model e minimizado no *PrepWizard*	95.86%	83.90%	1.10%
Swiss	Modelo construído no Swiss-Model	94.71%	86.90%	1.30%
Swiss *Preparation Wizard*	Modelo construído no Swiss-Model e minimizado no *PrepWizard*	95.86%	83.90%	1.10%
Maestro *energy-based*	Modelo construído no Maestro com o método *energy-based*	97.71%	86.10%	0.50%
Maestro *energy-based Preparation Wizard*	Modelo construído no Maestro com o método *energy-based* e minimizado no *PrepWizard*	97.71%	86.10%	0.50%
Maestro *knowledge-based*	Modelo construído no Maestro com o método *knowledge-based*	97.71%	83.20%	1.30%
Maestro *knowledge-based Preparation Wizard*	Modelo construído no Maestro com o método *knowledge-based* e minimizado no *PrepWizard*	95.88%	84.00%	1.30%

4.3 *RMSD* e *re-docking*

A fim de analisar a correlação entre o desvio das poses de *docking* molecular para a posição do ligante co-cristalizado (*RMSD*) e a força de ligação, foi plotado um gráfico com *RMSD vs* ΔG MM-GBSA para cada um dos modelos e ligantes. Em seguida, foi então traçada uma linha de tendência para o gráfico de *RMSD vs* ΔG MM-GBSA (valor de R^2 obtido na Tabela 2). Espera-se que quanto menor o desvio da pose com relação ao ligante co-cristalizado, mais negativo será o ΔG; dessa maneira, o modelo que apresenta maior R^2 e maior correlação entre as duas variáveis estaria mais próximo ao comportamento real. Além disso, constam também na Tabela 2 os valores médios de *RMSD* obtidos para as dez poses geradas para os cálculos de *docking* molecular.

Com essas análises foi possível concluir a etapa de definição do modelo, uma vez que o que apresentou melhor correlação em conjunto com melhores parâmetros geométricos-estruturais, tendo sido selecionado o modelo construído no software maestro pelo modo *knowledge-based*, sem a minimização do *Preparation Wizard*. Entretanto, como o software Glide aceita somente modelos preparados no *Preparation Wizard* e uma vez que a diferença entre o minimizado e o não minimizado foi pequena, optou-se por utilizar o minimizado.

Tabela 2: Resultados de R^2 de cada um dos modelos e seus respectivos ligantes

Modelo	**Ligante**	***R*²**	***Média dos RMSD***	**Modelo**	**Ligante**	***R*²**	***Média dos RMSD***
Energy based	ácido kójico	0.2233	1.00	L8B082 minimizado	ácido kójico	0.0289	1.52
Energy based	mimosina	0.6388	2.35	L8B082 minimizado	mimosina	0.1912	2.19
Energy based	tropolona	0.0485	2.25	L8B082 minimizado	tropolona	0.2743	2.72
Energy based minimizado	ácido kójico	0.0746	1.30	P14679	ácido kójico	0.5165	1.54
Energy based minimizado	mimosina	0.6995	2.98	P14679	mimosina	0.0200	2.58
Energy based minimizado	tropolona	0.1214	2.19	P14679	tropolona	0.0154	2.55
Knowledge based minimizado	ácido kójico	0.279	0.99	P14679 minimizado	ácido kójico	0.175	1.10
Knowledge based minimizado	mimosina	0.5938	2.71	P14679 minimizado	mimosina	0.1912	2.19
Knowledge based minimizado	tropolona	0.1816	2.57	P14679 minimizado	tropolona	0.2743	2.72
Knowledge based	ácido kójico	0.1503	1.21	Swiss minimizado	ácido kójico	0.0286	1.52
Knowledge based	mimosina	0.6857	2.44	Swiss minimizado	mimosina	0.1915	2.18
Knowledge based	tropolona	0.1836	2.27	Swiss minimizado	tropolona	0.2742	2.71
L8B082	ácido kójico	0.5165	1.54	Swiss	ácido kójico	0.5006	1.54
L8B082	mimosina	0.0024	2.30	Swiss	mimosina	0.0203	2.57
L8B082	tropolona	0.0154	2.55	Swiss	tropolona	0.0155	2.55

O melhor modelo apresentou os seguintes resultados para sua análise estrutural no programa PROCHECK (Figura 9) e Verify 3D (Figura 10). É importante ressaltar, contudo, que apesar de a estrutura de tirosinase humana não ter sua estrutura cristalina resolvida disponível e esses dados não representarem a proximidade com a estrutura real do modelo construído, são dados que apontam a qualidade estrutural do modelo.

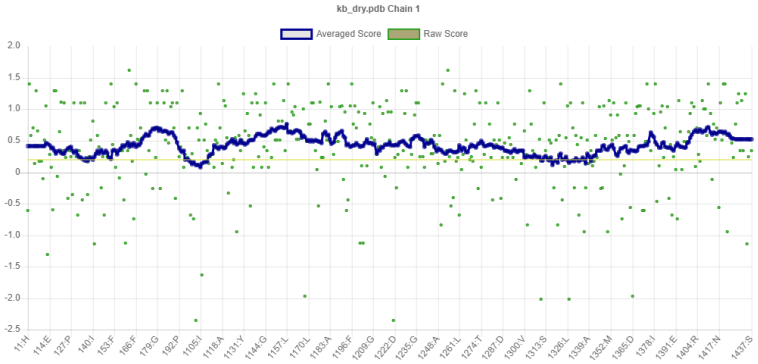


Figura 9: Gráfico do Verify 3D do modelo de homologia knowledge-based da tirosinase humana modelada usando 5M8L como modelo, onde 97,71% dos resíduos têm pontuação média 3D-1D >= 0,2

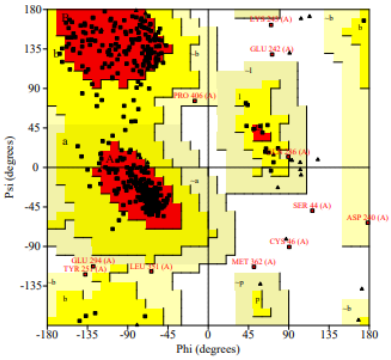


Figura 10: Gráfico de Ramachandran do modelo de homologia *knowledge based*. O percentual de resíduos na área mais favorável é de 83,20% e na área proibida é de 1,30%.

4.4 *Docking* molecular

Após definido o modelo de homologia a ser utilizado, foi dado o primeiro passo para a análise dos componentes de extrato de *Schinus Terebinthifolius*: o estudo de *docking* molecular. Para tanto, foram geradas 10 poses para cada um dos componentes **1-22**, com e sem água coordenada aos íons Cu^{2+}, em três programas (GOLD, Glide e Autodock Vina). Dessa forma, foi possível gerar poses usando diferentes algoritmos, aumentando as chances de se obter a mais verossímil.

Uma vez que que os parâmetros usados em cada um dos programas são diferentes, as poses geradas não podem ser comparadas usando apenas a pontuação de cada programa. Por isso foram realizados cálculos de ΔG MM-GBSA através do programa Prime do pacote Schroedinger 2020-3, com o objetivo de comparar as poses entre si. Os valores de ΔG MM-GBSA são apresentados na Tabela 3.

Tabela 3: Valor mais negativo de ΔG MM-GBSA, comparando os programas Glide, Autodock Vina e Gold para a proteína com e sem água coordenada aos átomos de Cu^{2+} dos ligantes 1-22.

Ligante	Δ*G* (kJ mol^{-1})	Programa	Proteína
Ácido kójico (1)	-13.77	Autodock	Com água
Metil galato (2)	-11.43	Autodock	Com água
Miricetina (3)	-33.50	Glide	Com água
Etil galato (4)	-11.28	Gold	Com água
Catequina (5)	-18.00	Glide	Com água
Quercetina ramnosídeo (6)	-31.16	Gold	Sem água
Miricetina ramnosídeo (7)	-27.54	Glide	Com água
Ácido elágico (8)	-26.06	Glide	Com água
Quercetina (9)	-31.29	Glide	Com água
Ácido cafeico (10)	-39.30	Gold	Sem água
Ácido siríngico (11)	-26.52	Glide	Sem água
Ácido coumárico (12)	-31.73	Gold	Sem água
1,6-Digaloil-ß-D-glucose (13)	-30.49	Glide	Com água
Ácido clorogênico (14)	-35.33	Gold	Sem água
Ácido gálico (15)	-39.65	Glide	Sem água
Ácido 5-p-coumaroilquínico (16)	-17.25	Gold	Com água
Luteolina (17)	-30.92	Glide	Sem água
Miricetina galactoside rhamnoside (18)	-35.02	Glide	Com água
Miricetina glucoronídeo (19)	-25.03	Gold	Com água
Miricetina galactosideo (20)	-34.78	Glide	Sem água
Robustaflavona (21)	-23.39	Glide	Com água

1,2,3,4,6-penta-O-galoil-glucose (22)	-61.11	Glide	Com água

O diagrama de interação proteína-ligante, como as poses 3D e Δ*G* das melhores poses de cada software e proteína, é mostrado no Anexo 3, Figuras 54 - 105. De forma geral, a maioria dos ligantes apresentou ligações de hidrogênio e interações π-π via anéis aromáticos. Apenas o ácido siríngico **11** (-26,52 kJ mol^{-1}), o ácido gálico **15** (-39,65 kJ mol^{-1}), o ácido coumárico **12** (-31,73 kJ mol^{-1}), o ácido clorogênico **14** (-35,33 kJ mol^{-1}) e o ácido cafeico **10** (-39,30 kJ mol^{-1}) apresentaram interações diretas com os íons cobre do sítio ativo responsáveis pela ação da enzima. Entretanto, para que ocorra a inibição, a molécula não precisa interagir com os íons Cu^{2+}, já que esta também pode acontecer bloqueando a entrada do sítio ativo.[76]

De acordo com a energia de ligação Δ*G* MM-GBSA, os melhores ligantes foram 1,2,3,4,6-penta-O-galoil-glucose **22** (-61,11 kJ mol^{-1}), ácido gálico **15** (-39,65 kJ mol^{-1}), e ácido cafeico **10** (-39,30 kJ mol^{-1}). Os diagramas de interação ligante-proteína (a, c, e) e figuras 3D do sítio ativo (b, c, f) são mostrados na Figura 11.

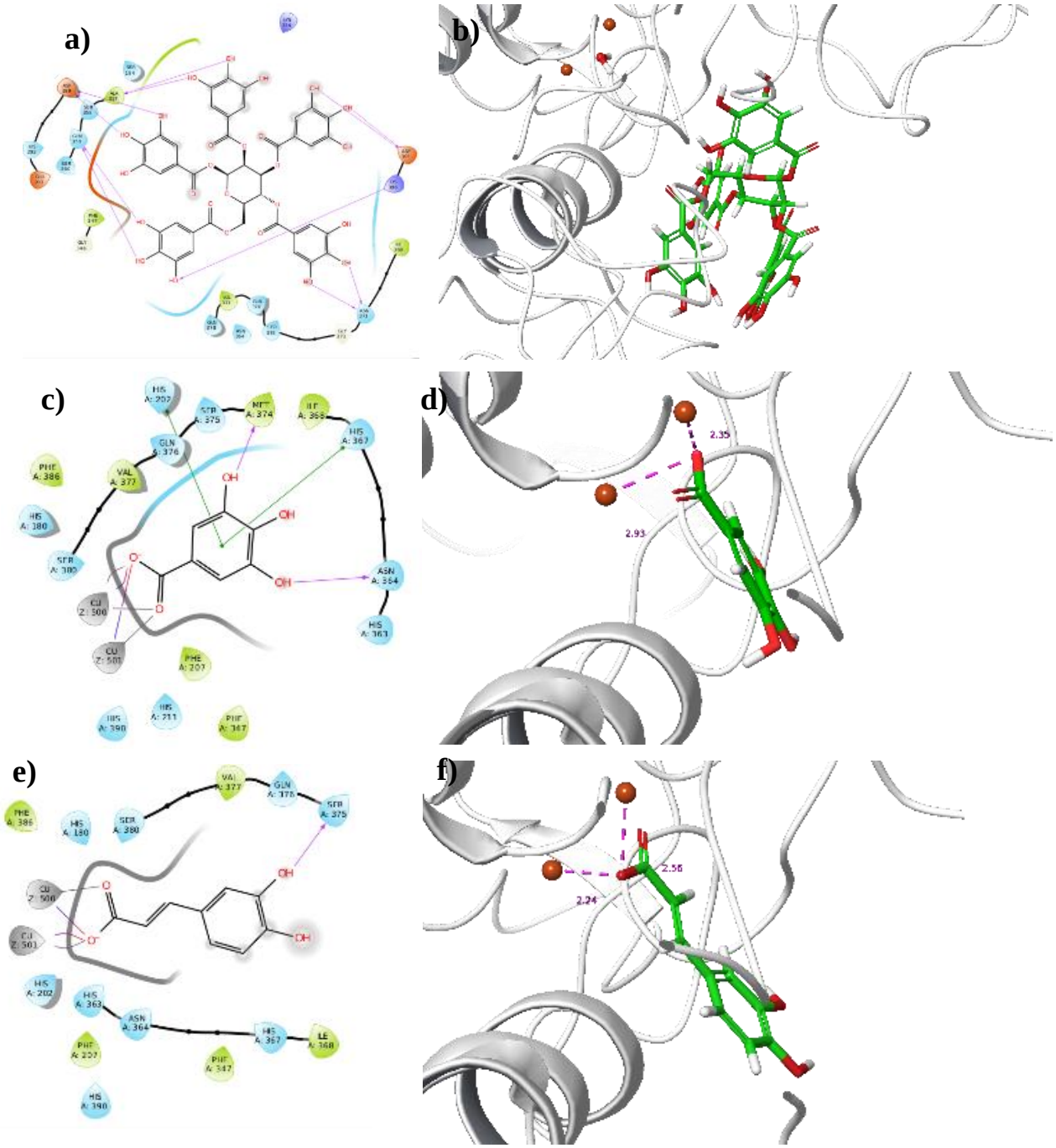


Figura 11: Diagrama de interação ligante-proteína (a, c, e) e vista 3D da pose de *docking* molecular Δ*G* MM-GBSA mais negativa (b, d, f) do 1,2,3,4,6-penta-O-galoil-glucose **22** (a, b), ácido gálico **15** (c, d) e ácido cafeico **10** (e, f). Legenda: Fitas brancas – proteína, tubos verdes – carbono, tubos vermelhos – oxigênio, tubos brancos – hidrogênio, esferas de cobre – Cu^{2+}, rosa – distância entre as interações ligante e cobre.

Os diagramas da Figura 11 indicam que **10** e **15** interagem com a enzima de forma análoga ao inibidor de referência, o ácido kójico, fazendo a coordenação oxigênio-cobre. Este resultado aponta para uma inibição competitiva, como o ácido kójico se ligar ao Cu^{2+} em vez de ao substrato, o que pode levar à oxidação dos ligantes. Na literatura, **10** é relatado experimentalmente como substrato,[77] enquanto **15** é relatado como inibidor da tirosinase de cogumelo (644,00±14,00 μg/mL).[78] Isso ocorre devido à atividade difenolase da oxi-tirosinase. Quando dois átomos de oxigênio se coordenam aos íons Cu^{2+}, os grupos hidroxila do anel de **10** se coordenam ao metal e sofrem oxidação para formar uma ortoquinona, agindo como um inibidor suicida.

O ligante **22** tem um modo de ligação diferente, pois não interage diretamente com os íons de cobre, mas seus múltiplos grupos hidroxila formam ligações de hidrogênio com ASP199, ALA357, ASP305, LYS306, ASN371 e GLN 359. Nesse sentido, foi relatada na literatura a atividade inibitória da monofenolase e da difenolase para o ligante **22** com uma IC50 42,65 ± 1,85 μg/mL.[78] Tal atividade inibitória está de acordo com os resultados de *docking* obtidos no presente trabalho.

4.5 Dinâmica molecular

Para entender melhor esses cálculos de dinâmica molecular, o RMSF (*Root Mean Square Fluctuation*) de *backbone*, proteína e cadeias laterais, o RMSD (C-alfa, ligante e sítio de ligação), o raio de giro da proteína e a distância entre o ligante e a proteína foram usados como parâmetros de qualidade. Os resultados completos estão no Anexo 3 Figura 54 - 105. A distância entre ligante e proteína permaneceu abaixo de 2,5 Ă para a maioria dos ligantes, exceto para **12**, **15**, **14** e **17**.

O RMSD (*Root Mean Square Deviation)*, desvio quadrático médio da raiz, é uma medida padrão da distância estrutural entre as coordenadas, de modo que mede a distância média entre um grupo de átomos (por exemplo, C-*alfa* de uma proteína). Se um RMSD calcula entre dois conjuntos de coordenadas atômicas - por exemplo, dois pontos de tempo da trajetória - o valor mede o quanto a conformação da proteína mudou. Neste caso, o RMSD do sítio de ligação permanece estável, abaixo de 1,5 nm, exceto para **1** , **6**, **9**, **10**, **11**, **14**, **15** e **19**. Além disso, **2** apresenta RMSD consideravelmente menor, indicando maior estabilidade.

Analisando as curvas do RMSD do C-*alfa*, pode-se observar baixos desvios em geral, indicando que o modelo de homologia construído para tirosinase humana é estável. O raio de giro (Rg), que é definido como a distribuição dos átomos de uma proteína em torno de seu eixo,

também contribuiu para essa interpretação, pois para todos os casos manteve uma variação de ~0,1 nm. Essa baixa variação, por sua vez, também corrobora com os resultados dos valores de RMSD dos ligantes, onde não houve desvios acima de 0,1 nm. Quando um composto ligante/chumbo se liga à proteína, a mudança conformacional também pode ser observada pelo raio de giro, nesse caso, foi observada estabilidade significativamente maior para **6, 14** e **15**.

Já a flutuação quadrática média (RMSF – *Root Mean Square Fluctuation*) mede o desvio médio de uma partícula (por exemplo, um resíduo de proteína) ao longo do tempo a partir de uma posição de referência (normalmente a posição média do tempo da partícula). Assim, o RMSF avalia como as porções da estrutura flutuam a partir de sua estrutura a média. Para os complexos proteína-ligante estudados, o RMSF permaneceu abaixo de 0,1 nm, também indicando alta estabilidade do sistema.

Por fim, todos os parâmetros apresentados anteriormente indicam que não houve grandes desvios ou movimentações incoerentes durante as dinâmicas. Para melhor avaliar a interação proteína-ligante propriamente dita, foram realizados cálculos de MM-PBSA em cada um dos frames.

4.6 MM-PBSA

No cálculo de MM-PBSA, o filme da dinâmica é quebrado em imagens estáticas conhecidas, os frames. Em cada uma delas são removidos os íons e o solvente, e o sistema proteína-ligante é inserido em um ambiente de solvente contínuo. Esses procedimentos diminuem o custo computacional para cálculo de energia de ligação proteína-ligante. Assim, foi feito o cálculo da média e desvio padrão para avaliação dos resultados, cujos valores obtidos foram utilizados para a dinâmica molecular de cada ligante.

Os valores absolutos calculados não possuem correlação direta com a realidade, por isso é necessário considerar um ligante referencial, no caso o ácido kójico (**1**) para avaliar a performance dos demais ligantes. Na Figura 12 é exibido o valor da energia de ligação entre ligante e proteína, a partir do qual observa-se que, na maioria dos casos, os ligantes apresentam uma interação desfavorável com os aminoácidos da proteína. Entretanto, na Figura 13, referente à a interação com os íons cobre do sítio ativo da enzima, percebem-se interações bem mais favoráveis.

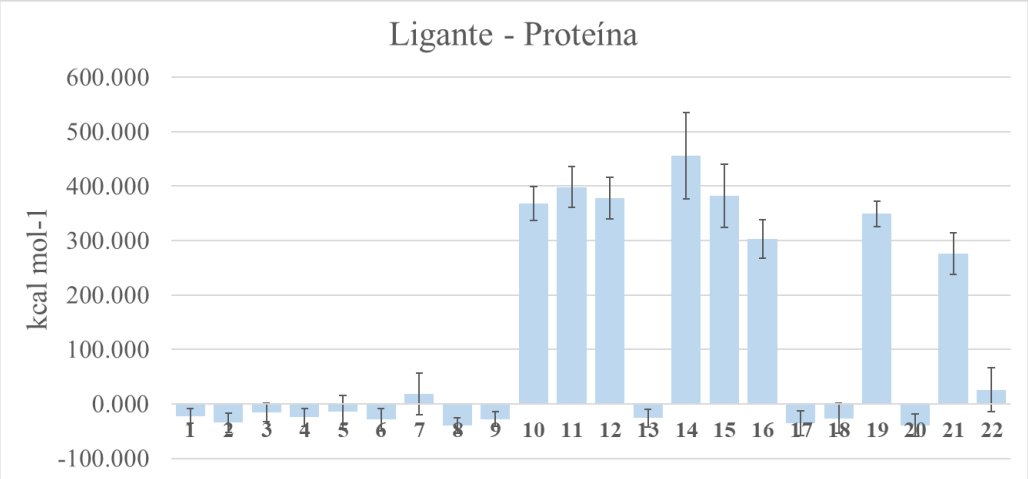


Figura 12: Gráfico com a média dos valores de MM-PBSA de interação ligante proteína, juntamente com desvio padrão ao longo da dinâmica de 100ns para os ligantes 1-22.

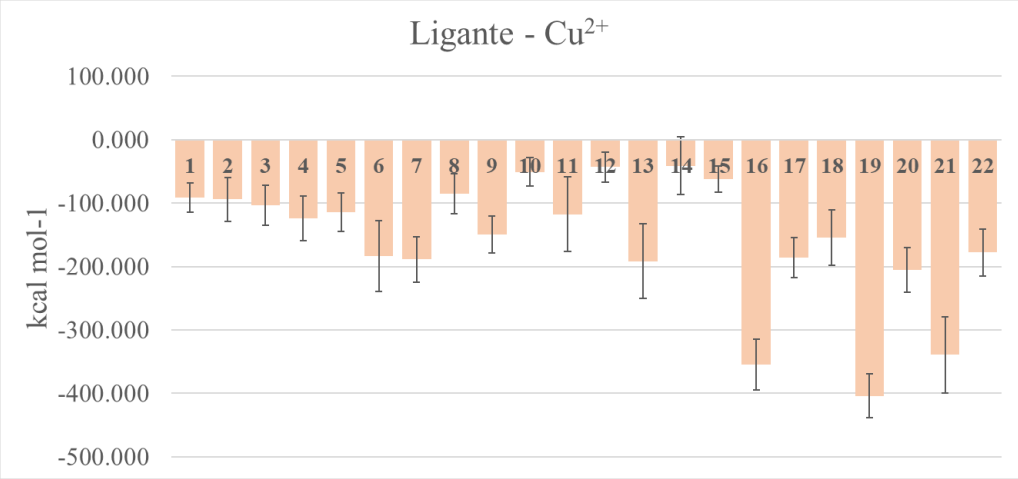


Figura 13: Gráfico com a média dos valores de MM-PBSA de interação ligante com os íons Cu^{2+}, juntamente com desvio padrão ao longo da dinâmica de 100 ns para os ligantes 1-22.

As interações acima foram somadas e seu erro propagado de forma a obter uma perspectiva da junção das duas interações, que pode ser observada na Figura 14. No entanto, não foi observada correlação entre os dados encontrados e os valores obtidos inicialmente através do *docking*, de forma que os ligantes que apresentaram excelentes resultados no *docking*, como ácido cafeico **10** e ácido gálico **15**, acabaram apresentando valores positivos de energia de ligação. Isso pode ser explicado pelo fato de que, neste cenário, está sendo avaliada uma interação dinâmica ao longo do tempo, que difere da avaliação estática de um cálculo de *docking* molecular.

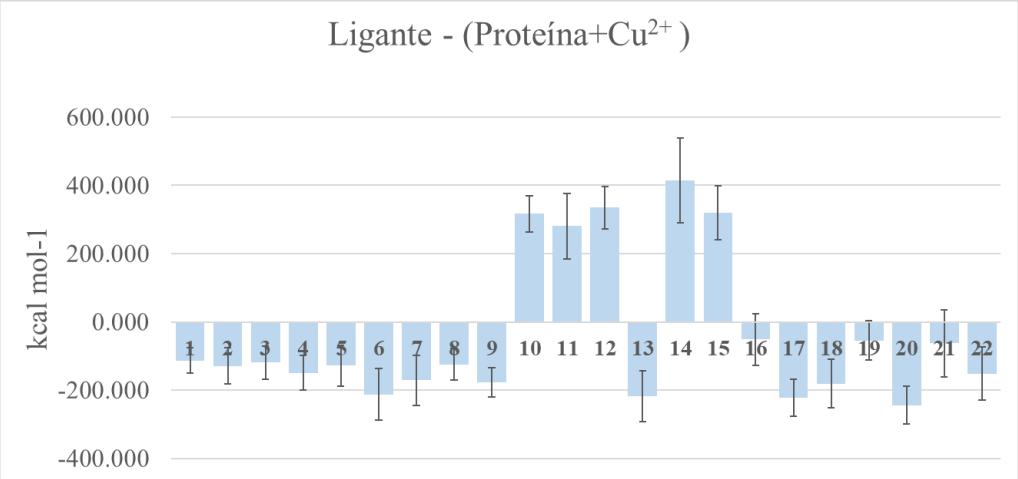


Figura 14: Gráfico com a média dos valores de MM-PBSA de interação ligante proteína somado a interação ligante com íons Cu^{2+}, juntamente com desvio padrão ao longo da dinâmica de 100 ns para os ligantes 1-22.

4.7 MIA-QSAR

Os três componentes do extrato de *Schinus Terebinthifolius* que apresentaram menor valor de Δ*G* MM-GBSA no cálculo de *docking* (ácido cafeico **10**, ácido gálico **15** e 1,2,3,4,6-penta-O-galoil-glucose **22**) tiveram seus pontos de contato com a proteína (grupos hidroxila) substituídos por F, Cl, Br, I, etil, metil e S-metil a fim de avaliar o quanto a variação de eletronegatividade e tamanho do grupo afetavam a energia de ligação. O átomo de F foi inicialmente escolhido, pois se trata de um substituinte comum na indústria farmacêutica para o desenvolvimento de novos medicamentos. Assim, seguiu-se a linha dos halogênios para avaliar tanto a mudança de tamanho quanto de eletronegatividade. No caso do grupo metila, era esperado que houvesse interações na qual a nuvem eletrônica atuasse como doadora, e no caso do etila esperava-se que os elétrons π fossem os responsáveis pelas interações. Analogamente, as hidroxilas que foram substituídas, assim como o grupo S-metila também foram avaliados, uma vez que o átomo de S pode ser usado como substituinte para o oxigênio, que se encontra um período acima na tabela periódica.

Esses novos ligantes foram submetidos ao procedimento de *docking* molecular, idêntico ao aplicado nos componentes do extrato inicialmente avaliados. O menor valor de Δ*G* obtido encontra-se na Tabela 4.

Tabela 4: Menor valor de Δ*G* MM-GBSA para os ligantes modificados utilizados na construção do modelo de MIA-QSAR.

Estrutura	ΔG	Estrutura	ΔG	Estrutura	ΔG
23	-25.34	42	-60.51	61	-41.61
24	-48.89	43	-45.78	62	-38.76
25	-24.58	44	-31.99	63	-40.32
26	-61.82	45	-43.17	64	-38.22
27	-53.76	46	-43.05	65	-3.37
28	-47.49	47	-65.14	66	-28.84
29	-64.29	48	-45.55	67	-21.50
30	-30.03	49	-74.15	68	-22.61
31	-51.33	50	-67.90	69	-26.09
32	-51.21	51	-34.43	70	-26.92
33	-23.95	52	-37.13	71	-21.06
34	-44.35	53	-33.45	72	-9.300
35	-36.34	54	-37.61	73	-28.47
36	-59.39	55	-33.51	74	-6.342
37	-41.71	56	-37.68	75	-27.73
38	-43.15	57	-35.02	76	-30.57
39	-31.75	58	-40.14	77	-23.94
40	-41.39	59	-40.47	78	-29.80
41	-36.83	60	-39.30		

Como se trata de uma grande quantidade de dados, o MIA-QSAR se mostra como uma excelente ferramenta para avaliar os resultados. Uma vez que as mudanças estruturais são evidenciadas nas imagens 2D construídas. O modelo mais confiável de correlação estrutura 2D e Δ*G* MM-GBSA foi obtido usando a eletronegatividade atômica de Pauling (ε), o que implica que a dependência eletrostática pode ser mais relevante para as energias livres de ligação do que para as interações relacionadas ao tamanho atômico.

Sendo assim, o modelo MIA-QSAR foi construído com quinze amostras para o conjunto de testes escolhidas por amostragem aleatória (10;23;24;31;37;43;44;47;49;58;63;67;68;74;75) e quatro variáveis latentes. Primeiramente, o modelo foi submetido a uma análise utilizando *studentized residues*[47] e não foram encontrados *outliers* (Figura 18), de forma que o modelo apresentou capacidade de modelagem confiável ($r^2 = 0,725$)[72] e com boa previsibilidade (r^2_{test} =0,606 e $r^2_m = 0,508$).[79] Além disso, o modelo pode ser considerado confiável ($^cr_p = 0,521$).[80] O valor $q^2 = 0,218$ obtido na validação cruzada interna *leave-one-out* (LOOCV) foi insuficiente ($q^2 < 0,5$).[81] Isso pode ser atribuído ao fato de que os substituintes são muito diferentes entre si, além de que o centro comum entre ligantes é pequeno. Dessa forma, quando uma molécula é

retirada da validação cruzada, ela é a única representativa do substituinte nessa posição e, consequentemente, o q^2 diminui.

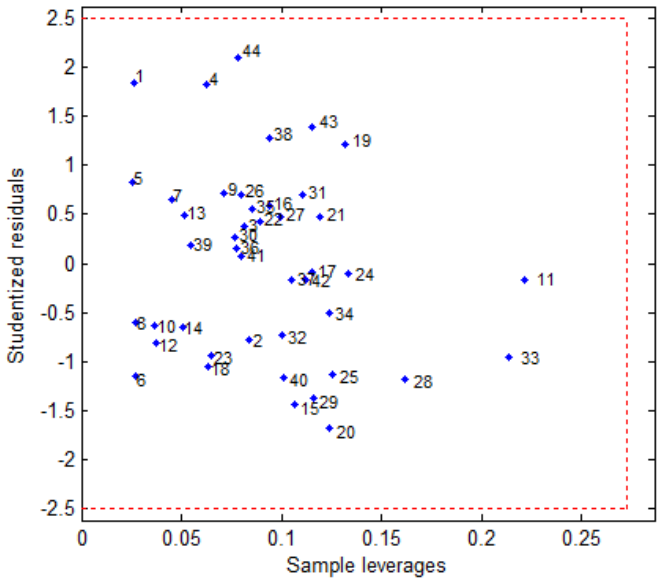


Figura 15: Gráfico de William baseado em *studentized residues* e *sample leverages* para detectar dados discrepantes.

Considerando-se o modelo obtido, a influência dos substituintes foi avaliada por MIA-plots (Figura 16)[82], que correspondem a mapas de contorno molecular que indicam: a) quanto (através da variável importância nos escores de projeção), e b) como (através da regressão PLS coeficientes) os substituintes moleculares afetam a propriedade medida. Conforme relatado pelas energias de ligação ΔG MM-GBSA, os compostos **49** (ΔG = -74,15 kJ mol^{-1}), **50** (ΔG = -67,90 kJ mol^{-1}), **47** (ΔG = -65,15 kJ mol^{-1}) e **29** (ΔG = -64,29 kJ mol^{-1}) apresentaram melhora relevante em relação ao ligante original (**22**), sugerindo que grupos de alto volume e baixa eletronegatividade diminuem a energia de ligação especialmente na posição 5 do anel. Os derivados dos ligantes **10** e **15**, por outro lado, não apresentaram melhoria significativa.

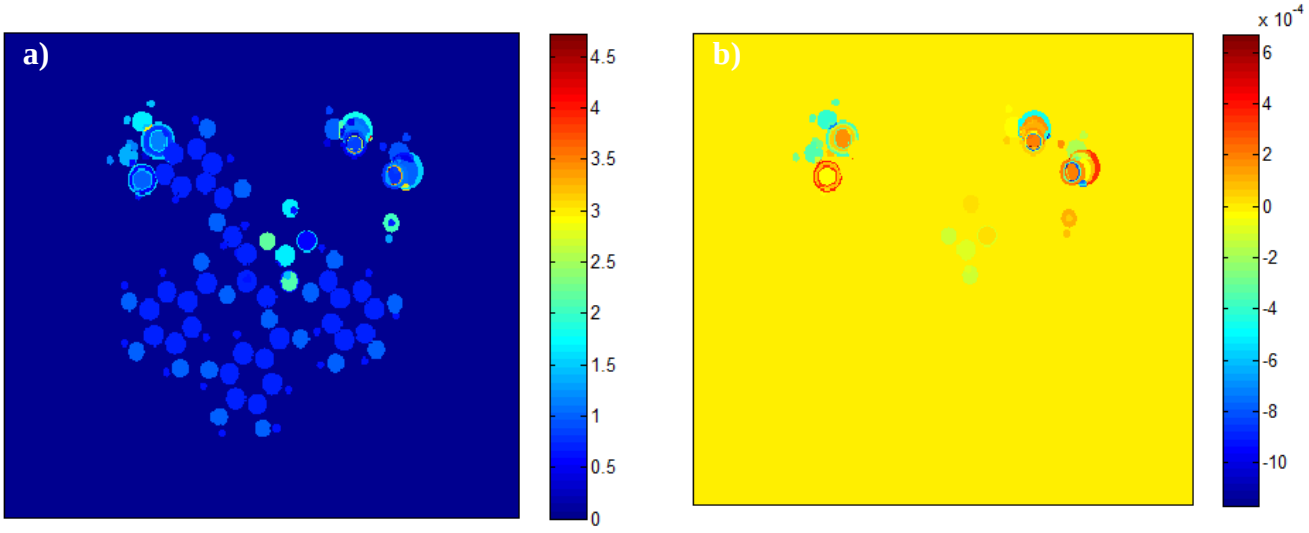


Figura 16: Gráfico MIA que mostra a importância da variável em diferentes regiões das estruturas sobrepostas (a) e os coeficientes de regressão PLS que mostram o impacto da variável (positivo ou negativo) nas estruturas sobrepostas (b).

Analisando-se os gráficos de MIA da Figura 16, o de projeção das pontuações de importância de variável (Figura 16a) mostra algumas áreas em azul claro no mapa, que são as posições onde os substituintes diminuem ligeiramente as energias de ligação Δ*G* MM-GBSA. Toda a região do mapa de contorno que corresponde aos análogos de **22** está destacada em azul claro, mostrando que, em geral, todas essas moléculas possuem um valor menor de Δ*G* . A região amarelo-claro, por sua vez, provavelmente corresponde aos grupos carboxila de **10** e **15**, pois não foi observada melhora quando esses grupos foram substituídos.

No gráfico de coeficientes de regressão PLS (Figura 16b), a área destacada em vermelho corresponde à parte superior dos ligantes derivados de **22**, o que sugere tanto que os substituintes modificados nesta posição têm um efeito redutor na variável medida, quanto que substituintes com Δ*G* mais negativos são de alto volume, uma vez que a região externa dos círculos no mapa de contorno é vermelha.

Conforme relatado pelas energias de ligação Δ*G* MM-GBSA, os compostos **49** (Δ*G* = -74,15 kJ mol^{-1}), **50** (Δ*G* = -67,90 kJ mol^{-1}), **47** (Δ*G* = -65,15 kJ mol^{-1}) e **29** (Δ*G* = -64,29 kJ mol^{-1}) apresentaram melhora relevante em relação ao ligante original, sugerindo que grupos volumosos e de baixa eletronegatividade diminuem a energia de ligação, principalmente na posição 5 do anel. Assim, é possível entender porque os derivados dos ligantes **10** e **15** não mostraram aumento significativo na afinidade de ligação.

4.8 Permeabilidade cutânea teórica

Considerando a relevância da permeabilidade cutânea teórica para este estudo, consideramos o coeficiente de permeabilidade (Kp), que relaciona o fluxo de soluto e a

concentração do gradiente através da membrana: quanto mais negativo o log Kp (com Kp em cm/s), menor a permeação cutânea da molécula.[74] Os resultados obtidos pelo software SWISS -ADME são mostrado na Tabela 5.

Em geral, as moléculas estudadas não apresentaram boa permeabilidade cutânea. Assim, os ligantes **5**, **8**, **9**, **10**, **11**, **12**, **15**, **17** e **21** foram ligeiramente mais positivos do que o inibidor de referência, ácido kójico **1**, indicando uma melhor permeabilidade teórica na pele. Ao realizar uma análise conjunta com os valores de *docking* e dinâmica molecular, é possível apontar que o componente com melhor aplicabilidade cosmética é o ligante luteolina, **17**.

Tabela 5:Valores de Log Kps para os componentes do extrato de Schinus terebinthifolius e ácido kójico.

Ligand	**LogKp (cm^{-1})**	**Ligand**	**LogKp (cm^{-1})**
Ácido kójico (1)	-7.83	Ácido coumárico (**12**)	-6.26
Metil galato (**2**)	-11.42	1,6-Digaloil-ß-D-glucose (**13**)	-9.43
Miricetina (**3**)	-7.83	Ácido clorogênico (**14**)	-8.76
Etil galato (**4**)	-11.28	Ácido gálico (**15**)	-6.84
Catequina (**5**)	-7.05	Ácido 5-p-coumaroilquínico (**16**)	-8.02
Quercetina ramnosídeo (**6**)	-8.42	Luteolina (**17**)	-6.62
Miricetina ramnosídeo (**7**)	-9.12	Miricetina galactoside rhamnoside (**18**)	-11.63
Ácido elágico (8)	-7.36	Miricetina glucoronídeo (**19**)	-9.14
Quercetina (**9**)	-7.82	Miricetina galactosídeo (**20**)	-9.51
Ácido cafeico (10)	-6.58	Robustaflavona (**21**)	-6.01
Ácido siríngico (**11**)	-6.77	1,2,3,4,6-penta-O-galoil-glucose (**22**)	-9.49

5 Discussão

Inicialmente, para a construção do modelo por homologia, foi utilizado o software Modeller, o qual, porém, teve que ser substituído visto que não apresentou um resultado muito satisfatório. Diante desse cenário, optou-se por testar outros softwares: o SWISS-MODEL e o Schroedinger 2020-3, cujos resultados foram mais coerentes do ponto de vista estrutural.

Uma vez que houve um grande número de modelos de homologia construídos, fez-se necessário selecionar ferramentas de simples testagem, mas que, ao mesmo tempo, fornecessem resultados acurados. Para tanto, foram escolhidos os softwares *web based* PROCHECK e Verify 3D, tanto por sua velocidade e simplicidade, quanto por serem amplamente usados na literatura.[83]

PROCHECK e o Verify 3D, contudo, não foram suficientes para definir qual seria o melhor modelo, tornando necessária a realização de análise de *re-docking* e cálculo de RMSD em relação aos ligantes co-cristalizados. Sendo assim, uma vez que a estrutura da proteína *template* se encontra disponível no PDB co-cristalizada aos ligantes ácido kójico, tropolona e mimosina, esses foram submetidos a cálculos de *docking*, gerando 10 poses no software Glide. Posteriormente, foi realizado o cálculo de RMSD das poses com relação aos ligantes co-cristalizados e, para cada uma dessas poses de *docking*, realizou-se o cálculo de ΔG (MM-GBSA).

A fim de avaliar a performance dos modelos nesse sentido, foram gerados gráficos de ΔG *vs* RMSD para cada um dos ligantes, de modo que o modelo cujas retas tivessem o R^2 mais alto, em teoria, corresponderia àquele com melhor correlação RMSD x ΔG.Vale lembrar que neste cálculo é esperado que quanto menor o RMSD, mais negativo o valor de ΔG, uma vez que o ligante se encontra mais próximo à sua posição de menor energia.

No tocante ao método, apesar de ainda não ter sido apresentado na literatura, ele consiste, basicamente, em metodologias análogas às usadas para a validação de protocolos de *docking*, mais especificamente quando a estrutura que se deseja analisar já se encontra em bases de dados cristalográficas como o PDB.[84] Durante essa etapa de análise de RMSD, percebeu-se a importância da presença de uma molécula de água coordenada aos átomos de Cu^{2+}, mas somente nos casos do ácido kójico e da mimosina, por isso, decidiu-se analisar o *docking* dos ligantes nos dois casos, com e sem a água coordenada.

Assim, o modelo escolhido para prosseguir com os cálculos foi aquele que apresentou melhor performance em ambas as avaliações com e sem água coordenada. No caso, o modelo

construído com o algoritmo *knowledge based* do *Structure Prediction Wizard*, do pacote Schroedinger 2020-3. Uma vez escolhido o modelo, foi realizado o cálculo de *docking* molecular para todos os compostos identificados no extrato de *Schinus Terebinthifolius*, usando os três programas Glide, Autodock Vina e Gold (por conta de seus algoritmos de ranqueamento serem diferentes), com e sem água coordenada. Assim, tornou-se possível obter uma maior diversidade de poses e, espera-se, maior acurácia em relação aos dados experimentais.

A grande maioria dos resultados de energia de livre de ligação ΔG MM-GBSA mais negativos foram encontrados pelo software Glide (13 resultados), seguido pelo Gold (7 resultados) e por fim pelo Autodock (2 resultados). De fato, o software Glide foi o que mais encontrou resultados, provavelmente por ter uma busca por poses mais refinadas, especialmente no modo XP Glide, no qual o software gera uma série de conformações internamente e as passa através de uma série de filtros. Esse método de pontuação do Glide tem como propósito eliminar falsos positivos e fornecer uma melhor correlação com dados reais, justificando seu melhor desempenho.[85] Nesse processo, centraliza-se o ligante em várias posições em uma grade de 1 Å e o gira em torno dos três ângulos de Euler, de forma que elimina modos de ligação improváveis.[86] Feito esse levantamento, há ainda uma etapa que consiste em uma avaliação de campo de força (OPLS-2005) baseada em uma grade de refinamento das poses; só então as poses finais são obtidas.

A partir das poses finais, pôde-se observar que a maioria dos ligantes apresentou menor valor de energia livre de ligação quando a proteína apresentava a água coordenada aos átomos de cobre. Isso ocorre porque a água acaba por fazer um papel de ponte entre o ligante e o metal; como a maioria dos ligantes interage através de grupos hidroxila e considerando que interações tipo ligação de hidrogênio são mais favoráveis do que interações metais–hidroxila, as ligações de hidrogênio entre moléculas de água resulta em melhores sobreposição e interação, já que o tamanho dos orbitais é mais próximo. As moléculas estudadas apresentam em sua maioria anéis aromáticos e grupos hidroxila, permitindo observar que, de forma geral, as interações apresentadas no cálculo de *docking* molecular são do tipo π-*stacking* e ligação de hidrogênio.

Os ligantes que apresentaram menor valor de energia livre de ligação foram: ácido cafeico **10**, ácido gálico **15** e 1,2,3,4,6-penta-O-galoil-glucose **22**. Como os dois primeiros apresentaram interação com os átomos de cobre responsáveis pela atividade oxidativa da enzima, ambas podem assumir um papel de inibição competitiva ou substrato, enquanto que o ligante **22** não interage com os átomos de cobre. Os três são reportados na literatura com atividade inibitória em ensaios de tirosinase não humana.

Na etapa seguinte, inicialmente se havia optado por realizar o estudo de dinâmica molecular dentro do pacote Schroedinger 2020-3, devido aos melhores resultados obtidos pelo Glide. Assim, foram calculadas dinâmicas utilizando-se o software Desmond[87] e um *script* de cálculo de energia livre por MM-GBSA para analisar as dinâmicas moleculares. Entretanto, como o desvio padrão apresentado foi alto, os resultados obtidos se tornaram inviáveis, o que levou à realização dos mesmos cálculos com o GROMACS - nesse caso, para a análise de energia livre, o *script* disponível utiliza o método MM-PBSA. Como desvio padrão apresentado foi consideravelmente menor, o estudo seguiu com esses resultados.

Esses resultados obtidos através da análise da dinâmica molecular, contudo, não apresentaram correlação com os resultados obtidos por *docking* (r^2= 0.0379), o que pode ser explicado pelo fato de que a dinâmica molecular analisa uma interação ao longo do tempo, não uma pose estática. Além disso, diversas poses de *docking* possuíam uma molécula de água realizando a coordenação dos átomos de cobre ao ligante, a qual acaba saindo do sítio ativo quando iniciada a dinâmica. Desse modo, foi analisada a correlação do Δ*G* MM-PBSA da dinâmica molecular com as melhores poses obtidas sem a presença de água no cálculo de *docking* e, apesar de ter havido um aumento no valor de r^2, este ainda não seria significativo (r^2=0.191). Por outro lado, um aumento significativo do valor de r^2 ocorre quando são correlacionados os valores de Δ*G* MM-PBSA com a eficiência do ligante (r^2=0.560), a qual corresponde ao valor de Δ*G* do *docking* molecular dividido pela massa molecular do ligante. A eficiência permite reduzir o viés observado entre moléculas maiores, consequentemente tendo um maior número de pontos de ligação para moléculas menores. Os valores de eficiência do ligante podem ser encontrados na Tabela 6.

Tabela 6: Eficiência dos ligantes, componentes do extrato de *Schinus Terebinthifolius*, calculada através da divisão do Δ*G* MM-GBSA obtido no docking pela massa molecular do ligante

Ligante	Eficiência do ligante
Ácido kójico (**1**)	-0,10
Metil galato (**2**)	-0,06
Miricetina (**3**)	-0,11
Etil galato (**4**)	-0,06
Catequina (**5**)	-0,06
Quercetina ramnosídeo (**6**)	-0,07
Miricetina ramnosídeo (**7**)	-0,06
Ácido elágico (**8**)	-0,09
Quercetina (**9**)	-0,10
Ácido cafeico (**10**)	-0,22
Ácido siríngico (**11**)	-0,13

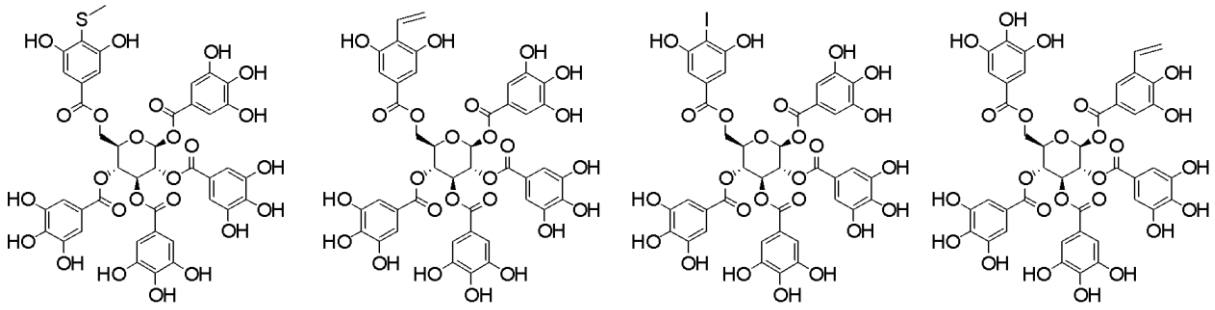
Ácido coumárico (**12**)	-0,19
1,6-Digaloil-ß-D-glucose (**13**)	-0,06
Ácido clorogênico (**14**)	-0,09
Ácido gálico (**15**)	-0,23
Ácido 5-p-coumaroilquínico (**16**)	-0,05
Luteolina (**17**)	-0,07
Miricetina galactoside rhamnoside (**18**)	-0,06
Miricetina glucoronídeo (**19**)	-0,05
Miricetina galactosideo (**20**)	-0,07
Robustaflavona (**21**)	-0,04
1,2,3,4,6-penta-O-galoil-glucose (**22**)	-0,06

O cálculo de Δ*G* MM-PBSA foi feito separadamente para as interações ligante–aminoácidos e ligante–cobre. Os ligantes em sua maioria apresentaram interações desfavoráveis com os aminoácidos da proteína que, por sua vez, eram compensadas pelas interações favoráveis com os metais no sítio ativo. Dentre os ligantes que apresentaram menor valor de energia livre de ligação, pode-se destacar: quercetina ramnosídeo **6**, 1,6-Digaloil-ß-D-glucose **13**, luteolina **17** e miricetina galactosídeo **20**.

A quercetina ramnosídeo **6** apresentou um IC_{50} = 98,24 ± 2,31 µM para tirosinase de cogumelo[88]; e ao luteolina **17**, que é apresentada na literatura como inibidor de tirosinase de cogumelo, obteve IC_{50} = 17,40 ± 0,62 µM. O mesmo estudo aponta seu derivado glicosilado, a luteolina 5-O-β-D-glucopiranosídeo, como um inibidor mais potente.[89] Considerando-se que **6**, **20** e **13** apresentam grupos glicosilados, de uma forma geral, sua presença pode melhorar a performance de flavonoides inibidores de tirosinase. No caso de 1,6-Digaloil-ß-D-glucose **13**, mais precisamente, não foram encontrados dados na literatura para fins de comparação, sendo então uma molécula promissora para testes experimentais, - testes estes que, apesar de previstos no presente projeto, não foram realizados em virtude das limitações impostas pela pandemia de COVID-19.

As análises de MIA-QSAR foram realizadas em paralelo com os cálculos de dinâmica molecular a partir dos compostos que apresentaram melhor resultado nos cálculos de *docking*. As variações dos substituintes foram definidas a fim de estudar o impacto de alterações na eletronegatividade e volume dos pontos de interação com a proteína. Assim, foram comparados os átomos/grupos F, Cl, Br, I, metila, etila e S-metila. Os resultados dos modelos apontaram que a eletronegatividade teve maior impacto do que o raio de van der Waals, uma vez que foi o descritor que gerou um modelo mais estatisticamente confiável em relação ao que utilizava o raio e eletronegatividade dividida pelo raio.

Além disso, foram obtidas energias livres de ligação, Δ*G* MM-GBSA, mais negativas que do ligante original para o caso de alguns derivados do composto **22**, sendo eles **49**, **50**, **47** e **29**, apresentados na Figura 17.



49 **50** **47** **29**

Figura 17: Compostos que apresentaram melhor Δ*G* MM-GBSA em relação ao ligante original **22**.

Neste momento pode-se observar um padrão que reforça a análise dos MIA-plots e o fato de o melhor modelo ter sido construído utilizando-se eletronegatividade: grupos pouco eletronegativos, de grande volume e alta polarizabilidade tornaram mais negativa a energia livre de ligação, em especial quando localizados na posição 5 do anel.

A fim de complementar este estudo, realizou-se análise de permeação cutânea, uma vez que uma das principais aplicações de inibidores de tirosinase é na indústria cosmética, para a qual a permeação do composto na pele é desejada. Os resultados obtidos na Tabela 5 não apontam melhor desempenho de nenhum dos compostos em termos de permeabilidade, isto é, em nenhum dos casos esta foi menos negativa que o inibidor de referência, o ácido kójico **1**. Já combinando essa análise com a etapa de dinâmica molecular, o composto que apresentou valor mais negativo de Δ*G* MM-PBSA e melhor permeabilidade cutânea foi a luteolina **17**. Como o 1,2,3,4,6-penta-O-galoil-glucose **22** é uma molécula grande e de baixa permeabilidade cutânea, pode ser interessante estudar seus derivados de menor energia livre de ligação (**49**, **50**, **47** e **29**) para aplicações na indústria alimentícia, evitando o escurecimento de frutas.

6 Conclusões

Os estudos de modelagem por homologia, *docking* molecular, dinâmica molecular e MIA-QSAR possibilitaram mapear as interações dos componentes do extrato de *Schinus terebinthifolius* e tirosinase humana, além de observar a relação estrutura-atividade desses compostos. A partir disso, os resultados obtidos por este trabalho sugerem a continuidade dos

estudos dos compostos presentes no extrato etanólico de *Schinus terebinthifolius*, uma vez que podem se tratar de candidatos interessantes para a avaliação de ensaios biológicos mais complexos, como toxicidade *in vitro*, culturas de melanócitos e, eventualmente, testes *in vivo*. Além disso, também é possível direcionar a planta para que produza majoritariamente compostos de interesse e assim produzir um extrato voltado especificamente para essa aplicação.

Apesar de hoje existirem uma série de moléculas sendo testadas para a inibição da tirosinase humana, é necessário levar adiante esses estudos para ensaios experimentais e clínicos em humanos, principalmente porque a grande maioria dos estudos avalia apenas a atividade inibidora em tirosinases derivadas de fungos. Nesse sentido, torna-se importante avaliar fatores como toxicidade, penetração cutânea, estabilidade e atividade *in vivo.* Com a latente necessidade de substâncias clareadoras que sejam de fato eficazes e cujos efeitos colaterais à saúde sejam reduzidos, é muito interessante estudar novas substâncias obtidas a partir da exploração de plantas, sobretudo considerando a crescente aderência do público consumidor de cosméticos a produtos naturais.

O uso de estudos computacionais para realizar um *screening* das moléculas pode ser aplicado com um teste de baixo custo para filtrar potenciais moléculas e aumentar a taxa de sucesso dos testes experimentais. Entretanto, sente-se falta de um protocolo mais definido para tais estudos a fim de possibilitar a comparação entre diferentes autores.

Após realizados todos os cálculos e análises dos resultados no presente trabalho, o composto que apresentou maior potencial para a aplicação cosmética inicialmente proposta foi a luteolina **17**. A partir desse resultado, seriam necessários teste *in vitro* e *in vivo* para confirmar o potencial de ação dessa molécula. Com essa perspectiva em mãos, podem ser propostas melhorias sintéticas que aperfeiçoem sua performance como inibidor, além da produção de extratos da planta com enfoque em maior quantidade desse composto no mesmo. Isso possibilitaria o uso de um extrato natural, na melhor aceitação mercadológica, com tecnologia e alta performance.

Por fim, pode ser interessante estudar os derivados de 1,2,3,4,6-penta-O-galoil-glucose **22**, que apresentaram menor energia livre de ligação (**49**, **50**, **47** e **29**). Apesar de o composto **22** ser uma molécula grande e de baixa permeabilidade cutânea - não apresentando boa aplicabilidade na indústria cosmética - sua aplicação na indústria alimentícia pode ser promissora para evitar as perdas decorrentes do escurecimento de frutas.

7 Bibliografia

1. Merad, M., Ginhoux, F. & Collin, M. Origin , homeostasis and function of Langerhans cells and other langerin- expressing dendritic cells. **8**, (2008).
2. Videira, I. F. dos S., Moura, D. F. L. & Magina, S. Mechanisms regulating melanogenesis*. *An. Bras. Dermatol.* **88**, 76–83 (2013).
3. Hearing, V. J. & Jiménez, M. Mammalian tyrosinase-The critical regulatory control point in melanocyte pigmentation. *Int. J. Biochem.* **19**, 1141–1147 (1987).
4. Qian, W. *et al.* Natural skin-whitening compounds for the treatment of melanogenesis (Review). *Exp. Ther. Med.* **20**, 173–185 (2020).
5. Agarwal, P., Singh, M., Singh, J. & Singh, R. P. *Microbial Tyrosinases: A Novel Enzyme, Structural Features, and Applications*. *Applied Microbiology and Bioengineering* (Elsevier Inc., 2019). doi:10.1016/b978-0-12-815407-6.00001-0.
6. Farthmann, B., Schmitz, S. & Orfanos, C. E. Photoprotection by total melanin content and the phenotype (eumelanin, phaeomelanin) in human melanoma cell lines. *Melanoma Res.* **6**, S64 (1996).
7. Desai, S. R. Hyperpigmentation therapy: A review. *J. Clin. Aesthet. Dermatol.* **7**, 13–17 (2014).
8. Otwell, W. S. & Iyengar, R. Inhibition of Enzymatic Browning in Foods and Beverages. *Crit. Rev. Food Sci. Nutr.* **32**, 253–273 (1992).
9. Loizzo, M. R., Tundis, R. & Menichini, F. Natural and Synthetic Tyrosinase Inhibitors as Antibrowning Agents: An Update. *Compr. Rev. Food Sci. Food Saf.* **11**, 378–398 (2012).
10. Draelos, Z. D. Skin lightening preparations and the hydroquinone controversy. *Dermatol. Ther.* **20**, 308–313 (2007).
11. Zolghadri, S. *et al.* A comprehensive review on tyrosinase inhibitors. *J. Enzyme Inhib. Med. Chem.* **34**, 279–309 (2019).
12. Fonseca-Santos, B., Antonio Corrêa, M. & Chorilli, M. Sustainability, natural and organic cosmetics: Consumer, products, efficacy, toxicological and regulatory considerations. *Brazilian J. Pharm. Sci.* **51**, 17–26 (2015).
13. Falcão, M. P. M. M., Oliveira, T. K. B., Ó, N. P. R. do, Sarmento, D. A. & Gadelha, N. C. Schinus terebinthifolius Raddi (Aroeira) e suas propriedades na Medicina Popular. *Rev. Verde Agroecol. e Desenvolv. Sustentável* **10**, 23 (2015).
14. Chave, P. Schinus terebinthifolius Raddi : chemical composition , biological properties and toxicity. 158–169 (2013).
15. Barbosa, L. C. A., Demuner, A. J., Clemente, A. D., De Paula, V. F. & Ismail, F. M. D. Seasonal variation in the composition of volatile oils from Schinus terebinthifolius Raddi. *Quim. Nova* **30**, 1959–1965 (2007).
16. Imagem. *https://www.greenmebrasil.com/morar/horta-e-jardim/4121-aroeira-medicinal-cosmetica-culinaria/.*

17. Sassi, A. Ben, Elayeb, A., Karaman, I., Marzouk, B. & Mastouri, M. Phytochemical Profile and Antiproliferative, Anti-tyrosinase, Antioxidant, and Antibacterial Potential of Schinus terebinthifolius Growing in Tunisia. *J. Herbs, Spices Med. Plants* **26**, 61–76 (2019).

18. Ceruks, M., Romoff, P., Fávero, O. A. & Lago, J. H. G. Constituíntes fenólicos polares de Schinus terebinthifolius raddi (Anacardiaceae). *Quim. Nova* **30**, 597–599 (2007).

19. Farag, S. F. Polyphenolic compounds from the leaves of Schinus terebinthifolius Raddi. *Bull. Pharm. Sci.* **31**, 319–329 (2008).

20. El-Massrry, K. F., El-Ghorab, A. H., Shaaban, H. A. & Shibamoto, T. Chemical compositions and antioxidant/antimicrobial activities of various samples prepared from Schinus terebinthifolius leaves cultivated in Egypt. *J. Agric. Food Chem.* **57**, 5265–5270 (2009).

21. Da Silva, A. P. *et al.* Tyrosinase inhibitory activity, molecular docking studies and antioxidant potential of chemotypes of Lippia origanoides (Verbenaceae) essential oils. *PLoS One* **12**, 1–10 (2017).

22. Terstappen, G. C. & Reggiani, A. In silico research in drug discovery. *Trends Pharmacol. Sci.* **22**, 23–26 (2001).

23. Tuffery, P. & Derreumaux, P. Flexibility and binding affinity in protein-ligand, protein-protein and multi-component protein interactions: Limitations of current computational approaches. *J. R. Soc. Interface* **9**, 20–33 (2012).

24. Waterhouse, A. *et al.* SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Res.* **46**, W296–W303 (2018).

25. Webb, B. & Sali, A. Comparative protein structure modeling using MODELLER. *Curr. Protoc. Bioinforma.* **2016**, 5.6.1-5.6.37 (2016).

26. Kopp, J. & Schwede, T. Automated protein structure homology modeling: A progress report. *Pharmacogenomics* **5**, 405–416 (2004).

27. Jacobson, M. P. *et al.* A Hierarchical Approach to All-Atom Protein Loop Prediction. *Proteins Struct. Funct. Genet.* **55**, 351–367 (2004).

28. Berman, H. M. *et al.* The protein data bank. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **58**, 899–907 (2002).

29. Nayeem, A. A comparative study of available software for high-accuracy homology modeling: From sequence alignments to structural models. *Protein Sci.* **15**, 808–824 (2006).

30. Pagadala, N. S., Syed, K. & Tuszynski, J. Software for molecular docking: a review. *Biophys. Rev.* **9**, 91–102 (2017).

31. Sylvetsky, N. Toward Simple, Predictive Understanding of Protein-Ligand Interactions: Electronic Structure Calculations on Torpedo Californica Acetylcholinesterase Join Forces with the Chemist's Intuition. *Sci. Rep.* **10**, 1–12 (2020).

32. Roy, K., Kar, S. & Das, R. N. *Other Related Techniques*. *Understanding the Basics of QSAR for Applications in Pharmaceutical Sciences and Risk Assessment* (2015). doi:10.1016/b978-0-12-801505-6.00010-7.

33. Pozzan, A. Molecular Descriptors and Methods for Ligand Based Virtual High Throughput Screening in Drug Discovery. *Curr. Pharm. Des.* **12**, 2099–2110 (2006).

34. Friesner, R. A. *et al.* Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *J. Med. Chem.* **47**, 1739–1749 (2004).

35. Morris, G. M. *et al.* Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J. Comput. Chem.* **19**, 1639–1662 (1998).

36. Jones, G., Willett, P., Glen, R. C., Leach, A. R. & Taylor, R. Development and validation of a genetic algorithm for flexible docking 1 1Edited by F. E. Cohen. *J. Mol. Biol.* **267**, 727–748 (1997).

37. Ylilauri, M. & Pentikäinen, O. T. MMGBSA as a tool to understand the binding affinities of filamin-peptide interactions. *J. Chem. Inf. Model.* **53**, 2626–2633 (2013).

38. Kollman, P. A. *et al.* Calculating structures and free energies of complex molecules: Combining molecular mechanics and continuum models. *Acc. Chem. Res.* **33**, 889–897 (2000).

39. Wang, L. *et al.* Accurate and reliable prediction of relative ligand binding potency in prospective drug discovery by way of a modern free-energy calculation protocol and force field. *J. Am. Chem. Soc.* **137**, 2695–2703 (2015).

40. Pu, C., Yan, G., Shi, J. & Li, R. Assessing the performance of docking scoring function, FEP, MM-GBSA, and QM/MM-GBSA approaches on a series of PLK1 inhibitors. *Medchemcomm* **8**, 1452–1458 (2017).

41. Ahinko, M., Niinivehmas, S., Jokinen, E. & Pentikäinen, O. T. Suitability of MMGBSA for the selection of correct ligand binding modes from docking results. *Chem. Biol. Drug Des.* **93**, 522–538 (2019).

42. Zheng, L., Alhossary, A. A., Kwoh, C. K. & Mu, Y. Molecular dynamics and simulation. *Encycl. Bioinforma. Comput. Biol. ABC Bioinforma.* **1**–**3**, 550–566 (2018).

43. Adcock, S. A. & McCammon, J. A. Molecular dynamics: Survey of methods for simulating the activity of proteins. *Chem. Rev.* **106**, 1589–1615 (2006).

44. Sharp, K. A. & Honig, B. Calculating total electrostatic energies with the nonlinear Poisson-Boltzmann equation. *J. Phys. Chem.* **94**, 7684–7692 (1990).

45. Tsui, V. & Case, D. A. Theory and applications of the Generalized Born solvation model in macromolecular simulations. *Biopolymers* **56**, 275–291 (2000).

46. Kumari, R., Kumar, R. & Lynn, A. G-mmpbsa -A GROMACS tool for high-throughput MM-PBSA calculations. *J. Chem. Inf. Model.* **54**, 1951–1962 (2014).

47. Barigye, S. J. & Puggina de Freitas, M. Ten Years of the MIA-QSAR Strategy. *Int. J. Quant. Struct. Relationships* **1**, 64–77 (2016).

48. Daina, A., Michielin, O. & Zoete, V. SwissADME: A free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Sci. Rep.* **7**, 1–13 (2017).

49. Lai, X., Wichers, H. J., Soler-Lopez, M. & Dijkstra, B. W. Structure of Human Tyrosinase Related Protein 1 Reveals a Binuclear Zinc Active Site Important for Melanogenesis. *Angew. Chemie - Int. Ed.* **56**, 9812–9815 (2017).

50. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283–291 (1993).

51. Luthy, R., Bowei, J. & Einsenberg, D. Verify3D: Assessment of protein models with three-dimensional profiles. *Methods Enzymol.* **277**, 396–404 (1997).

52. Venthur, H., Mutis, A., Zhou, J. J. & Quiroz, A. Ligand binding and homology modelling of insect odorant-binding proteins. *Physiol. Entomol.* **39**, 183–198 (2014).

53. Bateman, A. *et al.* UniProt: The universal protein knowledgebase in 2021. *Nucleic Acids Res.* **49**, D480–D489 (2021).

54. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).

55. Thompson, J. D., Higgins, D. G. & Gibson, T. J. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673–4680 (1994).

56. Bienert, S. *et al.* The SWISS-MODEL Repository-new features and functionality. *Nucleic Acids Res.* **45**, D313–D319 (2017).

57. Madhavi Sastry, G., Adzhigirey, M., Day, T., Annabhimoju, R. & Sherman, W. Protein and ligand preparation: Parameters, protocols, and influence on virtual screening enrichments. *J. Comput. Aided. Mol. Des.* **27**, 221–234 (2013).

58. Harder, E. *et al.* OPLS3: A Force Field Providing Broad Coverage of Drug-like Small Molecules and Proteins. *J. Chem. Theory Comput.* **12**, 281–296 (2016).

59. Uchikoga, N., Matsuzaki, Y., Ohue, M., Hirokawa, T. & Akiyama, Y. Re-Docking Scheme for Generating Near-Native Protein Complexes by Assembling Residue Interaction Fingerprints. *PLoS One* **8**, (2013).

60. Cousins, K. R. Computer review of ChemDraw ultra 12.0. *J. Am. Chem. Soc.* **133**, 8388 (2011).

61. Greenwood, J. R., Calkins, D., Sullivan, A. P. & Shelley, J. C. Towards the comprehensive, rapid, and accurate prediction of the favorable tautomeric states of drug-like molecules in aqueous solution. *J. Comput. Aided. Mol. Des.* **24**, 591–604 (2010).

62. Li, Y. *et al.* Comparative assessment of scoring functions on an updated benchmark: 1. compilation of the test set. *J. Chem. Inf. Model.* **54**, 1700–1716 (2014).

63. O'Boyle, N. M. *et al.* Open Babel: An Open chemical toolbox. *J. Cheminform.* **3**, 33 (2011).

64. Abraham MJ, van der Spoel D, Lindahl E, H. B. The GROMACS Development Team. *GROMACS User Man. 2019*.

65. Robertson, M. J., Tirado-Rives, J. & Jorgensen, W. L. Improved Peptide and Protein Torsional Energetics with the OPLS-AA Force Field. *J. Chem. Theory Comput.* **11**, 3499–3509 (2015).

66. Sousa Da Silva, A. W. & Vranken, W. F. ACPYPE - AnteChamber PYthon Parser interfacE. *BMC Res. Notes* **5**, 1–8 (2012).

67. Ribeiro, A. A. S. T., Horta, B. A. C. & De Alencastro, R. B. MKTOP: A program for automatic construction of molecular topologies. *J. Braz. Chem. Soc.* **19**, 1433–1435 (2008).

68. Berendsen, H. J. C., Postma, J. P. M., Van Gunsteren, W. F., Dinola, A. & Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **81**, 3684–3690 (1984).

69. Swope, W. C. *et al.* A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules : Application to small water clusters computer simulation method for the calculation of ~ uilibrium constants for the formation o. **637**, (1982).

70. Miller, B. R. *et al.* MMPBSA . py : An E ffi cient Program for End-State Free Energy Calculations. (2012).

71. Kuhn, B., Gerber, P., Schulz-Gasch, T. & Stahl, M. Validation and use of the MM-PBSA approach for drug discovery. *J. Med. Chem.* **48**, 4040–4048 (2005).

72. Nunes, C. A., Freitas, M. P., Pinheiro, A. C. M. & Bastos, S. C. Chemoface: a novel free user-friendly interface for chemometrics. *J. Braz. Chem. Soc.* **23**, 2003–2010 (2012).

73. Dennington, R. D.; Keith, T. A.; Millam, M. J. GaussView 5.0. (2008).

74. Ranjith, D. & Ravikumar, C. SwissADME predictions of pharmacokinetics and drug-likeness properties of small molecules present in Ipomoea mauritiana Jacq. *~ 2063 ~ J. Pharmacogn. Phytochem.* **8**, 2063–2073 (2019).

75. Egan, W. J., Merz, K. M. & Baldwin, J. J. Prediction of drug absorption using multivariate statistics. *J. Med. Chem.* **43**, 3867–3877 (2000).

76. Lai, X. & Wichers, H. J. Phenylthiourea Binding to Human. 1–11 (2020).

77. Micillo, R. *et al.* Conjugation with dihydrolipoic acid imparts caffeic acid ester potent inhibitory effect on dopa oxidase activity of human tyrosinase. *Int. J. Mol. Sci.* **19**, (2018).

78. Nithitanakool, S., Pithayanukul, P., Bavovada, R. & Saparpakorn, P. Molecular docking studies and anti-tyrosinase activity of thai mango seed kernel extract. *Molecules* **14**, 257–265 (2009).

79. Roy, K., Kar, S. & Ambure, P. On a simple approach for determining applicability domain of QSAR models. *Chemom. Intell. Lab. Syst.* **145**, 22–29 (2015).

80. Tropsha, A. Best Practices for QSAR Model Development, Validation, and Exploitation. *Mol. Inform.* **29**, 476–488 (2010).

81. Roy, K. *et al.* Some case studies on application of " r m 2 " metrics for judging quality of quantitative structure-activity relationship predictions: Emphasis on scaling of response data. *J. Comput. Chem.* **34**, 1071–1082 (2013).

82. Barigye, S. J., Duarte, M. H., Nunes, C. A. & Freitas, M. P. MIA-plot: a graphical tool for viewing descriptor contributions in MIA-QSAR. *RSC Adv.* **6**, 49604–49612 (2016).

83. Haddad, Y., Adam, V. & Heger, Z. Ten quick tips for homology modeling of high-resolution protein 3D structures. *PLOS Comput. Biol.* **16**, e1007449 (2020).

84. Sándor, M., Kiss, R. & Keseru, G. M. Virtual fragment docking by glide: A validation study on 190 protein-fragment complexes. *J. Chem. Inf. Model.* **50**, 1165–1172 (2010).

85. Friesner, R. A. *et al.* Extra precision glide: Docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. *J. Med. Chem.* **49**, 6177–6196 (2006).

86. Tripathi, S. K., Muttineni, R. & Singh, S. K. Extra precision docking, free energy calculation and molecular dynamics simulation studies of CDK2 inhibitors. *J. Theor. Biol.* **334**, 87–100 (2013).

87. Bowers, K. J. *et al.* Scalable algorithms for molecular dynamics simulations on commodity clusters. *Proc. 2006 ACM/IEEE Conf. Supercomput. SC'06* (2006) doi:10.1145/1188455.1188544.

88. Activities, A., Yang, K. & Li, D. Polyphenols from Melastoma normale. (2021).

89. Wagle, A. *et al.* Discovery of a Highly Potent Tyrosinase Inhibitor, Luteolin 5- O-β- d -Glucopyranoside, Isolated from Cirsium japonicum var. maackii (Maxim.) Matsum., Korean Thistle: Kinetics and Computational Molecular Docking Simulation. *ACS Omega* **3**, 17236–17245 (2018).

90. Marciane, M. *et al.* Schinus terebinthifolius : phenolic constituents and in vitro antioxidant , antiproliferative and in vivo anti-inflammatory activities. *Rev. Bras. Farmacogn.* **27**, 445–452 (2017).

ANEXO 1

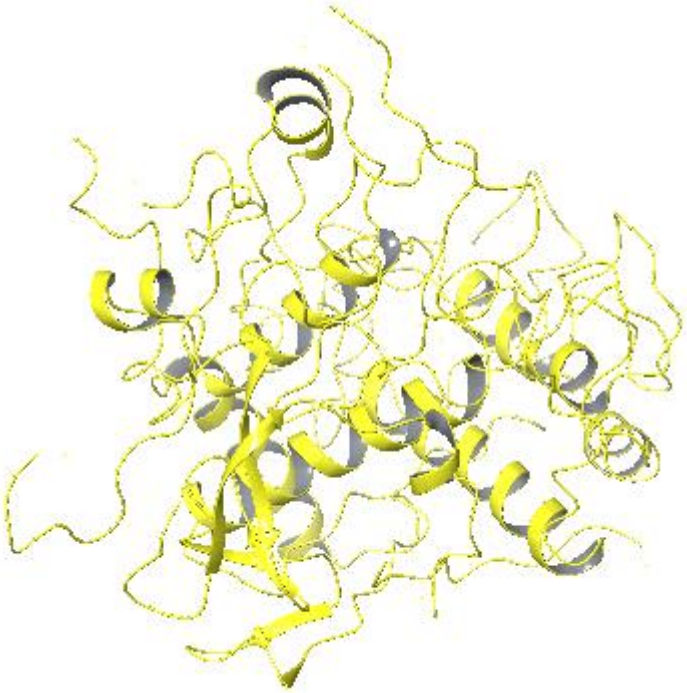


Figura 1: Modelo energy based - Maestro

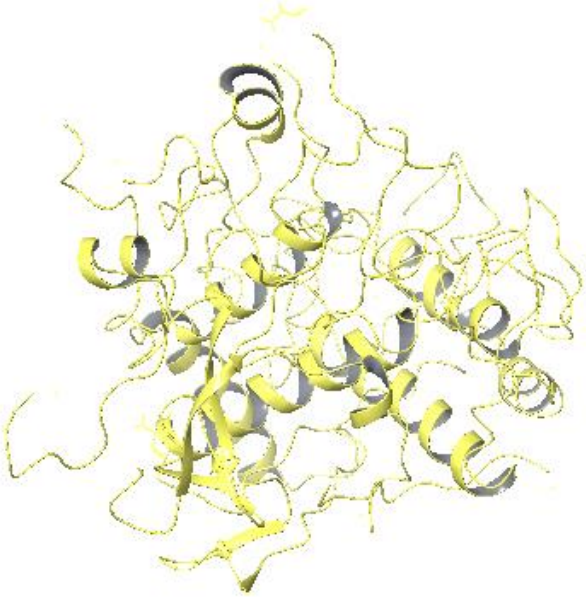


Figura 2: Modelo energy based – maestro após minimização.

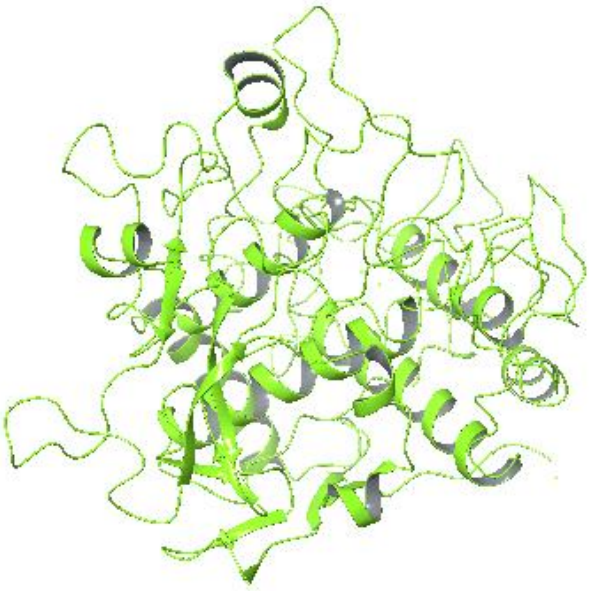


Figura 3: Modelo *knowledge based* - Maestro

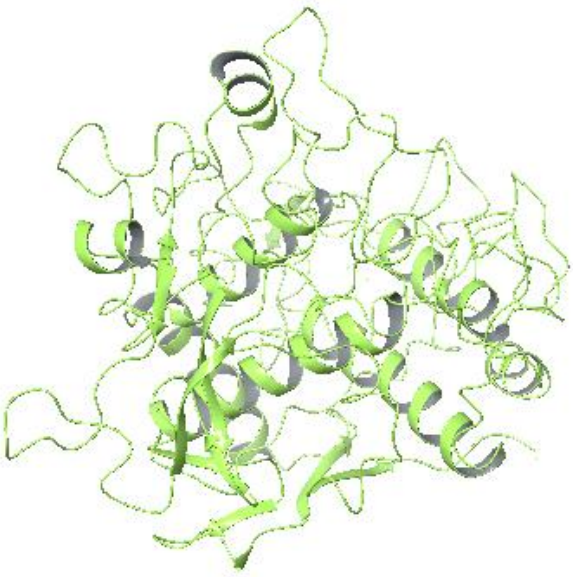


Figura 4:Modelo *knowledge based* – Maestro, após minimização

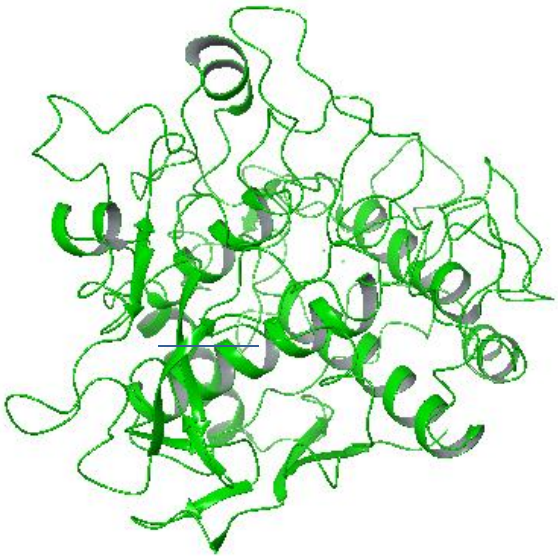


Figura 5: Modelo L8B082 (repositório) do Swiss-Model

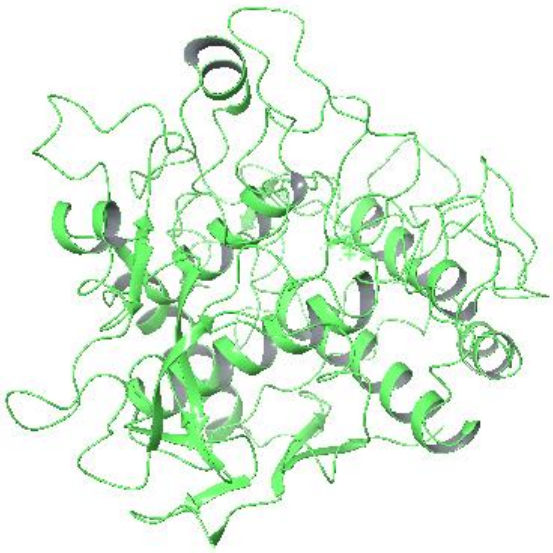


Figura 6: Modelo L8B082 (repositório) do Swiss-Model após minimização.

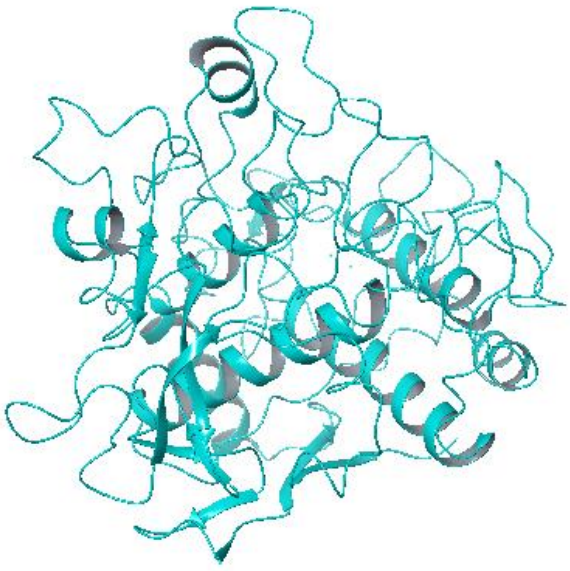


Figura 7: Modelo Swiss-Model

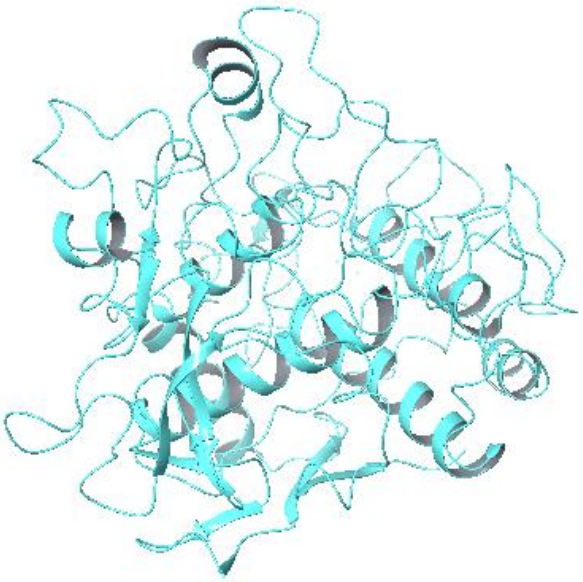


Figura 8: Modelo Swiss-Model após minimização

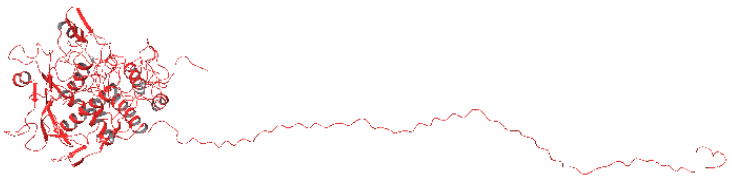


Figura 9: Modelo Modeller

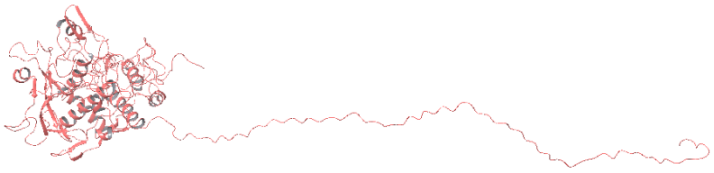


Figura 10: Modelo Modeller após minimização

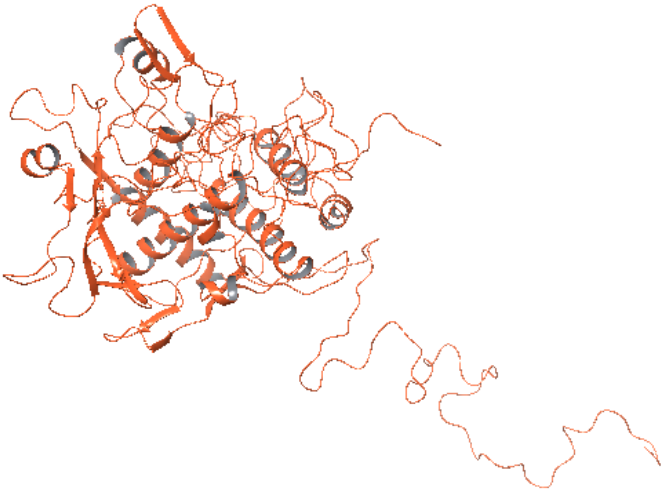


Figura 12: modelo Modeller após aplicado o *script* loop (100x)

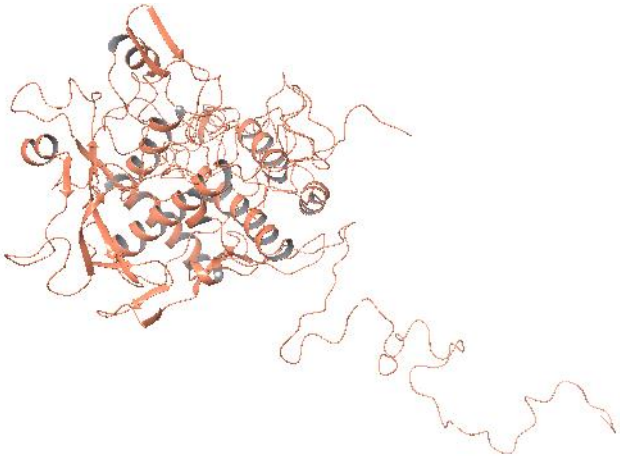


Figura 11: Modelo Modeller após aplicado *script* loop (100x) e minimizado

Figura 13: Modelo P14679 (repositório do Swiss-Model)

Figura 14: Modelo P14679 (repositório do Swiss-Model) após minimização

ANEXO 2

Compostos *docking* molecular

Tabela 7: O país da planta, o solvente utilizado na extração e a quantidade de cada componente do extrato de *Schinus terebinthifolius* de acordo com sua respectiva referência.

Referência	Ceruk et al[18]	El massry et al[20]	Farag et al[19]	Marciane et al[90]	Da Silva et al[21]
País	Brasil	Egito	Egito	Brasil	Brasil
Solvente	Etanol	Etanol	Acetona	Metanol	Metanol
Composto					
Ácido Kójico **(1)**					
Metil galato **(2)**	803.33 mg			28%	19.40 mg
Miricetina **(3)**	289.00 mg				
Etil Galato (4)	691.67 mg				
Catequin **(5)**		trace	9.5 mg		
Quercetina ramnosídeo **(6)**	112.00 mg				17.8 mg
Miricetina ramnosídeo **(7)**	172.00 mg				
Ácido elágico **(8)**		0.54 mg			
Quercetina **(9)**					22.8 mg
Ácido cafeico**(10)**		5.07 mg			
Ácido siríngico **(11)**		1.59 mg			
Ácido coumárico **(12)**		1.64 mg			
1,6-Digaloil-ß-D-glucose **(13)**			3.0 mg		
Ácido clorogênico **(14)**			37.0 mg		
Ácido gálico **(15)**		0.27 mg		4%	
Ácido 5-p-coumaroilquinico **(16)**			27.0 mg		
Luteolina **(17)**					26 mg
Miricetina ramnosídeo D- galactosídeo **(18)**			8.4 mg		
Miricetina glucoronídeo **(19)**			18.3 mg		
Miricetina galactosídeo **(20)**			8.0 mg		
Robustaflavona **(21)**					12.2 mg
1,2,3,4,6-penta-O-galloyl-glucopiranosídeo **(22)**				26%	25.4 mg

Metodologia de extração

Ceruk et al

Após secagem e moagem, 158 g de folhas de *S. terebinthifolius* foram submetidas à extração por maceração com etanol à temperatura ambiente. Após a remoção do solvente sob

vácuo, obtiveram-se 40,0 g de um material viscoso esverdeado. O extrato foi dissolvido em n-BuOH-H2O (1:1), a partir do que se verificou que os compostos ativos estavam concentrados na fase n-BuOH. Após evaporação do solvente, uma solução de AcOEt:H2O (1:1) foi adicionada a esta fase. A fase ativa de EtOAc (9,8 g) foi submetida à cromatografia de exclusão (Sephadex LH-20 com MeOH), resultando em 34 frações de 15 ml cada. Cada uma foi submetida à análise de RMN. Os espectros de Ressonância Magnética Nuclear foram registrados em um espectrômetro Brüker DPX-300 (1H: 300 MHz e 13C: 75 MHz) usando DMSO-d6 e metanol-d4 como solventes e o sinal do solvente residual como referência interna.

El-massry et al

Folhas de plantas frescas (200 g) foram embebidas em 500 ml de etanol 95% em temperatura ambiente por 2 dias no escuro. Após a filtração da solução, o solvente foi removido por evaporador rotativo sob pressão reduzida (95 mmHg) a 50°C. O extrato condensado foi armazenado no escuro a -5°C até novos experimentos.

A análise do extrato de etanol e dos compostos padrão foi realizada com um sistema de HPLC modelo Agilent 1100 equipado com um Phenomenex Aqua C18 (5 μm) 250 mm x 4,6 mm i.d. coluna (Torrance, CA) e um sistema de detecção de arranjo de diodos G1315A. A fase móvel consistia em 3,5% de ácido acético em água (A) e 3,5% de ácido acético em metanol (B). O modo gradiente foi inicialmente fixado em 20% B, então linearmente aumentado para 100% B em 40 min e mantido até o final (40 a 85 min). O volume de injeção para todas as amostras foi de 100 μl. A monitorização simultânea foi realizada a 280 nm a um caudal de 1 ml/min.

Farag et al

As folhas de *Schinus terebinthifolius Raddi* foram coletadas durante a fase de floração em julho de 2006 da Estação Experimental de Plantas Medicinais da Faculdade de Agricultura da Universidade de Assiut, Assiut, Egito. Folhas em pó secas ao ar (2,0 kg) foram extraídas com acetona aquosa a 70% (3 x 6 L) e os extratos combinados, filtrados e concentrados. O extrato foi suspenso em H2O (1 L) e particionado sucessivamente com Et2O (3 x 1 L), EtOAc (3 x 1 L) e nBuOH saturado com H2O (3 x 1 L). Cada fase foi concentrada sob pressão reduzida para dar a fração solúvel correspondente (17,1 g), (114,7 g) e (73,0 g), respectivamente, além de 14,5 g como resíduo solúvel em água.

Cerca de 63,0 g da fração n-BuOH foi aplicada a uma coluna de Diaion HP-20 e eluída sucessivamente com H2O, 20% MeOH, 40% MeOH, 60% MeOH e 100% MeOH para produzir 5 frações.

Marciane et al

Folhas inteiras de *S. terebinthifolius* foram coletadas em julho de 2000 no Rio de Janeiro, Brasil, com um exemplar (RB444.626) depositado no Herbário Jardim Botânico do Rio de Janeiro. O material vegetal foi moído e extraído exaustivamente com metanol por maceração estática. A solução foi filtrada e concentrada em rotaevaporador e o extrato foi suspenso em água-metanol (9:1) e particionado, por sua vez, com diclorometano (20 mL × 3) e acetato de etila (20 mL × 3). A fração de acetato de etila foi evaporada do extrato de metanol sob pressão reduzida para 9,36% (p/p) e, na sequência, analisada por HPLC (Shimadzu LC-10 ADVP) com um detector de arranjo de diodos (SPDM10AVP) e um autoinjetor (SIL-10ADVP). A análise foi realizada em uma coluna de supelcosil (LC18) e as seguintes condições do sistema solvente foram usadas: (A) acetonitrila/metanol (75:25) e (B) TFA 0,05%, vazão 1,0 ml/min, detecção a 220 nm .

Da Silva et al

As folhas de *Schinus terebinthifolius Raddi, Anacardiaceae,* foram coletadas no Jardim de Plantas Medicinais da Universidade Federal da Grande Dourados (22◦11 43,7S, 54◦56 08,5W e 430 m), em novembro de 2014.

As folhas (860 g) foram secas e extraídas por maceração com metanol, na sequência foram filtradas, concentradas sob pressão reduzida e liofilizadas para produzir o extrato metanólico (MEST) (42,7 g). O MEST (30 g) foi dividido com hexano, clorofórmio e acetato de etilo. A fração de clorofórmio (1,8 g) foi submetida a cromatografia em coluna (CC) de sílica gel. O fracionamento de parte da fração de acetato de etila (6,2 g) por CC em sílica gel foi realizado utilizando uma mistura de hexano/EtOAc e EtOAc/MeOH, em polaridade crescente. A fração hidrometanólica (13 g) foi purificada por CC sucessiva em Sephadex LH-20 usando H2O, H2O/MeOH 7:3–3:7 e MeOH como eluentes.

Os espectros 1H e 13CNMR foram coletados usando um espectrômetro Varian Mercury Plus BB operando a 300 MHz e 75,5 MHz usando CD3OD como solvente e tetrametilsilano (TMS) como padrão interno.

ANEXO 3

Neste anexo encontram-se as figuras com as poses de cada um dos programas, sendo o ligante em azul correspondente à melhor pose gerada pelo Glide, em amarelo, pelo Gold e em rosa, pelo Autodock; assim como o diagrama 2D de interação proteína ligante para a pose de ΔG mais negativo.

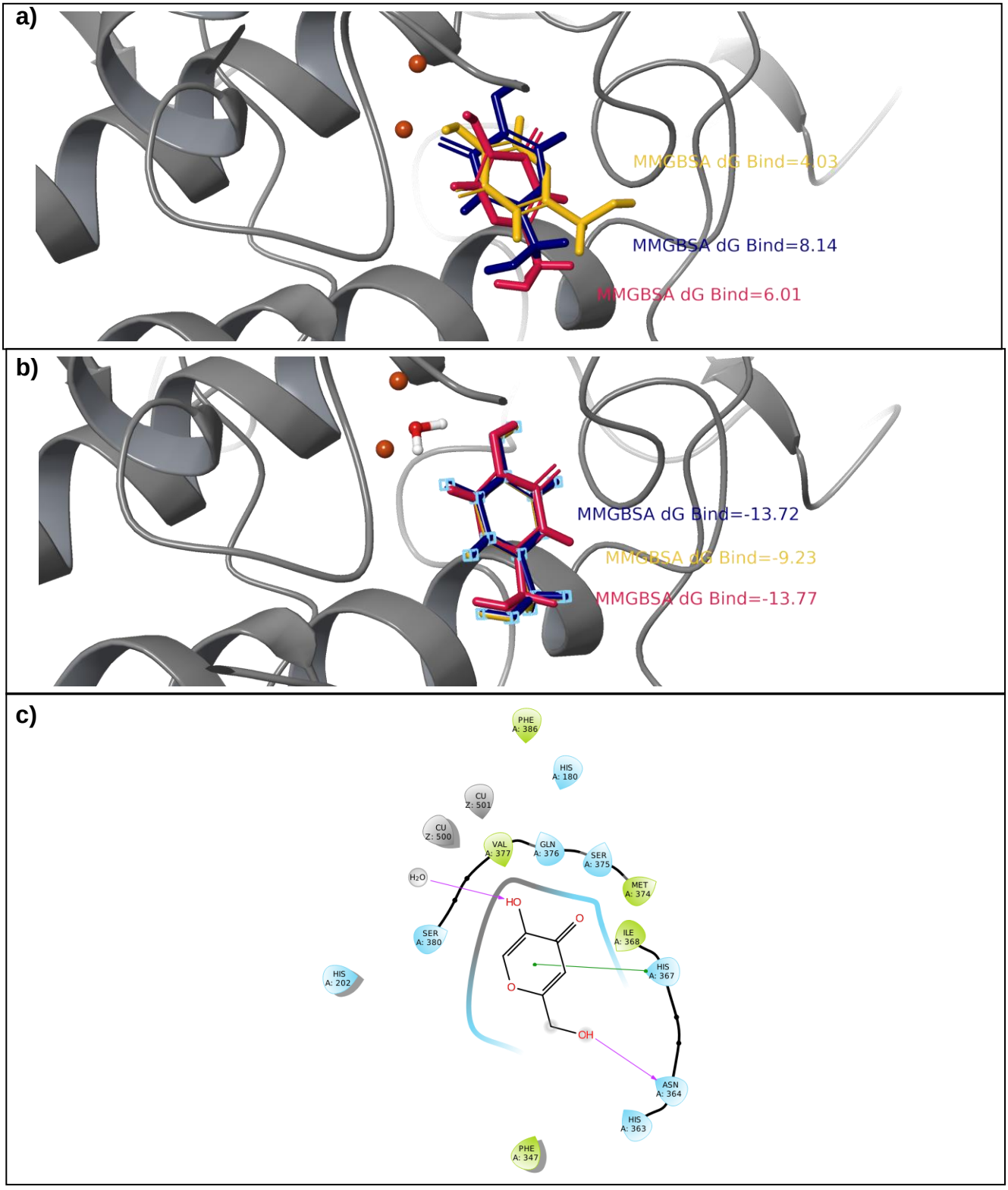


Figura 15:Vista 3D da pose de *docking* ΔG MM-GBSA mais negativa do ácido kójico (1) obtida por cada programa (Glide -azul, Gold -amarelo e Autodock- rosa) – a) sem água complexada com Cu^{2+} e b) com água complexada com Cu^{2+}. c) Diagrama de interação ligante da melhor pose - Glide com água ΔG = -13,76 kcal/mol

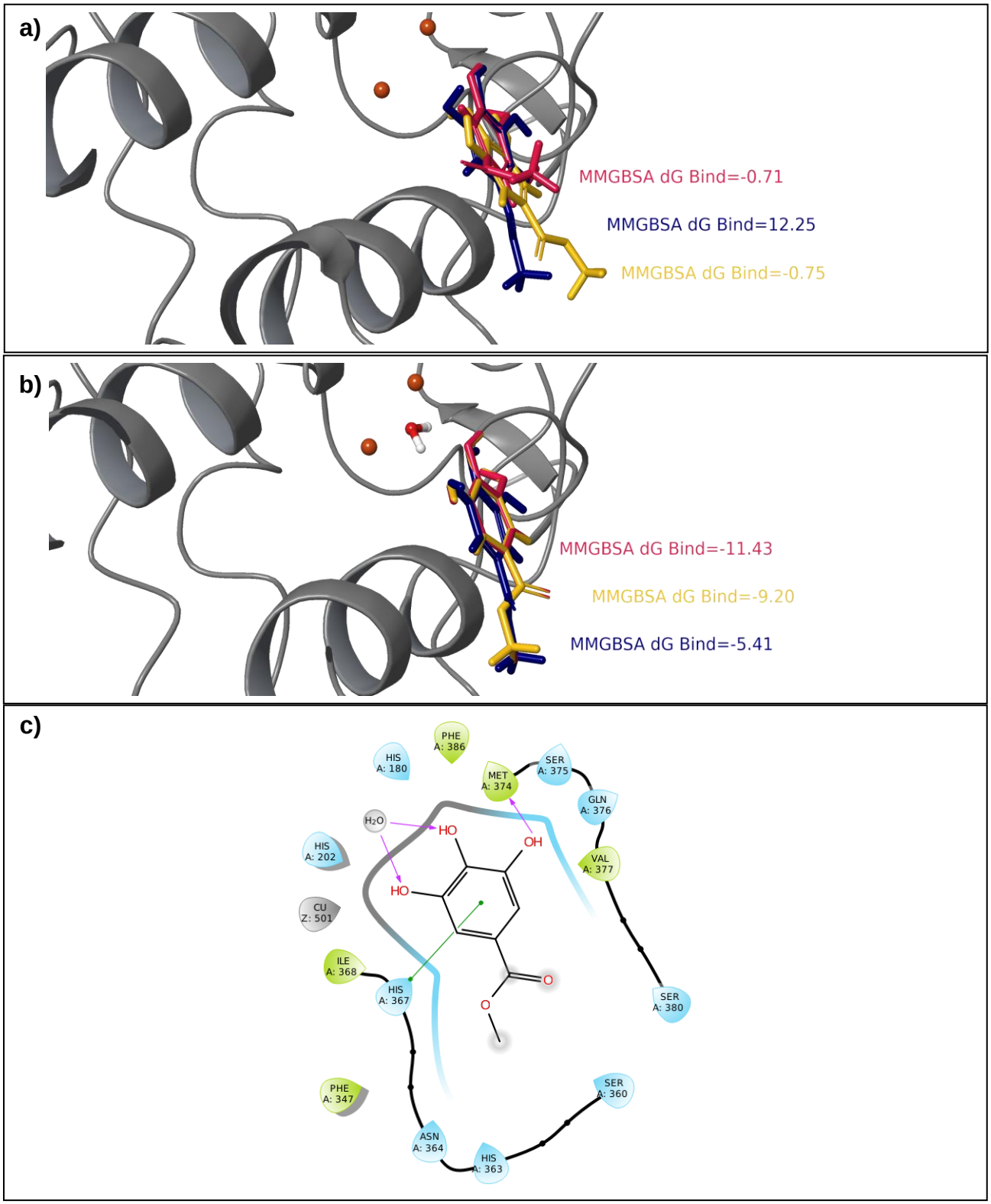


Figura 16: Vista 3D da pose de *docking* ΔG MM-GBSA mais negativa do metil galato (2) obtida por cada programa (Glide -azul, Gold -amarelo e Autodock- rosa) – a) sem água complexada com Cu^{2+} e b) com água complexada com Cu^{2+}. c) Diagrama de interação ligante da melhor pose - Autodock com água ΔG = -11,42 kcal/mol

:

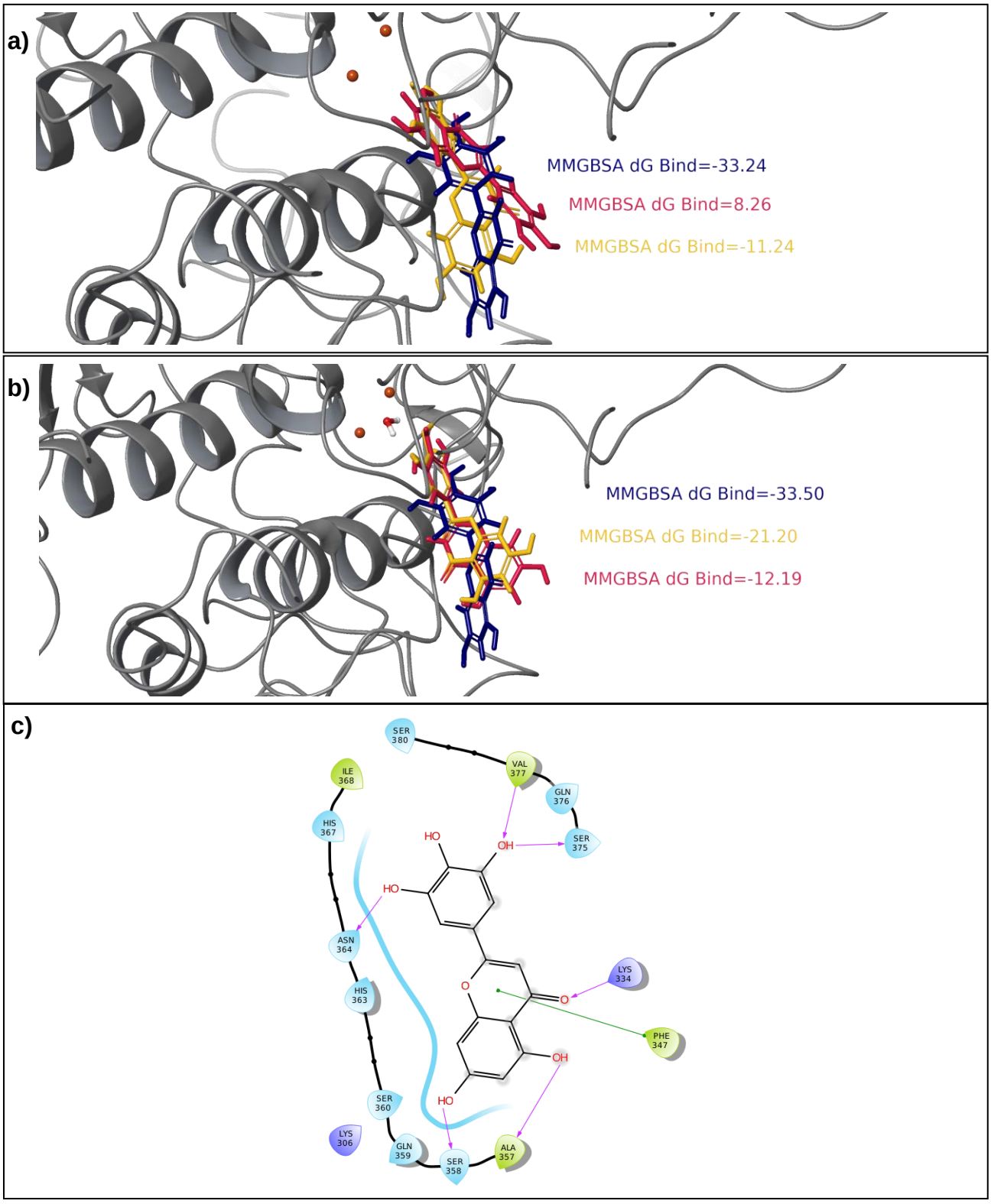


Figura 17:Vista 3D da pose de *docking* ΔG MM-GBSA mais negativa do miricetina (3) obtida por cada programa (Glide -azul, Gold -amarelo e Autodock- rosa) – a) sem água complexada com Cu^{2+} e b) com água complexada com Cu^{2+}. c) Diagrama de interação ligante da melhor pose - Glide com água ΔG = -33,49 kcal/mol

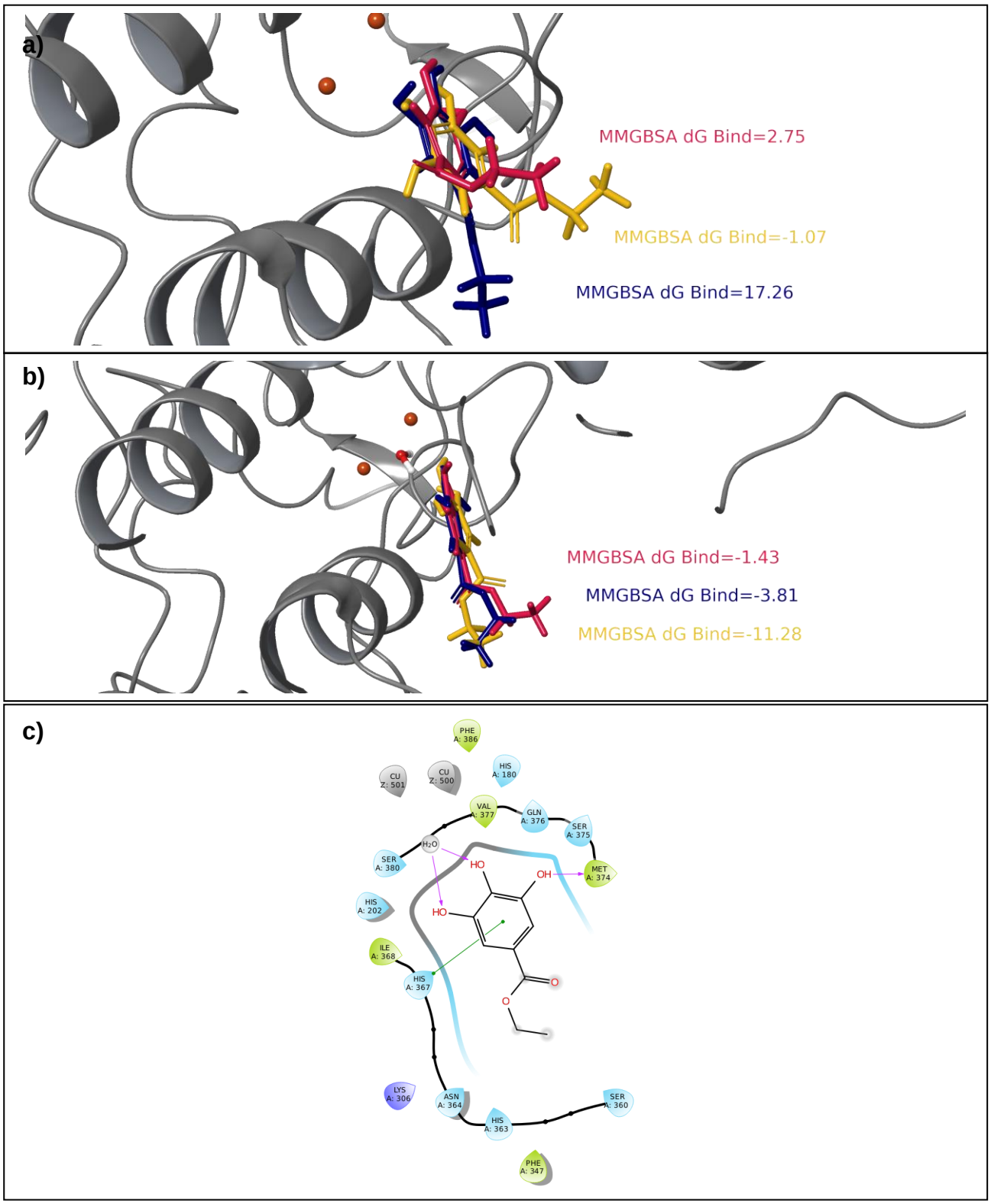


Figura 18: Vista 3D da pose de *docking* ΔG MM-GBSA mais negativa do etil galato (4) obtida por cada programa (Glide -azul, Gold -amarelo e Autodock- rosa) – a) sem água complexada com Cu^{2+} e b) com água complexada com Cu^{2+}. c) Diagrama de interação ligante da melhor pose - Glide sem água ΔG = -17,26 kcal/mol

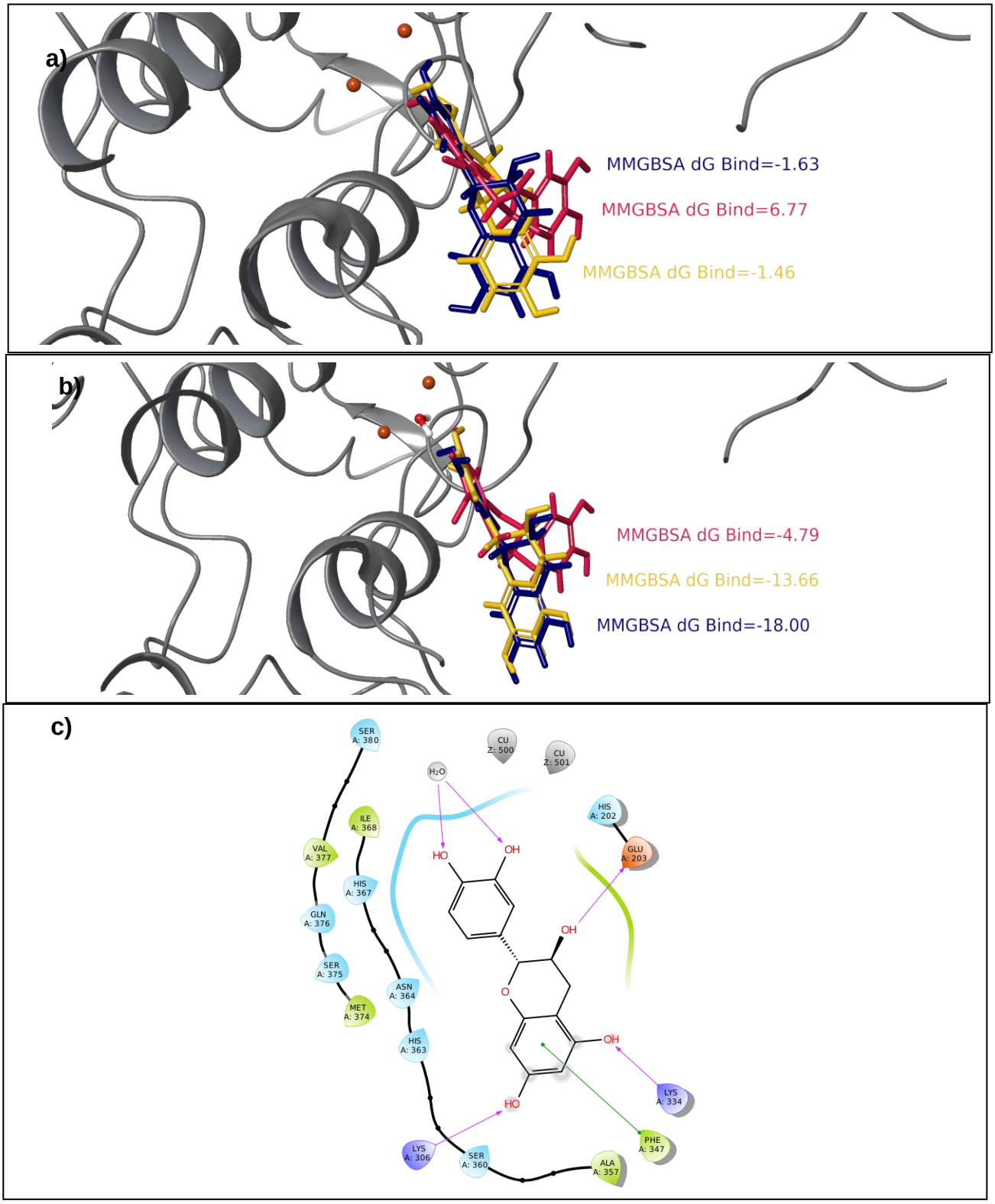


Figura 19: Vista 3D da pose de *docking* ΔG MM-GBSA mais negativa do catequina (5) obtida por cada programa (Glide -azul, Gold -amarelo e Autodock- rosa) – a) sem água complexada com Cu^{2+} e b) com água complexada com Cu^{2+}. c) Diagrama de interação ligante da melhor pose - Glide com água ΔG = -17,99 kcal/mol

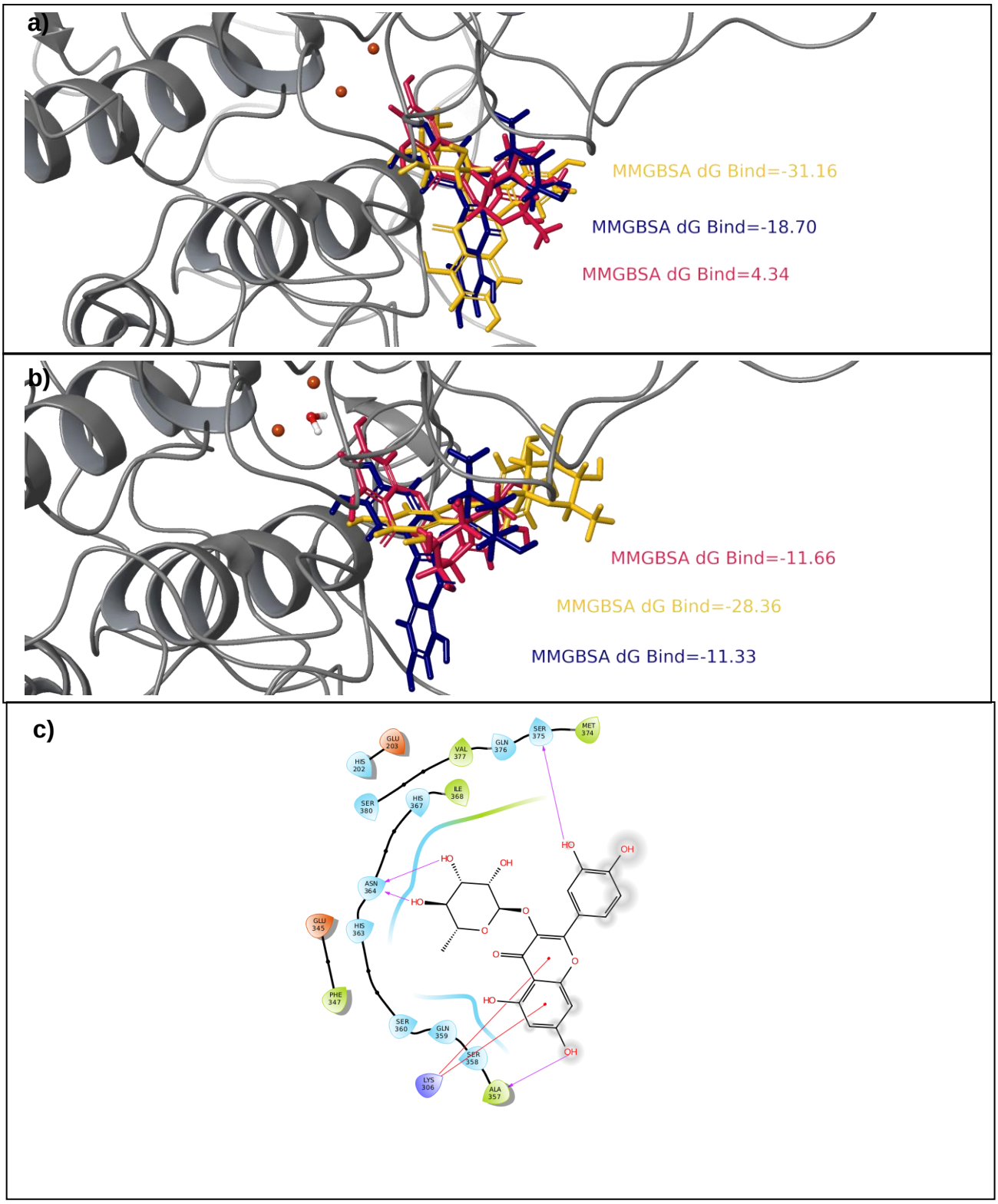


Figura 20: Vista 3D da pose de *docking* ΔG MM-GBSA mais negativa do quercetina ramnosídeo (6) obtida por cada programa (Glide -azul, Gold -amarelo e Autodock- rosa) – a) sem água complexada com Cu^{2+} e b) com água complexada com Cu^{2+}. c) Diagrama de interação ligante da melhor pose – Gold sem água ΔG = -31,16 kcal/mol

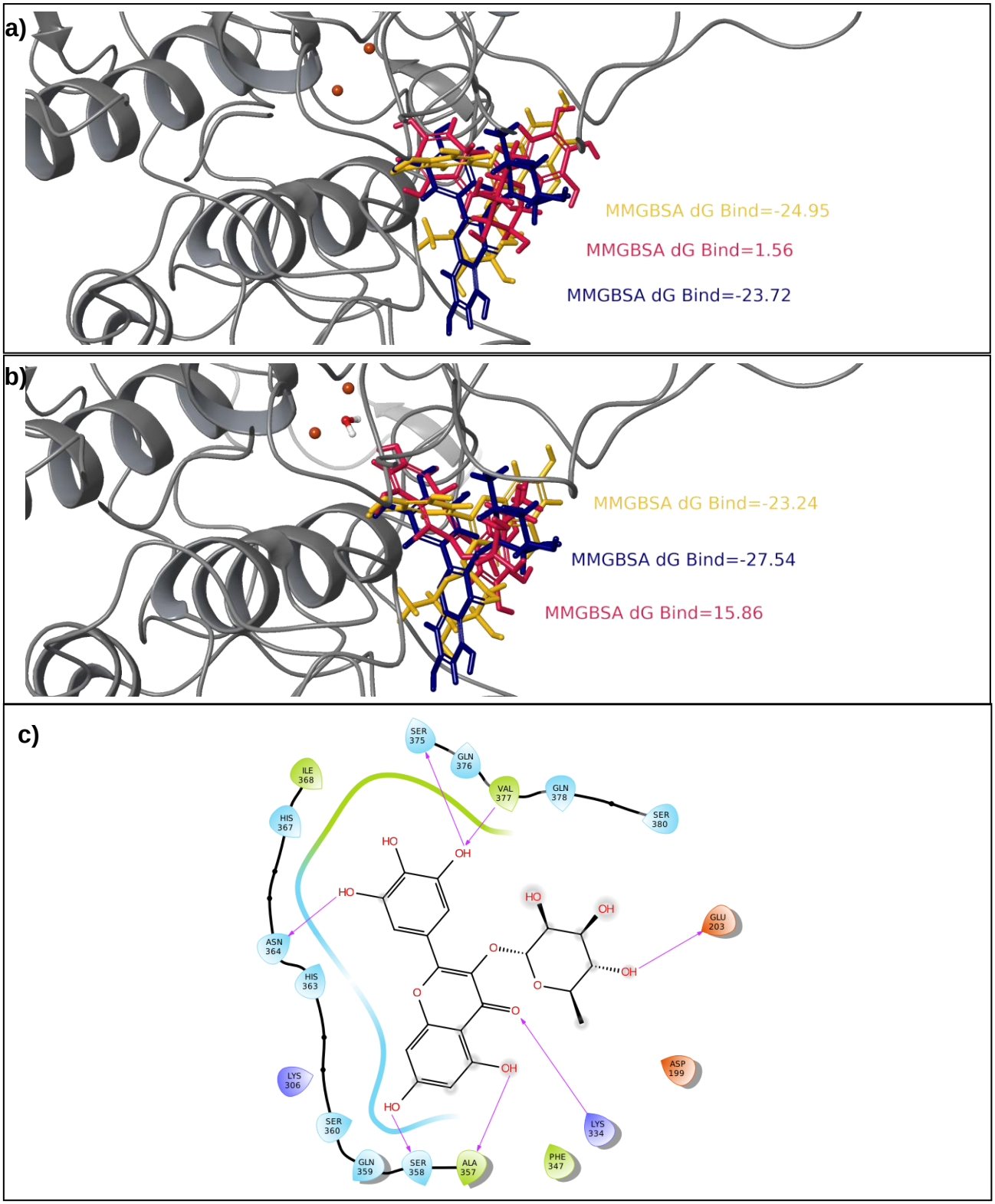


Figura 21: Vista 3D da pose de *docking* ΔG MM-GBSA mais negativa do miricetina ramnosídeo (7) obtida por cada programa (Glide -azul, Gold -amarelo e Autodock- rosa) – a) sem água complexada com Cu^{2+} e b) com água complexada com Cu^{2+}. c) Diagrama de interação ligante da melhor pose – Glide sem água ΔG = -27,71 kcal/mol

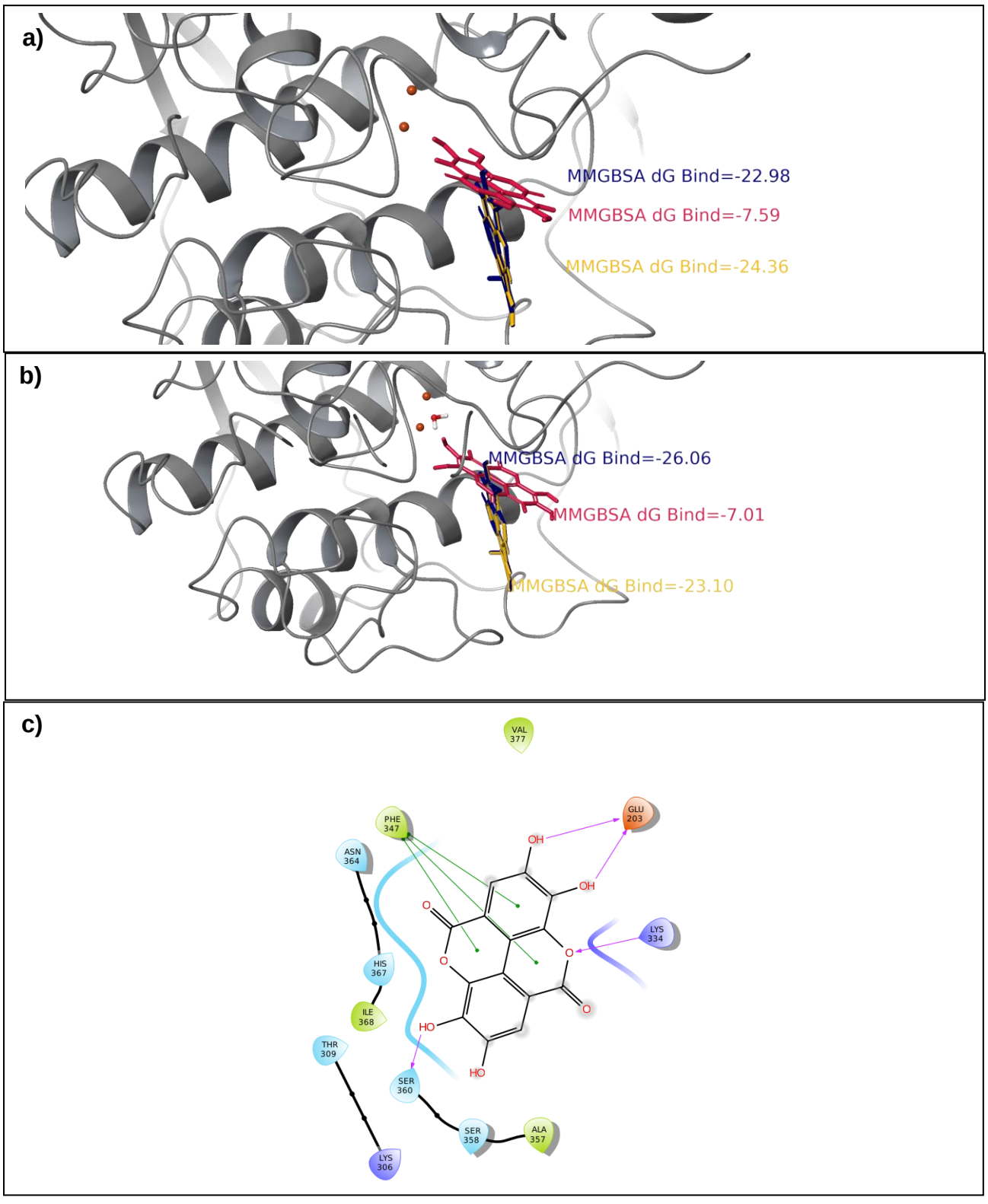


Figura 22: Vista 3D da pose de *docking* ΔG MM-GBSA mais negativa do ácido elágico (8) obtida por cada programa (Glide -azul, Gold -amarelo e Autodock- rosa) – a) sem água complexada com Cu^{2+} e b) com água complexada com Cu^{2+}. c) Diagrama de interação ligante da melhor pose – Glide com água ΔG = -26,05 kcal/mol

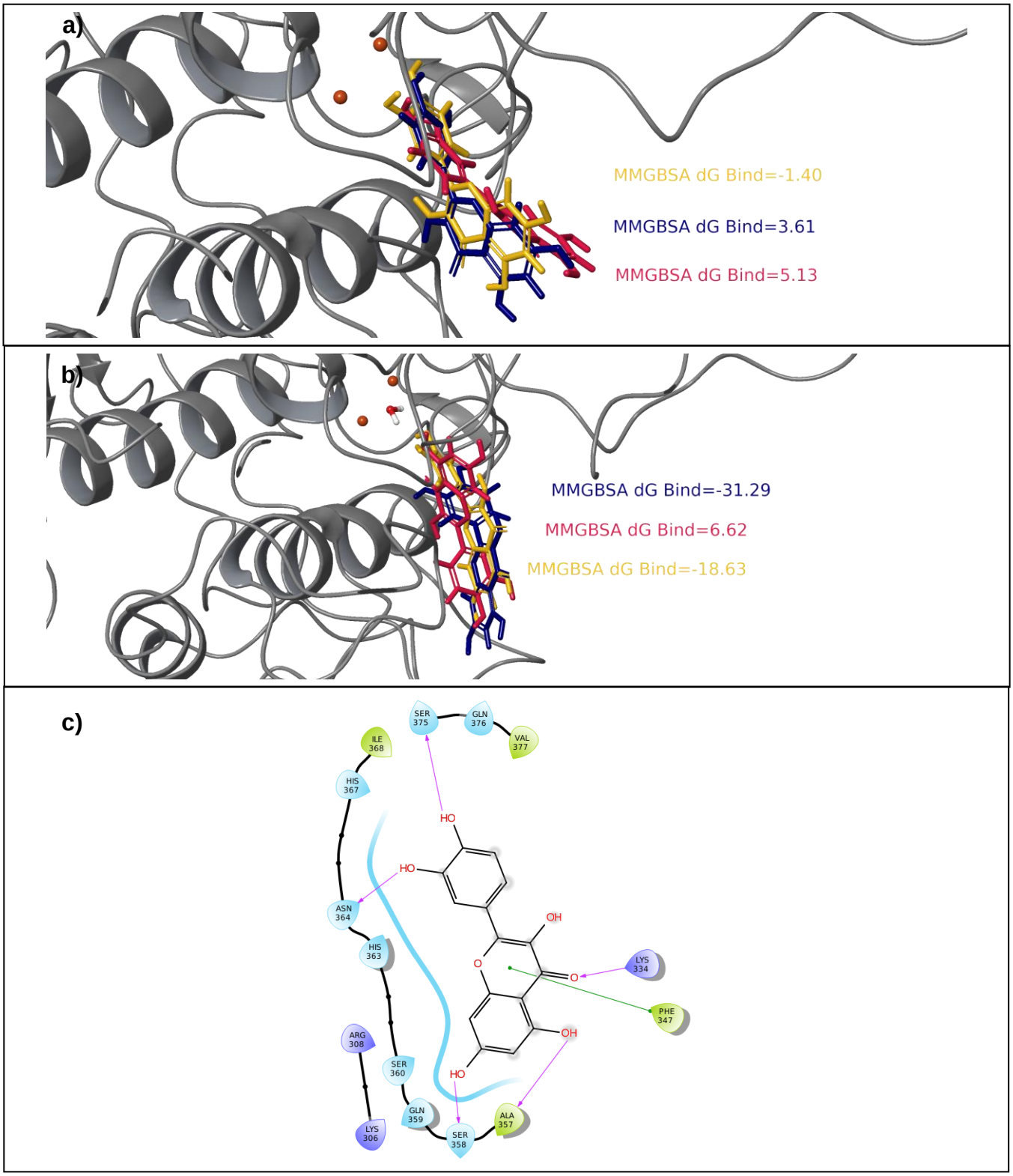


Figura 23: Vista 3D da pose de *docking* ΔG MM-GBSA mais negativa do quercetina (9) obtida por cada programa (Glide -azul, Gold -amarelo e Autodock- rosa) – a) sem água complexada com Cu^{2+} e b) com água complexada com Cu^{2+}. c) Diagrama de interação ligante da melhor pose – Glide com água ΔG = -31,28 kcal/mol

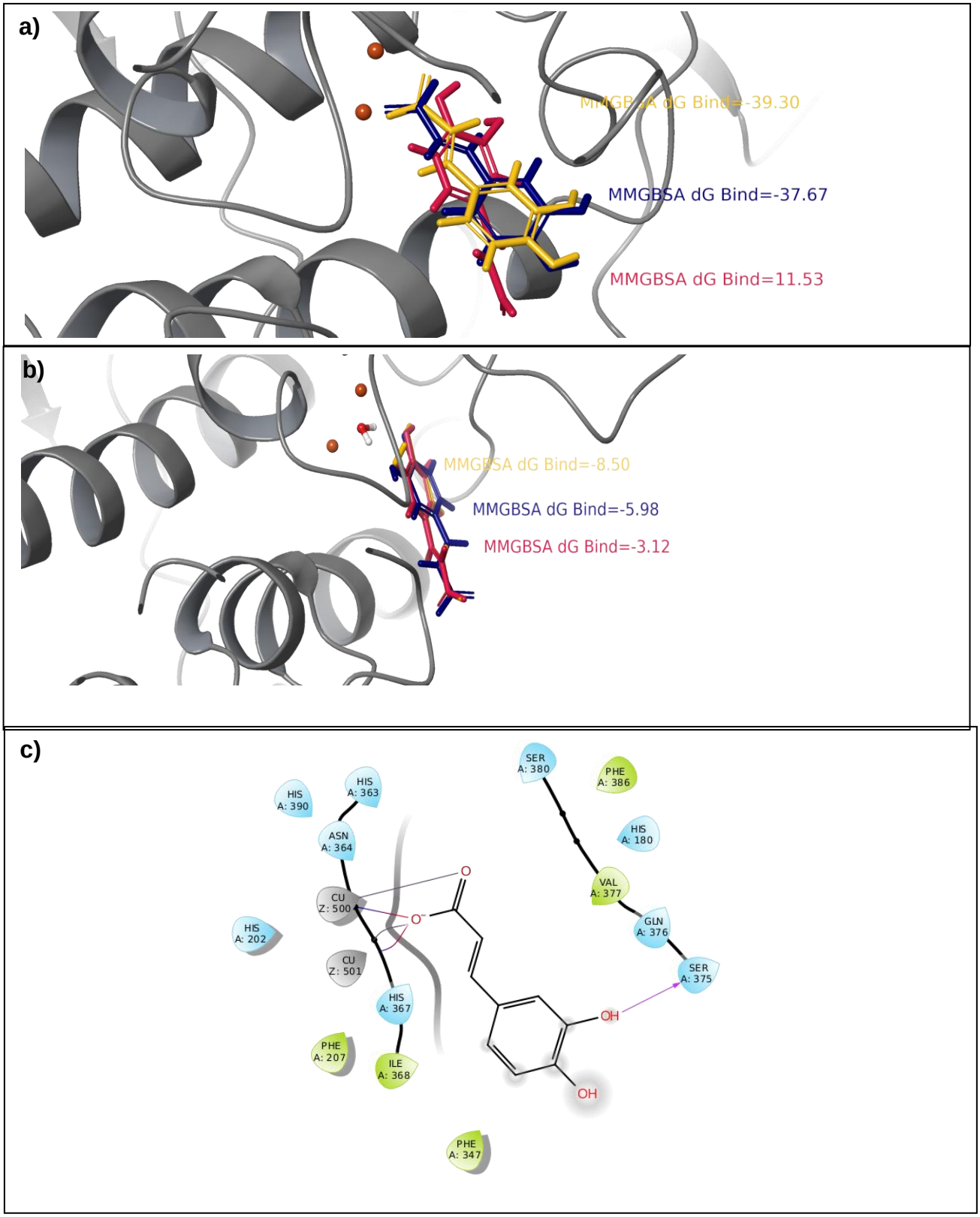


Figura 24: Vista 3D da pose de *docking* ΔG MM-GBSA mais negativa do ácido cafeico (10) obtida por cada programa (Glide -azul, Gold -amarelo e Autodock- rosa) – a) sem água complexada com Cu^{2+} e b) com água complexada com Cu^{2+}. c) Diagrama de interação ligante da melhor pose – Glide com água ΔG = -37,67 kcal/mol

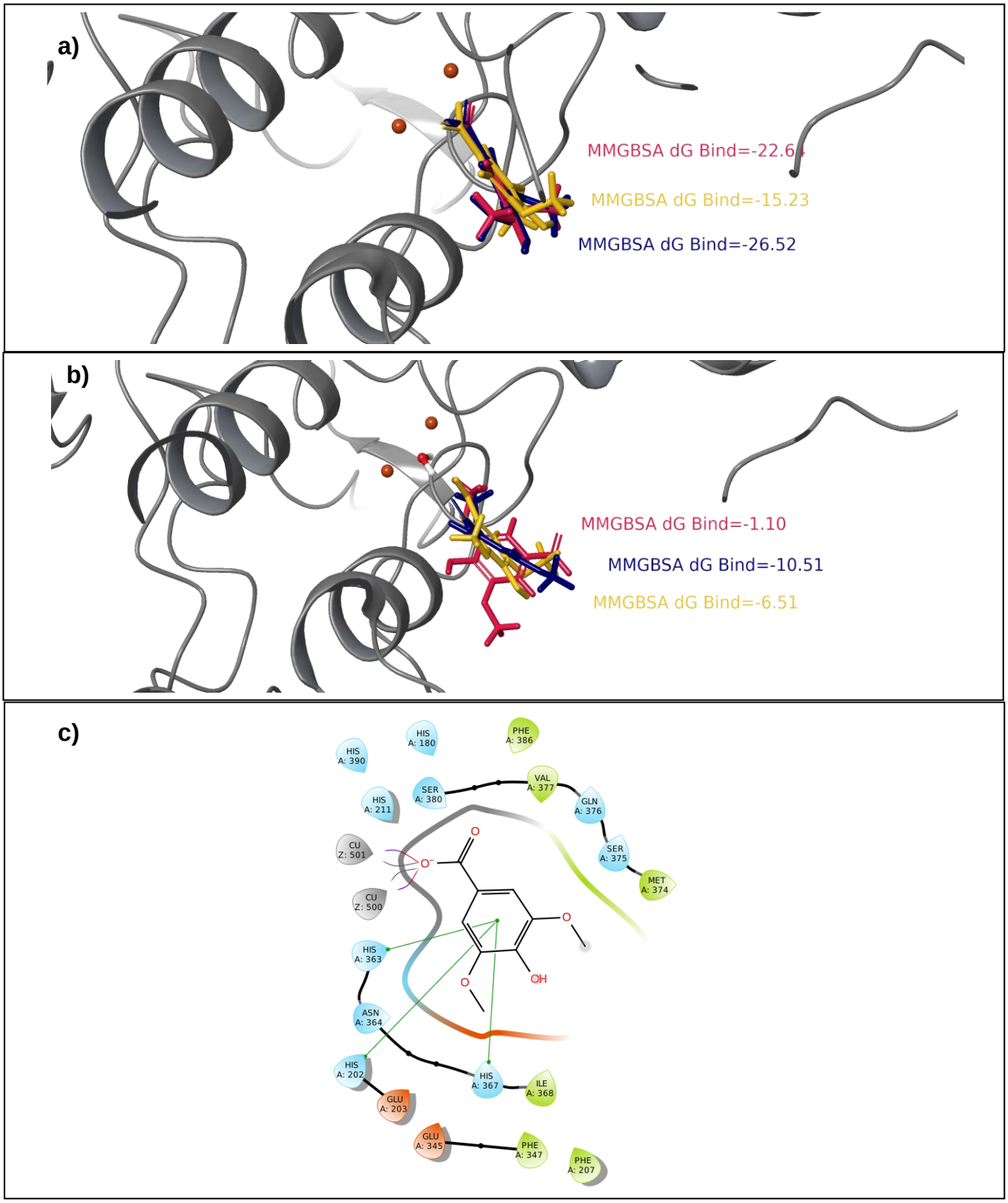


Figura 25: Vista 3D da pose de *docking* ΔG MM-GBSA mais negativa do ácido siríngico (11) obtida por cada programa (Glide -azul, Gold -amarelo e Autodock- rosa) – a) sem água complexada com Cu^{2+} e b) com água complexada com Cu^{2+}. c) Diagrama de interação ligante da melhor pose – Autodock sem água ΔG = -26,63 kcal/mol

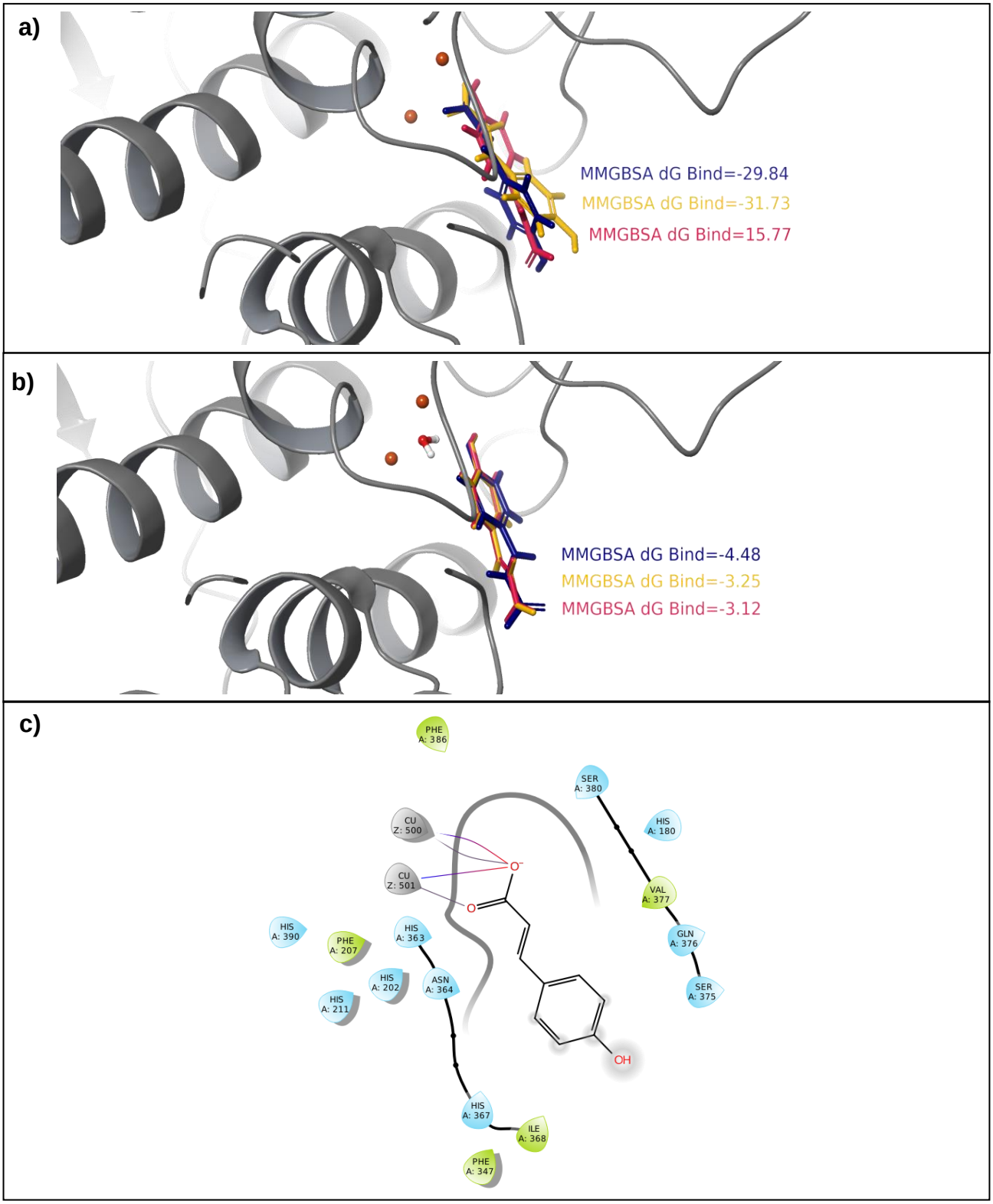


Figura 26: Vista 3D da pose de *docking* ΔG MM-GBSA mais negativa do ácido coumárico (12) obtida por cada programa (Glide -azul, Gold -amarelo e Autodock- rosa) – a) sem água complexada com Cu^{2+} e b) com água complexada com Cu^{2+}. c) Diagrama de interação ligante da melhor pose – Gold sem água ΔG = -31,73 kcal/mol

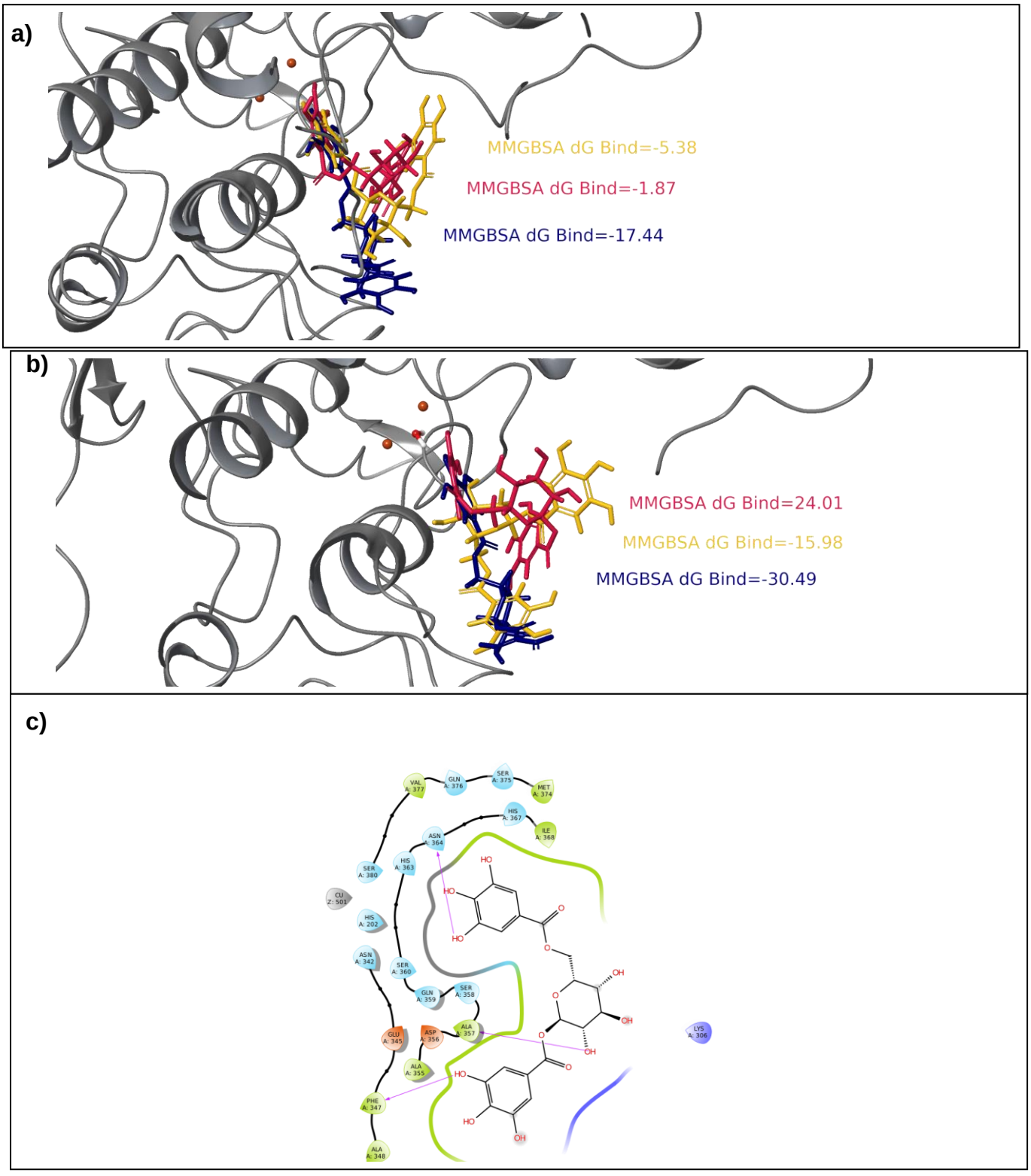


Figura 27: Vista 3D da pose de *docking* ΔG MM-GBSA mais negativa do 1,6-Digaloil-glucose (13) obtida por cada programa (Glide -azul, Gold -amarelo e Autodock- rosa) – a) sem água complexada com Cu^{2+} e b) com água complexada com Cu^{2+}. c) Diagrama de interação ligante da melhor pose – Glide com água ΔG = -30,48 kcal/mol

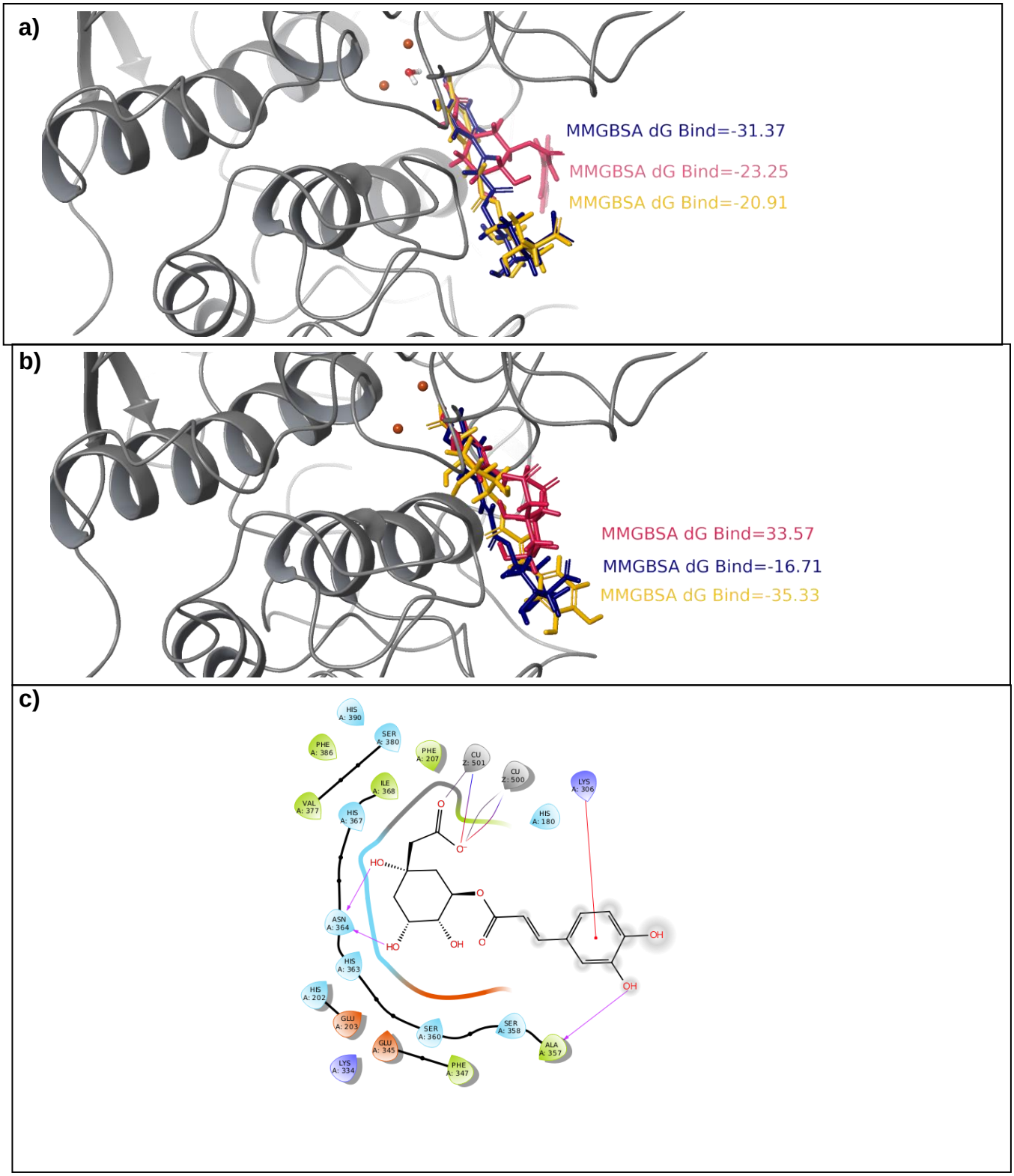


Figura 28: Vista 3D da pose de *docking* ΔG MM-GBSA mais negativa do ácido clorogênico (14) obtida por cada programa (Glide -azul, Gold -amarelo e Autodock- rosa) – a) sem água complexada com Cu^{2+} e b) com água complexada com Cu^{2+}. c) Diagrama de interação ligante da melhor pose – Gold sem água ΔG = -35,32 kcal/mol

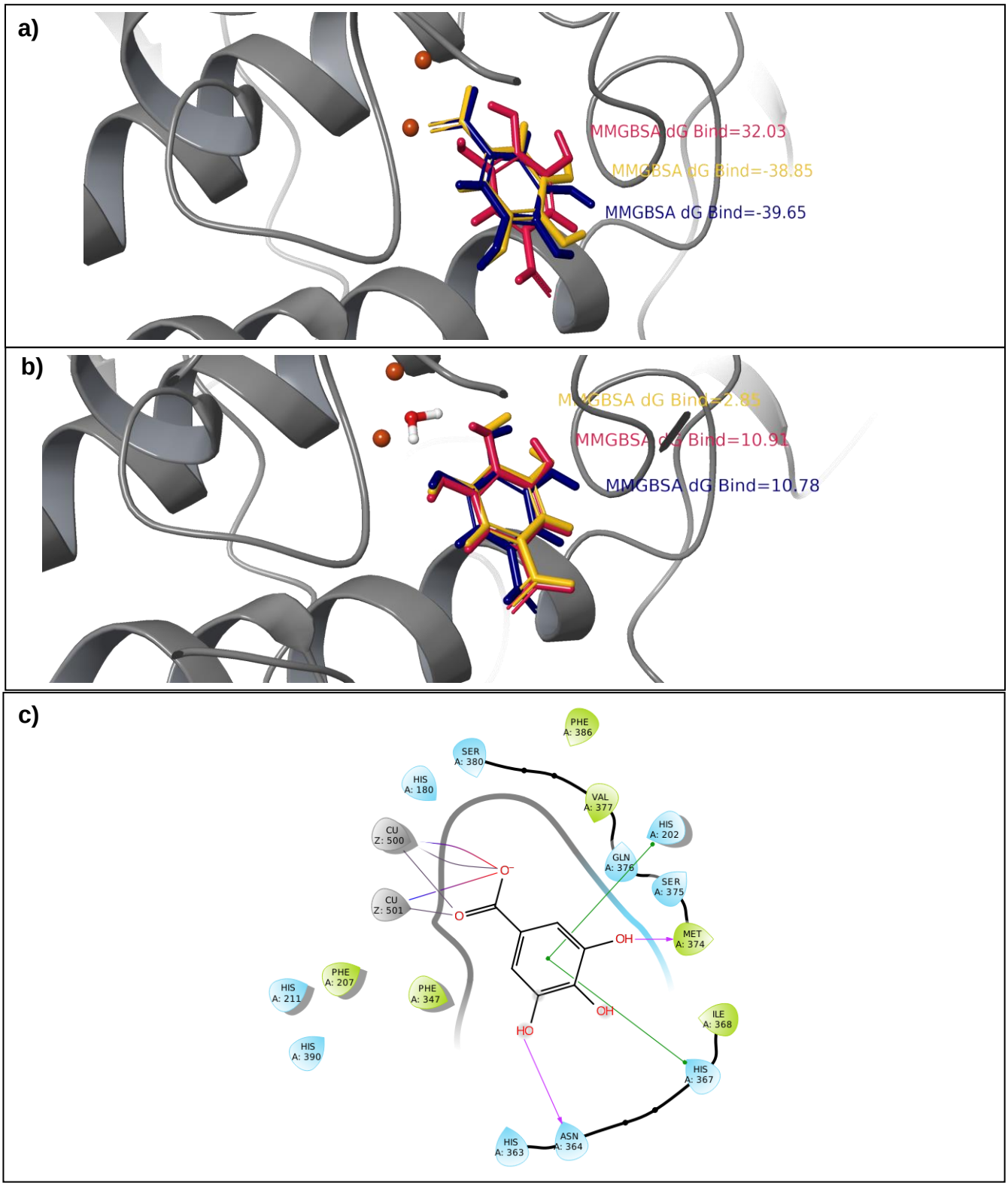


Figura 29: Vista 3D da pose de *docking* ΔG MM-GBSA mais negativa do ácido gálico (15) obtida por cada programa (Glide -azul, Gold -amarelo e Autodock- rosa) – a) sem água complexada com Cu^{2+} e b) com água complexada com Cu^{2+}. c) Diagrama de interação ligante da melhor pose – Gold sem água ΔG = -39,65 kcal/mol

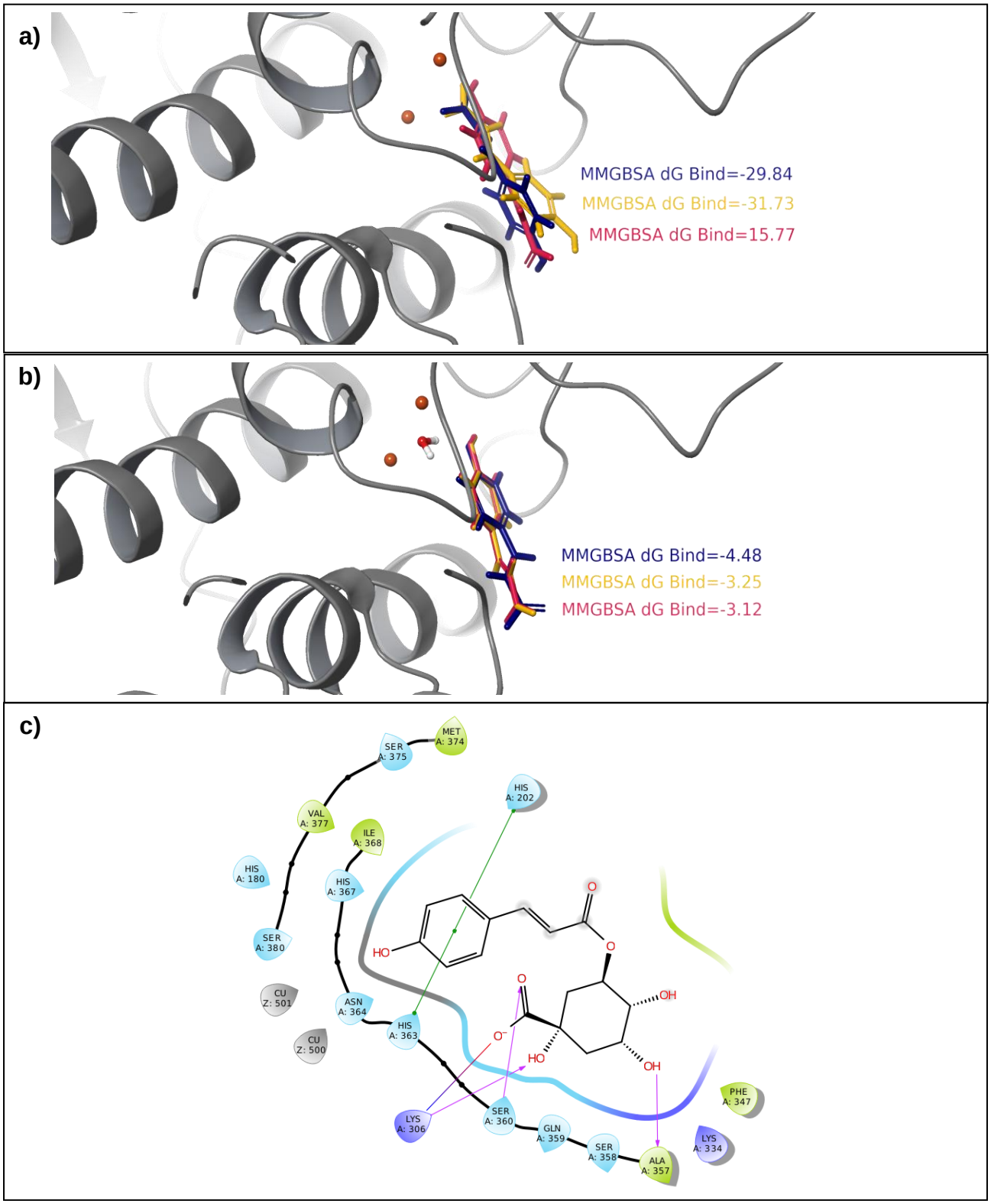


Figura 30: Vista 3D da pose de *docking* ΔG MM-GBSA mais negativa do ácido 5-p-coumaroilquinico (16) obtida por cada programa (Glide -azul, Gold -amarelo e Autodock- rosa) – a) sem água complexada com Cu^{2+} e b) com água complexada com Cu^{2+}. c) Diagrama de interação ligante da melhor pose – Glide com água ΔG = -15,57 kcal/mol

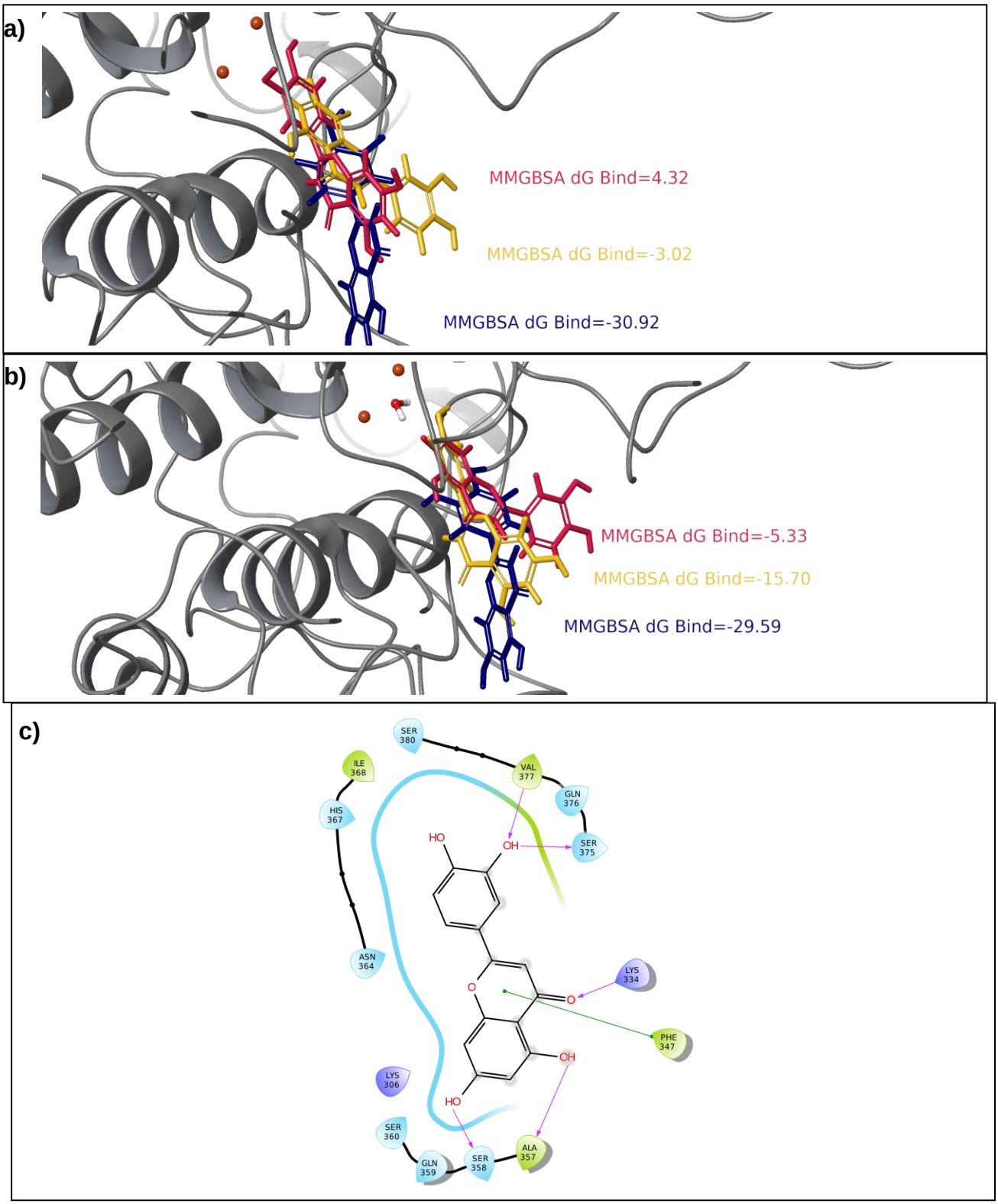


Figura 31: Vista 3D da pose de *docking* ΔG MM-GBSA mais negativa da luteolina (17) obtida por cada programa (Glide -azul, Gold -amarelo e Autodock- rosa) – a) sem água complexada com Cu^{2+} e b) com água complexada com Cu^{2+}. c) Diagrama de interação ligante da melhor pose – Glide sem água ΔG = -30,91 kcal/mol

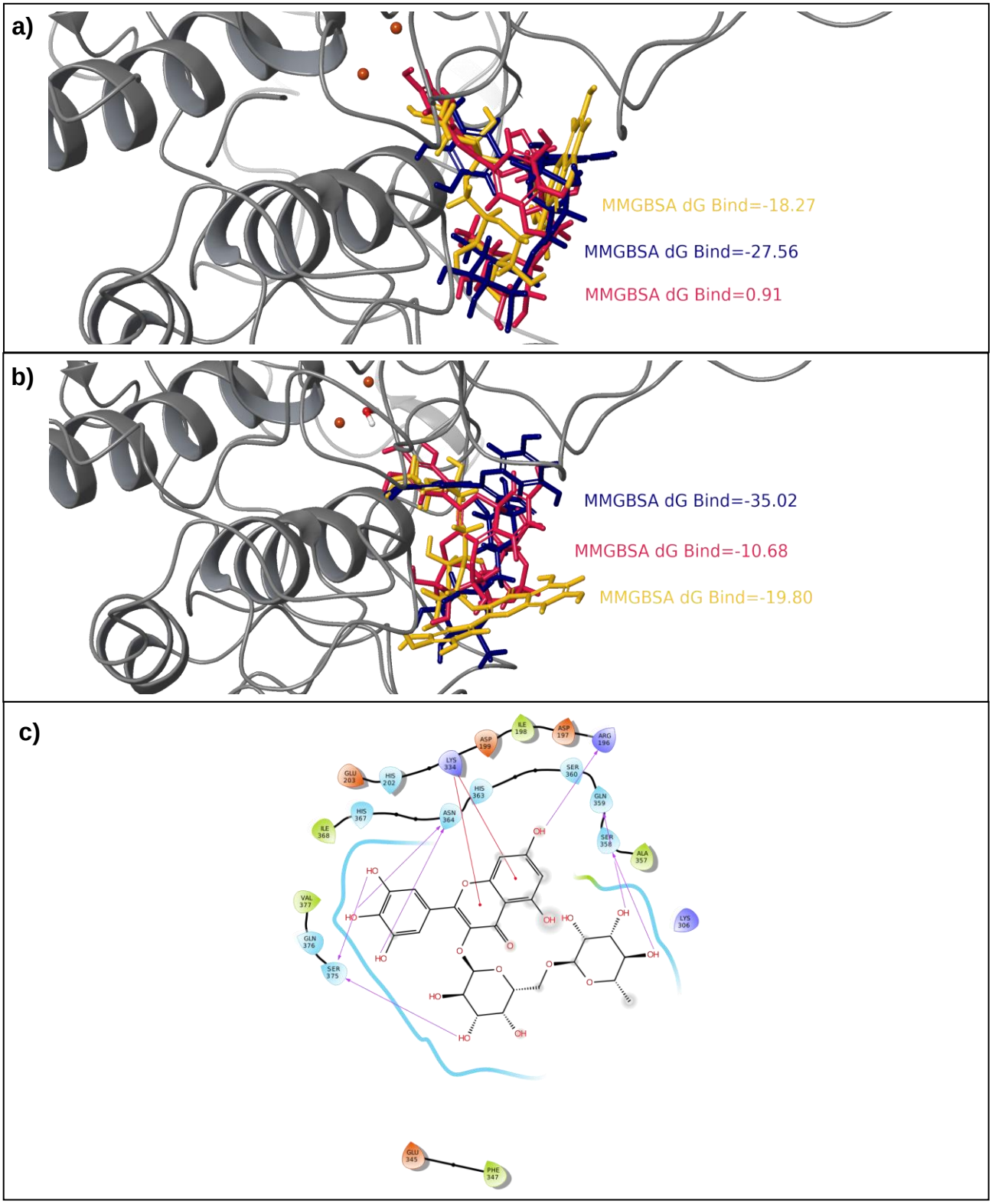


Figura 32: Vista 3D da pose de *docking* ΔG MM-GBSA mais negativa do miricetina ramnosídeo galactosídeo (18) obtida por cada programa (Glide -azul, Gold -amarelo e Autodock- rosa) – a) sem água complexada com Cu^{2+} e b) com água complexada com Cu^{2+}. c) Diagrama de interação ligante da melhor pose – Glide com água ΔG = -35,01 kcal/mol

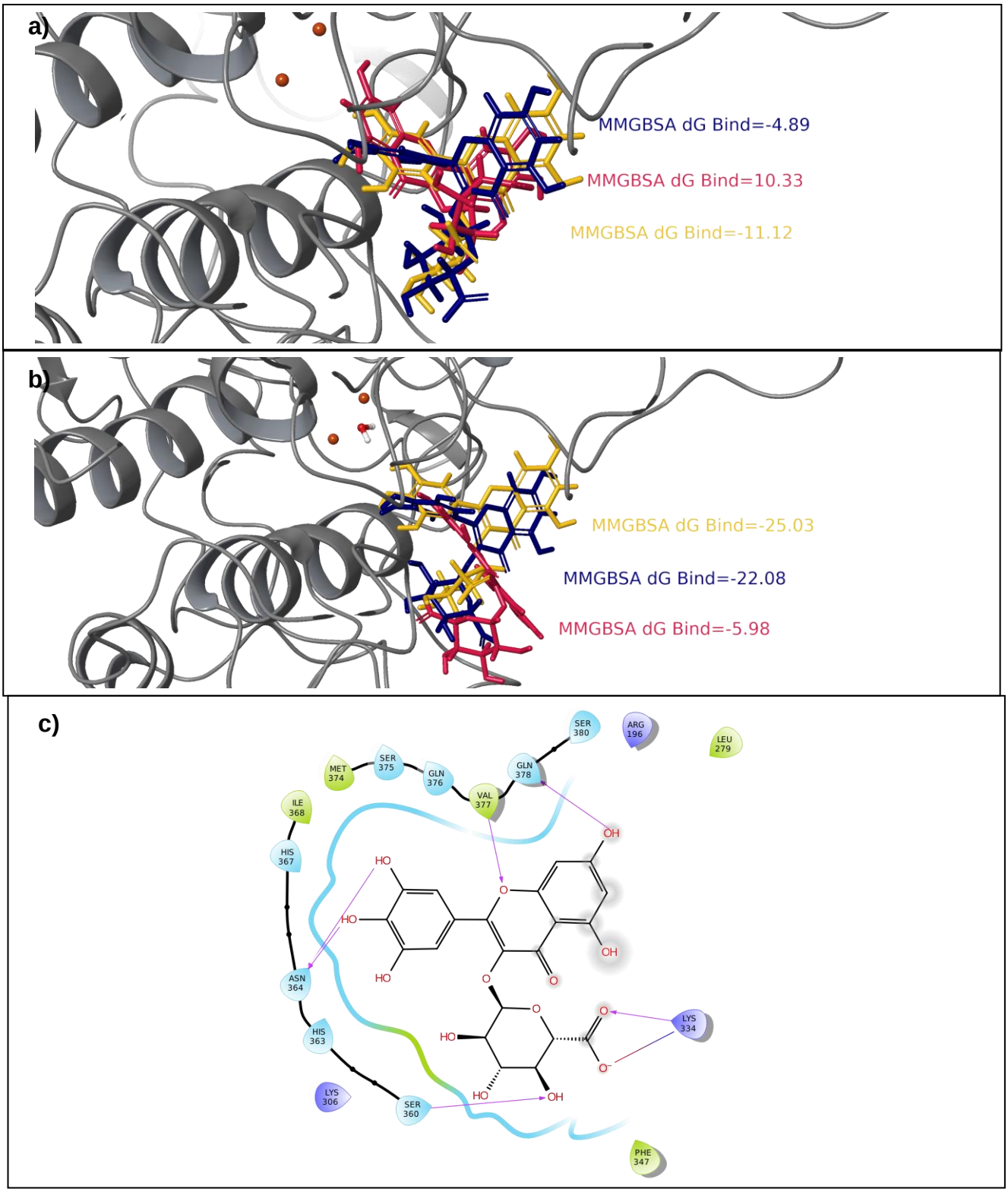


Figura 33: Vista 3D da pose de *docking* ΔG MM-GBSA mais negativa do miricetina glucoronídeo (19) obtida por cada programa (Glide -azul, Gold -amarelo e Autodock- rosa) – a) sem água complexada com Cu^{2+} e b) com água complexada com Cu^{2+}. c) Diagrama de interação ligante da melhor pose – Gold com água ΔG = -25,02 kcal/mol

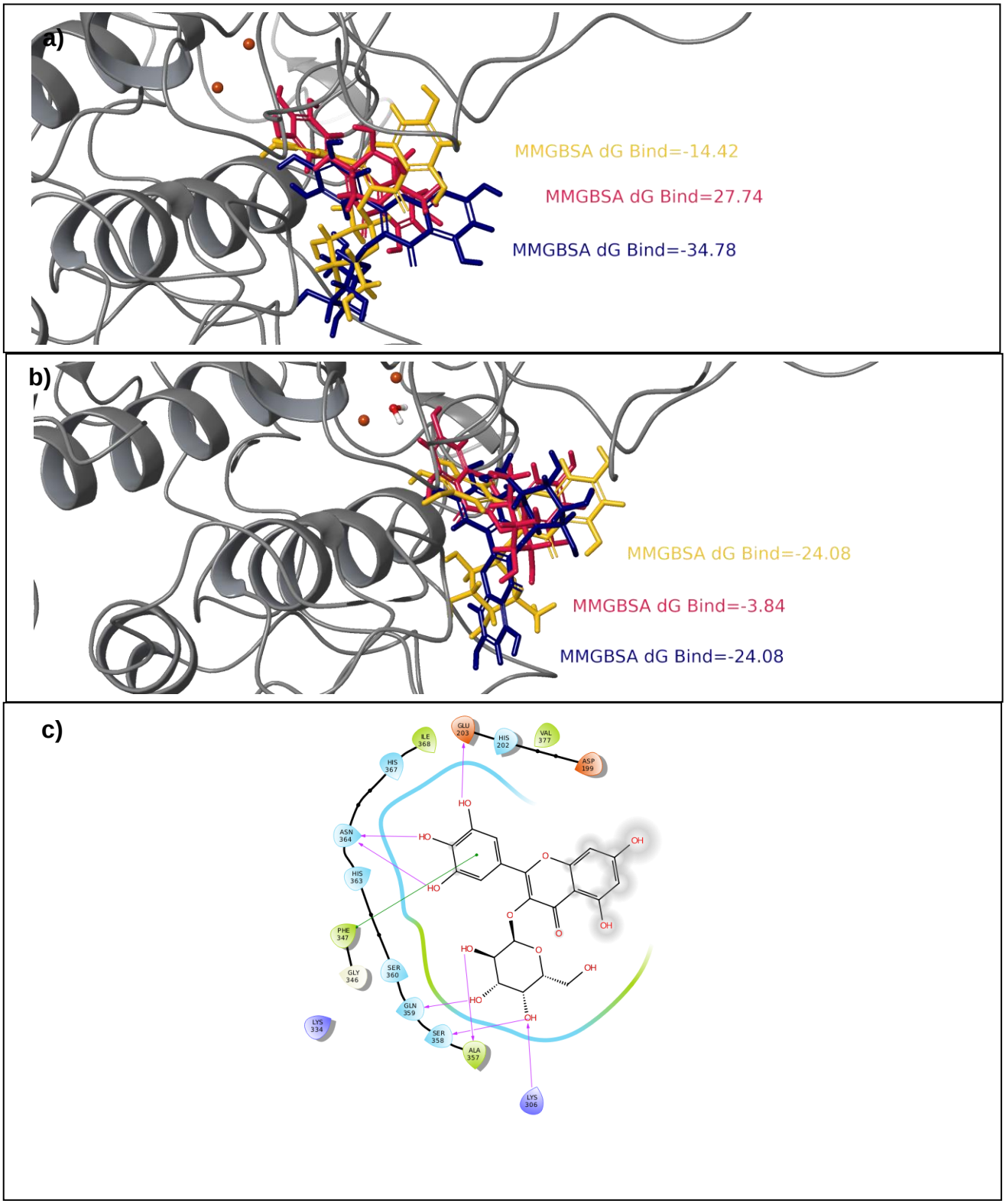


Figura 34: Vista 3D da pose de *docking* ΔG MM-GBSA mais negativa do miricetina galactosídeo (20) obtida por cada programa (Glide -azul, Gold -amarelo e Autodock- rosa) – a) sem água complexada com Cu^{2+} e b) com água complexada com Cu^{2+}. c) Diagrama de interação ligante da melhor pose – Glide sem água ΔG = -34,77 kcal/mol

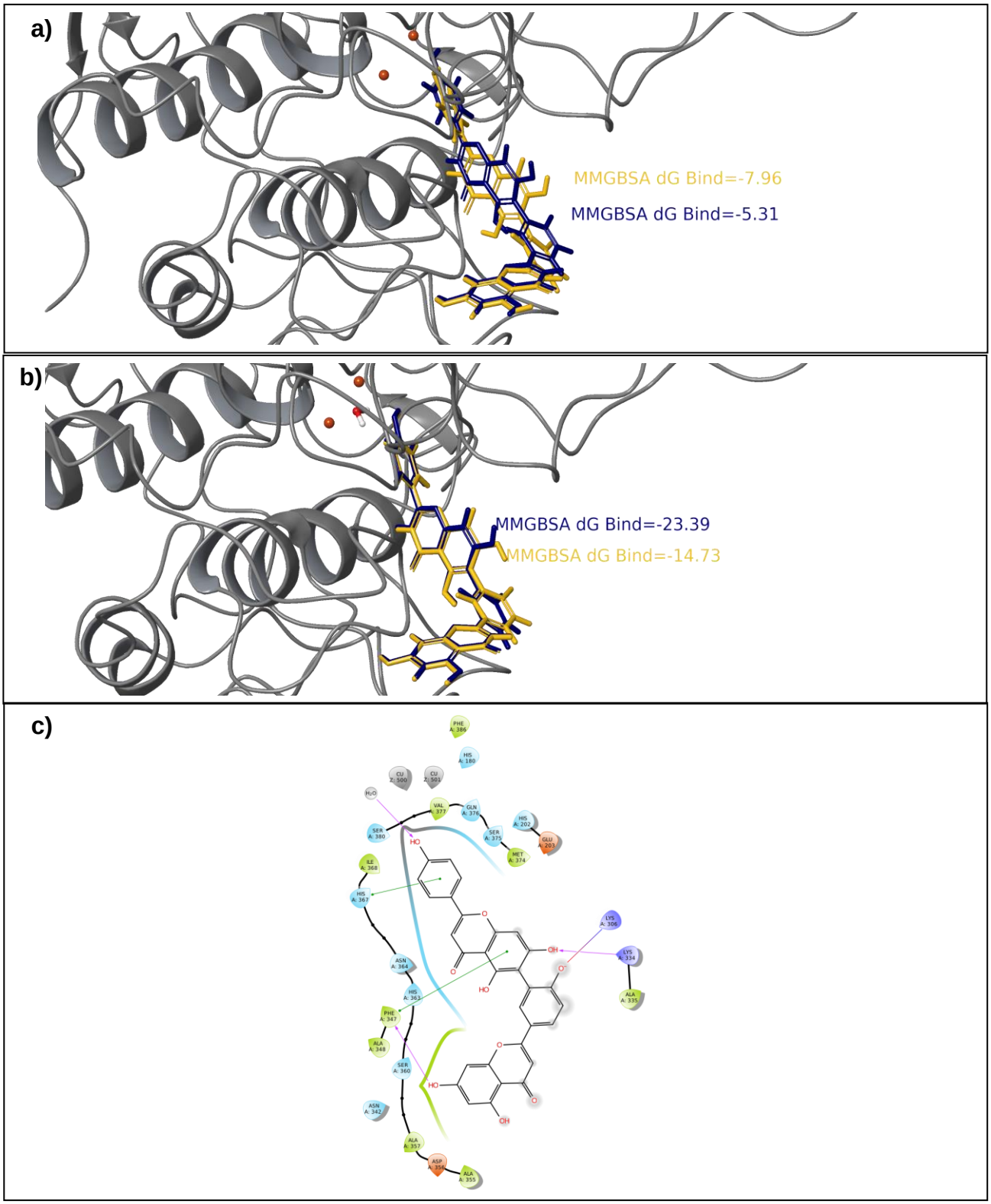


Figura 35: Vista 3D da pose de *docking* ΔG MM-GBSA mais negativa do robustaflavona (21) obtida por cada programa (Glide -azul, Gold -amarelo e Autodock- rosa) – a) sem água complexada com Cu^{2+} e b) com água complexada com Cu^{2+}. c) Diagrama de interação ligante da melhor pose – Glide com água ΔG = -23,38 kcal/mol

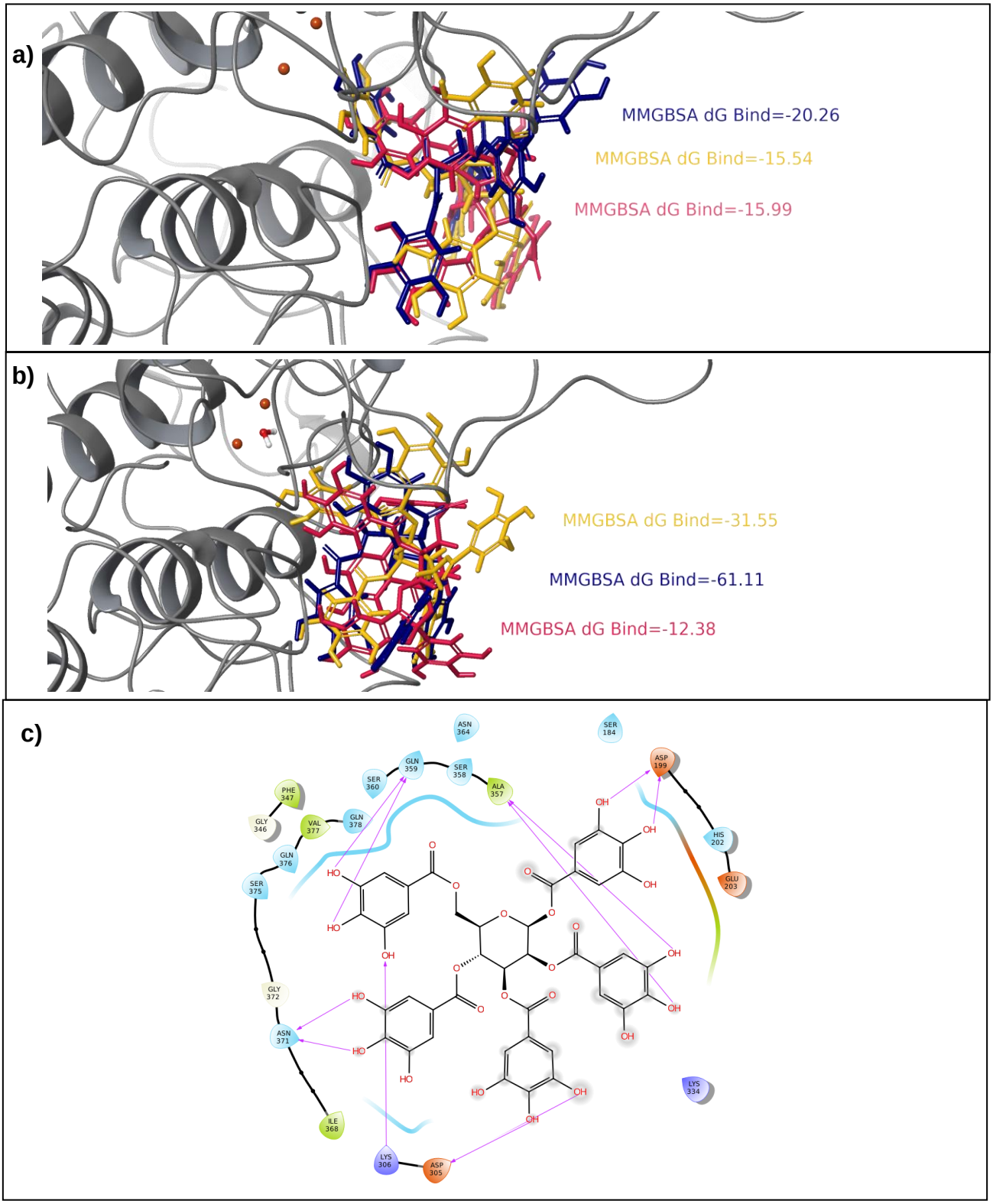


Figura 36: Vista 3D da pose de *docking* Δ*G* MM-GBSA mais negativa do 1,2,3,4,6-penta-O-galoil-glucopiranosideo (22) obtida por cada programa (Glide -azul, Gold -amarelo e Autodock- rosa) – a) sem água complexada com Cu^{2+} e b) com água complexada com Cu^{2+}. c) Diagrama de interação ligante da melhor pose – Glide com água ΔG = -61,11 kcal/mol

ANEXO 4

Nesse anexo constam as figuras das análises realizadas pós cálculos de dinâmica molecular.

Distância entre ligante e proteína

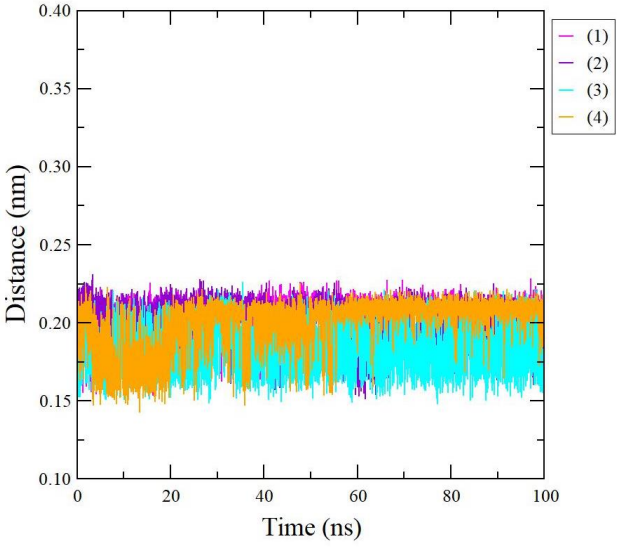


Figura 54:Distância entre proteína e ligantes, ácido kójico (1-magenta), metil galato (2-violeta), miricetina (3-ciano), etil galato (4-laranja).

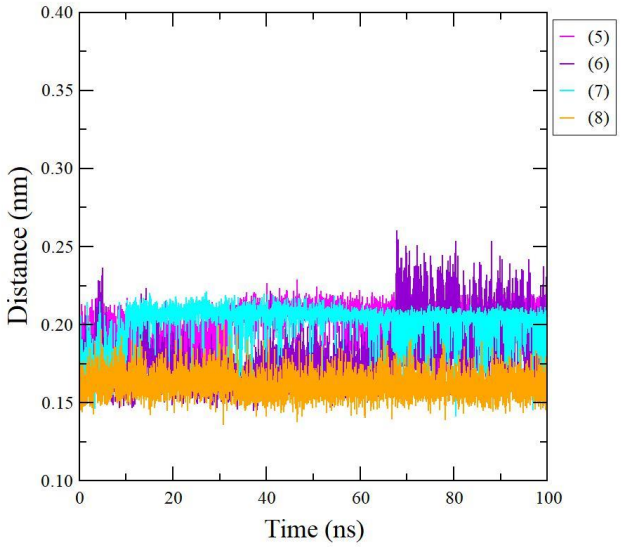


Figura 37: Distância entre proteína e ligantes, catequina (5-magenta), quercetina-3-O-α-ramnosídeo (6-violeta), miricetina-3-O-α-ramnosídeo (7-ciano), ácido elágico (8-laranja) .

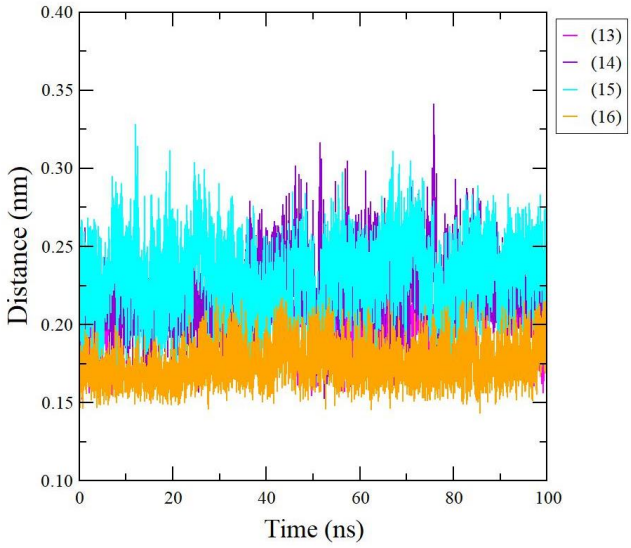


Figura 39: Distância entre proteína e ligantes, 1,6-digalloil-ß-D-glicose (13-magenta), ácido clorogênico (14-violeta), ácido gálico (15-ciano), ácido 5-p-cumaroilquínico (16-laranja) .

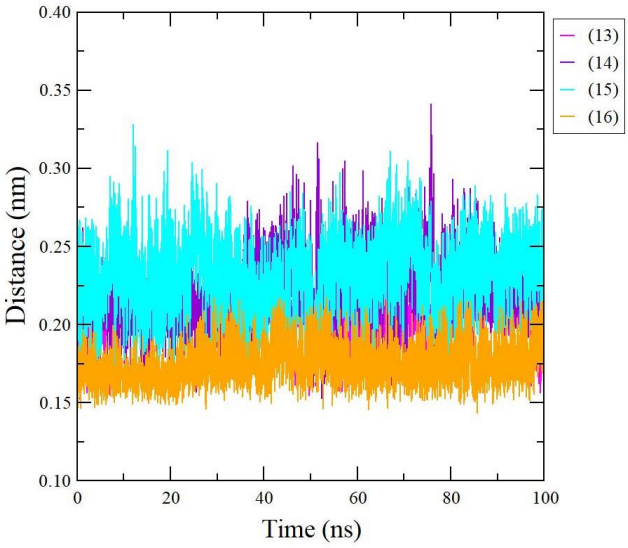


Figura 38: Distância entre proteína e ligantes, 1,6-digalloil-ß-D-glicose (13-magenta), ácido clorogênico (14-violeta), ácido gálico (15-ciano), ácido 5-p-cumaroilquínico (16-laranja) .

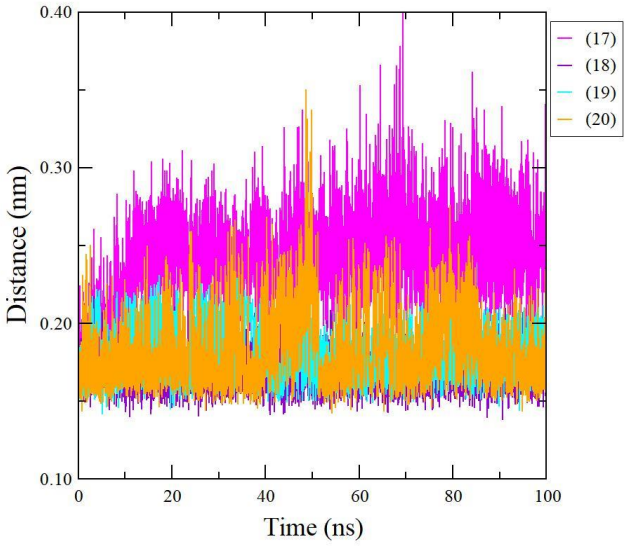


Figura 41: Distância entre proteínas e ligantes, luteolina (17-magenta), miricetina- ramnosídeo 1'''->6'' galactosídeo (18-violeta), miricetina glucoronídeo (19-ciano), miricetina galactosídeo (20-laranja).

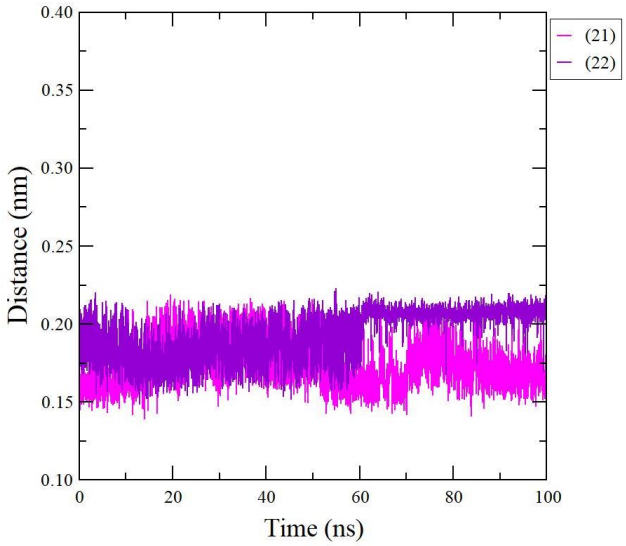


Figura 40:Distância entre proteína e ligantes, flavona robusta (21-magenta), 1,2,3,4,6-penta-O-galloil-glucopiranosídeo (22-violeta).

RMSD sitio de ligação

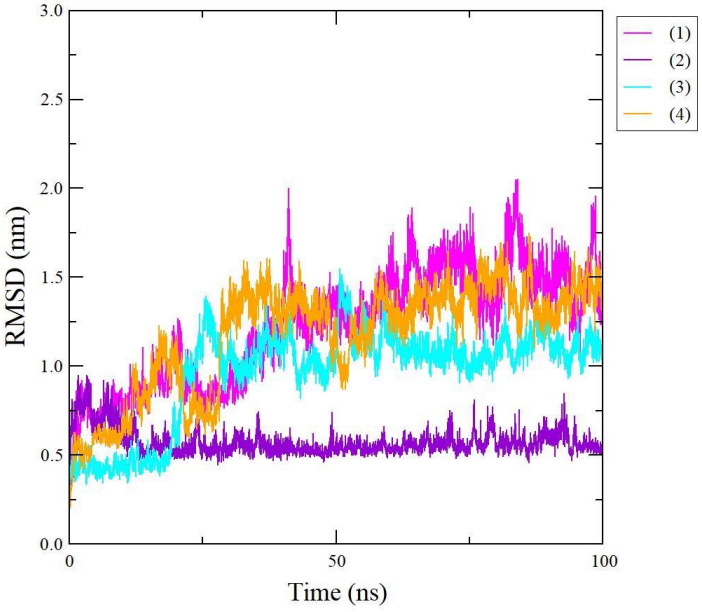


Figura 42:RMSD (nm) do sítio de ligação, a partir da proteína simulada com ácido kójico (1-magenta), metil galato (2-violeta), miricetina (3-ciano), etil galato (4-laranja).

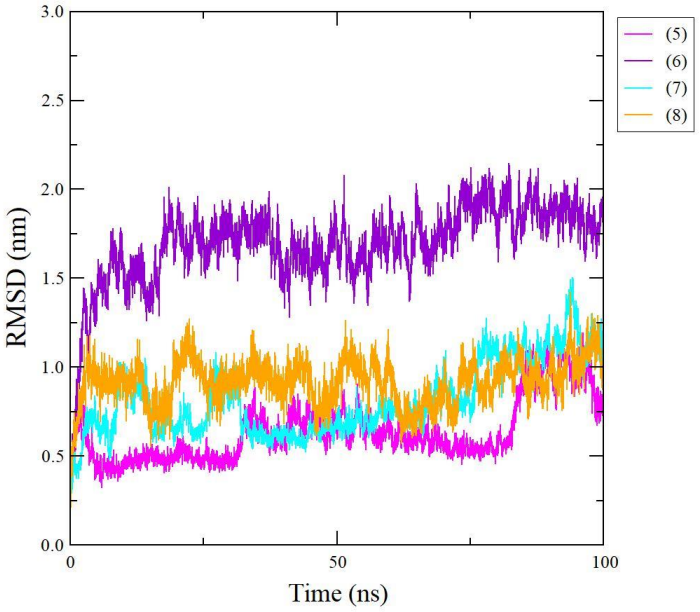


Figura 43:RMSD (nm) do sítio de ligação, da proteína simulada com catequina (5-magenta), quercetina-3-O-α-ramnosídeo (6-violeta), miricetina-3-O-α-ramnosídeo (7-ciano), ácido elágico (8-laranja).

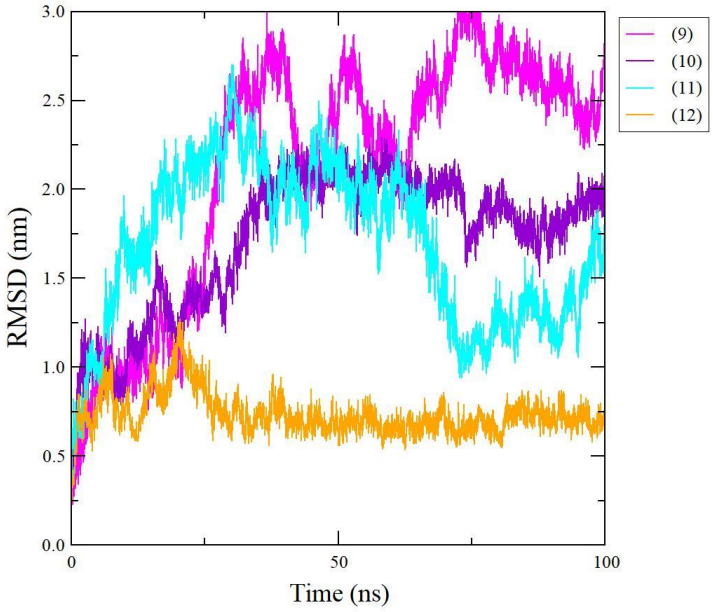


Figura 44: RMSD (nm) do sítio de ligação, a partir da proteína simulada com quercetina (9-magenta), ácido cafeico (10-violeta), ácido siríngico (11-ciano), ácido coumárico (12-laranja).

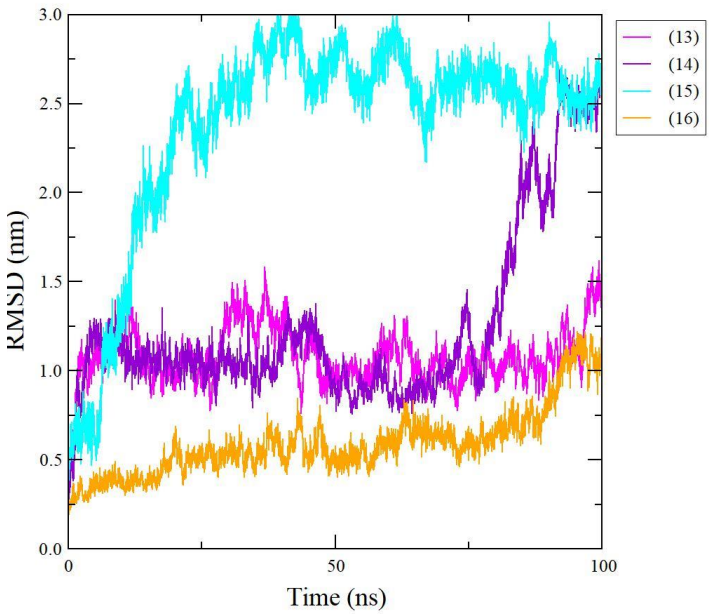


Figura 45: RMSD (nm) do sítio de ligação, da proteína simulada com 1,6-digalloil-ß-D-glicose (13-magenta), ácido clorogênico (14-violeta), ácido gálico (15-ciano), 5- ácido p-coumaroilquínico (16-laranja).

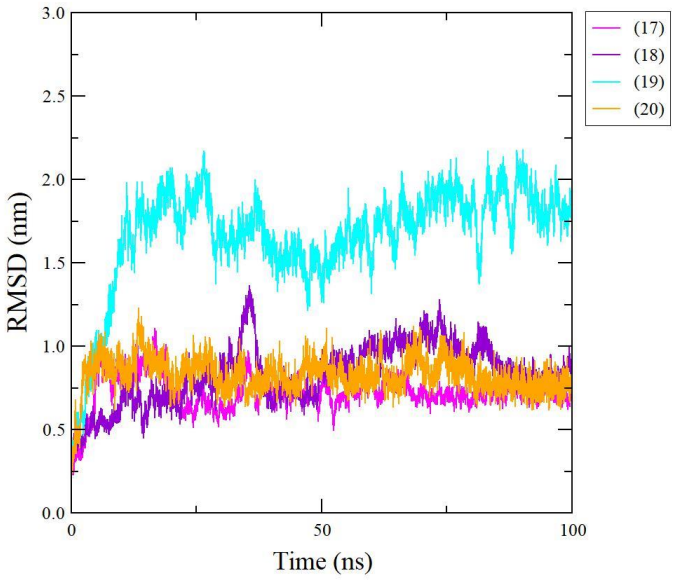


Figura 46: RMSD (nm) do sítio de ligação, da proteína simulada com luteolina (17-magenta), miricetina-ramnosídeo 1'"->6" galactosídeo (18-violeta), miricetina glucoronídeo (19-ciano), miricetina galactosídeo (20-laranja).

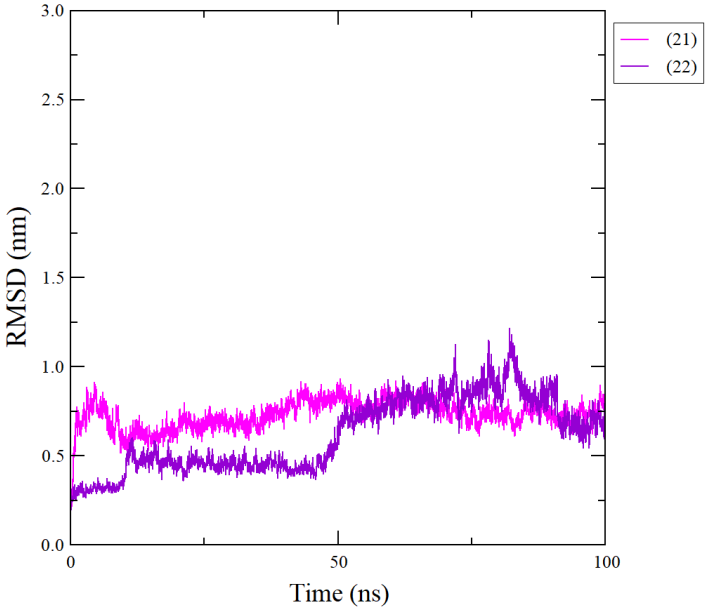


Figura 47: RMSD (nm) do sítio de ligação, da proteína simulada com robustaflavona (21-magenta), 1,2,3,4,6-penta-O-galloil-glucopiranosídeo (22-violeta).

RMSD C-alfa

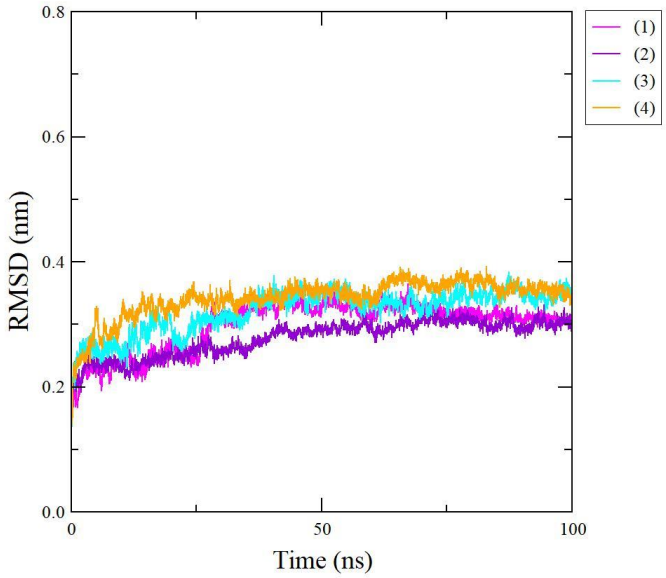


Figura 49: RMSD (nm) do C-alfa, a partir da proteína simulada com ácido kójico (1-magenta), metil galato (2-violeta), miricetina (3-ciano), etil galato (4-laranja).

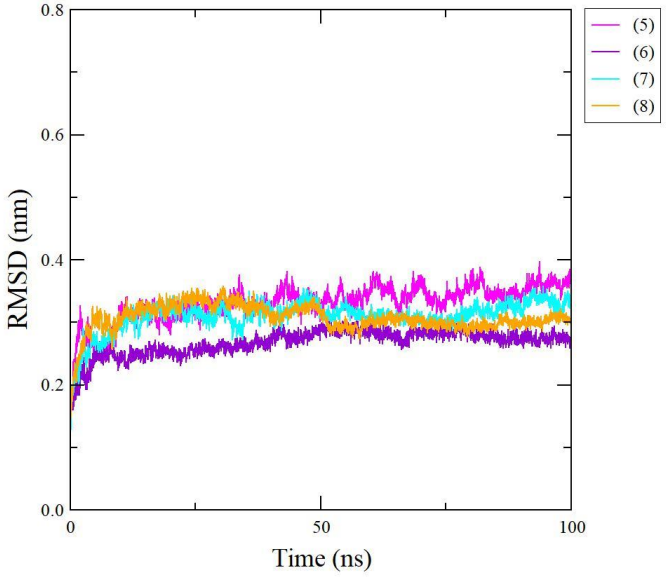


Figura 48: RMSD (nm) do C-alfa, da proteína simulada com catequina (5-magenta), quercetina-3-O-α-ramnosídeo (6-violeta), miricetina-3-O-α-ramnosídeo (7 -ciano), ácido elágico (8-laranja).

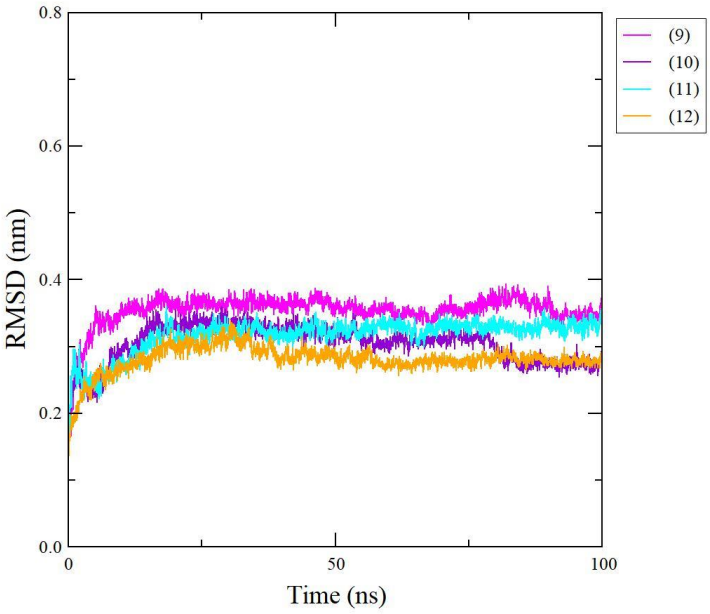


Figura 51:RMSD (nm) do C-alfa, a partir da proteína simulada com quercetina (ácido cafeico 9-magenta (10-violeta), ácido siríngico (11-ciano), ácido coumárico (12-laranja).

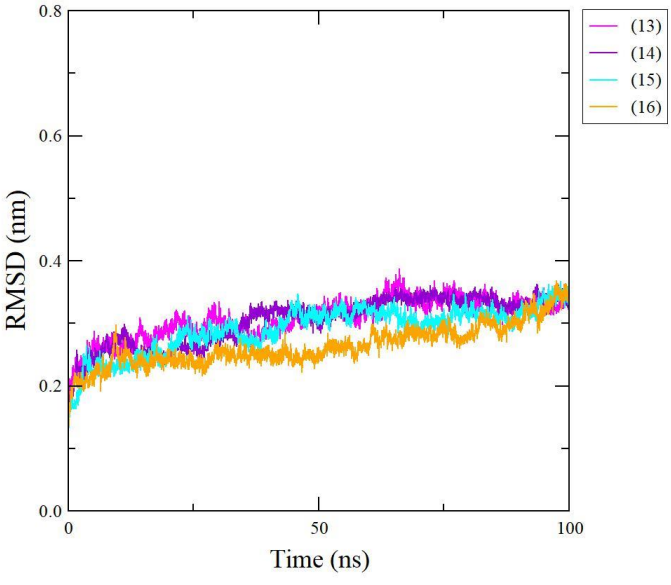


Figura 50*:* RMSD (nm) do C-alfa, da proteína simulada com 1,6-digalloil-ß-D-glicose (13-magenta), ácido clorogênico (14-violeta), ácido gálico (15-ciano), 5 ácido -p-coumaroilquínico (16-laranja).

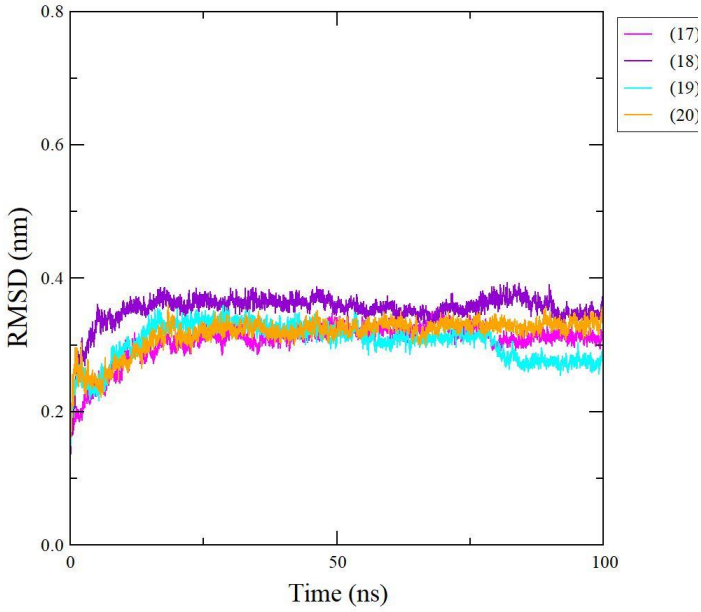


Figura 52: RMSD (nm) do C-alfa, da proteína simulada com luteolina (17-magenta), miricetina- ramnosídeo 1'"->6" galactosídeo (18-violeta), miricetina glucoronídeo (19-ciano), miricetina galactosídeo (20-laranja).

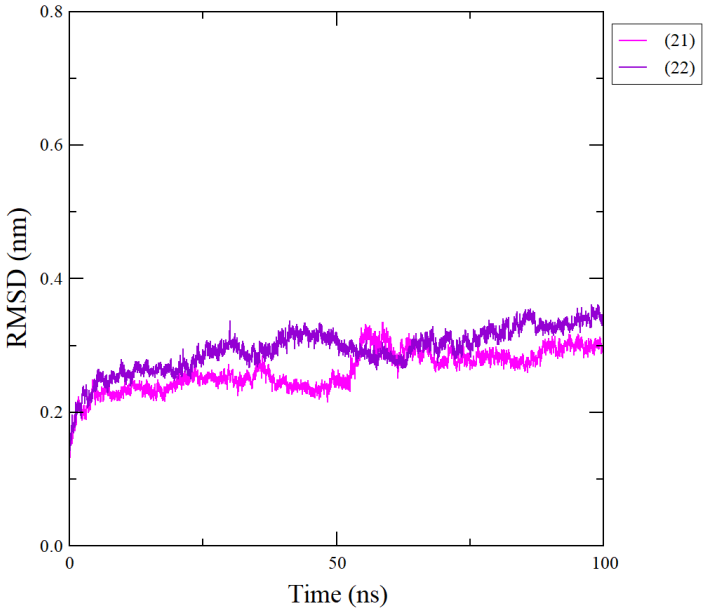


Figura 53: RMSD (nm) do C-alfa, da proteína simulada com robustaflavona (21-magenta), 1,2,3,4,6-penta-O-galloil-glucopiranosídeo (22-violeta).

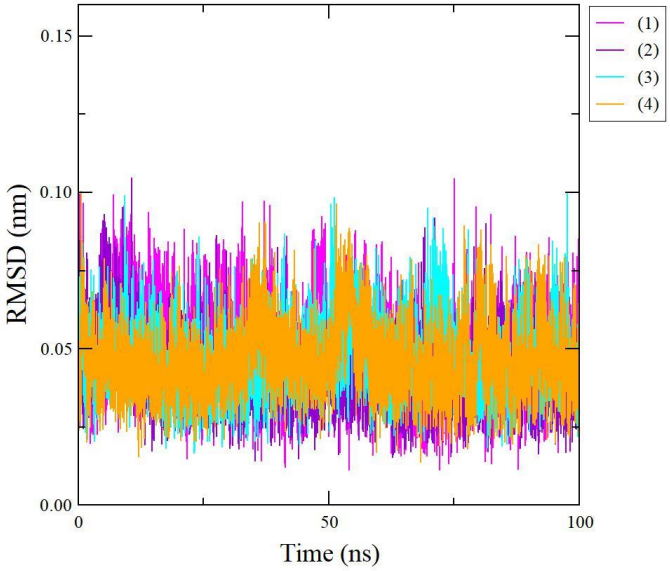


Figura 54: RMSD (nm) dos ligantes, ácido kójico (1-magenta), metil galato (2-violeta), miricetina (3-ciano), etil galato (4-laranja).

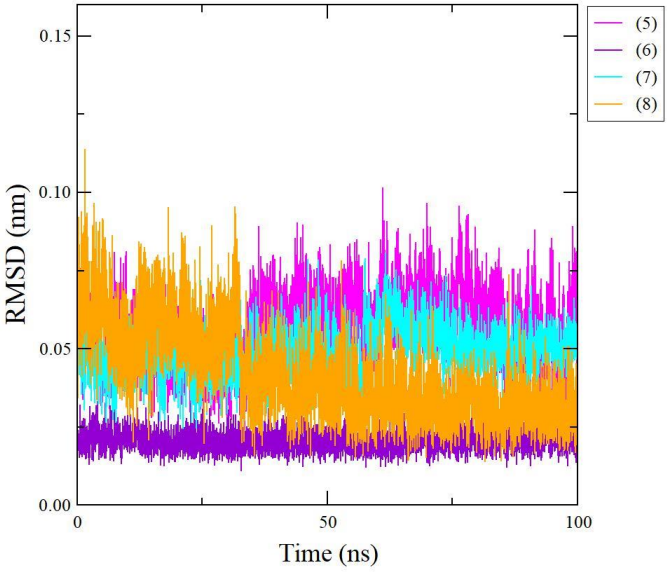


Figura 73: RMSD (nm) de ligantes, catequina (5-magenta), quercetina-3-O-α-ramnosídeo (6-violeta), miricetina-3-O-α-ramnosídeo (7-ciano), ácido elágico (8-laranja).

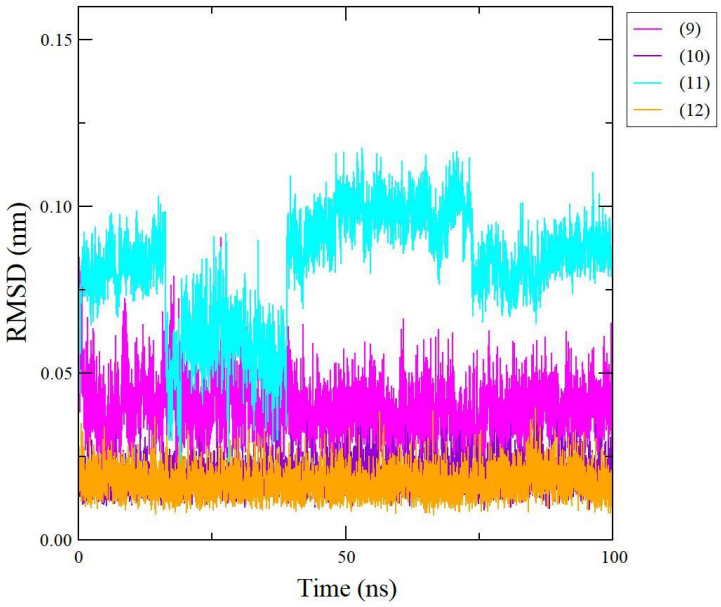


Figura 55: RMSD (nm) dos ligantes, quercetina (ácido cafeico 9-magenta (10-violeta), ácido siríngico (11-ciano), ácido coumárico (12-laranja).

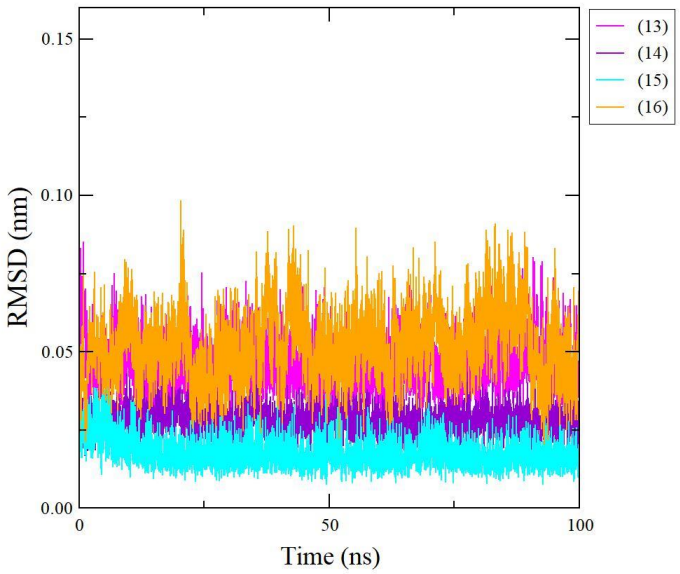


Figura 56: RMSD (nm) dos ligantes, da proteína simulada com 1,6-digalloil-ß-D-glucose (13-magenta), ácido clorogênico (14-violeta), ácido gálico (15-ciano), 5-p -ácido coumaroilquínico (16-laranja).

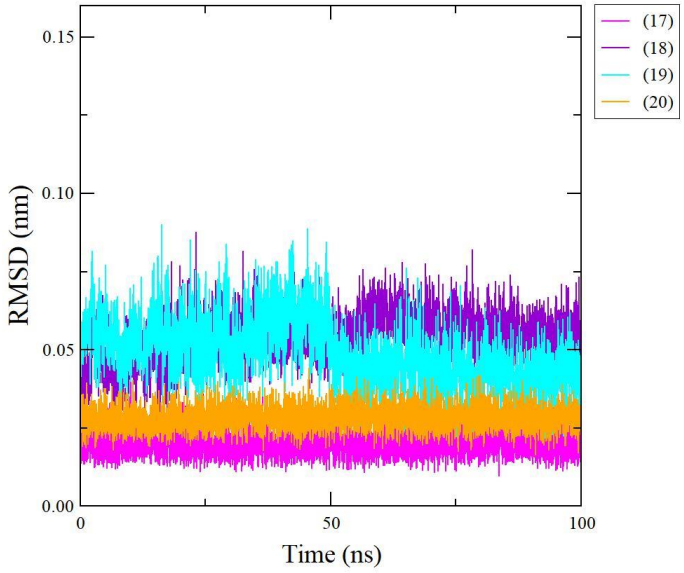


Figura 57: RMSD (nm) dos ligantes, luteolina (17-magenta), miricetina- ramnosídeo 1'"->6" galactosídeo (18-violeta), miricetina glucoronídeo (19-ciano), miricetina galactosídeo (20-laranja).

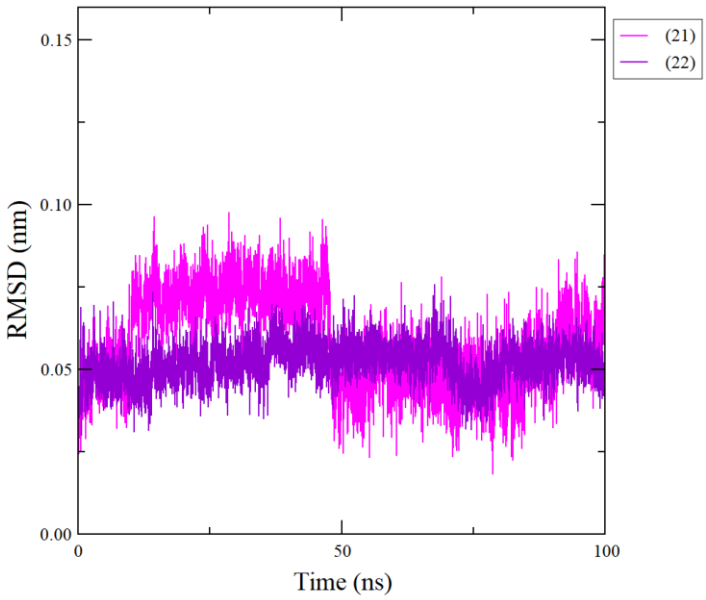


Figura 58: RMSD (nm) dos ligantes robustaflavona (21-magenta), 1,2,3,4,6-penta-O-galloil-glucopiranosídeo (22-violeta).

RMSF

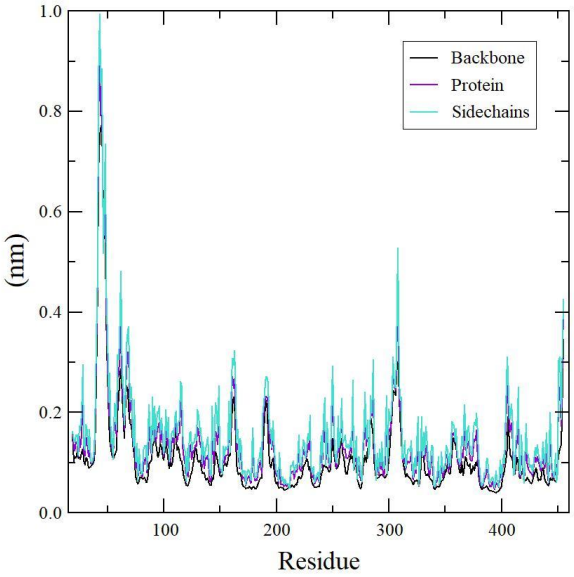


Figura 60: RMSF de *backbone*, proteína e cadeia lateral da simulação de dinâmica molecular com ácido kójico (1).

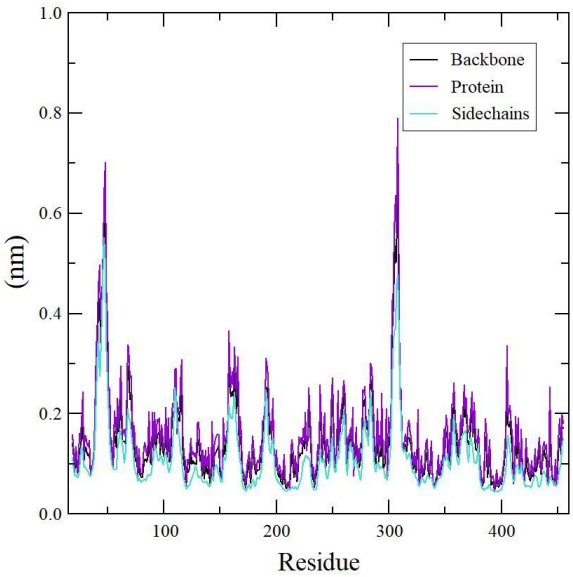


Figura 59: RMSF de *backbone*, proteína e cadeia lateral da simulação de dinâmica molecular com metil galato (2).

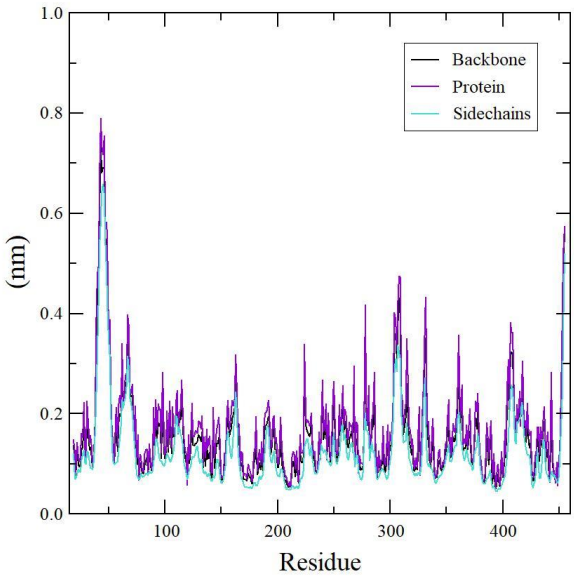


Figura 61: RMSF de *backbone*, proteína e cadeia lateral da simulação de dinâmica molecular com miricetina (3).

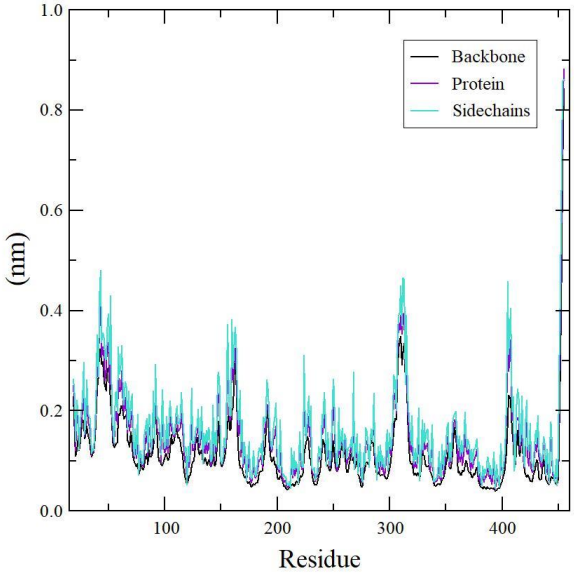


Figura 62: RMSF de *backbone*, proteína e cadeia lateral da simulação de dinâmica molecular com etil galato (4).

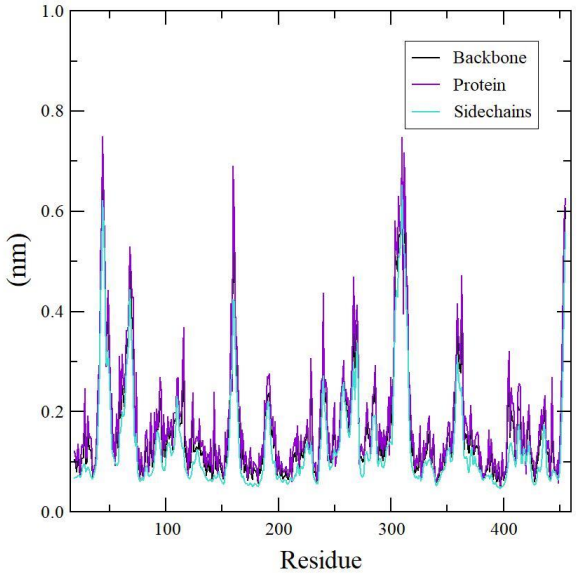


Figura 64: RMSF de *backbone*, proteína e cadeia lateral da simulação de dinâmica molecular com catequina (5).

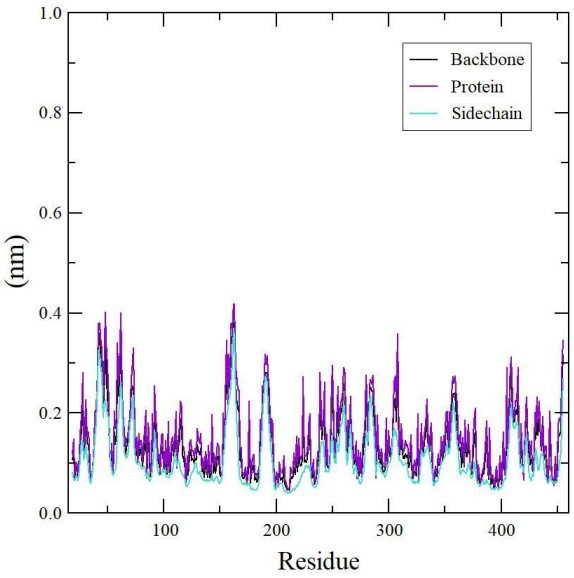


Figura 63: RMSF de *backbone*, proteína e cadeia lateral da simulação de dinâmica molecular com quercetina ramnosídeo (6)

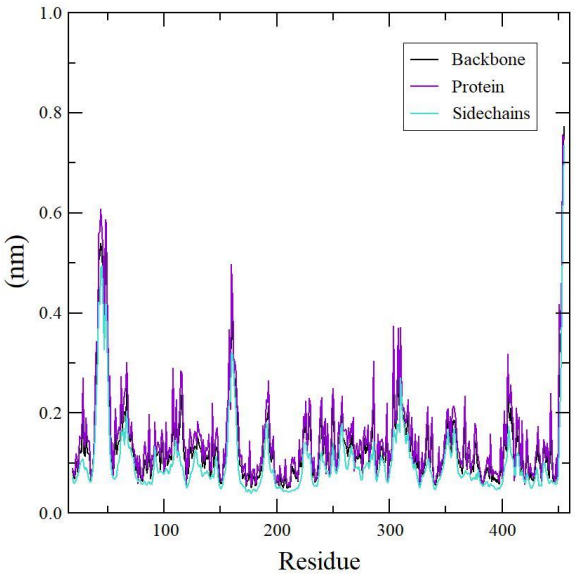


Figura 65: RMSF de *backbone*, proteína e cadeia lateral da simulação de dinâmica molecular com miricetina-3-O-α-ramnosídeo (7)

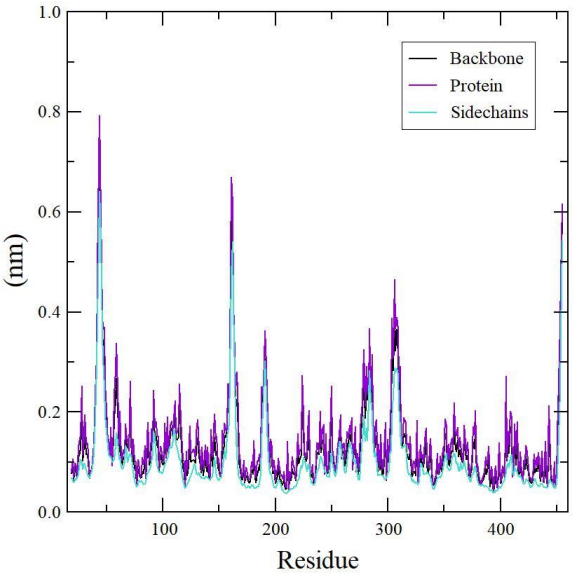


Figura 66 : RMSF de *backbone*, proteína e cadeia lateral da simulação de dinâmica molecular com ácido elágico (8)

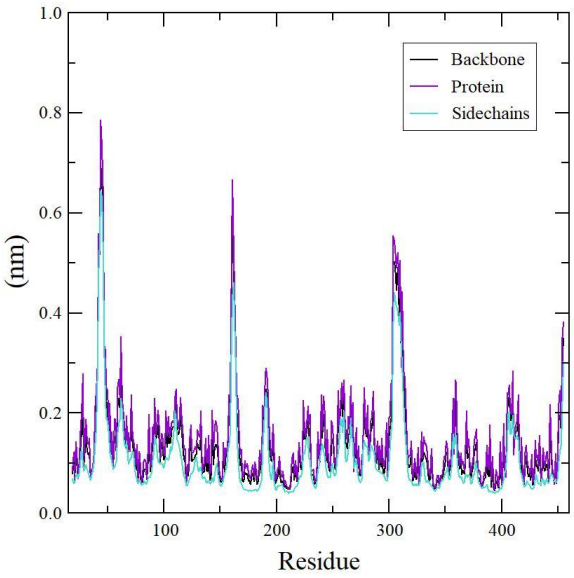


Figura 68: RMSF de *backbone*, proteína e cadeia lateral da simulação de dinâmica molecular com quercetina (9)

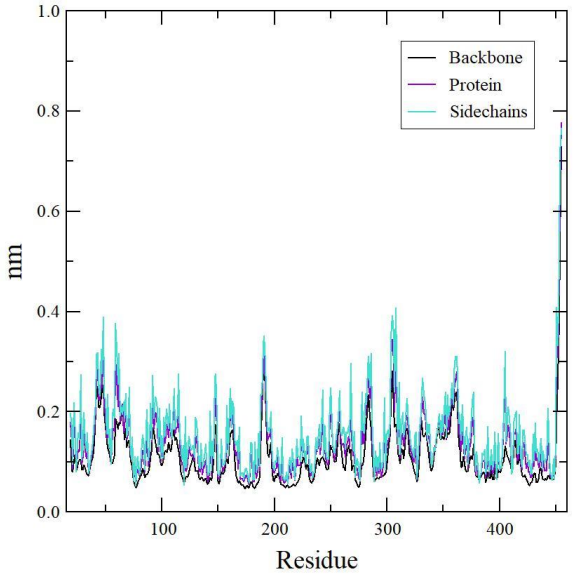


Figura 67: RMSF de *backbone*, proteína e cadeia lateral da simulação de dinâmica molecular com ácido cafeico (10)

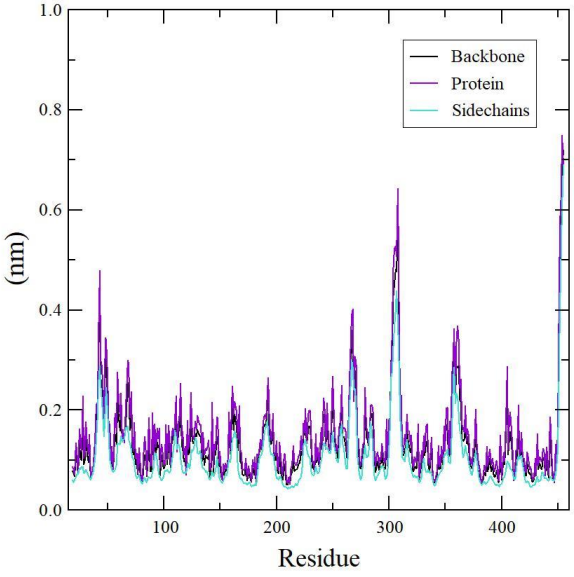


Figura 70: RMSF de *backbone*, proteína e cadeia lateral da simulação de dinâmica molecular com ácido siríngico (11)

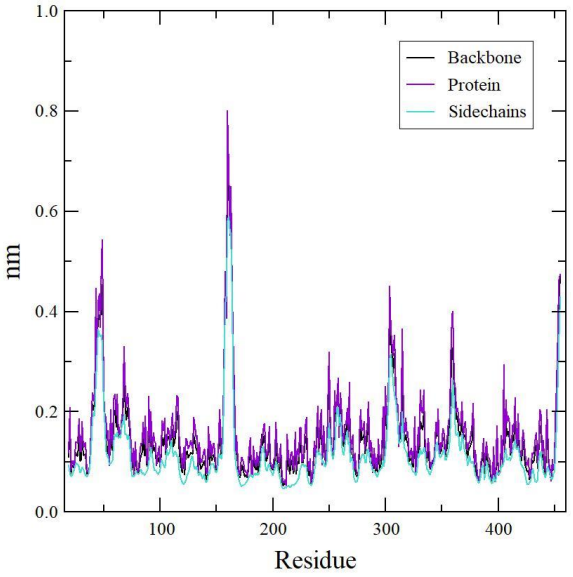


Figura 69: RMSF de *backbone*, proteína e cadeia lateral da simulação de dinâmica molecular com ácido coumárico (12)

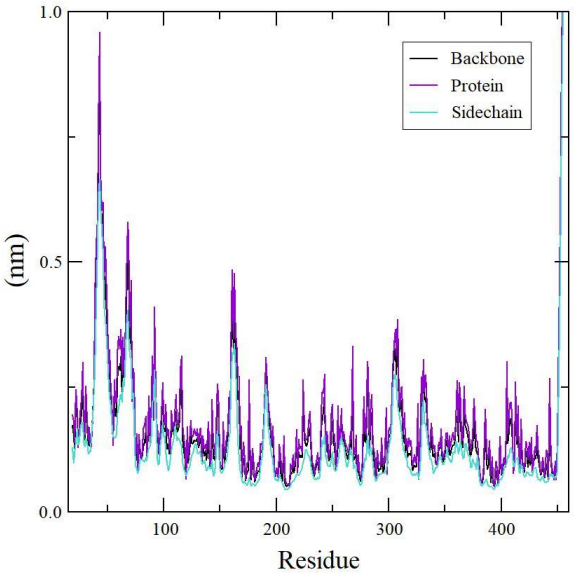


Figura 72: RMSF de *backbone*, proteína e cadeia lateral da simulação de dinâmica molecular com 1,6-digalloil-ß-D-glucose (13).

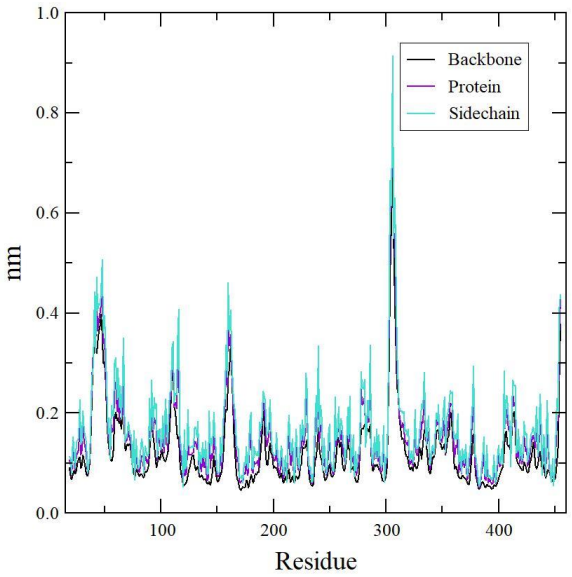


Figura 71: RMSF de *backbone*, proteína e cadeia lateral da simulação de dinâmica molecular com ácido clorogênico (14).

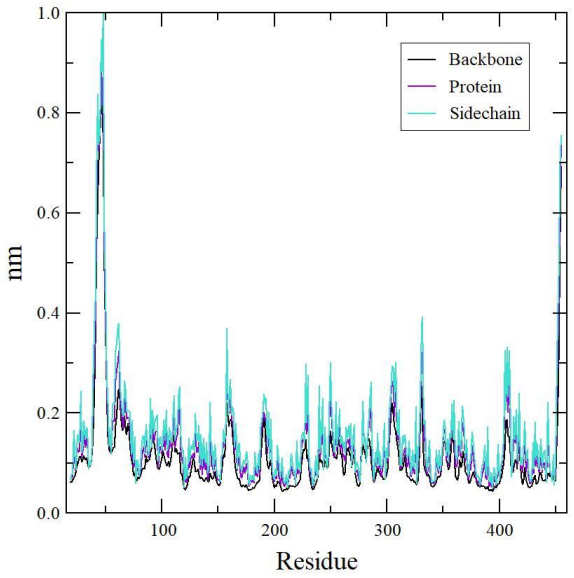


Figura 74: RMSF de *backbone*, proteína e cadeia lateral da simulação de dinâmica molecular com ácido gálico (15)

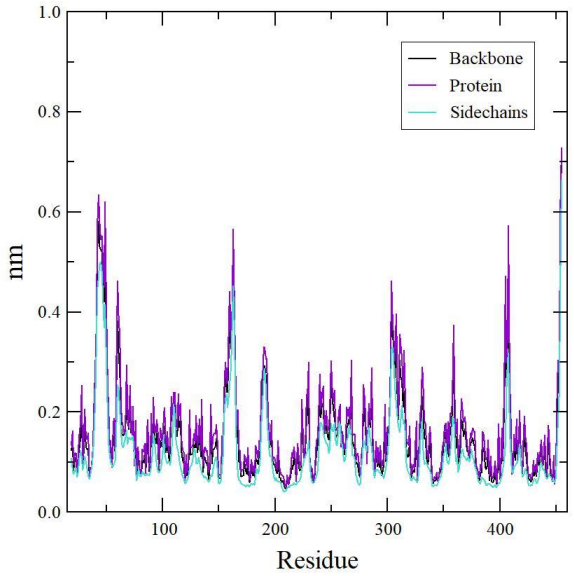


Figura 73: RMSF de *backbone*, proteína e cadeia lateral da simulação de dinâmica molecular com ácido 5-p-cumaroilquínico (16).

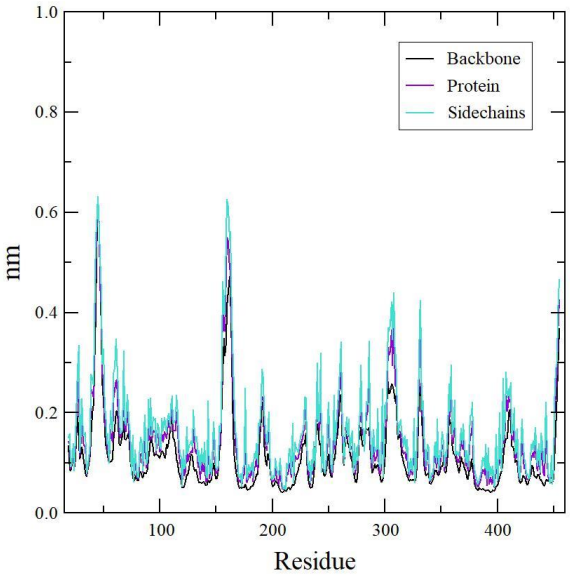


Figura 76: RMSF de *backbone*, proteína e cadeia lateral da simulação de dinâmica molecular com luteolina (17).

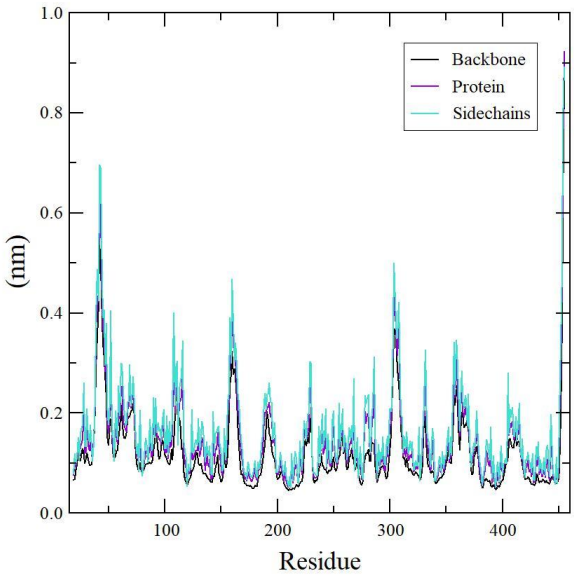


Figura 75: RMSF de backbone, proteína e cadeia lateral da simulação de dinâmica molecular com miricetina-ramnosídeo 1'"->6" galactosídeo (18).

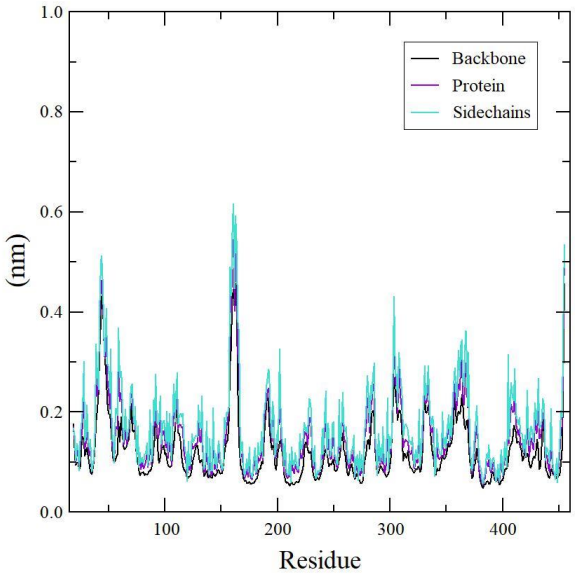


Figura 77: RMSF de backbone, proteína e cadeia lateral da simulação de dinâmica molecular com miricetina glucoronídeo (19).

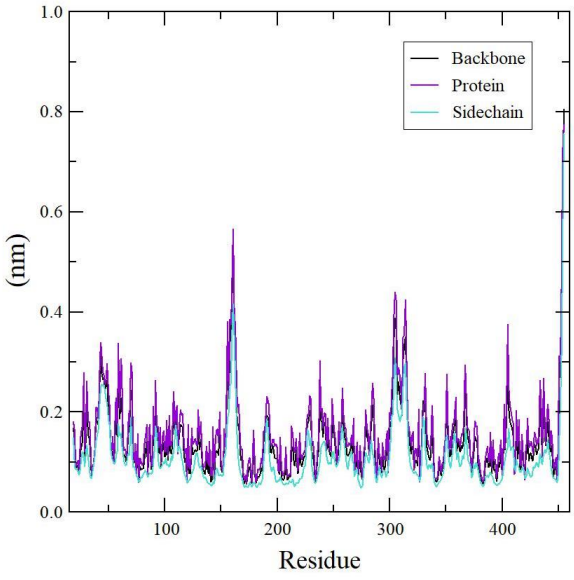


Figura 78: RMSF de backbone, proteína e cadeia lateral da simulação de dinâmica molecular com miricetina galactosídeo (20).

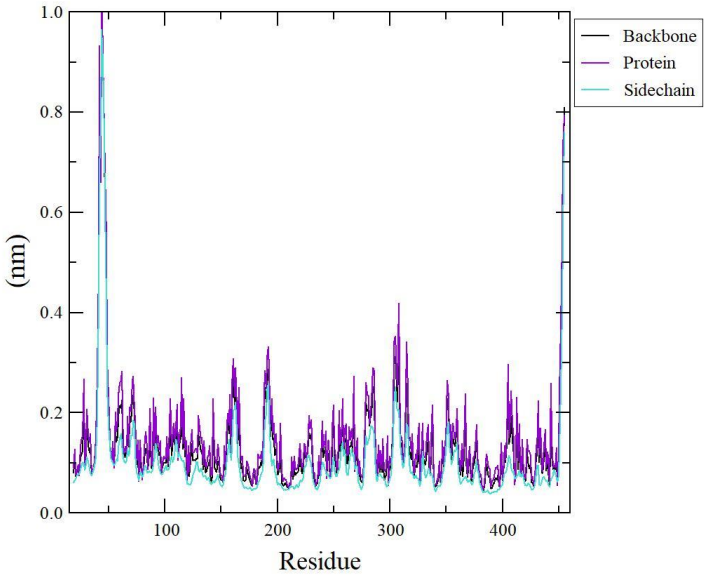


Figura 79: RMSF de *backbone*, proteína e cadeia lateral da simulação de dinâmica molecular com robustaflavona (21)

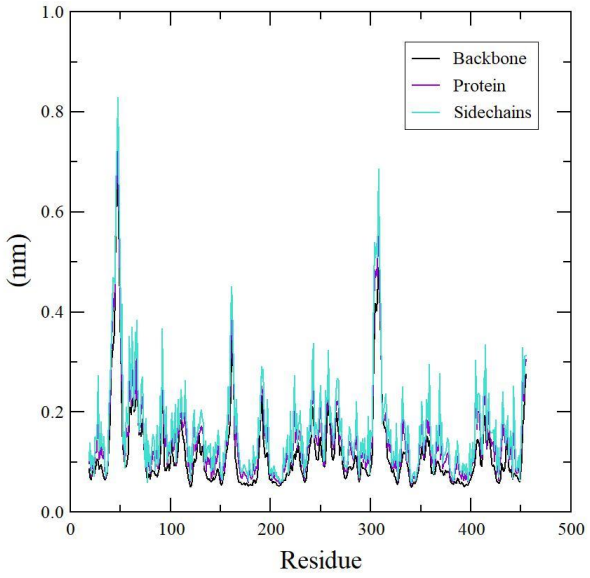


Figura 80: RMSF de *backbone*, proteína e cadeia lateral da simulação de dinâmica molecular com 1,2,3,4,6-penta-O-galloil-glucopiranosídeo (22-violeta).

Raio de giro (Rg)

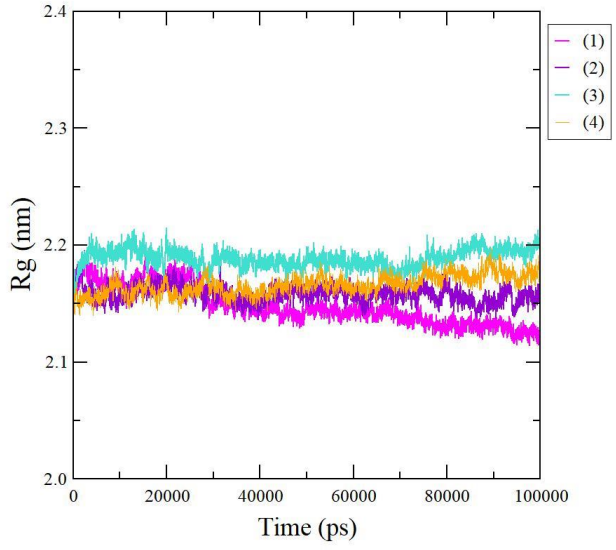


Figura 81: Raio de giro do ácido kójico (1-magenta), metil galato (2-violeta), miricetina (3-ciano), etil galato (4-laranja).

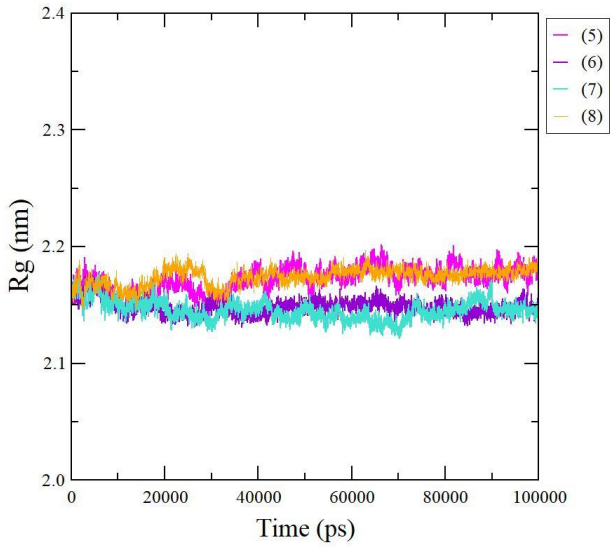


Figura 82: Raio de giro de catequina (5-magenta), quercetina-3-O-α-ramnosídeo (6-violeta), miricetina-3-O-α-ramnosídeo (7-ciano), ácido elágico (8-laranja).

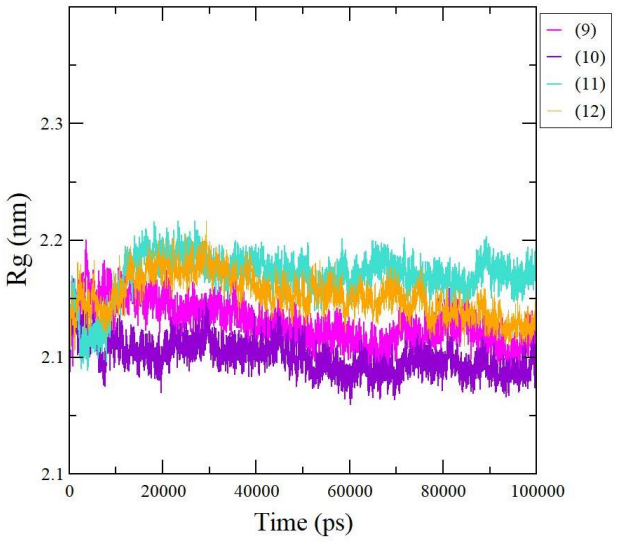


Figura 83: Raio de giro da quercetina (ácido cafeico 9-magenta (10-violeta), ácido siríngico (11-ciano), ácido coumárico (12-laranja).

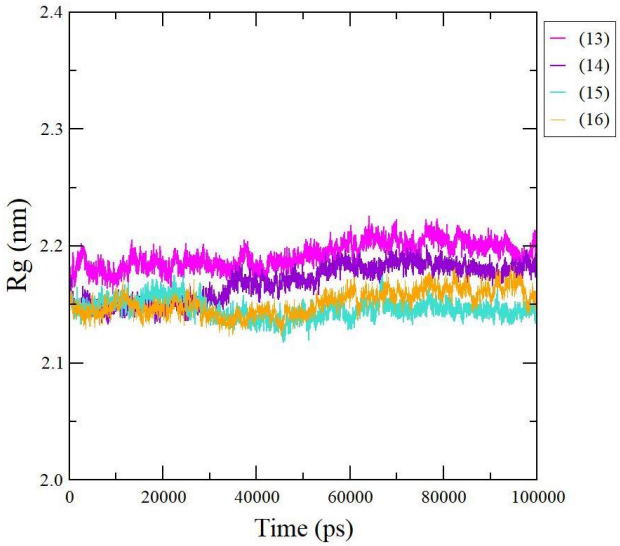


Figura 84: Raio de giro de 1,6-digalloil-ß-D-glicose (13-magenta), ácido clorogênico (14-violeta), ácido gálico (15-ciano), ácido 5-p-cumaroilquínico (16-laranja).

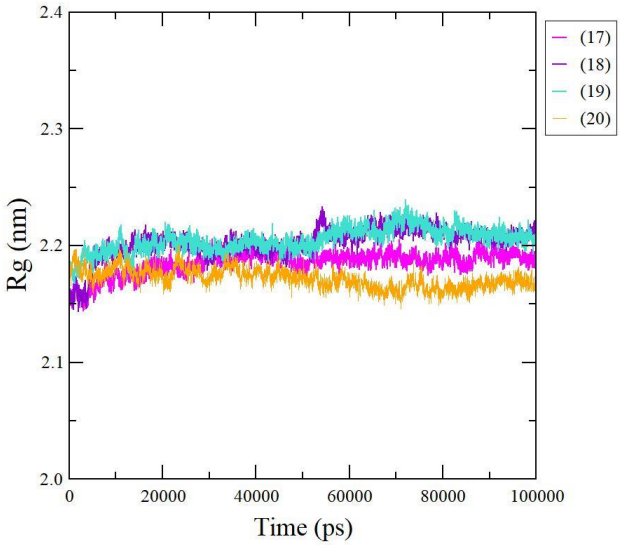


Figura 85: Raio de giro da luteolina (17-magenta), miricetina-ramnosídeo 1'''->6'' galactosídeo (18-violeta), miricetina glucoronídeo (19-ciano), miricetina galactosídeo (20-laranja).

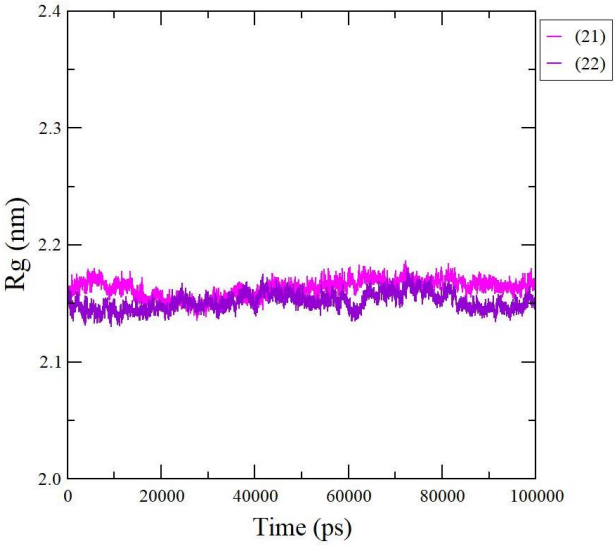


Figura 86: Raio de giro da robustaflavona (21-magenta) e com 1,2,3,4,6-penta-O-galloil-glucopiranosídeo (22-violeta).