



**UNIVERSIDADE FEDERAL DO CEARÁ**  
**CENTRO DE CIÊNCIAS**  
**DEPARTAMENTO DE FÍSICA**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM FÍSICA**  
**DOUTORADO EM FÍSICA**

**ANDRÉ HADAD CORTEZ**

**DESVENDANDO A LIGAÇÃO DA FRUTOSE E DA GLICOSE À ALBUMINA DO  
SORO HUMANO: MEDIÇÕES DE FLUORESCÊNCIA E CÁLCULOS DE  
DOCKING, DINÂMICA MOLECULAR E BIOQUÍMICA QUÂNTICA**

**FORTALEZA**

**2024**

ANDRÉ HADAD CORTEZ

DESVENDANDO A LIGAÇÃO DA FRUTOSE E DA GLICOSE À ALBUMINA DO  
SORO HUMANO: MEDIÇÕES DE FLUORESCÊNCIA E CÁLCULOS DE DOCKING,  
DINÂMICA MOLECULAR E BIOQUÍMICA QUÂNTICA

Tese apresentada ao Programa de Pós-Graduação em Física do Centro de Ciências da Universidade Federal do Ceará, como requisito parcial à obtenção do título de doutor em Física. Área de Concentração: Física da Matéria Condensada.

Orientador: Prof. Dr. Valder Nogueira Freire.

FORTALEZA  
2024

Dados Internacionais de Catalogação na Publicação  
Universidade Federal do Ceará  
Sistema de Bibliotecas  
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

---

C858d Cortez, André Hadad.

Desvendando a ligação da frutose e da glicose à albumina do soro humano: medições de fluorescência e cálculos de docking, dinâmica molecular e bioquímica quântica / André Hadad Cortez. – 2024.  
141 f. : il. color.

Tese (doutorado) – Universidade Federal do Ceará, Centro de Ciências, Programa de Pós-Graduação em Física, Fortaleza, 2024.

Orientação: Prof. Dr. Valder Nogueira Freire.

1. Albumina do Soro Humano (HSA). 2. Frutose. 3. Glicose. 4. Dinâmica molecular. 5. Bioquímica Quântica. I. Título.

CDD 530

---

ANDRÉ HADAD CORTEZ

DESVENDANDO A LIGAÇÃO DA FRUTOSE E DA GLICOSE À ALBUMINA DO  
SORO HUMANO: MEDIÇÕES DE FLUORESCÊNCIA E CÁLCULOS DE DOCKING,  
DINÂMICA MOLECULAR E BIOQUÍMICA QUÂNTICA

Tese apresentada ao Programa de Pós-Graduação em Física do Centro de Ciências da Universidade Federal do Ceará, como requisito parcial à obtenção do título de doutor em Física. Área de Concentração: Física da Matéria Condensada.

Aprovado em: 08/07/2024

BANCA EXAMINADORA

---

Prof. Dr. Valder Nogueira Freire (Orientador)  
Universidade Federal do Ceará (UFC)

---

Prof. Dr. Paulo de Tarso Cavalcante Freire  
Universidade Federal do Ceará (UFC)

---

Profª. Dra. Eveline Matias Bezerra  
Universidade Federal Rural do Semi-Árido (UFERSA)

---

Prof. Dr. Bruno Poti e Silva  
Instituto Federal de Educação, Ciência e Tecnologia do Ceará (IFCE)

---

Prof. Dr. Ewerton Wagner Santos Caetano  
Instituto Federal de Educação, Ciência e Tecnologia do Ceará (IFCE)



A Deus.

A minha família, Elaine e Arthur.

## AGRADECIMENTOS

Gostaria de expressar minha sincera gratidão ao Prof. Dr. Valder Nogueira Freire pela orientação excepcional ao longo deste trabalho.

Aos distinguidos membros da banca examinadora, reconheço o tempo dedicado, bem como suas valiosas contribuições e sugestões que enriqueceram este estudo.

Um agradecimento especial ao grupo de trabalho LabInitio, cuja contribuição foi fundamental para o progresso desta jornada, em particular ao Victor L.B. França.

Sou imensamente grato ao Prof. Dr. Hernandes F. Carvalho e ao seu grupo de pesquisa da Universidade Estadual de Campinas, pelas suas valiosas contribuições e orientações.

Também desejo expressar minha gratidão à Profa. Dra. Kellen Brunaldi e ao seu grupo de pesquisa na Universidade Estadual de Maringá (UEM).

Ao Programa de Pós-Graduação em Física da Universidade Federal do Ceará, manifesto meu reconhecimento pela infraestrutura proporcionada ao longo deste percurso acadêmico.

A todos os professores do curso de Pós-Graduação em Física da UFC, agradeço pela sólida formação acadêmica oferecida, assim como aos servidores do Departamento de Física pelo apoio constante.

À Universidade Federal do Ceará, expressei minha gratidão pelos recursos disponíveis, pelos laboratórios e pela sólida estrutura de pesquisa na área de modelagem molecular.

Agradeço a todos que contribuíram, de maneira direta ou indireta, para este importante marco em minha jornada acadêmica.

Por último, mas não menos importante, dedico um profundo agradecimento à minha família, em especial à minha esposa Elaine Vieira Hadad e ao meu filho Arthur Vieira Hadad, por seu apoio inabalável e por terem acreditado em mim durante todo o percurso.

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

“Mas tu não deves esquecer. Tu te tornas eternamente responsável por aquilo que cativas.”

O pequeno príncipe - Antoine de Saint-Exupéry

## RESUMO

Esta investigação examina a interação entre a albumina do soro humano (HSA) e várias formas de açúcar ( $\beta$ -D-frutofuranose (FRC),  $\alpha$ -D-glucopiranosose (GLC), ceto-D-frutose (FRO), Aldehydo-D-glicose (GLO), e hexane-1,2,3,4,5-pentol (D-glicose modificada) (GLOm)) utilizando espectroscopia de fluorescência, simulações de *docking* molecular, dinâmica molecular, *clusters* conformacionais de proteínas (EnGens), fracionamento molecular com capas conjugadas (MFCC) e análise bioquímica quântica. Analisamos aspectos moleculares e quânticos, descobrindo energias de interação entre átomos de açúcar e aminoácidos. A energia total de interação considera a fragmentação das proteínas, a decomposição energética e a energia de interação numa perspectiva ascendente. A dinâmica molecular revela que a Aldehydo-D-glicose (GLO) não modificada escapa aos locais de ligação da HSA, explicando a glicação gradual. Estudamos de forma pioneira o mecanismo de ligação da HSA com glicose e frutose numa proporção de 1:1 utilizando longas simulações de dinâmica molecular. Os resultados sugerem que a forma transitória GLOm tem uma maior propensão para o *Sudlow Site I* do que a glicose não modificada, crucial para a glicação LYS195. As tendências de interação de FRO e GLOm deslocam-se para uma cavidade FA7 mais profunda, perto do seu centro. Esta abordagem elucida eficazmente os mecanismos de ligação de pequenas moléculas, de acordo com resultados experimentais anteriores.

**Palavras-chave:** albumina do soro humano (HSA); frutose; glicose; docking; dinâmica molecular; EnGens; fracionamento molecular com capas conjugadas (MFCC); bioquímica quântica; glicação gradual.

## ABSTRACT

This research examines the interaction between human serum albumin (HSA) and various sugar forms ( $\beta$ -D-fructofuranose (FRC),  $\alpha$ -D-glucopyranose (GLC), Keto-D-fructose (FRO), Aldehydo-D-gucose (GLO), and hexane-1,2,3,4,5-pentol modified D-gucose (GLO<sub>m</sub>)) using fluorescent spectroscopy, molecular docking simulations, molecular dynamics, protein conformational clusters (EnGens), molecular fractionation with conjugate caps (MFCC) and quantum biochemistry analysis. We analyze molecular and quantum aspects, uncovering interaction energies between sugar atoms and amino acids. Total interaction energy considers protein fragmentation, energetic decomposition, and interaction energy from a bottom-up perspective. Molecular dynamics reveal that unmodified Aldehydo-D-gucose (GLO) escapes HSA binding sites, explaining gradual glycation. We pioneer studying HSA's binding mechanism with glucose and fructose in a 1:1 ratio using long molecular dynamics simulations. Results suggest the transitional GLO<sub>m</sub> form has a higher Sudlow site I propensity than unmodified glucose, crucial for LYS195 glycation. FRO and GLO<sub>m</sub> interaction tendencies move toward a deeper FA7 cavity, near its center. This approach effectively elucidates small molecule binding mechanisms, consistent with previous experimental results.

**Keywords:** human serum albumin (HSA), fructose, glucose, docking; molecular dynamics, EnGens, molecular fractionation with conjugate caps (MFCC), quantum biochemistry, gradual glycation.

## LISTA DE FIGURAS

Figura 1 –	Estruturas da albumina do soro humano (HSA), frutose e glicose. (A) A estrutura global previamente determinada da HSA com os seus subdomínios identificados por cores diferentes. Estrutura química para (B) $\beta$ -D-Frutofuranose (FRC), (C) $\alpha$ -D-Glucopiranosose (GLC), (D) Keto-D-Frutose (FRO), (E) Aldeído-D-Glicose (GLO) e (F) hexane-1,2,3,4,5-pentol (D-Glicose) (GLOm) .....	16
Figura 2 –	Espectros de emissão de fluorescência. HSA titulada com glicose (A) e frutose (B). 10 $\mu$ M de HSA em tampão PBS (pH 7,4). Concentração de glicose e frutose: 0,2, 0,4, 0,6, 1 e 1,4 mM. Excitação a 280 nm. Fendas de excitação e emissão de 10 nm. Temperatura de 25 °C .....	28
Figura 3 –	<i>Docking</i> cego baseado em modelos moleculares, listando as duas poses com a energia de ligação mais baixa (kcal.mol <sup>-1</sup> ). (A) FRC, (B) GLC, (C) FRO, (D) GLO e (E) GLOm .....	29
Figura 4 –	Desvio raiz-quadrado médio de todos os átomos pesados (RMSD) em função do tempo dos complexos HSA:Monossacarídeos. FRC (A), GLC (B), FRO (C), GLO (D), GLOm (E) e simulação repetida GLO (F) .....	32
Figura 5 –	C-alfa RMSD dos subdomínios IA, IB, IIA, IIB, IIIA e IIIB da HSA. HSA ligado a formas (A) fechadas e (B) abertas (incluindo glicose linear modificada - GLOm) .....	33
Figura 6 –	Distâncias entre dois pares de resíduos que demonstram o movimento em tesoura da HSA ao longo do tempo. Distâncias entre TYR195 e ASP451 nos complexos (A) HSA::FRC, (B) HSA::GLC, (C) HSA::FRO, e (D) HSA::GLOm. (E) Aproximação e separação dos subdomínios IIB e IIIB em (E) HSA::FRC, (F) HSA::GLC, (G) HSA::FRO e (H) HSA::GLOm representados pelas distâncias entre os resíduos VAL116 e ASP510. HSA::FRC <sub>C</sub> (HSA::FRC <sub>D</sub> ) e HSA::FRO <sub>C</sub> (HSA::FRO <sub>D</sub> ) são representados pela cor verde (azul), enquanto as linhas ciano (salmão escuro) indicam HSA::GLC <sub>C</sub> (HSA::GLC <sub>D</sub> ) e HSA::GLO <sub>mC</sub> (HSA::GLO <sub>mD</sub> ) .....	36

Figura 7 –	Reorientação do anel indol TRP214 (W214) em IIA da HSA ligado a frutose fechada (A, B), glucopiranosose (C, D), ceto-D-frutose (E, F) e glicose linear modificada (G, H) .....	37
Figura 8 –	Descrição da mobilidade e dos contatos do complexo biológico HSA::GLO. RMSD dos subdomínios de HSA para (A) HSA::GLO <sub>D</sub> e (B) HSA::GLO <sub>C</sub> . Flutuação das distâncias (C) LYS195-ASP451 (K195-D451) e (D) VAL116-ASP510 (V116-D510) em ensaios de dinâmica molecular. (E-F) Número e (G-H) detalhes das ligações de hidrogênio formadas entre resíduos de GLO e HSA .....	38
Figura 9 –	(A-B) Alta mobilidade de GLO não modificada representada pelo alinhamento das conformações de HSA::GLO a cada 20 ns e deslocamento de TRP214 em (C) HSA::GLO <sub>D</sub> e (D) HSA::GLO <sub>C</sub> , onde as conformações inicial e final são coloridas em verde e cinza claro, respectivamente .....	39
Figura 10 –	Flexibilidade da frutose e da glicose não ligadas à HSA. Ilustrado pela representação de diferentes ocupações de (A, B) FRC, (C, D) FRO, (E, F) GLC e (G, H) GLO <sub>m</sub> no espaço conformacional obtido através da dinâmica molecular .....	41
Figura 11 –	Ligações de hidrogênio HSA::Monossacarídeo em função do tempo. (A, B) Contatos observados nas simulações realizadas utilizando a forma fechada da frutose. (C, D) Ligações H encontradas nos complexos compostos por frutose linear. (E, F) Ligações H realizadas por GLC. (G, H) Número de ligações de hidrogênio formadas pela forma linear modificada da glicose .....	42
Figura 12 –	Análise detalhada das ligações de hidrogênio detectadas nos sistemas HSA::Monossacarídeo ao longo do tempo de simulação. Para (A) HSA::FRC <sub>D</sub> , (B) HSA::FRC <sub>C</sub> , (C) HSA::FRO <sub>D</sub> , (D) HSA::FRO <sub>C</sub> , (E) HSA::GLC <sub>D</sub> , (F) HSA::GLC <sub>C</sub> , (G) HSA::GLO <sub>mD</sub> , (H) HSA::GLO <sub>mC</sub> . O eixo y representa os resíduos de HSA, enquanto o eixo x corresponde ao tempo de simulação .....	44
Figura 13 –	Comparação da ocupação final de cada (A, B) FRC, (C, D) FRO, (E, F) GLC, e (G, H) GLO <sub>m</sub> após 100 ns de dinâmica molecular. Alguns	

	resíduos localizados na região superior do subdomínio IIA estão representados como linhas .....	46
Figura 14 –	A energia total de interação entre os açúcares e os resíduos de HSA em função do raio do local de ligação. (A) FRC <sub>D</sub> /FRC <sub>C</sub> ; (B) GLC <sub>D</sub> /GLC <sub>C</sub> ; (C) FRO <sub>D</sub> /FRO <sub>C</sub> ; (D) GLO <sub>D</sub> ; (E) GLO <sub>mD</sub> /GLO <sub>mC</sub> .....	48
Figura 15 –	O painel BIRD exibe a energia de interação dos resíduos com as regiões moleculares para os <i>clusters</i> de energia mais baixa e mais alta. (A) <i>cluster</i> FRC <sub>C</sub> conf. 2; (B) <i>cluster</i> FRC <sub>C</sub> conf. 1; (E) <i>cluster</i> GLC <sub>C</sub> conf. 2; (F) <i>cluster</i> final GLC <sub>D</sub> . Interações não covalentes num raio de 10 Å da molécula (C) <i>cluster</i> FRC <sub>C</sub> conf. 2; (D) <i>cluster</i> FRC <sub>C</sub> conf. 1; (G) <i>cluster</i> GLC <sub>C</sub> conf. 2; (H) <i>cluster</i> final GLC <sub>D</sub> .....	51
Figura 16 –	Densidade eletrônica. (A, B) forma fechada e (C, D) forma aberta de frutose antes e depois da dinâmica molecular .....	56
Figura 17 –	Comparação da densidade eletrônica inicial e final. (A, B) glucopiranosose e (C, D) glicose de cadeia longa .....	57
Figura 18 –	O painel BIRD exibe a energia de interação dos resíduos com as regiões moleculares para os <i>clusters</i> de energia mais baixa e mais alta. (A) <i>Cluster</i> FROC conf. 0; (B) FRO <sub>D</sub> final conf.; (E) GLO <sub>D</sub> <i>cluster</i> conf. 0; (F) GLO <sub>D</sub> <i>cluster</i> conf. 2. Interações não covalentes num raio de 10 Å da molécula (C) FRO <sub>C</sub> <i>cluster</i> conf. 0; (D) FRO <sub>D</sub> conf. final; (G) GLO <sub>D</sub> <i>cluster</i> conf. 0; (H) GLO <sub>D</sub> <i>cluster</i> conf. 2 .....	59
Figura 19 –	O painel BIRD exibe a energia de interação dos resíduos com as regiões moleculares para os <i>clusters</i> de energia mais baixa e mais alta. (A) <i>cluster</i> GLO <sub>mC</sub> conf. 1; (B) GLO <sub>mD</sub> <i>conf. final</i> . Interações não covalentes num raio de 10 Å da molécula (C) GLO <sub>mC</sub> <i>cluster</i> conf. 1; (D) GLO <sub>mD</sub> <i>conf. final</i> .....	61
Figura B1 –	Fluxo de trabalho do CB-Dock2. <i>Docking</i> cego com integração de modelos homólogos .....	98
Figura B2 –	Área de submissão de proteínas no CB-Dock2. Avisos importantes para Otimização da Preparação .....	99
Figura B3 –	Modalidades de tarefa no CB-Dock2. <i>Search Cavities</i> e <i>Auto Blind Docking</i> .....	101
Figura B4 –	Remoção de resultados no CB-Dock2. Gerenciando tarefas .....	102



Figura B5 – Detecção e análise de cavidades com base na estrutura no CB-Dock2. Exploração estrutural e docking molecular .....	103
Figura B6 – Detecção e análise de cavidades com base em modelos no CB-Dock2. Exploração estrutural e docking molecular .....	104
Figura B7 – <i>Docking</i> cego baseado na estrutura no CB-Dock2. Análise estrutural e visualização 3D dos modos de ligação .....	105
Figura B8 – <i>Docking</i> cego baseado em modelos no CB-Dock2. Análise estrutural e visualização 3D dos modos de ligação .....	106
Figura B9 – Análise integrativa de docking cego no CB-Dock2. Combinação de métodos estrutural e baseado em modelos para predição de poses de ligação .....	107
Figura D1 – Fluxograma do Processo Autoconsistente. Solução das Equações de Kohn-Sham em DFT .....	115

## LISTA DE TABELAS

Tabela 1 –	Complexos de açúcares complexados com HSA utilizados nas simulações .....	22
Tabela 2 –	Energia de interação de todas as simulações, com valor médio, desvio padrão ( $\sigma$ ) e coeficiente de variação ( $C_v$ ) .....	49
Tabela A1 –	Resíduos, distância (Å), energia de interação ( $\text{kcal.mol}^{-1}$ ), átomo da HSA, átomo da $\text{FRC}_C$ ( <i>cluster conf. 2</i> ) .....	77
Tabela A2 –	Resíduos, distância (Å), energia de interação ( $\text{kcal.mol}^{-1}$ ), átomo da HSA, átomo da $\text{FRC}_C$ ( <i>cluster conf. 1</i> ) .....	79
Tabela A3 –	Resíduos, distância (Å), energia de interação ( $\text{kcal.mol}^{-1}$ ), átomo da HSA, átomo da $\text{GLC}_C$ ( <i>cluster conf. 2</i> ) .....	81
Tabela A4 –	Resíduos, distância (Å), energia de interação ( $\text{kcal.mol}^{-1}$ ), átomo da HSA, átomo da $\text{GLC}_D$ (conf. final) .....	83
Tabela A5 –	Resíduos, distância (Å), energia de interação ( $\text{kcal.mol}^{-1}$ ), átomo da HSA, átomo da $\text{FRO}_C$ ( <i>cluster conf. 0</i> ) .....	85
Tabela A6 –	Resíduos, distância (Å), energia de interação ( $\text{kcal.mol}^{-1}$ ), átomo da HSA, átomo da $\text{FRO}_D$ (conf. final) .....	87
Tabela A7 –	Resíduos, distância (Å), energia de interação ( $\text{kcal.mol}^{-1}$ ), átomo da HSA, átomo da $\text{GLO}_D$ ( <i>cluster conf. 0</i> ) .....	89
Tabela A8 –	Resíduos, distância (Å), energia de interação ( $\text{kcal.mol}^{-1}$ ), átomo da HSA, átomo da $\text{GLO}_D$ ( <i>cluster conf. 2</i> ) .....	91
Tabela A9 –	Resíduos, distância (Å), energia de interação ( $\text{kcal.mol}^{-1}$ ), átomo de HSA, átomo da $\text{GLO}_{mC}$ ( <i>cluster conf. 1</i> ) .....	93
Tabela A10 –	Resíduos, distância (Å), energia de interação ( $\text{kcal.mol}^{-1}$ ), átomo da HSA, átomo da $\text{GLO}_{mC}$ (conf. final) .....	95

## LISTA DE ABREVIATURAS E SIGLAS

CGenFF	Charmm General Force Field
DFT	Density Functional Theory
DNP	Double Numerical Plus Polarization
EnGens	Ensemble Generation
FRC	$\beta$ -D-Frutofuranose
FRO	Keto-D-Frutose
GGA	Generalized Gradient Approximation
GLC	$\alpha$ -D-Glucopiranosose
GLO	Aldeído-D-Glicose
GLOm	hexane-1,2,3,4,5-pentol (D-Glicose modificada)
GMM	Gaussian Mixture Models
HSA	Human Serum Albumin
K	Aminoácido Lisina
MD	Molecular Dynamics
MFCC	Molecular Fractionation with Conjugate Caps
PBE	Perdew-Burke-Ernzerhof
PBS	Solução de Tampão Salina-Fosfato
PDB	Protein Data Bank
pH	Potencial Hidrogeniônico
QM/MM	Quantum Mechanics/Molecular Mechanics
RMSD	Root Mean Square Deviation
SCF	Self Consistent Field
TRP	Aminoácido Triptofano
TS	Tkatchenko-Scheffler
UMAP	Uniform Manifold Approximation And Projection
VMD	Visual Molecular Dynamics

## SUMÁRIO

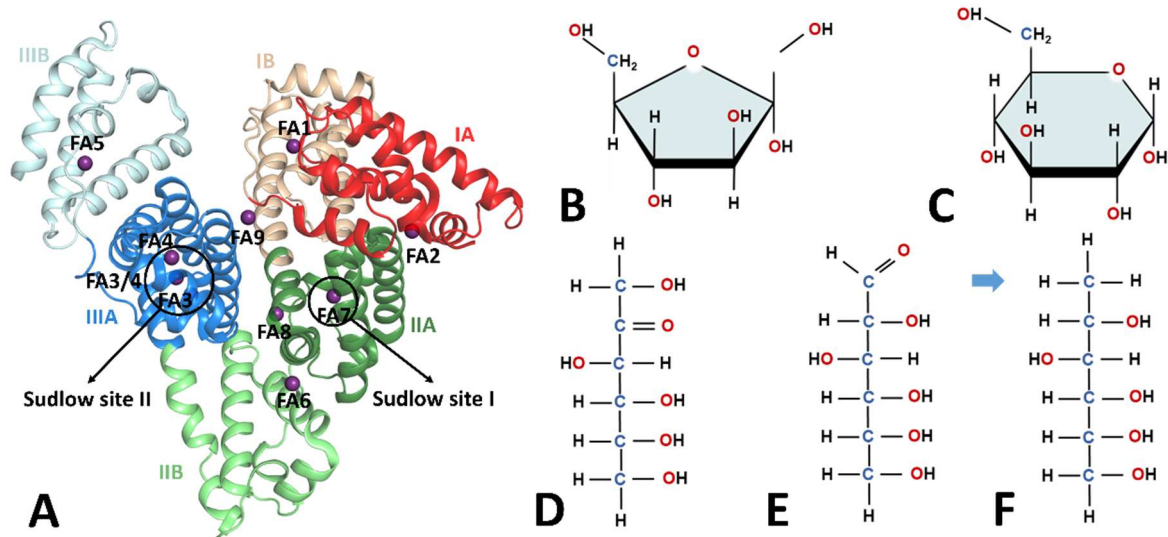
1	INTRODUÇÃO .....	15
2	MATERIAL E MÉTODOS .....	20
2.1	Espectroscopia de Fluorescência .....	20
2.2	Aquisição e preparação das estruturas de HSA, FRC, FRO, GLC, GLO e GLOm .....	20
2.3	<i>Docking</i> molecular .....	21
2.4	Dinâmica molecular (MD) .....	22
2.5	Geração de <i>clusters</i> conformacionais representativos de proteínas .....	24
2.6	Fracionamento molecular com capas conjugadas (MFCC) e bioquímica quântica .....	25
3	RESULTADOS E DISCUSSÃO .....	28
3.1	Supressão da fluorescência .....	28
3.2	<i>Docking</i> molecular .....	29
3.3	Dinâmica molecular .....	30
3.4	Bioquímica quântica e detalhes atômicos das ligações de FRC/FRO/GLC/GLO/GLOm com a HSA .....	47
3.4.1	<i>Interações FRC::HSA no Sudlow Site I</i> .....	50
3.4.2	<i>Interações GLC::HSA no Sudlow Site I</i> .....	53
3.4.3	<i>Interações FRO::HSA no interior do Sudlow Site I</i> .....	58
3.4.4	<i>Interações GLO/GLOm::HSA no Sudlow Site I</i> .....	60
4	CONCLUSÃO .....	64
	REFERÊNCIAS .....	67
	APÊNDICE A – TABELAS .....	77
	APÊNDICE B – <i>DOCKING</i> : CB-DOCK2 .....	97
	APÊNDICE C – DINÂMICA MOLECULAR (DM) .....	108
	APÊNDICE D – TEORIA DO FUNCIONAL DA DENSIDADE (DFT) .....	111
	ANEXO A – ARTIGO PUBLICADO .....	119

## 1 INTRODUÇÃO

A albumina do soro humano (HSA), que constitui até 60% do total das proteínas séricas, desempenha um papel fundamental na facilitação do transporte de ligantes endógenos e exógenos através de interações não covalentes reversíveis (CARTER; HO, 1994; PETERS, 1995). A importância da HSA reside na sua notável capacidade de ligação, que lhe permite interagir com vários fármacos, como a varfarina, o ibuprofeno e a indometacina (ANGUIZOLA *et al.*, 2013; PETITPAS *et al.*, 2001). É importante ressaltar que mudanças no ambiente patológico ou fisiológico podem potencialmente induzir alterações na conformação de uma proteína, conseqüentemente afetando a eficiência de ligação dos ligantes e levando ao comprometimento funcional da HSA (AWANG *et al.*, 2017; PONGPRAYOON; GLEESON, 2014). Entre as múltiplas deficiências encontradas na HSA, a presença de glicação emergiu como um fator significativo, conforme descrito por Jana *et al.* (2016).

Com uma disposição notável de 585 aminoácidos distribuídos por três domínios helicoidais homólogos (I-III), cada um dos quais compreende dois subdomínios (A e B), a HSA é um componente indispensável em vários processos cruciais, como a absorção, distribuição, metabolismo e excreção de pequenas moléculas. A organização estrutural da HSA é composta por nove locais de ligação distintos nos seus domínios, especificamente concebidos para acomodar ácidos graxos de cadeia longa e média (FA). É de salientar que os sítios primários de ligação de fármacos estão situados nos sítios hidrofóbicos dos subdomínios II-A (comumente referido como *Sudlow Site I* ou FA7) e III-A (conhecido como *Sudlow Site II* ou FA3) (ASCENZI *et al.*, 2015; FANALI *et al.*, 2012), conforme ilustrado na Figura 1A. A palavra "*Sudlow*" designa dois locais específicos de ligação de pequenas moléculas na albumina sérica, identificados pelo bioquímico britânico Gerald Sudlow na década de 1970. Sudlow e sua equipe descobriram que a albumina sérica humana possui dois locais principais de ligação, agora conhecidos como *Sudlow Site I* e *Sudlow Site II* (SUDLOW; BIRKETT; WADE, 1975).

Figura 1 - Estruturas da albumina do soro humano (HSA), frutose e glicose. (A) A estrutura global previamente determinada da HSA com os seus subdomínios identificados por cores diferentes. Estrutura química para (B)  $\beta$ -D-Fructofuranose (FRC), (C)  $\alpha$ -D-Glucopiranosose (GLC), (D) Keto-D-Frutose (FRO), (E) Aldeído-D-Glicose (GLO) e (F) hexane-1,2,3,4,5-pentol (D-Glicose modificada) (GLOm).



Os hidratos de carbono, amplamente reconhecidos como frutose (FRC/FRO) e glicose (GLC/GLO) (Figura 1B-1E), são considerados as principais fontes fisiológicas de carbono, um componente crucial que assume uma posição fulcral num vasto conjunto de processos fisiológicos e bioquímicos inerentes à existência de organismos vivos. A glicose, uma aldohexose, é o açúcar predominante na circulação sanguínea e a frutose, um isómero cetohexose da glicose, apresenta características muito próprias. É digno de nota que as manifestações de cadeia aberta de D-frutose e D-glicose, comumente denotadas como FRO e GLO, abrangem uma gama diversificada de configurações estruturais que abrangem as formas cetohexose e aldohexose como suas respectivas representações (QI; TESTER, 2019). Curiosamente, investigações preliminares conduzidas por nossos estudos forneceram indicações da eficiência limitada de ligação entre a glicose aberta e a Albumina de Soro Humano (HSA), levando a uma exploração completa de sua interação. Como resultado, essa investigação acabou culminando em uma modificação estrutural da molécula de glicose, que é visualmente retratada na Figura 1F.

O FRC/FRO e o GLC/GLO são eficientemente assimilados e metabolizados por vários tecidos, em particular o fígado, e estão implicados em condições relacionadas com o metabolismo dos monossacarídeos, enquanto uma ingestão elevada de hidratos de carbono está associada a uma série de doenças (ASP, 1994). Em particular, o consumo de bebidas ricas em

frutose está associado a uma maior suscetibilidade à obesidade, à diabetes de tipo 2 e às doenças cardiovasculares (MALIK; HU, 2015). Além disso, a investigação salientou o impacto dos níveis elevados de glicose no plasma sobre a capacidade de ligação da HSA (CHAVES; SOARES; CAMPOS DE OLIVEIRA, 2021).

A intrincada interação entre os açúcares redutores e as proteínas é desvendada por modificações químicas, especificamente a glicação não enzimática. Esta modificação pós-translacional, que resulta de uma reação de condensação entre o grupo carbonila de um açúcar redutor e o grupo amino de uma proteína, produz produtos finais de glicação avançada (BAI *et al.*, 2012; KHALIFAH; BAYNES; HUDSON, 1999; REYNOLDS, 1963; THORNALLEY; LANGBORG; MINHAS, 1999). As proteínas glicadas, como a HSA, apresentam discrepâncias funcionais em comparação com as suas homólogas não glicadas (BROWNLEE, 1991) e contribuem para os danos tecidulares a longo prazo associados à fisiopatologia da diabetes (TRÜEB *et al.*, 1980).

Neste cenário abrangente, os avanços na determinação da estrutura molecular, facilitados pelo uso da cristalografia de raios X (BHATTACHARYA; GRÜNE; CURRY, 2000; GHUMAN *et al.*, 2005; PETITPAS *et al.*, 2001; ZUNSZAIN *et al.*, 2003, 2008) e da espectroscopia de RMN (HAMILTON, 2013; KRENZEL; CHEN; HAMILTON, 2013; SIMARD *et al.*, 2006), é complementada pela aplicação de técnicas de biologia computacional que permitem o estudo de complexos proteína-fármaco para além do âmbito das interações moleculares (GRINTER; ZOU, 2014; KITCHEN *et al.*, 2004). Estas técnicas incluem a dinâmica molecular clássica (GANESAN; COOTE; BARAKAT, 2017; SMITH; STERNBERG, 2002; ZHAO; CAFLISCH, 2015), a mecânica molecular (PEARLMAN *et al.*, 1995; RAPPE *et al.*, 1992) e métodos quânticos semi-empíricos (CHRISTENSEN *et al.*, 2016; THIEL, 2014). Cálculos de bioquímica quântica utilizando a teoria do funcional da densidade (DFT) (ANTONY; GRIMME, 2006, 2012; BECKE, 1992; EHRLICH; MOELLMANN; GRIMME, 2013; FOX *et al.*, 2014; JONES, 2015; MARDIROSSIAN; HEAD-GORDON, 2017) combinados com estratégias de fracionamento molecular (COLLINS; BETTENS, 2015; GORDON *et al.*, 2012a; JIANG; MA; JIANG, 2006; LIU; HERBERT, 2016; SAHA; RAGHAVACHARI, 2015; SÖDERHJELM; AQUILANTE; RYDE, 2009; WANG *et al.*, 2013a; ZHANG; ZHANG, 2003) são precisas para estimar as energias de interação.

A teoria do funcional da densidade (DFT) é uma ferramenta fundamental no arsenal da modelagem molecular, especialmente quando se trata de investigar e prever as interações entre ligantes e proteínas. A DFT permite calcular as propriedades eletrônicas de sistemas complexos,

incluindo moléculas isoladas e biomacromoléculas, a partir da distribuição da densidade eletrônica. Isso é particularmente relevante para entender como um ligante se liga a uma proteína, pois a DFT pode fornecer detalhes precisos sobre a energia de interação e as forças de atração ou repulsão que governam o comportamento dinâmico do complexo. Ao combinar a DFT com outras técnicas de modelagem molecular, é possível não apenas prever a estrutura e a estabilidade dos complexos ligante-proteína, mas também explorar a reatividade química e a eficiência do processo de ligação. Esta abordagem integrada é crucial para o design racional de novos fármacos e para a otimização de suas propriedades farmacológicas. Para um aprofundamento sobre a Teoria do Funcional da Densidade (DFT), consulte o Apêndice D.

As estruturas cristalinas tridimensionais da HSA ligada à frutose ou à glicose forneceram informações sobre a ligação das formas de cadeia fechada (GLC/FRC) e de cadeia aberta (GLO/FRO) no *Sudlow Site I* (FA7). A GLC estava localizada na parte inferior do sítio, enquanto a GLO estava ligada covalentemente a LYS195 (WANG *et al.*, 2013b). O LYS195 era o principal local para as reações da glicose e do HSA (NASIRI *et al.*, 2010). Não foi registrada qualquer ligação covalente entre a frutose e a HSA (WANG *et al.*, 2013b). As simulações de dinâmica molecular delinearam ainda mais a dinâmica destas interações, destacando os diferentes graus de movimento do domínio que afetam a afinidade de ligação fármaco/soluto. Embora investigações teóricas anteriores tenham realizado exames exaustivos da interação entre a HSA e diversos sacarídeos empregando uma relação estequiométrica 1:2 (HSA simultaneamente associada às configurações fechada e aberta do açúcar), optamos por concentrar os nossos esforços computacionais numa relação estequiométrica 1:1 (HSA ligada à configuração fechada ou aberta do açúcar). Esta decisão é motivada pelo mecanismo de abertura do anel desses açúcares, que requer que uma molécula de açúcar ocupe inicialmente a parte inferior do *Sudlow Site I*, enquanto uma segunda molécula de açúcar, posicionada na entrada do *Sudlow Site I*, recebe um próton de LYS199 (K199), facilitando a abertura subsequente do anel. Além disso, estudos teóricos/estruturais anteriores demonstraram a fuga de açúcares dos bolsos HSA (AWANG *et al.*, 2021; PONGPRAYOON; MORI, 2018), destacando a importância de uma inspeção 1:1.

O estudo da influência de diferentes formas de glicose e frutose antes das reações de glicação se manifestarem desempenha um papel crucial na compreensão das fases iniciais deste processo e da forma como estas etapas afetam a dinâmica das proteínas. A investigação destas fases iniciais pode fornecer informações significativas sobre os processos fundamentais e possíveis intervenções. Os estudos teóricos dedicados aos efeitos das diferentes manifestações



da glicose e da frutose têm o potencial de contribuir significativamente para o desenvolvimento de biomarcadores. Este progresso é particularmente relevante em doenças como a diabetes, fornecendo ferramentas para a detecção precoce e a monitorização eficaz da progressão da doença. Além disso, uma compreensão completa dos impactos destes açúcares pode ter um significado importante no avanço de medicamentos especializados que visam manipular as interações proteína-ligante.

Na nossa investigação preliminar, iniciámos um exame da interação não covalente entre FRO, FRC, GLO e GLC e a albumina de soro humano (HSA). Notavelmente, as nossas simulações computacionais de dinâmica molecular não revelaram um acoplamento significativo da GLO à HSA. Esta observação levou a uma abordagem inovadora: a modificação da molécula de GLO. Eliminando a ligação dupla com o oxigênio e substituindo-a por hidrogênio, criámos uma versão modificada denominada hexane-1,2,3,4,5-pentol, que optamos por chamar de D-glicose modificada (GLOm). Esta alteração estrutural visou a ligação covalente conhecida entre GLO e HSA, especificamente o local do carbono anteriormente ocupado pela ligação dupla de oxigênio. Notavelmente, após a implementação desta modificação, GLOm mostrou uma interação pronunciada com HSA, demonstrando a importância potencial desta adaptação estrutural na modulação da dinâmica de ligação entre derivados de glicose e HSA. Esta investigação será levada a cabo através de uma abordagem abrangente que combina a utilização de medidas de espectroscopia de fluorescência com a integração de simulações de *docking* e dinâmica molecular, *clusters* conformacionais de proteínas (EnGens), fracionamento molecular com capas conjugadas (*Molecular Fractionation with Conjugate Caps* - MFCC) e análise de bioquímica quântica.

## 2 MATERIAL E MÉTODOS

### 2.1 Espectroscopia de fluorescência

As medições de fluorescência foram efetuadas com um espectrômetro de fluorescência Perkin Elmer LS45 (Perkin-Elmer, Reino Unido) equipado com uma célula de quartzo de 1 cm. O HSA (isento de ácidos graxos, Sigma) foi excitado ( $\lambda_{exc}$ ) a 280 nm e a emissão registrada entre 300 e 500 nm, com as larguras das bandas de excitação e emissão fixadas em 10 nm cada. Os espectros de fluorescência foram obtidos a uma concentração fixa de HSA (10  $\mu$ M em 2,5 mL de solução salina tamponada com fosfato (PBS), pH 7,4) titulada com uma alíquota de 10  $\mu$ L de uma solução de D-glicose ou D-frutose em solução de tampão salina-fosfato (PBS - pH 7,4). A concentração final dos monossacarídeos foi de 0,2, 0,4, 0,6, 1 e 1,4 mM. Todas as medições foram efetuadas em triplicata a 25 °C. A concentração de HSA foi determinada espectrofotometricamente utilizando o coeficiente de extinção molar de 35.700 M<sup>-1</sup> cm<sup>-1</sup>, obtido de estudos anteriores (LEGGIO; GALANTINI; PAVEL, 2008).

### 2.2 Aquisição e preparação das estruturas de HSA, FRC, FRO, GLC, GLO e GLOm

As estruturas tridimensionais dos complexos FRC/FRO:HSA e GLC/GLO:HSA foram obtidas do Protein Data Bank com os identificadores 4IW1 (resolução de 2,56 Å) e 4IW2 (resolução de 2,41 Å), respectivamente (WANG *et al.*, 2013b). A remoção de água e pequenos ligantes da estrutura cristalina de HSA foi realizada usando o *Discovery Studio 2016* da *Dassault Systems BIOVIA*. Para calcular os estados de protonação dos resíduos de HSA a pH 7,4 e para completar os seus hidrogênios, utilizamos o servidor *web ProteinPrepare*, conforme descrito por Martínez-Rosell *et al.*, 2017. Além disso, as estruturas tridimensionais de GLC, GLO, GLOm, FRC e FRO foram submetidas à análise de carga em pH 7,4 usando *MarvinSketch 18.24* (*Marvin Beans Suite, ChemAxon*), seguido de otimização de geometria no *Discovery Studio 2016*.

Como mencionado acima, as nossas simulações computacionais de dinâmica molecular não revelaram um acoplamento significativo da GLO à HSA. Considerando que a formação da ligação covalente GLO:LYS195 resulta na libertação de um átomo de oxigênio ligado ao carbono 1 da glicose linear (NASIRI *et al.*, 2010), decidimos excluir a ligação dupla (C1=O) proveniente da abertura da piranose e substituí-la por átomos de hidrogênio. Esta abordagem

foi implementada para avaliar se esta modificação pré-glicação resultaria num complexo mais estável em comparação com a versão não modificada GLO.

### 2.3 *Docking* molecular

Decidimos realizar simulações de dinâmica molecular da albumina do soro humano (HSA) em duplicata para explorar diferentes perspectivas sobre a interação da proteína com açúcares. No entanto, para garantir uma abordagem abrangente, decidimos evitar utilizar o mesmo ligante como ponto de partida em ambas as simulações. Por conseguinte, na primeira simulação, decidimos utilizar o ligante derivado de dados cristalográficos para representar uma interação mais realista com a HSA com base em estruturas experimentais. Na segunda simulação, decidimos utilizar o resultado da acoplagem molecular como ligante, proporcionando uma abordagem computacional que tem em conta a complementaridade e a afinidade entre os açúcares e a HSA. Esta escolha estratégica visa proporcionar uma visão mais abrangente das interações moleculares envolvidas, justificando assim os cálculos de *docking* molecular como um passo crucial para enriquecer a nossa compreensão da dinâmica molecular da HSA em conjunto com diferentes açúcares.

Neste trabalho, implementamos o algoritmo CB-Dock2 para realizar o *docking* cego automatizado proteína-ligante, seguindo as quatro etapas descritas na literatura (LIU *et al.*, 2022). A primeira fase consistiu na entrada de dados, onde fornecemos os ficheiros PDB da proteína de interesse e do ligante, ambos derivados de dados cristalográficos. O CB-Dock2 efetuou então o processamento dos dados, gerou um modelo da cavidade de ligação da proteína e criou poses iniciais para o ligante. Na terceira etapa, realizou-se a deteção da cavidade e a acoplagem. Nessa etapa, o CB-Dock2 inicia a pesquisa de complexos conhecidos, utilizando dois pipelines paralelos: acoplagem cega baseada na estrutura e acoplagem baseada em modelos, para posteriormente calcular a energia de interação entre cada pose do ligante e a cavidade de ligação. Finalmente, a quarta etapa envolveu a visualização e análise dos resultados, utilizando as ferramentas fornecidas pelo CB-Dock2 para explorar a estrutura da proteína, analisar a energia de interação entre as moléculas e examinar a pose do ligante na cavidade de ligação da proteína.

*Docking* cego refere-se a uma técnica de *docking* molecular em que a posição exata do sítio de ligação do ligante (neste caso, glicose ou frutose) na proteína alvo (HSA, Albumina Sérica Humana) não é conhecida a priori. Em vez de se focar em um local específico predefinido

para o *docking*, o "*docking* cego" explora toda a superfície da proteína para encontrar os potenciais sítios de ligação do ligante.

Para mais detalhes sobre o CB-Dock2, incluindo a descrição completa de seu funcionamento e metodologia utilizada no *docking* cego automatizado proteína-ligante, consulte o Apêndice B deste trabalho.

## 2.4 Dinâmica molecular (MD)

O GROMACS 2021.2 foi utilizado para fundir os ligantes e o HSA na caixa de água e para efetuar simulações de minimização de energia, NVT, NPT e MD. As topologias do HSA foram determinadas usando o campo de força CHARMM36 (BEST *et al.*, 2012). Os parâmetros das coordenadas cristalográficas atômicas e as posições de *docking* molecular com os valores RMSD mais baixos foram fornecidos por CGenFF (VANOMMESLAEGHE *et al.*, 2009) e GROMACS 2021.2. Foram então construídos os seguintes complexos:

Tabela 1 - Complexos de açúcares complexados com HSA utilizados nas simulações.

	<i>FRC</i>	<i>FRO</i>	<i>GLC</i>	<i>GLO</i>	<i>GLOm</i>
<i>Crystallographic (C)</i>	FRC <sub>C</sub> :HSA	FRO <sub>C</sub> :HSA	GLC <sub>C</sub> :HSA	GLO <sub>C</sub> :HSA	GLO <sub>mC</sub> :HSA
<i>Docking (D)</i>	FRC <sub>D</sub> :HSA	FRO <sub>D</sub> :HSA	GLC <sub>D</sub> :HSA	GLO <sub>D</sub> :HSA	GLO <sub>mD</sub> :HSA

Com base nas estruturas de HSA previamente preparadas. Estes complexos foram inseridos separadamente em diferentes caixas de água (com base no modelo TIP3) com uma concentração iônica de 0,15 M (Na<sup>+</sup> e Cl<sup>-</sup>). Os complexos gerados foram minimizados energeticamente por um algoritmo de gradiente descendente, e essas simulações foram interrompidas quando a força máxima dos sistemas era inferior a 1.000 kJ.mol<sup>-1</sup>.nm<sup>-1</sup>. Em seguida, foram efetuados 5.000 passos de simulações NVT para equilibrar a temperatura dos sistemas a aproximadamente 300 K, utilizando um termostato de reescalonamento da velocidade, o termostato V-rescale. Em seguida, 5.000 etapas de simulações NPT baseadas no par de pressão isotrópica de Berendsen foram realizadas como uma continuação da NVT, usando o barostato de Berendsen em conjunto com o termostato V-rescale para manter a temperatura e a pressão do sistema durante a fase de equilíbrio NPT (BERENDSEN *et al.*, 1984). A integração temporal das simulações NVT e NPT foi realizada utilizando um algoritmo *Leap-Frog* com um passo de tempo de 2 fs, possibilitado pela combinação do algoritmo LINCS e da repartição da massa de hidrogênio. A preferência pelo algoritmo *Leap-Frog* em detrimento

de outros é motivada pela sua notável estabilidade na conservação de energia, simplicidade de implementação e eficácia na simulação de sistemas complexos. O corte de *Verlet* foi utilizado para calcular as interações não ligadas, com um corte de distância de 1,0 nm para as interações de Van Der Waals e de Coulomb no espaço real. Ao mesmo tempo, foram utilizadas condições de fronteira periódicas e o esquema Ewald de malha de partículas com parâmetros predefinidos para tratar a parte eletrostática de longo alcance das interações de Coulomb. Em seguida, foram efetuadas simulações MD de 100 ns para cada complexo.

O desvio médio quadrático (RMSD) das posições atômicas de HSA, FRC/FRO e GLC/GLO/GLOm também foi calculado para todas as conformações de dinâmica molecular nas suas amostras computacionais utilizando o módulo *gmx rms* GROMACS. Os resultados RMSD foram a média entre duas simulações MD (Crystal e Redocking). As coordenadas atômicas finais obtidas por este procedimento foram depois utilizadas para todos os cálculos de interação baseados em interações não ligadas. O VMD 1.9.3 foi utilizado para efetuar uma análise detalhada da mobilidade do subdomínio HSA e das ligações de hidrogênio formadas entre os açúcares e a albumina durante as simulações. A distância doador-aceitador (3,0 Å) e o ângulo de corte (20°) foram os parâmetros utilizados como critérios para o cálculo das ligações de hidrogênio.

Os resultados obtidos revelaram um comportamento interessante da GLO em relação à HSA, evidenciando que a molécula não possuía uma capacidade de ligação significativa com a proteína em uma proporção de 1:1. No entanto, uma modificação específica na estrutura molecular da glicose, que envolveu a eliminação da dupla ligação com o oxigênio O1 e a saturação das valências com hidrogênio, provocou uma mudança notável no cenário de interação. Propomos que ocorra uma reação química no sítio de ligação, resultando na transformação da molécula de glicose em hexane-1,2,3,4,5-pentol. É importante destacar que essa molécula, que ocorre naturalmente em plantas da espécie *Cnidium*, apresentou uma interação mais eficaz com a HSA em comparação com a glicose.

Para obtenção dos gráficos 2D das interações entre as moléculas e os resíduos da HSA, foi utilizado o software *Discovery Studio Receptor-Ligand Interactions*, que permitiu visualizar as interações sob uma perspectiva estática. Com base num ligante de qualidade média e preservando as interações de hidrogênio, foram explorados os tipos de interações a uma distância de 4 Å da molécula. Os critérios para definir as ligações de hidrogênio envolveram a medição de distâncias e ângulos entre átomos específicos, com uma distância D-A máxima de 3,4 Å. As relações angulares foram limitadas a 90 e 180 graus para diferentes combinações de

átomos. As ligações de hidrogênio com a água também foram consideradas quando a molécula de água formava uma ponte entre diferentes moléculas que satisfaziam os critérios.

O formalismo da Dinâmica Molecular (DM) é essencial para simular o movimento de partículas em sistemas físicos. Seguindo a equação de movimento de Newton para uma partícula (Eq. 1):

$$F_i(t) = m_i \alpha_i \quad \text{Eq. 1}$$

onde  $F_i(t)$  representa a força atuando sobre uma partícula específica no sistema no tempo  $t$ ,  $m_i$  é a massa da partícula e  $\alpha_i$  é sua aceleração.

O potencial de interação entre os átomos no sistema é modelado pelo campo de força, que inclui a energia potencial total  $V(r)$ , abrangendo ligações químicas, interações de van der Waals e Coulomb, descrita pela equação (Eq. 2):

$$V(r) = \sum V_{lig.}(r) + \sum V_{angular}(\theta) + \sum V_{diedro}(\phi) + \sum V_{vdw} + \sum V_{elet.} \quad \text{Eq. 2}$$

Essa abordagem permite a modelagem computacional do comportamento dinâmico das partículas no sistema, levando em conta tanto interações covalentes quanto não covalentes. Para mais detalhes sobre o formalismo e os métodos utilizados na Dinâmica Molecular, consulte o apêndice C.

## 2.5 Geração de *clusters* conformacionais representativos de proteínas

O método EnGens (abreviatura de *ensemble generation*) para gerar conformações representativas de proteínas a partir da dinâmica molecular divide-se em três etapas principais: importação dos resultados da dinâmica molecular, os resultados são importados sob a forma de arquivos de trajetórias que contêm as coordenadas espaciais dos átomos da proteína em cada instante de tempo; filtragem dos resultados da dinâmica molecular, removendo trajetórias incompletas ou com erros; e reconstrução das estruturas conformacionais, que é realizada utilizando o algoritmo de média, que calcula a média das coordenadas espaciais dos átomos em cada instante de tempo (CONEV *et al.*, 2023).

Após a reconstrução das estruturas conformacionais, é utilizado o algoritmo *Uniform Manifold Approximation and Projection* (UMAP) para reduzir a dimensionalidade das estruturas. O UMAP é um algoritmo não linear que preserva as relações estruturais entre as estruturas conformacionais (TROZZI; WANG; TAO, 2021). Ao reduzir a dimensionalidade, as

estruturas conformacionais podem ser visualizadas num espaço de baixa dimensão, o que facilita a análise e a interpretação dos resultados.

Para otimizar a eficiência computacional, neste caso particular optamos por estudar três conformações representativas, ou seja, o algoritmo UMAP identifica três grupos de estruturas conformacionais que são significativamente diferentes entre si. Estes grupos de estruturas conformacionais são então submetidos ao algoritmo *Gaussian Mixture Models* (GMM) (HAND; MCLACHLAN; BASFORD, 1989), que é um algoritmo de aprendizagem automática probabilística que assume que os dados provêm de uma mistura de distribuições normais. O algoritmo GMM é então utilizado para refinar o agrupamento inicial efetuado pelo UMAP. Esta combinação de algoritmos permite identificar conformações representativas da flexibilidade da proteína. Uma vez obtidos os resultados, foram efetuados cálculos de fracionamento molecular com capas conjugadas (MFCC) e análises bioquímicas quânticas sobre as três conformações representativas fornecidas pelo EnGens, bem como sobre a conformação final encontrada na dinâmica molecular.

## **2.6 Fracionamento molecular com capas conjugadas (MFCC) e bioquímica quântica**

Com base nos resultados obtidos, a análise prosseguiu com cálculos de fracionamento molecular com capas conjugadas (*Molecular Fractionation with Conjugate Caps* - MFCC) das três conformações representativas geradas por EnGens, para além do estudo da conformação final identificada por dinâmica molecular. Este método permitiu a avaliação das energias de interação entre as moléculas mais relevantes dos complexos estudados utilizando a teoria do funcional da densidade (DFT) (HOHENBERG; KOHN, 1964; JONES, 2015; SHAM; KOHN, 1966).

A abordagem MFCC permite um estudo mais aprofundado de muitos resíduos de uma proteína com um tempo de cálculo reduzido, mantendo a exatidão (CHEN; ZHANG; ZHANG, 2005; HE; ZHANG, 2005; ZHANG; ZHANG, 2003). Para limitar o número de resíduos a analisar sem perder interações críticas, foram consideradas as interações não covalentes num raio de 10 Å da molécula de açúcar e as moléculas de água num raio de 2,5 Å de cada resíduo de HSA e açúcar. É de salientar que, devido ao elevado custo computacional, apenas na conformação final da dinâmica molecular foram consideradas as moléculas de água. Além disso, os cálculos MFCC e DFT efetuados para conformações representativas fornecidas pelo EnGens não tiveram em conta os efeitos de solvatação. Esta abordagem otimiza a eficiência

computacional do estudo e permite que os cálculos sejam efetuados num período de tempo razoável utilizando os recursos computacionais disponíveis.

A energia de interação entre o  $i$ -ésimo resíduo de HSA de interesse ( $R_i$ ) e FRC/FRO/GLC/GLO/GLOm foi calculada do seguinte modo. Em primeiro lugar, foram definidos quatro sistemas auxiliares: ( $S_1$ )  $R_i$  mais os seus respectivos tampões conjugados  $C_{i-1}$  e  $C_{i+1}$  (os tampões estão ligados às porções amina e carboxil de  $R_i$  e refletem o ambiente eletrônico da sua vizinhança) e FRC/FRO/GLC/GLO/GLOm; ( $S_2$ )  $R_i$  mais os seus tampões conjugados apenas; ( $S_3$ ) os tampões conjugados de  $R_i$  e FRC/FRO/GLC/GLO/GLOm; ( $S_4$ ) os tampões conjugados de  $R_i$  apenas. Para cada sistema, a energia total  $E_{S_i}$  foi avaliada por um cálculo de *single point* utilizando formalismo DFT (que determina a energia de um sistema mantendo-se as posições atômicas fixas). Todas as ligações pendentes foram passivadas com átomos de hidrogênio e as tampas foram formadas a partir dos resíduos imediatamente adjacentes de  $R_i$  na cadeia proteica,  $C_{i\pm 1} = R_{i\pm 1}$ . A energia de interação entre FRC/FRO/GLC/GLO/GLOm e cada resíduo de aminoácido foi dada por:

$$E(\text{FRC/FRO/GLC/GLO/GLOm} - R_i) = E_{S_1} - E_{S_2} - E_{S_3} + E_{S_4} \quad (\text{Eq. 3})$$

Foram efetuados cálculos DFT-SP para investigar de forma detalhada as subcomponentes da Eq. 3 e descrever as energias de interação entre ligantes e proteínas. Para tal, foi utilizado o código DMol3 (Delley, 2000) do pacote de programas Biovia Accelrys Materials Studio. Foi escolhido um conjunto de bases *Double Numerical plus Polarization* (DNP) para expandir os orbitais Kohn-Sham de todos os elétrons. Para descrição da energia de correlação e troca, foi utilizada a aproximação do gradiente generalizado (sigla em inglês, GGA), conforme descrito por *Perdew-Burke-Ernzerhof* (PBE) (PERDEW; BURKE; ERNZERHOF, 1996), junto com o esquema de correção de dispersão *Tkatchenko-Scheffler* (GGA + TS) (TKATCHENKO; SCHEFFLER, 2009). O raio de corte orbital (distância crítica estabelecida para limitar a interação espacial entre orbitais atômicos ao calcular as propriedades eletrônicas de um sistema) foi fixado em 4 Å, e o limiar de convergência do campo autoconsistente (SCF) foi calibrado para  $10^{-6}$  Ha. A solvatação da água foi implementada através do modelo de solvatação implícito COSMO (KLAMT; SCHÜÜRMAN, 1993) com uma constante dielétrica de 40 ( $\epsilon=40$ ). Para compreender as interações entre solutos e solventes e seu impacto nas propriedades das estruturas, a constante dielétrica tem sido usada para descrever a polarização eletrostática (COSSI *et al.*, 2003). Neste estudo, utilizamos uma constante dielétrica de 40, considerada adequada para caracterizar o ambiente eletrostático em sistemas proteicos.



A constante dielétrica  $\epsilon_r(\omega)$ , que quantifica a capacidade de um material de reter cargas elétricas, desempenha um papel fundamental na simulação de fenômenos eletrostáticos dentro das proteínas, afetando sua configuração geral, resiliência e reatividade química. A constante dielétrica relativa  $\epsilon_r(\omega)$  é determinada comparando a permissividade de uma substância  $\epsilon(\omega)$  com a do vácuo  $\epsilon_0$ . Definida como:

$$\epsilon_r(\omega) = \frac{\epsilon(\omega)}{\epsilon_0} \quad \text{Eq. 4}$$

Apesar de ser rotulado como “constante”,  $\epsilon_r(\omega)$  está sujeito a flutuações com base em fatores como frequência do campo elétrico, temperatura, pressão e condições ambientais (MCKEEN, 2012).

Dentro da bioquímica quântica, uma compreensão profunda das interações eletrostáticas é vital para elucidar o comportamento e o papel das proteínas. No entanto, o cálculo dessas interações representa um desafio significativo devido à estrutura intrincada e aos diversos atributos eletrostáticos das proteínas (PERUTZ, 1978; WARSHEL *et al.*, 2006). Vários segmentos de uma proteína, incluindo resíduos de aminoácidos, cavidades e interações com solventes, exibem comportamentos dielétricos distintos que requerem incorporação em estruturas computacionais. Não existe uma constante dielétrica universal que seja universalmente aplicável a todas as proteínas (GORDON *et al.*, 2012b; WARSHEL; RUSSELL; CHURG, 1984). Estudos revelam um amplo espectro de valores, dependendo da propriedade específica sob investigação e do modelo escolhido. Notavelmente, uma constante dielétrica de 40 ( $\epsilon = 40$ ) é recomendada em estudos com foco em sistemas proteicos. Esse valor é considerado adequado para caracterizar o meio eletrostático desses sistemas e apresentar uma representação satisfatória das interações eletrostáticas (MORAIS *et al.*, 2020).

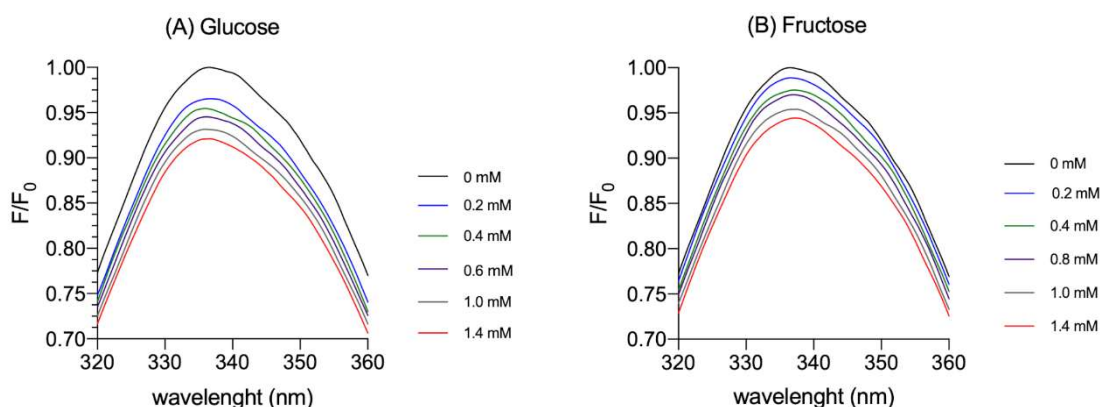
Finalmente, utilizando a base acima mencionada, as correções, o corte orbital (4 Å) e os ajustamentos do SCF, foram efetuados cálculos do potencial eletrostático (isosuperfícies) para as conformações inicial e final do açúcar em todos os complexos HSA::Açúcar.

### 3 RESULTADOS E DISCUSSÃO

#### 3.1 Supressão da fluorescência

A GLC e a FRC provocaram uma atenuação dependente da concentração da fluorescência intrínseca da HSA, com um efeito ligeiramente superior da GLC em comparação com a FRC (figura 2). Estes resultados indicam claramente a ligação de ambos os monossacarídeos à HSA. Embora a HSA apresente 40 resíduos aromáticos (31 fenilalanina, 18 tirosina e apenas um resíduo de triptofano) (ZHANG *et al.*, 2015), a sua fluorescência intrínseca é principalmente atribuída ao resíduo TRP214 encontrado na superfície do subdomínio II (sítio FA8) (LV *et al.*, 2022; TAYEH; RUNGASSAMY; ALBANI, 2009).

Figura 2 - Espectros de emissão de fluorescência. HSA titulada com glicose (A) e frutose (B). 10  $\mu$ M de HSA em tampão PBS (pH 7,4). Concentração de glicose e frutose: 0,2, 0,4, 0,6, 1 e 1,4 mM. Excitação a 280 nm. Fendas de excitação e emissão de 10 nm. Temperatura de 25 °C.



Os espectros também não apresentaram qualquer mudança no pico máximo de emissão de HSA (aproximadamente a 340 nm). Portanto, não ocorreu nenhuma alteração na hidrofobicidade do microambiente em torno do TRP214 após a ligação do GLC e do FRC ao HSA, consistente com estudos anteriores (SERGIO *et al.*, 2021).

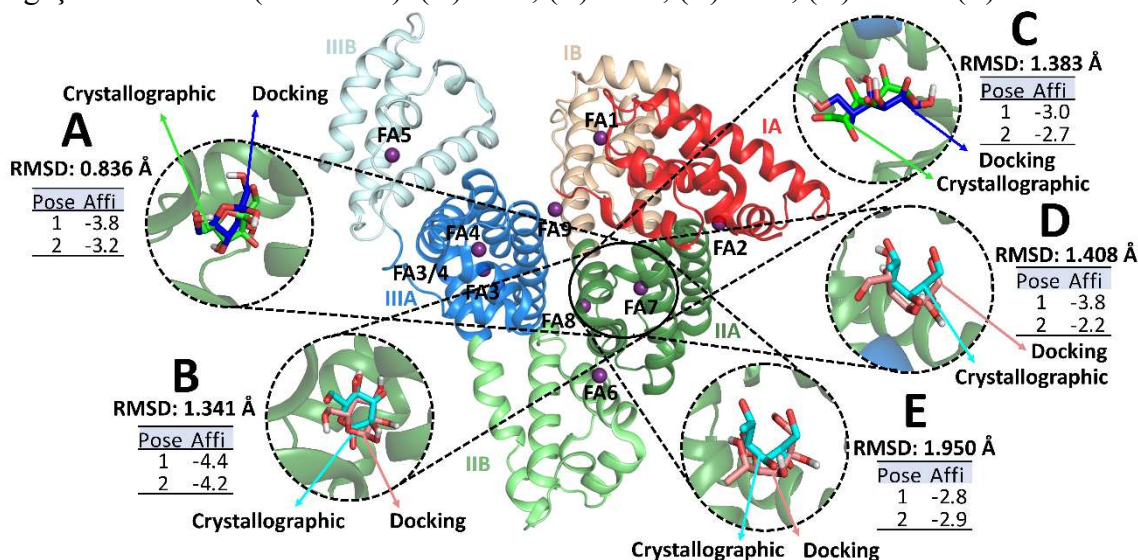
Os nossos resultados de extinção de fluorescência estão de acordo com um estudo anterior que relatou uma ligação fraca e espontânea de GLC e FRC a HSA (CHAVES; SOARES; CAMPOS DE OLIVEIRA, 2021). Os autores também mostraram que a interação de GLC e FRC com HSA ocorreu através de uma associação de estado fundamental (mecanismo de extinção de fluorescência estática) e foi principalmente governada por ligações de hidrogênio, forças de Van Der Waals, bem como efeitos hidrofóbicos relacionados com a libertação da camada de hidratação após a ligação do ligante, destacando a importância de investigar estas interações a nível molecular (CHAVES; SOARES; CAMPOS DE OLIVEIRA, 2021).

Como será melhor discutido na seção 3.4.4, as conformações de maior afinidade, de acordo com a bioquímica quântica, foram os derivados da glicose. Embora a diferença observada não alcance significância estatística, é plausível sugerir que esse padrão possa estar associado a uma influência ligeiramente mais pronunciada da glicose na fluorescência da HSA em comparação com a frutose.

### 3.2 Docking molecular

A ligação de numerosas pequenas moléculas às albuminas de soro humano e bovino já foi estudada utilizando *docking* molecular e métodos computacionais com resultados promissores (LI *et al.*, 2010; LV *et al.*, 2022; SERGIO *et al.*, 2021; SHAMSI; AHMED; BANO, 2018; WANG *et al.*, 2019; WANI *et al.*, 2021). Aqui, usamos testes de *docking* cego com CB-Dock2 para prever posições FRC/GLC/FRO/GLO/GLOm semelhantes às conformações cristalográficas relatadas anteriormente (WANG *et al.*, 2019). O CB-Dock2 produziu resultados de *docking* cego baseados em modelos muito próximos dos modos de ligação do cristal (FRC/FRO/GLC/GLO) - ver Figura 3.

Figura 3 - *Docking* cego baseado em modelos moleculares, listando as duas poses com a energia de ligação mais baixa ( $\text{kcal.mol}^{-1}$ ). (A) FRC, (B) GLC, (C) FRO, (D) GLO e (E) GLOm.



O RMSD entre as poses de *docking* cego baseados em modelos (posição 1) e as estruturas cristalinas (Figura 3) foi inferior a 2 Å. Estes resultados CB-Dock2 estão em total concordância com o estudo cristalográfico que revelou uma ocupação perto do fundo do *Sudlow Site I* por estas formas fechadas de açúcar (WANG *et al.*, 2013b).

Os valores de CB-Dock2 indicaram uma maior afinidade de ligação do GLC (-4,4 e -4,2 kcal.mol<sup>-1</sup>) para o *Sudlow Site I* (FA7) do que a do FRC (-3,8 e -3,2 kcal.mol<sup>-1</sup>) (Figura 3A e 3B). Estes resultados contrastam com uma investigação anterior de *docking* orientado para o *Sudlow Site I*, que sugeriu valores de *docking* mais favoráveis para FRC:HSA (18,4) em vez de GLC:HSA (17,2) (CHAVES; SOARES; CAMPOS DE OLIVEIRA, 2021). No caso da FRO, as pontuações de *docking* obtidas no *Sudlow Site I* foram de -3,0 kcal.mol<sup>-1</sup> e -2,7 kcal.mol<sup>-1</sup>. Além disso, o RMSD entre as posições de *docking* e as estruturas cristalinas foi de 1,383 Å (Figura 3C).

Para GLO e GLOm, o CB-Dock2 também apresentou resultados satisfatórios, com resultados de *docking* no *Sudlow Site I* de -3,8 e -2,2 kcal.mol<sup>-1</sup> para GLO (Figura 3D) e -2,8 e -2,9 kcal.mol<sup>-1</sup> para GLOm (Figura 3E). O *docking* de GLO e GLOm produziu RMSDs de 1,408 Å e 1,950 Å, respectivamente (Figura 3D-3E). Embora as moléculas GLO e GLOm sejam ligeiramente diferentes da molécula GLO no cristal, achamos interessante realizar o RMSD para efeitos de comparação.

Em conclusão, o CB-Dock2 provou ser eficaz na descrição e previsão das interações entre FRC, GLC, FRO, GLO, GLOm e HSA. Os resultados do *docking* cego baseado em modelos mostraram concordância com as estruturas cristalinas e pontuações de *docking* favoráveis, indicando uma maior afinidade de ligação. Por conseguinte, o CB-Dock2 revela-se uma abordagem adequada para futuras investigações das interações entre FRC, GLC, FRO, GLO, GLOm e HSA.

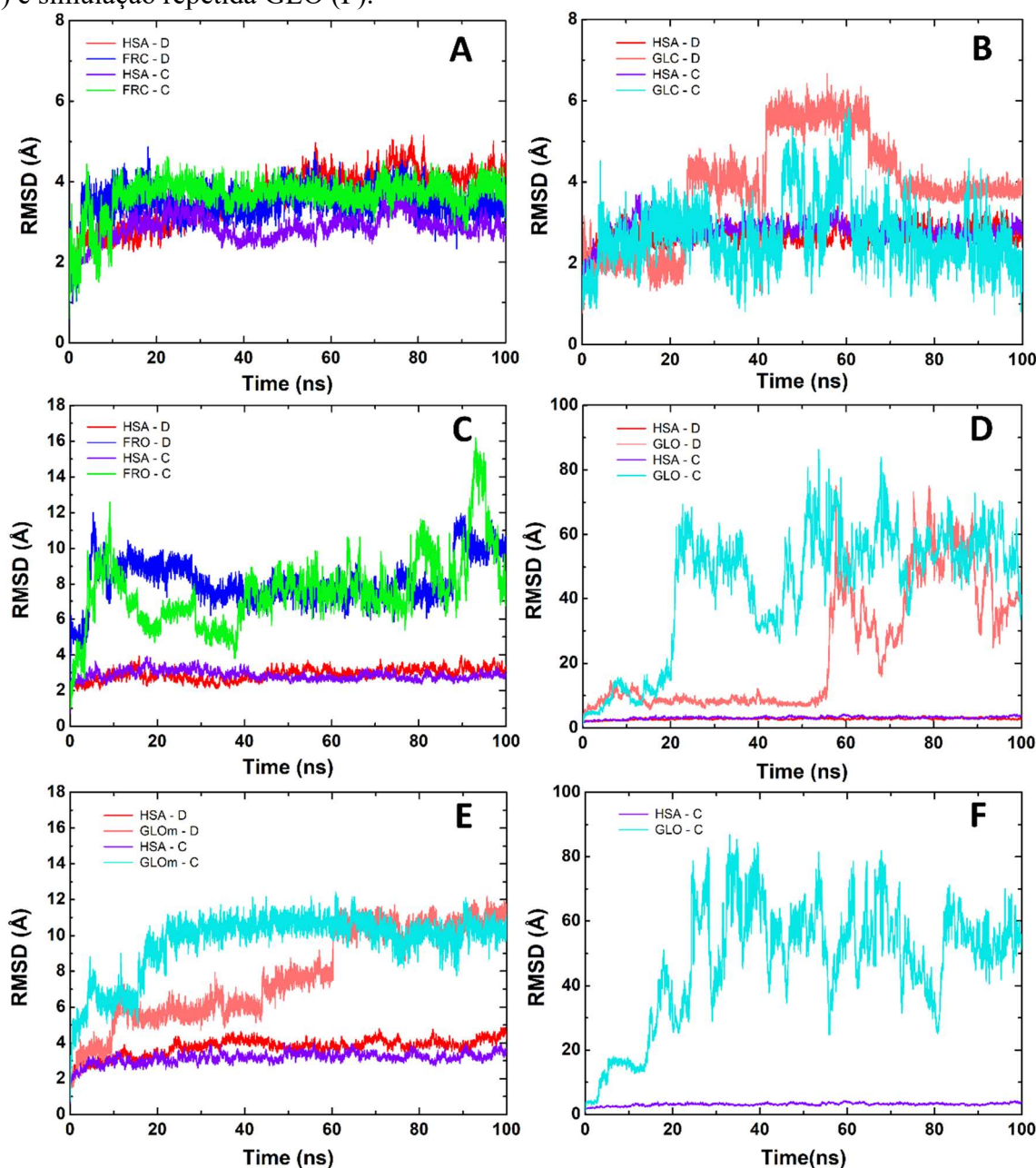
### 3.3 Dinâmica molecular

A dinâmica molecular (MD) tem sido amplamente aplicada a estudos detalhados de proteínas complexadas a uma variedade de pequenas moléculas (GELPI *et al.*, 2015; HOLLINGSWORTH; DROR, 2018; HU *et al.*, 2023), incluindo a ligação de monossacarídeos a HSA (AWANG *et al.*, 2017, 2021; NASIRI *et al.*, 2010; PONGPRAYOON; MORI, 2018). Identificamos locais de interação de ligação por *docking* cego baseado em modelos (Figura 3). Estudos cristalográficos e espectroscópicos apontaram para *Sudlow Site I* como o principal local de ligação para frutose e glicose (BARZEGAR *et al.*, 2007; CHAVES; SOARES; CAMPOS DE OLIVEIRA, 2021; SZKUDLAREK *et al.*, 2016; WANG *et al.*, 2019; ZAMAN *et al.*, 2018). As simulações de dinâmica molecular foram realizadas em duplicata usando dois ligantes diferentes. O primeiro ligante foi derivado de dados cristalográficos, enquanto o segundo

ligante foi obtido a partir de *docking* molecular (topo 1). A combinação dos resultados destas simulações permitiu obter uma imagem mais completa da interação entre os ligantes e a proteína. A realização de simulações em duplicata é uma boa prática para aumentar a confiabilidade dos resultados. As simulações podem ser influenciadas por fatores aleatórios, como a posição inicial das moléculas. Para reduzir a probabilidade de os resultados serem influenciados por estes fatores, podem ser efetuadas duas simulações com condições iniciais diferentes.

Nas simulações MD, o desvio quadrático médio medido (RMSD) das coordenadas atômicas numa molécula pode ser utilizado como um indicador temporal da estabilidade estrutural (MARUYAMA *et al.*, 2023). Todos os valores RMSD dos átomos pesados foram calculados para avaliar a estabilidade global do HSA e do açúcar ao longo das simulações. O HSA apresenta uma estabilidade consistente em todas as simulações, sem diferenças significativas nas variações de RMSD da albumina ligada a diferentes açúcares - ver Figura 4. O  $\Delta$ RMSD-HSA encontrado aqui é consistente com medições anteriores de RMSD C-alfa de HSA simultaneamente ligado a formas de açúcar fechadas e lineares de frutose (glicose), que geralmente se encaixam nas faixas de  $\sim 2$ -4 Å ou  $\sim 3$ -5 Å (NOCEDAL; WRIGHT, 1960; PONGPRAYOON; MORI, 2018). Em comparação com a albumina ligada a açúcares numa proporção de 1:2, a estrutura global da albumina aqui estudada (1:1) não parece apresentar alterações estruturais significativas.

Figura 4 - Desvio raiz-quadrado médio de todos os átomos pesados (RMSD) em função do tempo dos complexos HSA:Monossacarídeos. FRC (A), GLC (B), FRO (C), GLO (D), GLOm (E) e simulação repetida GLO (F).

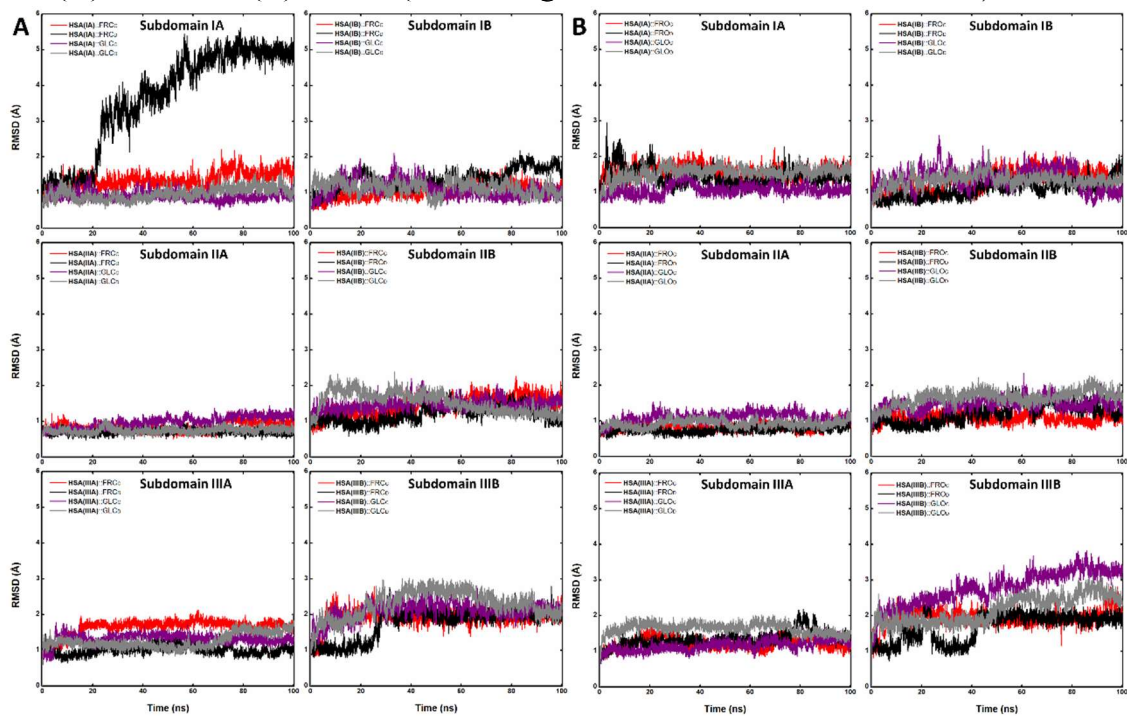


Os valores RMSD dos átomos C-alfa indicam que a ligação da frutofuranose resultou no maior grau de flexibilidade no subdomínio IA, Figura 5. Os açúcares de cadeia longa parecem induzir níveis mais elevados de variação estrutural nos subdomínios IB e IIIB, Figura 5. Além disso, os subdomínios IIA e IIIA, que contêm o *Sudlow Site I* e o *Sudlow Site II*, foram os mais rígidos em todas as simulações. Considerando estudos computacionais anteriores que investigaram o modo não ligado de monossacarídeos com HSA, Awang et al., 2017 identificaram rigidez no *Sudlow Site I* e maior flutuação no *Sudlow Site II* de pGHSA (HSA



pré-glicado em uma proporção de 1: 2). Estudos identificaram baixos níveis de flutuação em ambos os subdomínios IIA e IIIA (Pongprayoon & Mori, 2018). Awang et al., 2021 relataram alta flexibilidade no subdomínio IA dos complexos HSA:FRC+FRO e rigidez reduzida no IIA do complexo HSA:FRC-FRO número 1 (denominado FHSA\_1). Além disso, todas estas medições mostram que o subdomínio IIIB é altamente flexível na presença de açúcares nos seus modos não ligados. Estes resultados também sugerem que a baixa rigidez encontrada no subdomínio IA do complexo HSA::FRC-FRO se deve à influência da forma fechada da frutose em vez da sua forma linear.

Figura 5 - C-alfa RMSD dos subdomínios IA, IB, IIA, IIB, IIIA e IIIB da HSA. HSA ligado a formas (A) fechadas e (B) abertas (incluindo glicose linear modificada - GLOm).



Verificamos que os diferentes açúcares não resultaram em diferenças significativas nas flutuações de proximidade entre os subdomínios IB e IIIA da HSA (representados pela distância LYS195-ASP451), que variaram geralmente entre 6 Å e 12 Å, Figura 6A-6D. As distâncias entre VAL116 e ASP510 variaram principalmente entre 8 Å e 34 Å, revelando o movimento tipo tesoura entre os subdomínios superiores IB e IIIB da albumina, Figura 6E-6H. Além disso, o movimento em tesoura observado em todos os 1 sistemas de HSA ligados a monossacarídeos é semelhante e consistente com descobertas anteriores de 2 estudos de albumina ligada a açúcar (Nocedal & Wright, 1960; Sittivanichai et al., 2023; Sousa et al., 2016).

Embora os nossos resultados de extinção de fluorescência indicam que os açúcares não ligados à HSA não alteram a polaridade do microambiente polar do TRP214, decidimos examinar as alterações posicionais do TRP214 antes e depois das simulações no estado de HSA ligado a 1 açúcar. Como resultado, observamos que a ligação de açúcar fechado e linear à HSA conduziu a orientações semelhantes do TRP214 - ver Figura 7. Em comparação com as orientações estruturais de TRP214 observadas nas simulações de HSA glicada ao longo de 500 ns, as modificações de TRP214 após a ligação de açúcar a HSA em modo não covalente e razão molecular 1:1 mostraram uma semelhança notável (SITTIWANICHAJ *et al.*, 2023). Estas incluem a realocização do anel de indol numa orientação paralela à hélice alfa.

Em contraste com a albumina, os ligantes de açúcar exibiram uma instabilidade significativa ao longo das simulações, exceto para o FRC, que apresentou uma gama de RMSD semelhante à da proteína. Durante os 100 ns de simulação, a variação total de RMSD para FRC (*docking* e cristalografia) variou de 0,4 Å a 4,8 Å, para GLC (*docking* e cristalografia) de 0,7 Å a 6,6 Å, para FRO de 1,0 Å a 16,1 Å, para GLO de 1,0 Å a 86,3 Å e para GLO<sub>m</sub> de 0,7 Å a 12,4 Å (Figura 4). Nos últimos 50 ns, FRC<sub>D</sub> e FRC<sub>C</sub> mostraram  $\Delta$ RMSD de 2,4 Å e 1,7 Å, respectivamente (Figura 4A), GLC<sub>D</sub> e GLC<sub>C</sub> mostraram  $\Delta$ RMSD de 3,2 Å e 5,1 Å (Figura 4B), FRO<sub>D</sub> e FRO<sub>C</sub> mostraram  $\Delta$ RMSD de 6,1 Å e 10,8 Å (Figura 4C), GLO<sub>D</sub> e GLO<sub>C</sub> mostraram  $\Delta$ RMSD de 68,7 Å e 52,7 Å (Figura 4D), e GLO<sub>mD</sub> e GLO<sub>mC</sub> mostraram  $\Delta$ RMSD de 5,3 Å e 4,6 Å (Figura 4E). Estes resultados indicam que os monossacarídeos são estruturalmente menos estáveis do que a HSA em todas as simulações, com a GLC a apresentar uma maior instabilidade estrutural em comparação com a FRC.

Para confirmar os nossos resultados, repetimos a simulação de dinâmica molecular para GLO (Figura 4F) utilizando as mesmas condições. Esta repetição foi efetuada para responder às preocupações sobre os RMSDs drasticamente alterados (68,7 Å e 52,7 Å) na estrutura GLO.

Os resultados da simulação repetida (Figura 4F) são consistentes com os resultados anteriores. Durante os 100 ns de simulação, o RMSD total do complexo GLO<sub>D</sub>/HSA variou de 1,0 Å a 86,8 Å. Nos últimos 50 ns, o RMSD do complexo variou 57,2 Å. Estes resultados confirmam que GLO apresenta uma instabilidade estrutural significativa quando interage com HSA. O RMSD (Figura 4D) e a análise da ligação de hidrogênio indicam que a GLO se dissocia da HSA, sugerindo que esta conformação não é ideal para a interação com a HSA, exceto se se ligar numa proporção de 1:2 e tiver o efeito de uma piranose, como proposto anteriormente (WANG *et al.*, 2013b). As Figuras 8 e 9 fornecem mais pormenores sobre a ligação e dissociação de GLO::HSA, a mobilidade do domínio e a deslocação de TRP214. É interessante



notar que a GLO<sub>m</sub> com a ligação dupla excluída e valências saturadas com hidrogênio apresenta uma maior estabilidade conformacional ao longo das simulações. Uma vez que as GLOs escapam dos sítios de ligação da HSA, apenas efetuamos cálculos bioquímicos quânticos da conformação final (100ns) para esta estrutura GLO<sub>m</sub> modificada (Figura 4E). O comportamento da ligação de hidrogênio indica que a saída completa ocorreu entre 40-60 em ambas as simulações moleculares. Estes resultados confirmam o mecanismo previamente proposto no estudo cristalográfico de HSA::monossacarídeos, uma vez que esta GLO modificada se encontra no estado imediatamente pré-glicado de glicose linear (WANG *et al.*, 2013b), sugerindo que este intermediário tem mais afinidade pelo *Sudlow Site I* do que a sua conformação anterior com C1=O1. Embora as reações químicas sejam geralmente processos rápidos (femtosegundos) (POTTER *et al.*, 1992), acreditamos que estes dados indicam que a perda de C1=O1 é um passo crítico no mecanismo de glicação HSA::GLO.

Figura 6 - Distâncias entre dois pares de resíduos que demonstram o movimento em tesoura da HSA ao longo do tempo. Distâncias entre TYR195 e ASP451 nos complexos (A) HSA::FRC, (B) HSA::GLC, (C) HSA::FRO, e (D) HSA::GLOm. (E) Aproximação e separação dos subdomínios IIB e IIIB em (E) HSA::FRC, (F) HSA::GLC, (G) HSA::FRO e (H) HSA::GLOm representados pelas distâncias entre os resíduos VAL116 e ASP510. HSA::FRC<sub>C</sub> (HSA::FRC<sub>D</sub>) e HSA::FRO<sub>C</sub> (HSA::FRO<sub>D</sub>) são representados pela cor verde (azul), enquanto as linhas ciano (salmão escuro) indicam HSA::GLC<sub>C</sub> (HSA::GLC<sub>D</sub>) e HSA::GLO<sub>mC</sub> (HSA::GLO<sub>mD</sub>).

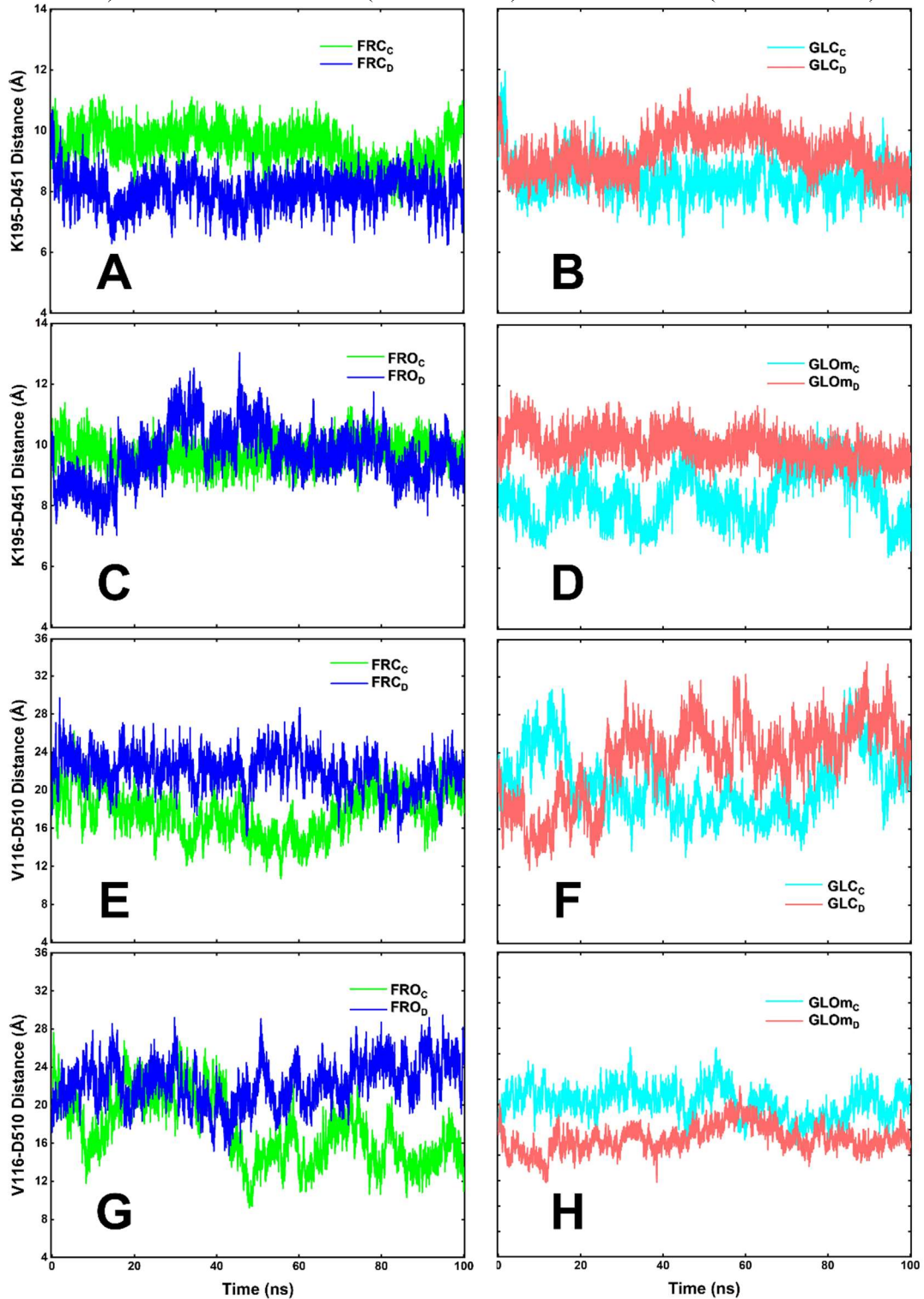


Figura 7 - Reorientação do anel indol TRP214 (W214) em IIA da HSA ligado a frutose fechada (A, B), glucopiranosose (C, D), ceto-D-frutose (E, F) e glicose linear modificada (G, H).

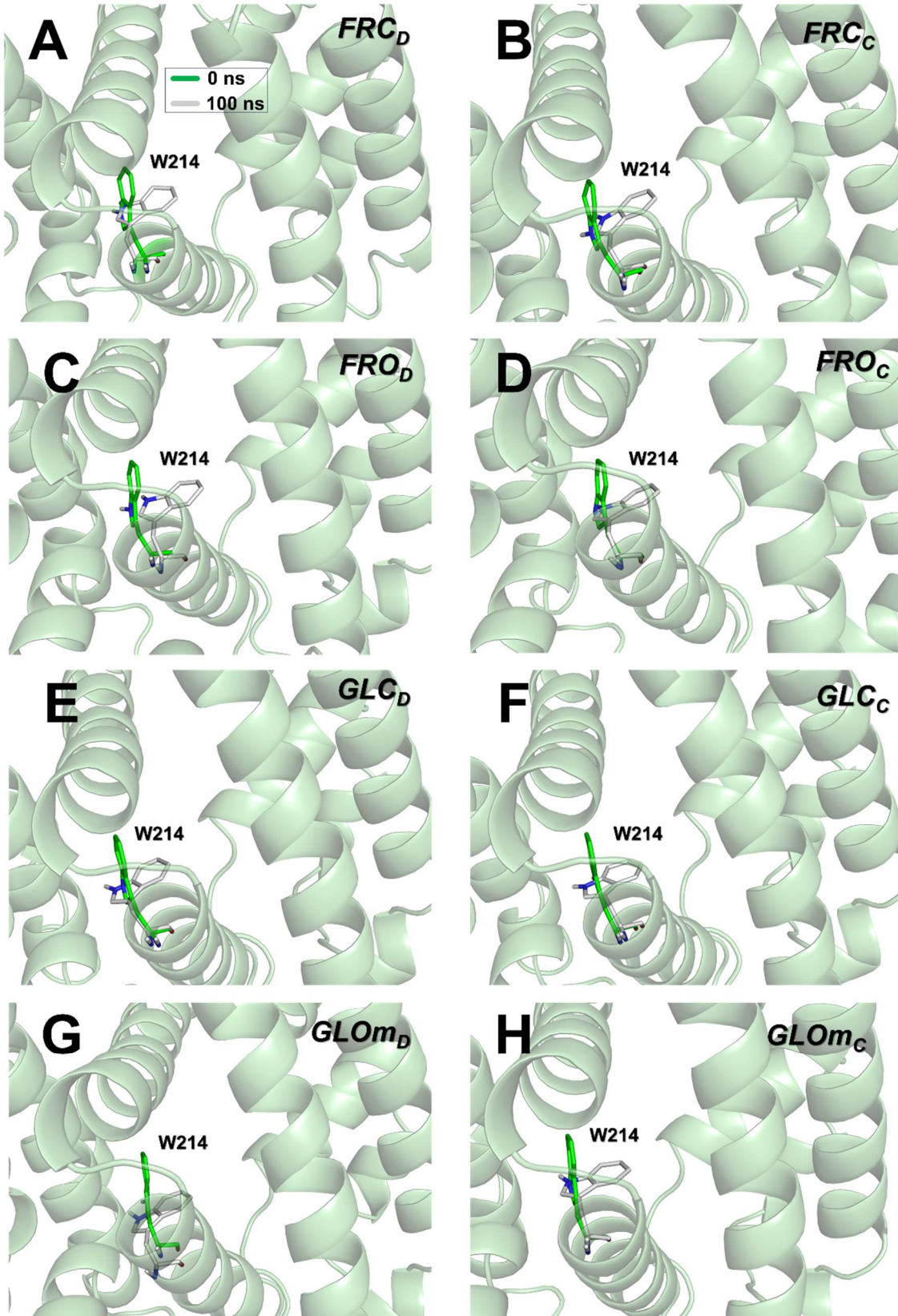




Figura 8 - Descrição da mobilidade e dos contatos do complexo biológico HSA::GLO. RMSD dos subdomínios de HSA para (A) HSA::GLO<sub>D</sub> e (B) HSA::GLO<sub>C</sub>. Flutuação das distâncias (C) LYS195-ASP451 (K195-D451) e (D) VAL116-ASP510 (V116-D510) em ensaios de dinâmica molecular. (E-F) Número e (G-H) detalhes das ligações de hidrogênio formadas entre resíduos de GLO e HSA.

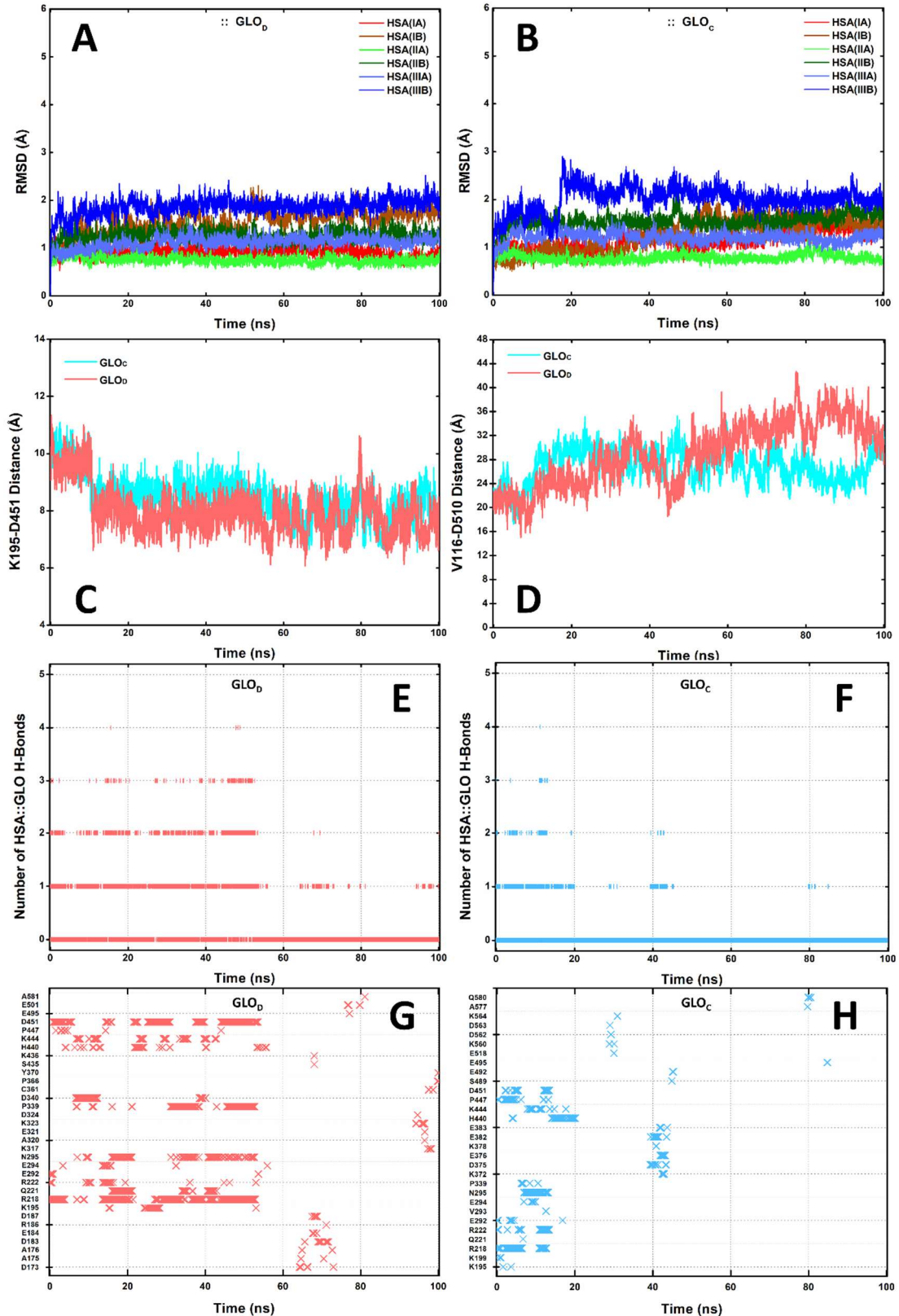
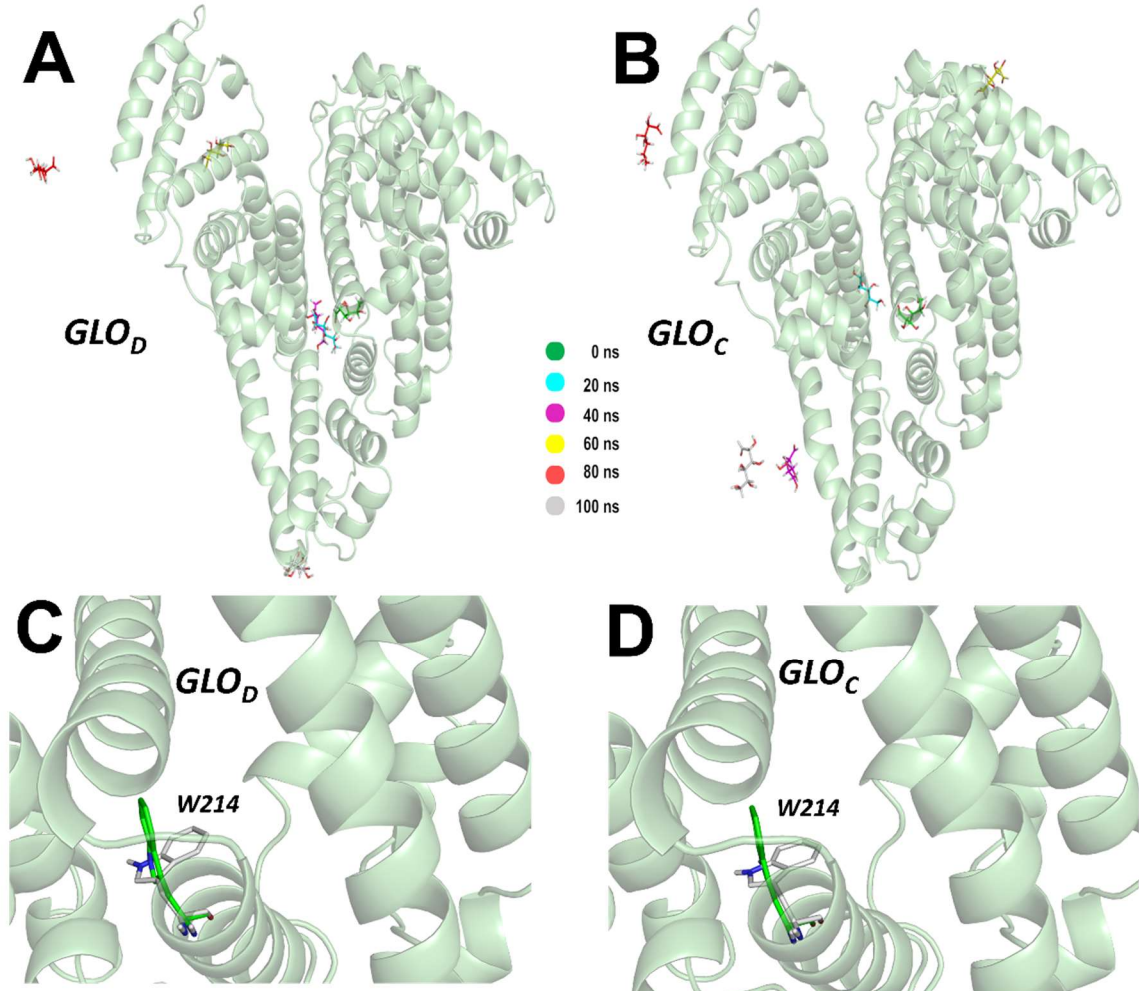


Figura 9 - (A-B) Alta mobilidade de GLO não modificada representada pelo alinhamento das conformações de HSA::GLO a cada 20 ns e deslocamento de TRP214 em (C) HSA::GLO<sub>D</sub> e (D) HSA::GLO<sub>C</sub>, onde as conformações inicial e final são coloridas em verde e cinza claro, respectivamente .



Os resultados de MD são apoiados pelos nossos dados de extinção de fluorescência (Figura 2) e são consistentes com um estudo de Chaves et al., 2021, que relatou fracas capacidades de ligação de frutose e glicose a HSA (CHAVES; SOARES; CAMPOS DE OLIVEIRA, 2021). Para FRO e GLO<sub>m</sub>, os resultados mostraram uma maior instabilidade estrutural em comparação com FRC e GLC. As análises estruturais das deslocções dos açúcares perto do subdomínio IIA indicaram claramente que os açúcares lineares apresentavam maior flexibilidade do que as formas fechadas - ver Figura 10. Além disso, todos os açúcares formaram menos de 5 ligações de hidrogênio com a albumina (1:1), reforçando as fracas interações entre os açúcares e a albumina - ver Figura 11. Tanto a FRC (~0,50 ligações H) como a FRO (~0,46) apresentaram um número médio ligeiramente superior de contatos fortes com a HSA do que a GLC/GLO<sub>m</sub> (~0,35/~0,34). Dado que a albumina humana é conhecida por ser

promíscua na ligação a uma vasta gama de ligantes (FANALI *et al.*, 2012) e não forma interações fortes com frutose e glicose numa razão molar de 1:1, os nossos resultados são consistentes com o fato bem documentado de que os açúcares não são ligantes primários para esta proteína (AWANG *et al.*, 2017, 2021; CHAVES; SOARES; CAMPOS DE OLIVEIRA, 2021).



Figura 10 - Flexibilidade da frutose e da glicose não ligadas à HSA. Ilustrado pela representação de diferentes ocupações de (A, B) FRC, (C, D) FRO, (E, F) GLC e (G, H) GLOm no espaço conformacional obtido através da dinâmica molecular.

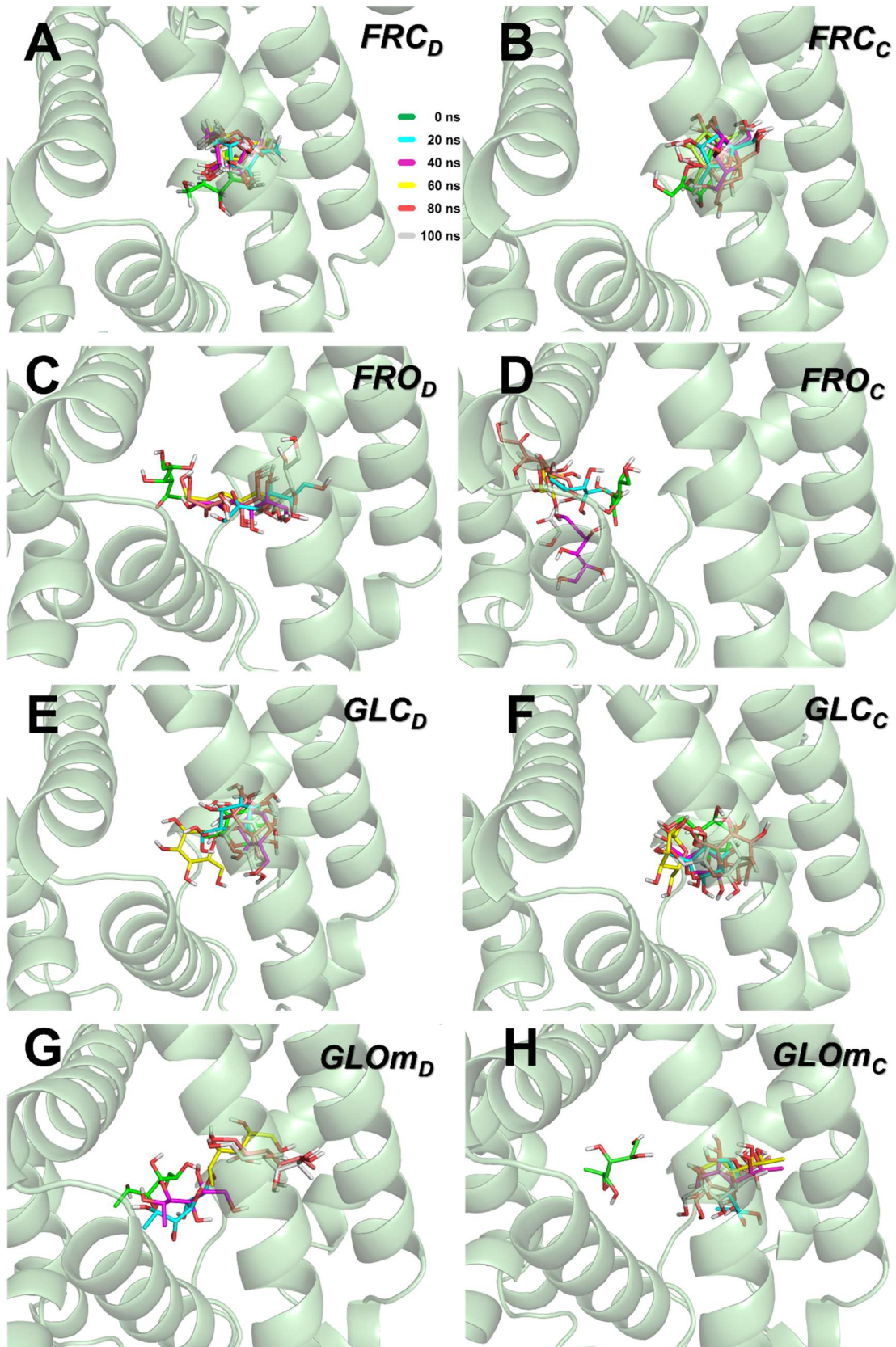
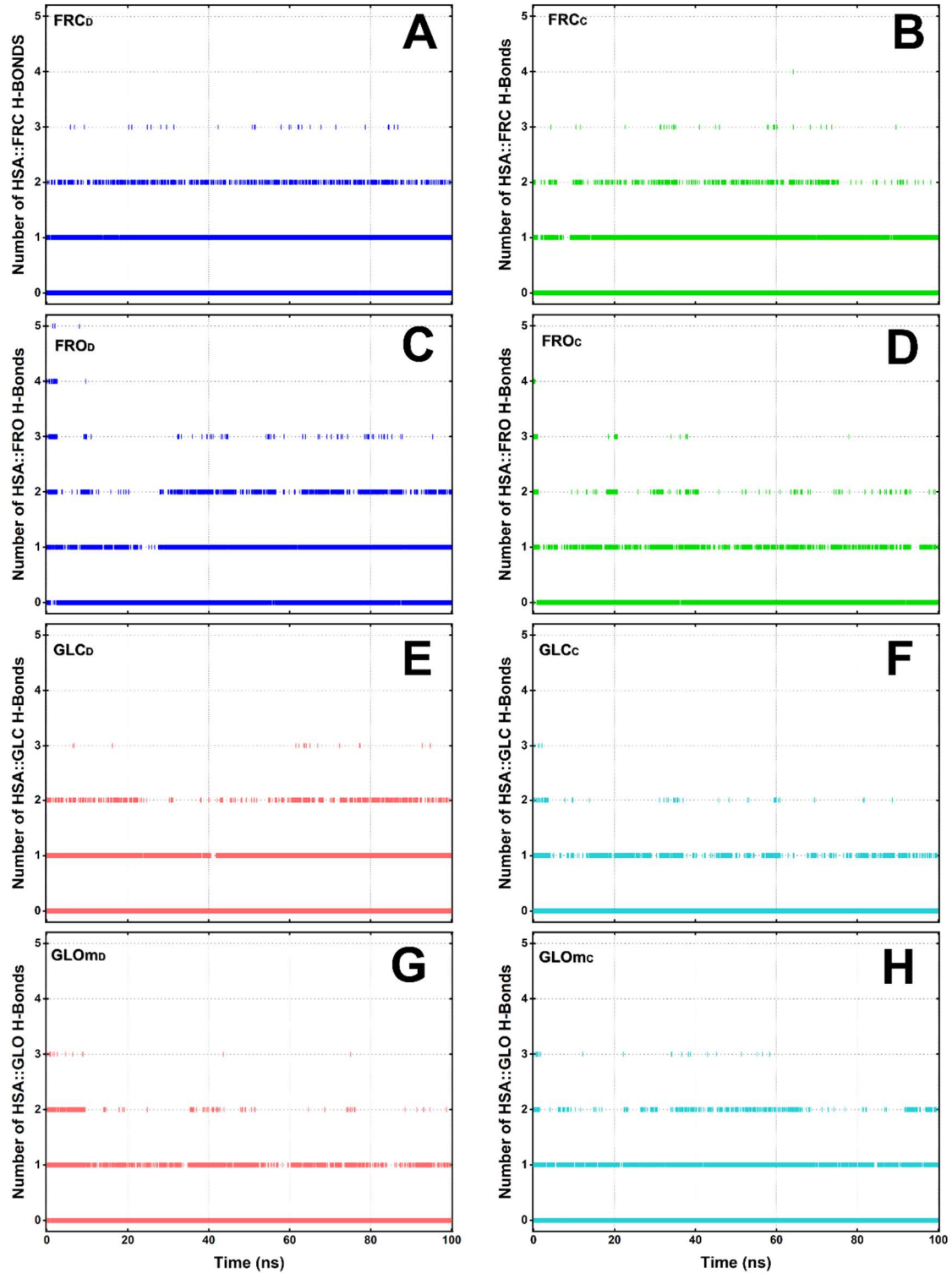


Figura 11 - Ligações de hidrogênio HSA::Monossacarídeo em função do tempo. (A, B) Contatos observados nas simulações realizadas utilizando a forma fechada da frutose. (C, D) Ligações H encontradas nos complexos compostos por frutose linear. (E, F) Ligações H realizadas por GLC. (G, H) Número de ligações de hidrogênio formadas pela forma linear modificada da glicose.



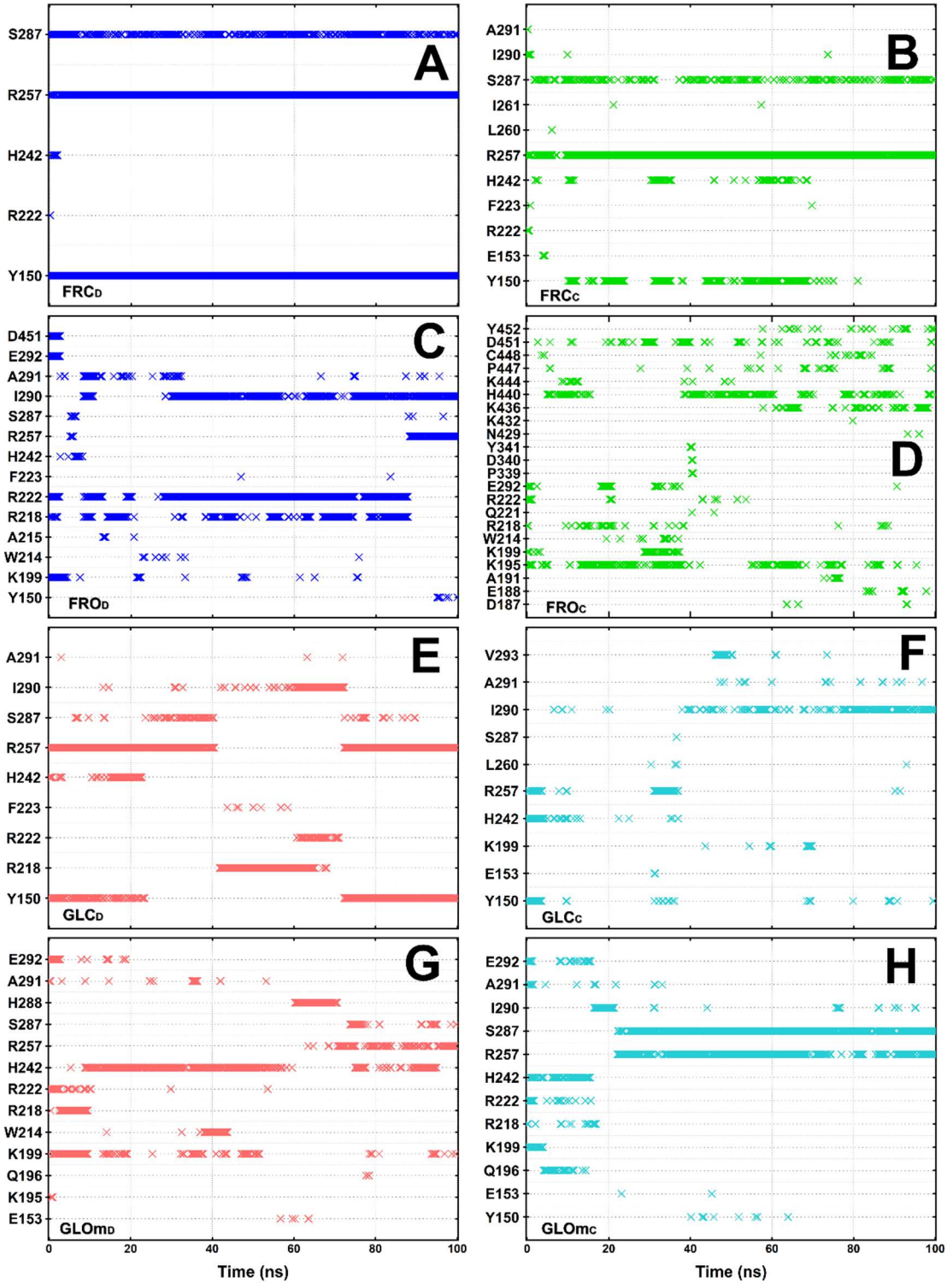


A instabilidade estrutural de GLO/FRO é consistente com relatórios teóricos anteriores (AWANG *et al.*, 2017, 2021), que sugerem que os açúcares lineares são mais suscetíveis à solvatação do que suas respectivas formas fechadas. Isso pode ser atribuído aos efeitos de solvatação e à ausência de estruturas de furanose/piranosose nas profundezas do *Sudlow Site I*. Consequentemente, os açúcares lineares foram capazes de se mover da entrada do *Sudlow Site I* para a direção superior direita do sítio (Figura 10C, 10G e 10H), exceto para FRO<sub>D</sub>, que mudou para o subdomínio IIIA (Figura 10D). A este respeito, Pongprayoon & Mori, 2018 relataram que ambas as formas de galactose foram capazes de escapar da vizinhança do *Sudlow Site I*, um evento não observado em HSA::GLC/GLO. Awang *et al.*, 2021 detectaram a fuga de galactose dos sítios da HSA e observaram dois comportamentos distintos de FRO em torno da vizinhança IIA: um deslocamento da entrada do *Sudlow Site I* para o seu canto superior direito, bem como oscilação em torno da mesma posição. Além disso, (AWANG *et al.*, 2017) observaram que o GLC permaneceu preso no *Sudlow Site I*, enquanto o GLO mostrou um maior grau de flutuação. Curiosamente, verificou-se que a forma fechada da glicose ligada à HSA glicada se deslocava em direção à porta do *Sudlow Site I* (SITTIWANICHAJ *et al.*, 2023).

De forma consistente, a ausência do dímero formado entre os açúcares lineares e fechados não resultou na saída dos açúcares fechados. Curiosamente, uma análise combinada das propriedades dinâmicas dos monossacarídeos ligados à albumina nas razões 1:1 e 1:2 sugere que a frutose fechada fica mais bem presa no *Sudlow Site I* na ausência de FRO-FRC, uma vez que a fuga de FRC foi detectada na HSA ligada a 2-frutose (AWANG *et al.*, 2021). Além disso, este fato é consistente com o mecanismo anteriormente descrito para a fuga de GLC do *Sudlow Site I* inferior na HSA glicada (SITTIWANICHAJ *et al.*, 2023).

Adicionalmente, o FRC apresentou a ligação mais linear no *Sudlow Site I*, conforme representado pelas interações permanentes com os resíduos Y150, R222 e S287 (Figura 12). Inicialmente, todas as formas lineares, exceto a FROC, estabeleceram fortes contatos com os resíduos R218, R222 e R292 localizados no *Sudlow Site I*, mas estas interações foram interrompidas a ~5-20 ns (Figura 12). Subsequentemente, todas as formas lineares mostraram uma tendência para interagir com resíduos mais profundos no *Sudlow Site I*, tais como Y150, R257 e S287 (Figura 12). Esta rápida mudança no padrão de ligação de hidrogênio (~15-30 ns) também sugere que os açúcares lineares têm uma maior afinidade para o *Sudlow Site I*, como mostra a Figura 10.

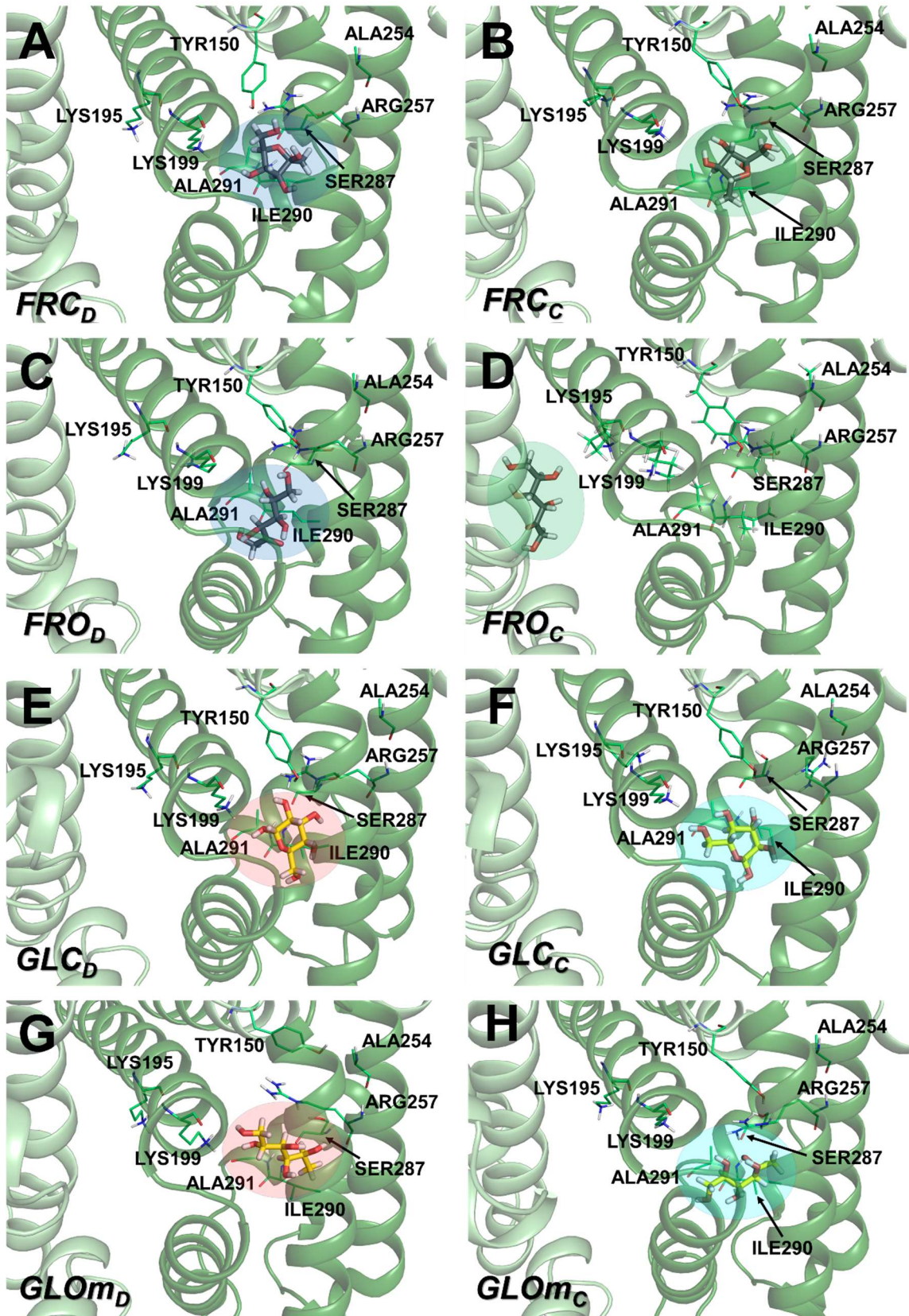
Figura 12 - Análise detalhada das ligações de hidrogênio detectadas nos sistemas HSA::Monossacarídeo ao longo do tempo de simulação. Para (A) HSA::FRC<sub>D</sub>, (B) HSA::FRC<sub>C</sub>, (C) HSA::FRO<sub>D</sub>, (D) HSA::FRO<sub>C</sub>, (E) HSA::GLC<sub>D</sub>, (F) HSA::GLC<sub>C</sub>, (G) HSA::GLO<sub>mD</sub>, (H) HSA::GLO<sub>mC</sub>. O eixo y representa os resíduos de HSA, enquanto o eixo x corresponde ao tempo de simulação.



Como descobrimos que a ausência de piranose no *Sudlow Site I* levou a uma mudança maciça de  $GLO_m$  para um bolso mais interno do subdomínio IIA (Figura 10 e 12), nossos resultados parecem estar de acordo com estudos anteriores que destacam o papel crítico da piranose para a glicação de HSA (CHAVES; SOARES; CAMPOS DE OLIVEIRA, 2021; WANG *et al.*, 2013b). O comportamento peculiar observado na  $FRO_C$  sugere que a frutose linear tem maior flexibilidade do que a glicose linear a uma temperatura de 300 K, o que pode ajudar a explicar a ausência da ligação covalente  $FRO_C::LYS195$  no estudo cristalográfico anterior (WANG *et al.*, 2013b).

Os resultados do MD são consistentes com estudos teóricos anteriores que avaliam esses complexos biológicos (AWANG *et al.*, 2021; CASTRO-ALVAREZ; COSTA; VILARRASA, 2017; NASIRI *et al.*, 2010). Por exemplo, a glicose fechada parece interagir aleatoriamente com resíduos carregados dentro do *Sudlow Site I* (PONGPRAYOON; MORI, 2018), o que está em boa concordância com  $\Delta RMSD-GLO/GLO_m$  (Figura 4) e a ausência de ligação de hidrogênio permanente com HSA (Figura 12). Para mais pormenores sobre a ocupação aleatória das formas fechada e aberta, ver as Figuras 9 e 13. Além disso, a HSA glicada apresenta uma tendência para reter os monossacarídeos fechados no *Sudlow Site I*, o que foi particularmente observado na HSA de Amadori (SITTIWANICHAJ *et al.*, 2023). Por conseguinte, a nossa hipótese subjacente é que o comportamento aleatório da GLC perto do *Sudlow Site I* (Figura 10) está relacionado com os diferentes resultados obtidos com amostras de HSA glicada utilizadas para investigar alterações de afinidade associadas à glicação de HSA (ANGUIZOLA *et al.*, 2013; GHOSH; KISHORE, 2022; JOSEPH; HAGE, 2010; MOEINPOUR *et al.*, 2016).

Figura 13 - Comparação da ocupação final de cada (A, B) FRC, (C, D) FRO, (E, F) GLC, e (G, H) GLOm após 100 ns de dinâmica molecular. Alguns resíduos localizados na região superior do subdomínio IIA estão representados como linhas.



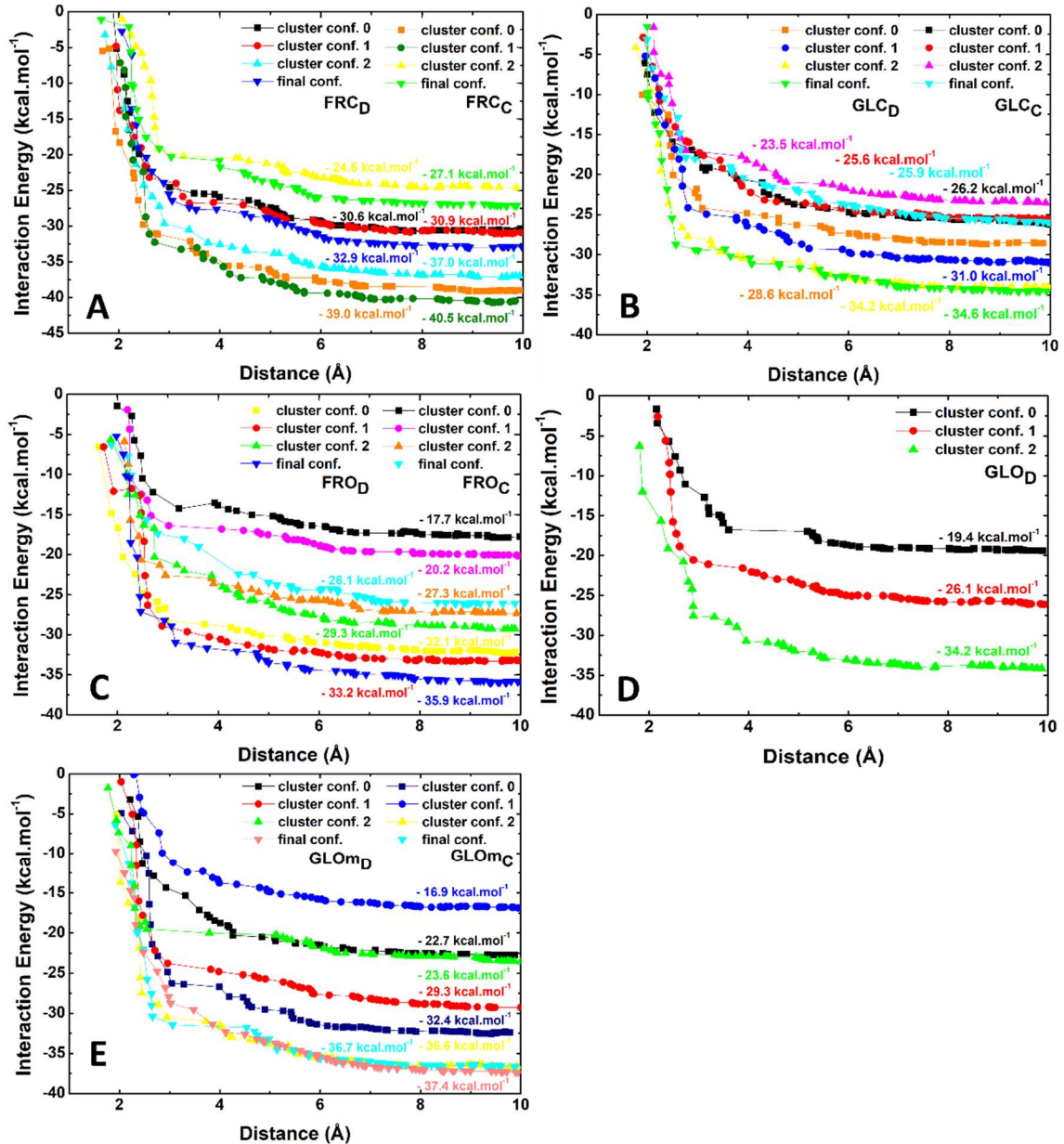


### 3.4 Bioquímica quântica e detalhes atômicos das ligações de FRC/FRO/GLC/GLO/GLO<sub>m</sub> com a HSA

Chaves et al., (2021) realizaram observações que sugerem que a interação entre a glicose e a frutose ocorre predominantemente através de mecanismos hidrofóbicos, eletrostáticos e de ligação de hidrogênio. Tendo em conta este cenário, utilizamos cálculos baseados em DFT em conjunto com o esquema MFCC para realizar uma investigação mais abrangente das interações não ligadas entre os monossacarídeos e os resíduos de HSA, considerando cada conjunto de três conformações representativas geradas por EnGens, bem como a conformação final resultante da dinâmica molecular.

As Figuras 14 a 19 fornecem informações pormenorizadas sobre a energia de interação, os modos de ligação e as interações entre FRC/FRO/GLC/GLO/GLO<sub>m</sub> e HSA na estrutura final de cada uma das simulações de dinâmica molecular aqui efetuadas. A convergência da energia de interação nas quatro amostras ocorreu perto do raio de 8 Å do centro de FRC/FRO/GLC/GLO/GLO<sub>m</sub>, indicando que os resíduos localizados a mais de 8 Å não contribuem significativamente para a estabilidade destes sistemas (Figura 14). Embora um estudo anterior tenha efetuado cálculos quânticos para estimar as energias das interações entre açúcares e HSA (AWANG *et al.*, 2017), limitaram os seus cálculos a uma região menor que 4 Å, que se estende do centro GLC/GLO em direção aos resíduos de albumina em direção aos resíduos de albumina. A bioquímica quântica revelou que os resíduos para além deste raio são cruciais para a captura dos açúcares no HSA. Como já foi referido, efetuamos 8 simulações MFCC/bioquímica quântica para cada simulação. Por exemplo, efetuamos 4 cálculos do complexo FRC<sub>C</sub> e 4 do FRC<sub>D</sub>. De cada uma destas 4 simulações, 3 foram geradas pelo método EnGens e 1 pela conformação final resultante da dinâmica molecular. É importante notar que, para o complexo GLO, foram obtidos somente 3 resultados para o GLO<sub>D</sub>, todos pelo método EnGens. Além disso, não foram obtidos resultados para a conformação final de GLO<sub>D</sub> ou GLO<sub>C</sub>. Isso ocorreu porque ambas as moléculas saíram do sítio de ligação da HSA. A saída dessas moléculas do sítio de ligação é mostrada na Figura 4 do RMSD.

Figura 14 - A energia total de interação entre os açúcares e os resíduos de HSA em função do raio do local de ligação. (A) FRC<sub>D</sub>/FRC<sub>C</sub>; (B) GLC<sub>D</sub>/GLC<sub>C</sub>; (C) FRO<sub>D</sub>/FRO<sub>C</sub>; (D) GLO<sub>D</sub>; (E) GLO<sub>mD</sub>/GLO<sub>mC</sub>.



A energia média de interação entre o FRC e a HSA atingiu um valor de  $-32,8 \text{ kcal.mol}^{-1}$  com um desvio padrão de  $5,29 \text{ kcal.mol}^{-1}$ . O coeficiente de variação, que é de 16,1%, indica que a energia de interação tem um grau de variabilidade relativamente elevado. Por outro lado, a energia média de interação entre GLC e HSA foi de  $-28,7 \text{ kcal.mol}^{-1}$  com um desvio padrão de  $3,89 \text{ kcal.mol}^{-1}$ . O coeficiente de variação de 13,6% sugere que a energia de interação é menos variável em comparação com a da frutose (como se mostra na Tabela 2). Estes resultados indicam que o FRC tem uma maior afinidade pelo *Sudlow Site I* do que o GLC. No entanto, o

GLC apresenta uma variação mais estável do que o FRC. Por outras palavras, em média, os valores da energia de interação do GLC estão mais próximos da média do que os do FRC. Curiosamente, a frutose aberta e fechada mostrou mais afinidade pela HSA do que as glicoses aberta e fechada (Tabela 2). No entanto, em comparação com as energias de interação obtidas nos nossos estudos anteriores de bioquímica quântica da HSA (DANTAS *et al.*, 2015; FRANÇA *et al.*, 2022), os açúcares abertos e fechados parecem ter afinidades mais fracas com a HSA. Isso é consistente com estudos anteriores que mostram que os açúcares fechados não são ligantes preferenciais para HSA (CHAVES; SOARES; CAMPOS DE OLIVEIRA, 2021; PONGPRAYOON; MORI, 2018).

Tabela 2 - Energia de interação de todas as simulações, com valor médio, desvio padrão ( $\sigma$ ) e coeficiente de variação ( $C_v$ ).

	Energia de interação								Valor médio	$\sigma$	$C_v$
	kcal.mol <sup>-1</sup>								kcal.mol <sup>-1</sup>	kcal.mol <sup>-1</sup>	
FRC	-24,6	-27,1	-30,6	-30,9	-32,9	-37,0	-39,0	-40,5	-32,8	5,29	16,1%
GLC	-23,5	-25,6	-25,9	-26,2	-28,6	-31,0	-34,2	-34,6	-28,7	3,89	13,6%
FRO	-17,7	-20,2	-26,1	-27,3	-29,3	-32,1	-33,2	-35,9	-27,7	5,90	21,3%
GLO	-19,4	-26,1	-34,2	---	---	---	---	---	-26,6	6,05	22,8%
GLOm	-16,9	-22,7	-23,6	-29,3	-32,4	-36,6	-36,7	-37,4	-29,5	7,18	24,4%

A energia de interação média para FRO foi de -27,7 kcal.mol<sup>-1</sup>, como se mostra na Figura 14C e na Tabela 2. O desvio padrão para este valor foi calculado como sendo 5,90 kcal.mol<sup>-1</sup>, dando um coeficiente de variação de 21,3%. Isto sugere que a FRO interage de forma consistente com a HSA, embora exista alguma variabilidade nos resultados. O sistema GLO apresentou uma energia de interação média de -26,6 kcal.mol<sup>-1</sup> com um desvio padrão de 6,05 kcal.mol<sup>-1</sup>. No entanto, em alguns casos, a GLO não interagiu com HSA (Figura 14D e Tabela 2). Isto pode indicar que a GLO tem uma maior tendência do que a FRO para se separar da HSA durante a interação. A GLOm apresentou o valor médio mais elevado de energia de interação de -29,5 kcal.mol<sup>-1</sup>, juntamente com o maior desvio padrão de 7,18 kcal.mol<sup>-1</sup>, resultando num coeficiente de variação de 24,4% (Figura 14E e Tabela 2). Isto indica que GLOm tem a interação mais forte com HSA, mas também a maior variação nos resultados. Notavelmente, os cálculos de perturbação livre mostraram anteriormente uma maior afinidade de HSA para glicose em comparação com frutose (PONGPRAYOON; MORI, 2018). É

importante ressaltar que nossos resultados são consistentes com estudos anteriores (AWANG *et al.*, 2017; WANG *et al.*, 2013b), que relataram uma ligação covalente GLO: LYS195.

Além disso, a bioquímica quântica sugere que as piranoses (formas cíclicas de carboidratos com um anel hexagonal, composto por cinco carbonos e um de oxigênio) de açúcar simples isoladas não desempenham um papel crucial na ligeira diferença entre as afinidades HSA:FRC e HSA:GLC obtidas através de experiências espectroscópicas que investigaram a ligação de monossacarídeos a HSA (CHAVES; SOARES; CAMPOS DE OLIVEIRA, 2021), o que poderia ser atribuído de forma mais consistente tanto ao maior número de ligações de hidrogênio HSA::Frutose (Figura 11) como aos contatos permanentes estabelecidos entre FRC e HSA (Figura 12). No entanto, a observação de uma afinidade média mais elevada da HSA para a frutose de cadeia longa do que para a glicose linear parece contradizer o efeito ligeiramente mais pronunciado da glicose na fluorescência da HSA em comparação com a frutose (Figura 2).

A baixa afinidade de ligação de ambos os monossacarídeos à HSA indica uma fraca capacidade de ligação dos hidratos de carbono à albumina. Vários fatores contribuem para isso, como o tamanho pequeno dos monossacarídeos, o tamanho grande do *Sudlow Site I* e o elevado efeito de solvatação dentro da cavidade. Os fatores mencionados facilitam o acesso fácil dos monossacarídeos às vias metabólicas (YANG *et al.*, 2007a), realçando a sua função primária no fornecimento de energia, tal como referido anteriormente por Chaves *et al.*, (2021).

### **3.4.1 Interações FRC::HSA no Sudlow Site I**

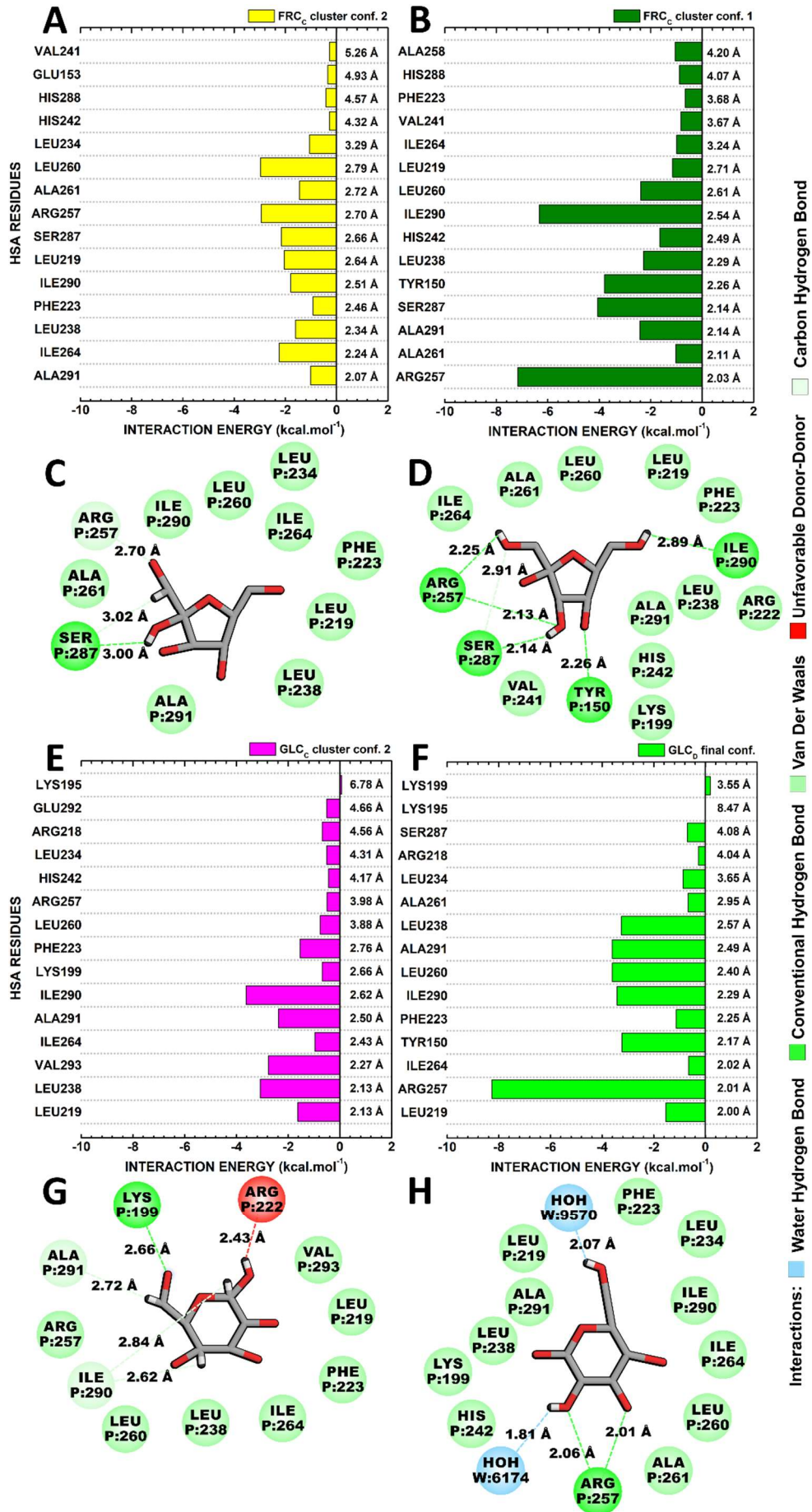
Pesquisa cristalográfica anterior identificou similaridades significativas entre os modos de ligação FRC/GLC e FRO/GLO no *Sudlow Site I*. Essa observação foi corroborada por redes de interações comparáveis estabelecidas entre FRC/GLC e os cinco resíduos de HSA::TYR150, ARG222, LEU238, ARG257 e ALA291. Adicionalmente, foram detectados contatos intensos entre FRO/GLO e os resíduos LYS195, LYS199, ARG218, ARG222 e ALA291 (WANG *et al.*, 2013b).

Para compreender melhor a natureza destas interações, nesta seção analisamos as conformações de menor e maior afinidade de cada complexo apresentado na Figura 14. A conformação de afinidade mais baixa (FRC<sub>C</sub> *cluster conf.* 2), que inclui os resíduos LEU260 (-2,97 kcal.mol<sup>-1</sup>), ARG257 (-2,94 kcal.mol<sup>-1</sup>) e ILE264 (-2,24 kcal.mol<sup>-1</sup>), apresentou as energias de interação mais elevadas com o FRC, sugerindo que estes resíduos podem



desempenhar um papel crítico na ligação do FRC ao HSA (Figura 15A). Em particular, o resíduo SER287 ( $-2,16 \text{ kcal.mol}^{-1}$ ) formou uma ligação de hidrogênio e uma ligação carbono-hidrogênio com o FRC, enquanto o resíduo ARG257 formou uma ligação carbono-hidrogênio com a HSA. Os outros resíduos formaram interações de Van Der Waals (Figura 15C).

Figura 15 - O painel BIRD exibe a energia de interação dos resíduos com as regiões moleculares para os clusters de energia mais baixa e mais alta. (A) cluster  $\text{FRC}_C$  conf. 2; (B) cluster  $\text{FRC}_C$  conf. 1; (E) cluster  $\text{GLC}_C$  conf. 2; (F) cluster final  $\text{GLC}_D$ . Interações não covalentes num raio de 10 Å da molécula (C) cluster  $\text{FRC}_C$  conf. 2; (D) cluster  $\text{FRC}_C$  conf. 1; (G) cluster  $\text{GLC}_C$  conf. 2; (H) cluster final  $\text{GLC}_D$ .



Por outro lado, a conformação de maior afinidade (FRC<sub>C</sub> *cluster conf.* 1), que inclui os resíduos ARG257 (-7,16 kcal.mol<sup>-1</sup>), ILE290 (-6,34 kcal.mol<sup>-1</sup>) e SER287 (-4,06 kcal.mol<sup>-1</sup>), apresentou as energias de interação mais elevadas com o FRC (Figura 15B). Isto sugere que estes resíduos podem ter uma afinidade particularmente elevada para o FRC, possivelmente devido a cargas específicas ou a interações hidrofóbicas. Neste contexto, o resíduo ARG257 formou duas ligações de hidrogênio com o FRC, enquanto os resíduos SER287, TYR150 e ILE290 (-3,80 kcal.mol<sup>-1</sup>) formaram, cada um, uma ligação de hidrogênio com o FRC. Os outros resíduos representados na Figura 15D formaram interações de Van Der Waals.

É digno de nota que vários resíduos críticos (13 dos 15 resíduos), tais como LEU260, ARG257, ILE264, SER287, LEU219, ILE290, LEU238, ALA261, ALA291, PHE223, HIS288, HIS242 e VAL241, foram identificados tanto nas conformações de menor como de maior afinidade. Isto sugere que estes resíduos podem desempenhar um papel fundamental na modulação da interação entre o FRC e o HSA, independentemente das condições energéticas.

A bioquímica quântica também revelou que o FRC interage fracamente com o *Sudlow Site I*. Verificamos que os resíduos não polares são os principais responsáveis pela energia de interação HSA::FRC. Além disso, os nossos resultados quânticos identificaram novos resíduos do subdomínio IIA, tais como ALA261, ILE264 e LEU219, como potencialmente essenciais para manter o FRC ligado ao HSA. A Figura 15 e as Tabelas A1-A2 do Apêndice A fornecem mais pormenores sobre as interações calculadas.

Embora as interações de Van Der Waals tenham mediado a maioria dos contatos HSA::FRC, também foram detectadas ligações de hidrogênio nas conformações. O baixo número de ligações de hidrogênio (Figura 15C e 15D) é consistente com a baixa energia de interação entre HSA e FRC, a instabilidade estrutural discutida acima e relatórios teóricos anteriores do complexo HSA::FRC (AWANG *et al.*, 2021). Além disso, os resíduos que circundam o FRC<sub>D</sub> neste estudo são semelhantes aos descobertos por cristalografia (WANG *et al.*, 2013b) e ligeiramente diferentes dos resíduos de bolso revelados por uma abordagem teórica anterior (AWANG *et al.*, 2021).

### 3.4.2 Interações GLC::HSA no *Sudlow Site I*

Os resultados da GLC são coerentes com as semelhanças notáveis anteriormente referidas entre os modos de ligação do FRC e da GLC no *Sudlow Site I*. Os cálculos DFT efetuados nas

coordenadas atômicas da GLC revelaram que 13 resíduos eram compartilhados entre os 15 resíduos com maior afinidade (15E-15F).

A análise estrutural e a bioquímica quântica revelaram que a conformação de afinidade mais baixa (cluster GLC<sub>C</sub> conf. 2), incluindo os resíduos ILE290 (-3,63 kcal.mol<sup>-1</sup>), LEU238 (-3,08 kcal.mol<sup>-1</sup>), VAL293 (-2,76 kcal.mol<sup>-1</sup>) e LYS199 (-0,69 kcal.mol<sup>-1</sup>), apresentava interações significativas com o GLC (Figura 15E). Em particular, o LYS199 formou uma ligação de hidrogênio com o GLC, o ALA291 (-2,38 kcal.mol<sup>-1</sup>) estabeleceu uma ligação carbono-hidrogênio com o GLC e o ILE290 formou duas ligações carbono-hidrogênio com o GLC. Os outros resíduos formaram interações de Van Der Waals (Figura 15G).

Por outro lado, a conformação de maior afinidade (GLC<sub>D</sub> conf. final 2), que inclui os resíduos ARG257 (-8,27 kcal.mol<sup>-1</sup>), ALA291 (-3,61 kcal.mol<sup>-1</sup>) e LEU260 (-3,61 kcal.mol<sup>-1</sup>), apresentaram interações significativas com o GLC (Figura 15F). Neste contexto, o resíduo ARG257 formou duas ligações de hidrogênio com HSA e duas ligações de hidrogênio entre GLC e moléculas de água. Os restantes resíduos formaram ligações de Van Der Waals (Figura 15H).

É de salientar que vários resíduos críticos, tais como ALA291, ARG218, ARG257, ILE264, ILE290, LEU219, LEU234, LEU238, LEU260, LYS195, LYS199 e PHE223, foram identificados tanto nas conformações de menor como de maior afinidade; estes resíduos podem desempenhar um papel fundamental na modulação da interação entre a GLC e o *Sudlow Site I* da HSA. Além disso, os resíduos que formam os subsítios SAL e AZT foram expressivamente responsáveis pelas interações não ligantes de HSA com GLC (Figura 15E-15F) (G. SUDLOW; D. J. BIRKETT; D. N. WADE, 1976; YANG *et al.*, 2007b). As tabelas A3 e A4 do Apêndice A apresentam em pormenor estas e outras interações fracas calculadas num raio de 10 Å a partir do centro do GLC.

Consistente com cálculos quânticos anteriores entre GLC e HSA (AWANG *et al.*, 2017), também encontramos uma energia de interação significativa no contato GLC:: ARG257. A bioquímica quântica revelou resíduos adicionais no subdomínio IIA, incluindo LEU219, LEU260, ILE264 e ILE290, que podem desempenhar um papel fundamental na retenção de GLC preso ao sítio FA7. Além disso, os nossos resultados computacionais sugerem fortemente que a presença de um açúcar linear na entrada do *Sudlow Site I* é fundamental para a afinação da glicose revelada pela cristalografia e subsequentemente confirmada por cálculos de dinâmica molecular e de energia de interação (AWANG *et al.*, 2021; NASIRI *et al.*, 2010; WANG *et al.*, 2013b).

A influência de LYS195 e LYS199 na abertura do anel de glicose e subsequente estabelecimento covalente com LYS195 foi revelada por estudos cristalográficos de HSA::monossacarídeo (WANG *et al.*, 2013b). Portanto, para obter mais informações sobre esses resíduos cruciais, acreditamos que é essencial investigá-los ainda mais. Apesar do papel crucial de LYS195 e LYS199 na glicação de HSA, os nossos resultados de bioquímica quântica mostram que a GLC tem uma afinidade fraca e até repulsiva por estes resíduos. Especificamente, a GLC liga-se fracamente ao LYS199 com energias de interação de  $-0,69$  kcal.mol<sup>-1</sup> (cluster GLC<sub>C</sub> conf. 2) e  $0,19$  kcal.mol<sup>-1</sup> (GLC<sub>D</sub> final conf. 2). Os nossos resultados também mostram que o GLC estava distante do LYS195 em ambas as amostras e tinha energias de interação negligenciáveis. Além disso, a nossa investigação revelou que a interação entre o GLC e o LYS199 é facilitada pelas forças de Van Der Waals, como se mostra na Figura 15G e 15H. Estes resultados não são contraditórios, uma vez que a glicação de HSA é um processo bastante longo que demora dias a semanas a concluir-se (BARNABY *et al.*, 2011; SATTARAHMADY *et al.*, 2007).

Além disso, as isosuperfícies de potencial eletrostático calculadas sugerem que as expressões finais (100 ns) da glucopiranosose têm os valores mais extremos de ocupação eletrônica (Figura 16 e 17). Isto pode indicar que a GLC é mais reativa do que a FRC e tem maior probabilidade de ser convertida na sua forma linear (WANG *et al.*, 2013b). Além disso, a fraca ocupação eletrostática em torno de C1 confirma este átomo como um importante ponto de ancoragem para a glicação (WANG *et al.*, 2013b).

Figura 16 - Densidade eletrônica. (A, B) forma fechada e (C, D) forma aberta de frutose antes e depois da dinâmica molecular.

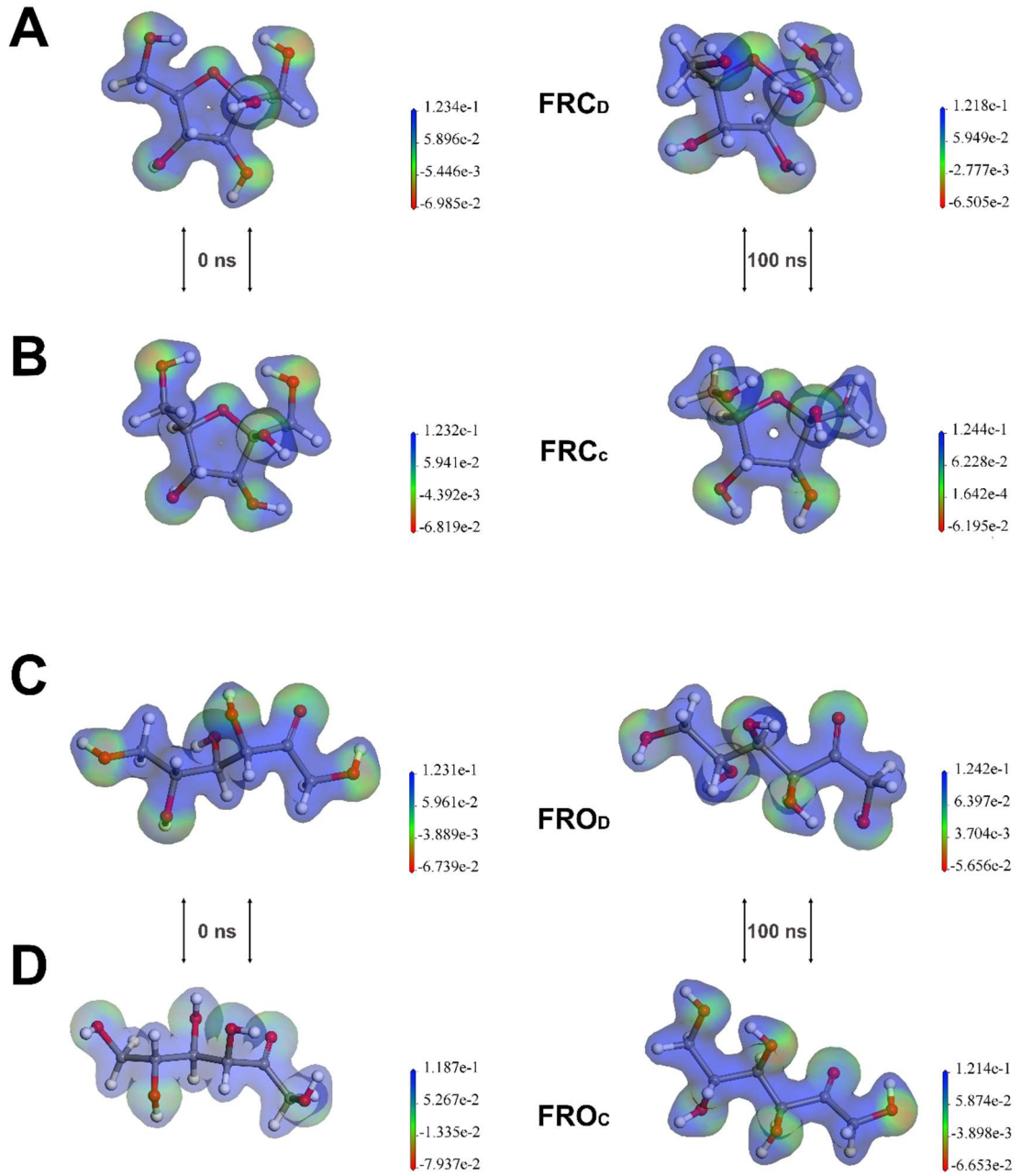
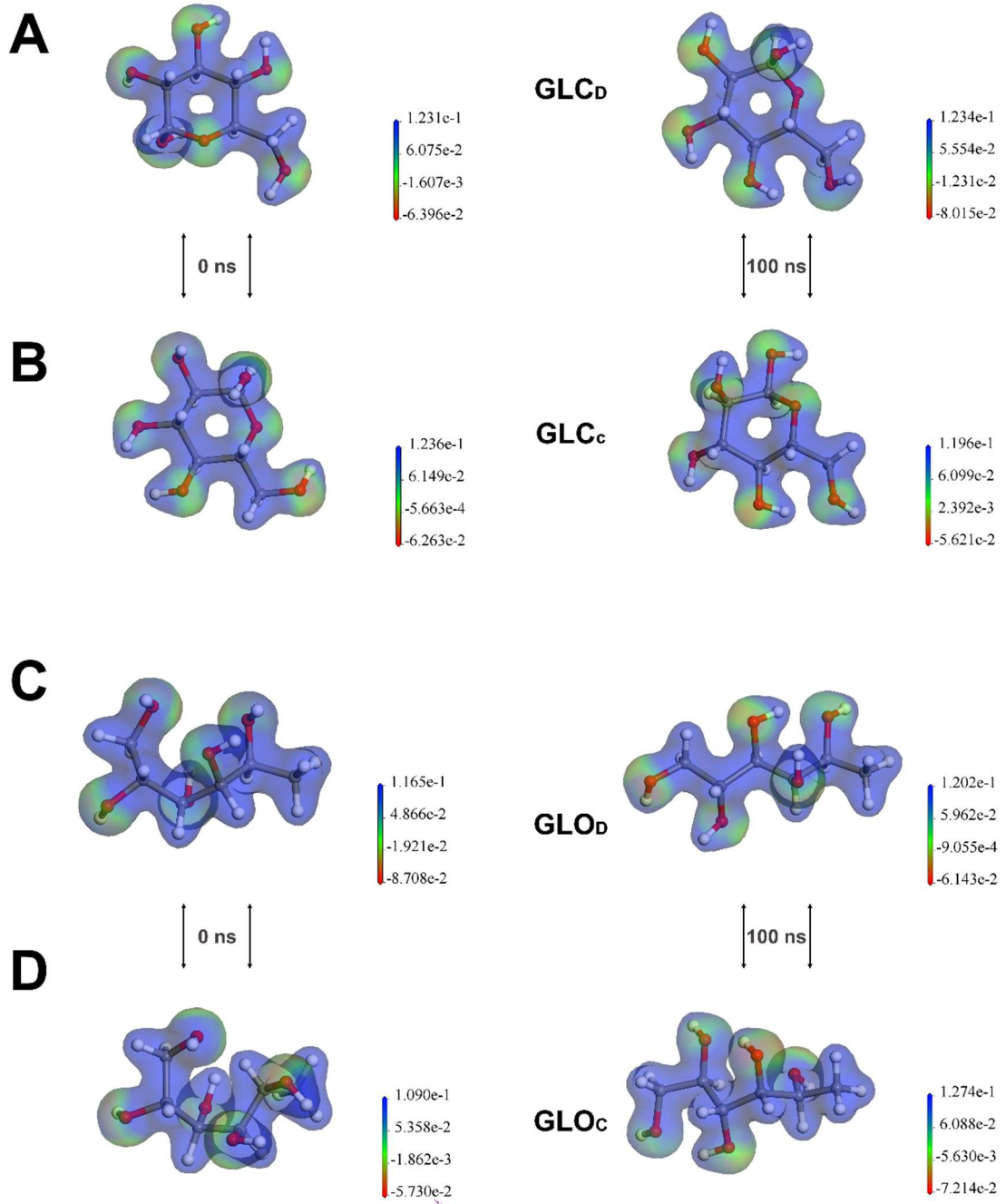


Figura 17 - Comparação da densidade eletrônica inicial e final. (A, B) glucopirranose e (C, D) glicose de cadeia longa.



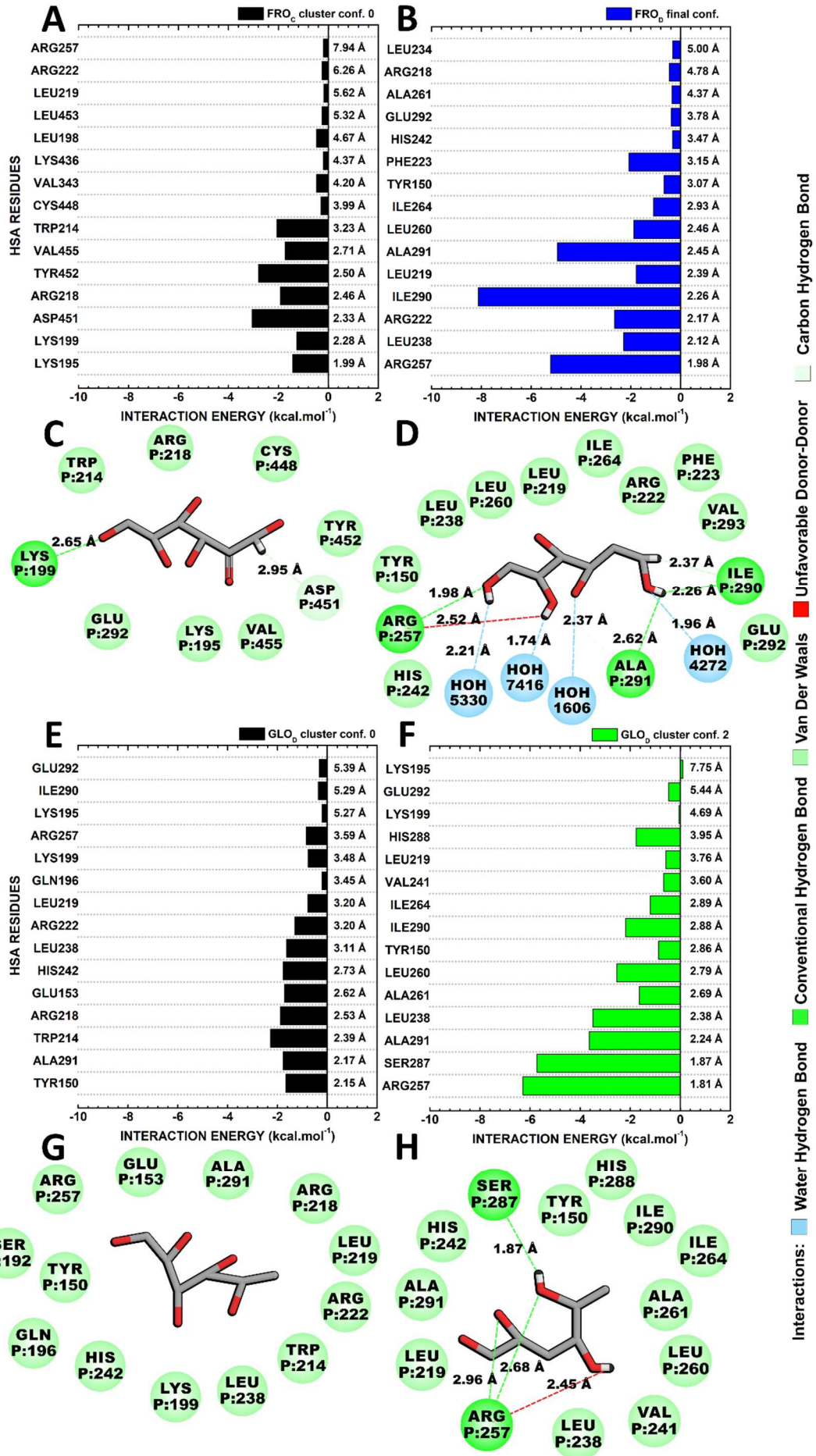
### 3.4.3 Interações FRO::HSA no interior do Sudlow Site I

A interação da molécula de FRO com os resíduos da albumina do soro humano (HSA) foi estudada através da análise das conformações de menor e maior afinidade. Os resultados obtidos revelam uma série de interações significativas que contribuem para a ligação da FRO à HSA.

Para a FRO, observamos resultados semelhantes aos obtidos para a FRC em termos de interação com a HSA. A distribuição das energias de interação mostrou diferenças significativas entre os dois casos. A conformação de menor afinidade (FRO<sub>C</sub> *cluster conf.* 0) apresentou as energias de interação mais elevadas, com os resíduos TRP214, ASP451 e TYR452 identificados como os principais contribuintes para estas interações, com energias de interação de  $-2,07 \text{ kcal.mol}^{-1}$ ,  $-3,05 \text{ kcal.mol}^{-1}$  e  $-2,80 \text{ kcal.mol}^{-1}$ , respectivamente (Figura 18A). Em particular, o resíduo LYS199 ( $-1,27 \text{ kcal.mol}^{-1}$ ) formou uma ligação de hidrogênio com FRO, enquanto o resíduo ASP451 formou uma ligação carbono-hidrogênio com FRO. Os outros resíduos formaram interações de Van Der Waals (Figura 18C).

Figura 18 - O painel BIRD exibe a energia de interação dos resíduos com as regiões moleculares para os clusters de energia mais baixa e mais alta. (A) Cluster FRO<sub>C</sub> *conf.* 0; (B) FRO<sub>D</sub> final *conf.*; (E) GLO<sub>D</sub> *cluster conf.* 0; (F) GLO<sub>D</sub> *cluster conf.* 2. Interações não covalentes num raio de 10 Å da molécula (C) FRO<sub>C</sub> *cluster conf.* 0; (D) FRO<sub>D</sub> *conf.* final; (G) GLO<sub>D</sub> *cluster conf.* 0; (H) GLO<sub>D</sub> *cluster conf.* 2.





Por outro lado, a conformação de maior afinidade (FRO<sub>C</sub> *cluster conf.* 0) também mostrou interações significativas. Os principais participantes no processo foram determinados como sendo os resíduos ILE290 (-8,12 kcal.mol<sup>-1</sup>), ARG257 (-5,23 kcal.mol<sup>-1</sup>) e ALA291 (-4,95 kcal.mol<sup>-1</sup>) (Figura 18B). O resíduo ARG257 formou uma ligação de hidrogênio e uma ligação doador-donor desfavorável com FRO. Além disso, os resíduos ALA291 e ILE290 estabeleceram ligações de hidrogênio individuais com FRO, enquanto o resíduo ALA291 também criou uma ligação carbono-hidrogênio com FRO. Foram também observadas quatro ligações de hidrogênio entre FRO e moléculas de água. Os outros resíduos formaram ligações de van der Waals (Figura 18D).

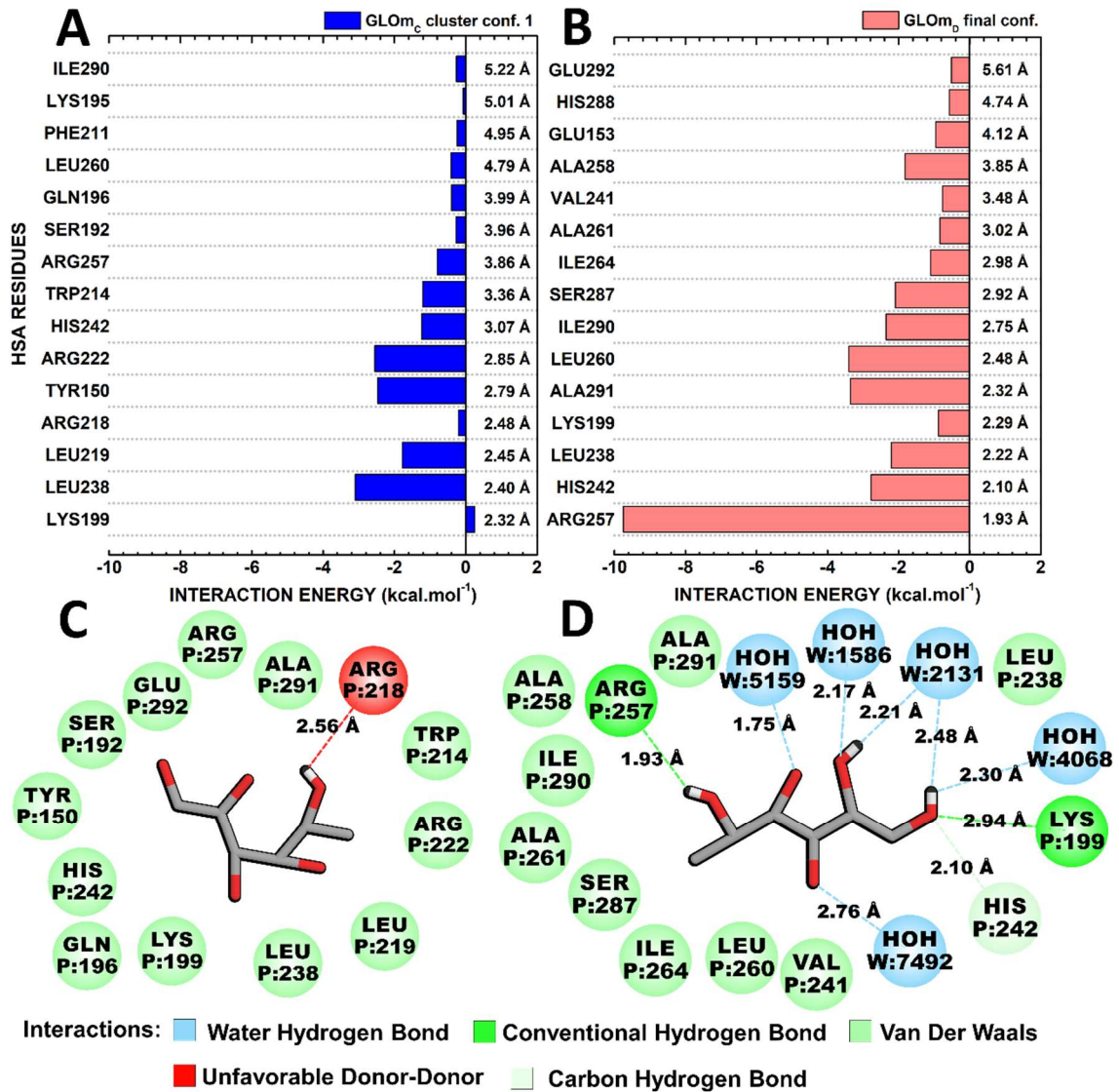
A análise dos resultados revela a presença de apenas quatro resíduos críticos sobrepostos na conformação de menor e maior afinidade. Estes resíduos de aminoácidos, nomeadamente ARG218, ARG222, ARG257 e LEU219, parecem ter uma função crítica na interação entre FRO e HSA, independentemente da disposição específica da molécula de FRO (Figura 18A-18B). A presença destes resíduos críticos sobrepostos sugere que podem ser alvos potenciais para a modulação da interação FRO-HSA. Observamos um efeito notável da água explícita nas interações facilitadas por resíduos polares, em contraste com as interações facilitadas por resíduos não polares. Em particular, a bioquímica quântica e as interações não ligadas detectadas sugerem que a FRO é mais suscetível aos efeitos de solvatação do que a FRC (Figura 18D). Essa maior suscetibilidade por um açúcar linear é consistente com descobertas teóricas anteriores envolvendo complexos HSA::Closed-Open (AWANG *et al.*, 2017, 2021; PONGPRAYOON; MORI, 2018; SERGIO *et al.*, 2021). Os resíduos localizados em torno de FRO<sub>D</sub> neste estudo foram ligeiramente diferentes dos encontrados por cristalografia (ARG218, ARG222 e ALA291). No entanto, os resíduos em torno de FRO<sub>D</sub> eram mais semelhantes aos encontrados pela cristalografia FRC (TYR150, ARG222, LEU238, ARG257 e ALA291) (WANG *et al.*, 2013b). Para obter detalhes adicionais sobre as interações calculadas, consulte as Tabelas A4-A5 no Apêndice A.

#### **3.4.4 Interações GLO/GLO<sub>m</sub>::HSA no Sudlow Site I**

Os cálculos DFT efetuados nas conformações com menor e maior afinidade entre HSA e GLO ou GLO<sub>m</sub> revelaram que 9 e 8 resíduos eram partilhados entre os 15 resíduos com maior afinidade (Figura 18E-18F, 19A-19B). A bioquímica quântica mostrou que a forma linear da glicose tem a maior afinidade de ligação para o *Sudlow Site I*, apesar da ausência de um átomo

polar (devido à remoção do C=O). Esta observação parece confirmar o papel dominante da glicose linear no estabelecimento da glicação (WANG *et al.*, 2013b). Além disso, a análise estrutural (Figura 19) e os cálculos de energia de interação (Figura 10) mostraram que, de acordo com a dinâmica molecular, GLO<sub>m</sub> é retido no *Sudlow Site I*. Além disso, os resíduos que formam os subsítios SAL e AZT (G. SUDLOW; D. J. BIRKETT; D. N. WADE, 1976; YANG *et al.*, 2007b) foram expressivamente responsáveis pelas interações não covalentes de HSA com GLO<sub>m</sub> (Figura 19).

Figura 19 - O painel BIRD exibe a energia de interação dos resíduos com as regiões moleculares para os clusters de energia mais baixa e mais alta. (A) cluster GLO<sub>mC</sub> *conf. 1*; (B) GLO<sub>mD</sub> *conf. final*. Interações não covalentes num raio de 10 Å da molécula (C) GLO<sub>mC</sub> *cluster conf. 1*; (D) GLO<sub>mD</sub> *conf. final*.



Como já foi referido, para o complexo HSA::GLO apenas se obtiveram 3 resultados para GLO<sub>D</sub>, todos eles obtidos pelo método EnGens. Além disso, não se obtiveram resultados para a conformação final GLO<sub>D</sub> ou GLO<sub>C</sub> porque estas moléculas foram libertadas da albumina, como referido anteriormente.

A conformação de menor afinidade (GLO<sub>D</sub> *cluster conf.* 0) revelou o envolvimento de vários resíduos, tais como TRP214 (-2,28 kcal.mol<sup>-1</sup>), ARG218 (-1,88 kcal.mol<sup>-1</sup>) e ALA291 (-1,77 kcal.mol<sup>-1</sup>), que interagiram de forma mais atrativa com a molécula GLO (Figura 18E). É interessante notar que todos estes resíduos estavam envolvidos em interações de van der Waals (Figura 18G).

A conformação de maior afinidade (GLO<sub>D</sub> *cluster conf.* 2) também envolveu vários resíduos. Os resíduos ARG257 (-6,29 kcal.mol<sup>-1</sup>), SER287 (-5,72 kcal.mol<sup>-1</sup>) e ALA291 (-3,64 kcal.mol<sup>-1</sup>) interagiram com a molécula GLO (Figura 18F). O ARG257 esteve envolvido em diferentes tipos de interação, formando duas ligações de hidrogênio e uma interação doador-donor desfavorável com HSA e o resíduo SER287 formou uma ligação de hidrogênio com GLO. Os outros resíduos formaram ligações de van der Waals (Figura 18H).

Nomeadamente, vários resíduos críticos como ALA291, ARG257, GLU292, ILE290, LEU219, LEU238, LYS195, LYS199 e TYR150 foram identificados nas conformações de menor e maior afinidade do complexo GLO::HSA.

Embora todos os açúcares lineares tenham formado ligações de hidrogênio com LYS199 em momentos diferentes da simulação, GLO<sub>mD</sub> formou uma ligação de hidrogênio e uma energia de interação ligeiramente atrativa (~-0,80 kcal.mol<sup>-1</sup>) com este resíduo, tal como em FRO. De notar que esta interação ocorre precisamente entre o átomo O6 de GLO<sub>m</sub> e a cadeia lateral de LYS199 (Figura 19D), sugerindo que O6 pode também desempenhar um papel fundamental no passo de pré-glicação, para além do O5 indicado anteriormente no estudo cristalográfico. Estes dados parecem sugerir que O6 pode ser crítico para estabilizar o GLO<sub>m</sub> perto de LYS195 e LYS199, o que parece ser consistente com a ligação covalente previamente detectada entre a glicose linear e LYS195 (WANG *et al.*, 2013b). Os locais com a menor quantidade de energia (GLO<sub>mC</sub> *cluster conf.* 1) exibiram as energias de interação mais substanciais. Entre essas interações, os resíduos LEU238, ARG222 e TYR150 foram identificados como os principais contribuintes, com energias de interação de -3,11 kcal.mol<sup>-1</sup>, -2,55 kcal.mol<sup>-1</sup> e -2,46 kcal.mol<sup>-1</sup>, respectivamente. Em particular, observamos que o resíduo ARG222 formou ligações doador-donor desfavoráveis com GLO<sub>m</sub>. Os outros resíduos formaram interações de van der Waals (Figura 19A-19C).

Por outro lado, a conformação de maior afinidade ( $GLO_{mC}$  *cluster conf.* 1), que inclui os resíduos ARG257 (-9,75 kcal.mol<sup>-1</sup>), LEU260 (-3,41 kcal.mol<sup>-1</sup>) e ALA291 (-3,36 kcal.mol<sup>-1</sup>), foi identificada como sendo a que mais contribui para estas interações. Em particular, os resíduos ARG257 e LYS199 formaram uma ligação de hidrogênio com  $GLO_m$  e o resíduo HIS242 formou uma ligação carbono-hidrogênio com  $GLO_m$ . Além disso, foram detectadas seis ligações de hidrogênio entre HSA e moléculas de água (Figura 19B-19D). Os outros resíduos formaram interações de van der Waals. A Figura 19 e as Tabelas A9-A10 do Apêndice A apresentam todos os pormenores destas e de outras interações fracas calculadas num raio de 10 Å a partir do centro do  $GLO_m$ .

Em comparação com o detalhamento quântico da interação  $GLO_m::HSA$  usando cálculos ONIOM, detectamos contatos distintos massivos (AWANG *et al.*, 2017). Além disso, observamos que essas interações são semelhantes a dos açúcares de anel fechado - veja as Figuras 7 e 9. Este comportamento revela o papel fundamental da piranose na manutenção do  $GLO_m$  na entrada do *Sudlow Site I* no estado pré-glicado da albumina. No entanto, os nossos resultados não podem explicar totalmente o papel da piranose/furanose na glicação. Em vista disso, estudos futuros usando uma abordagem QM/MM acoplada à dinâmica molecular podem fornecer uma melhor compreensão da função das piranoses no mecanismo de abertura do anel de glicose e, conseqüentemente, na glicação (JAÑA *et al.*, 2018; PAN; LIU; LIU, 2013).

## 4 CONCLUSÃO

Este trabalho relata os resultados da análise mecânica molecular e quântica dos complexos  $\beta$ -D-Fructofuranose::HSA,  $\alpha$ -D-Glucopiranosose::HSA, Keto-D-Frutose::HSA e Aldehydo-D-Glicose modificado::HSA a nível molecular e energético, mapeando e calculando as energias de interação entre os átomos de açúcar dos resíduos de aminoácidos para avaliar a energia total de interação através de conjuntos conformacionais representativos, fragmentação de proteínas, decomposição energética e uma perspectiva de energia de interação ascendente. A Aldehydo-D-Glicose não modificada também foi sondada através de ensaios de dinâmica molecular, que revelaram a sua fuga dos sítios de ligação da HSA e do volume. Acreditamos que este comportamento ajuda a explicar porque é que a glicação é um evento muito demorado que pode ocorrer numa janela de dias a semanas. Tanto quanto sabemos, este é o primeiro estudo teórico a utilizar concomitantemente dinâmica molecular e conformações representativas baseadas no HSA::Monossacarídeo com o objetivo de avaliar o mecanismo de ligação do HSA às formas de glicose e frutose numa proporção de 1:1. A análise bioquímica quântica revelou que as afinidades de todos os complexos HSA::Monossacarídeo apresentam uma elevada flutuação. Por conseguinte, é necessário examinar as conformações representativas para explorar a ligação entre a HSA e os açúcares de uma forma confiável. Além disso, este aspeto é clarificado pela descrição dos tipos de interação identificados nestes resíduos.

Curiosamente, mostramos que um possível intermediário de reação, nomeadamente GLO<sub>m</sub>, era mais estável e propenso a ligar-se perto do *Sudlow Site I* do que a glicose não modificada. Mostramos que esta conformação de transição, imediatamente antes da formação da ligação covalente Glicose::LYS195, tem uma afinidade mais elevada para o *Sudlow Site I* do que a glicose de cadeia longa não modificada, que deixou a HSA. Estudos futuros de QM/MM poderão ser úteis para descrever estes intermediários e a sua propensão para doar elétrons em torno de LYS195. Além disso, a FRO e a GLO<sub>m</sub> apresentaram uma tendência para se deslocarem da entrada do *Sudlow Site I* para uma cavidade mais profunda perto do centro da FA7. Consideramos que este comportamento de movimentação confirma a importância da ocupação do *Sudlow Site I* para o ajuste adequado que resulta na glicação do LYS195, como proposto anteriormente pelos autores do estudo cristalográfico. No entanto, tal como referido nos estudos de dinâmica molecular 1:2 anteriores, as nossas simulações clássicas em combinação com a abordagem de bioquímica quântica também sugerem que LYS195 e LYS199

não podem mediar fortes interações não-ligadas com GLC, apesar do seu papel essencial na glicação de HSA.

Os resultados de fluorescência mostram que tanto a glicose (GLC) quanto a frutose (FRC) suprimem a fluorescência intrínseca da HSA de forma dependente da concentração, com a GLC apresentando um efeito ligeiramente maior. A ausência de mudança no pico máximo de emissão (aproximadamente 340 nm) sugere que a hidrofobicidade do microambiente ao redor do resíduo TRP214 não foi alterada pela ligação dos monossacarídeos. Esses achados estão em linha com estudos anteriores, que indicam uma ligação fraca e espontânea de GLC e FRC à HSA, mediada por ligações de hidrogênio, forças de Van der Waals e efeitos hidrofóbicos. A análise de conformações de maior afinidade sugere que a glicose tem uma influência ligeiramente mais pronunciada na supressão da fluorescência da HSA. Esses resultados contribuem para o entendimento das interações entre HSA e monossacarídeos, fornecendo uma base para futuras pesquisas sobre a modulação da função proteica por ligantes.

Foi observado um padrão consistente e semelhante de deslocamento do TRP214 em todas as amostras de albumina, incluindo as inicialmente ligadas à glicose linear não modificada. Assim, o deslocamento observado do TRP214 não parece ser responsável pelas diferenças modestas na capacidade de inibição da fluorescência da HSA pela frutose em comparação com a glicose.

As energias de interação dos complexos HSA::monossacarídeo são menos atrativas em comparação com as energias de interação de HSA::atrazina-pesticida e HSA::ibuprofeno-droga obtidas anteriormente utilizando protocolos semelhantes. Isto indica que esta classe de hidratos de carbono não tem uma capacidade significativa de ligação à albumina, confirmando a disponibilidade do monossacarídeo para o metabolismo celular. Além disso, verificamos que a energia de interação mais atrativa dos complexos de frutose e o número médio mais elevado de ligações de hidrogênio FRC/FRO a HSA do que as estabelecidas em GLC/GLO<sub>m</sub>::HSA podem estar correlacionados com a afinidade de ligação previamente relatada por espectroscopia (CHAVES; SOARES; CAMPOS DE OLIVEIRA, 2021). Em conjunto, estes resultados sugerem que o estudo da albumina como proteína transportadora, combinando a mecânica molecular e quântica, pode ser uma ferramenta poderosa para descrever os seus mecanismos de ligação a uma vasta gama de pequenas moléculas, tanto do ponto de vista dinâmico como energético, uma vez que estão em total concordância com relatórios experimentais anteriores.

Para futuras pesquisas, uma análise detalhada do processo de mudança química do ligante GLO para o ligante GLO<sub>m</sub> (hexane-1,2,3,4,5-pentol) pode ser explorada através de estudos

avançados de QM/MM. Esses estudos poderiam revelar não apenas os intermediários de reação, como GLOm, mas também investigar em profundidade sua estabilidade relativa, cinética de ligação e interações específicas com resíduos como LYS195. Isso proporcionaria uma compreensão dos mecanismos moleculares envolvidos na modificação química da glicose e sua relevância para processos bioquímicos e fisiológicos.



## REFERÊNCIAS

- ANGUIZOLA, J.; MATSUDA, R.; BARNABY, O. S.; HOY, K. S.; WA, C.; DEBOLT, E.; KOKE, M.; HAGE, D. S. Review: Glycation of human serum albumin. **Clinica Chimica Acta**, v. 425, p. 64–76, out. 2013.
- ANTONY, J.; GRIMME, S. Density functional theory including dispersion corrections for intermolecular interactions in a large benchmark set of biologically relevant molecules. **Physical Chemistry Chemical Physics**, v. 8, n. 45, p. 5287, 2006.
- ANTONY, J.; GRIMME, S. Fully ab initio protein-ligand interaction energies with dispersion corrected density functional theory. **Journal of Computational Chemistry**, v. 33, n. 21, p. 1730–1739, 5 ago. 2012.
- ASCENZI, P.; DI MASI, A.; FANALI, G.; FASANO, M. Heme-based catalytic properties of human serum albumin. **Cell Death Discovery**, v. 1, n. 1, p. 15025, 7 set. 2015.
- ASP, N. Nutritional classification and analysis of food carbohydrates. **The American Journal of Clinical Nutrition**, v. 59, n. 3, p. 679S-681S, mar. 1994.
- AWANG, T.; NIRAMITRANON, J.; JAPRUNG, D.; SAPARPAKORN, P.; PONGPRAYOON, P. Investigating the binding affinities of fructose and galactose to human serum albumin: simulation studies. **Molecular Simulation**, v. 47, n. 9, p. 738–747, 13 jun. 2021.
- AWANG, T.; WIRIYATANAKORN, N.; SAPARPAKORN, P.; JAPRUNG, D.; PONGPRAYOON, P. Understanding the effects of two bound glucose in Sudlow site I on structure and function of human serum albumin: theoretical studies. **Journal of Biomolecular Structure and Dynamics**, v. 35, n. 4, p. 781–790, 12 mar. 2017.
- BAI, X.; WANG, Z.; HUANG, C.; WANG, Z.; CHI, L. Investigation of Non-Enzymatic Glycosylation of Human Serum Albumin Using Ion Trap-Time of Flight Mass Spectrometry. **Molecules**, v. 17, n. 8, p. 8782–8794, 25 jul. 2012.
- BARNABY, O. S.; CERNY, R. L.; CLARKE, W.; HAGE, D. S. Comparison of modification sites formed on human serum albumin at various stages of glycation. **Clinica Chimica Acta**, v. 412, n. 3–4, p. 277–285, jan. 2011.
- BARZEGAR, A.; MOOSAVI-MOVAHEDI, A.; SATTARAHMADY, N.; HOSSEINPOUR-FAIZI, M.; AMINBAKHSH, M.; AHMAD, F.; SABOURY, A.; GANJALI, M.; NOROUZI, P. Spectroscopic Studies of the Effects of Glycation of Human Serum Albumin on L-Trp Binding. **Protein & Peptide Letters**, v. 14, n. 1, p. 13–18, 1 jan. 2007.
- BECKE, A. D. Density-functional thermochemistry. I. The effect of the exchange-only gradient correction. **The Journal of Chemical Physics**, v. 96, n. 3, p. 2155–2160, 1 fev. 1992.
- 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. **The Journal of Chemical Physics**, v. 81, n. 8, p. 3684–3690, 15 out. 1984.

- BEST, R. B.; ZHU, X.; SHIM, J.; LOPES, P. E. M.; MITTAL, J.; FEIG, M.; MACKERELL, A. D. Optimization of the Additive CHARMM All-Atom Protein Force Field Targeting Improved Sampling of the Backbone  $\phi$ ,  $\psi$  and Side-Chain  $\chi_1$  and  $\chi_2$  Dihedral Angles. **Journal of Chemical Theory and Computation**, v. 8, n. 9, p. 3257–3273, 11 set. 2012.
- BHATTACHARYA, A. A.; GRÜNE, T.; CURRY, S. Crystallographic analysis reveals common modes of binding of medium and long-chain fatty acids to human serum albumin 1 Edited by R. Huber. **Journal of Molecular Biology**, v. 303, n. 5, p. 721–732, nov. 2000.
- BROWNLEE, M. Glycosylation Products as Toxic Mediators of Diabetic Complications. **Annual Review of Medicine**, v. 42, n. 1, p. 159–166, fev. 1991.
- CARTER, D. C.; HO, J. X. Structure of Serum Albumin. *Em*: [s.l: s.n.]p. 153–203.
- CASTRO-ALVAREZ, A.; COSTA, A.; VILARRASA, J. The Performance of Several Docking Programs at Reproducing Protein–Macrolide-Like Crystal Structures. **Molecules**, v. 22, n. 1, p. 136, 17 jan. 2017.
- CHAVES, O. A.; SOARES, M. A. G.; CAMPOS DE OLIVEIRA, M. C. Monosaccharides interact weakly with human serum albumin. Insights for the functional perturbations on the binding capacity of albumin. **Carbohydrate Research**, v. 501, p. 108274, mar. 2021.
- CHEN, X.; ZHANG, Y.; ZHANG, J. Z. H. An efficient approach for *ab initio* energy calculation of biopolymers. **The Journal of Chemical Physics**, v. 122, n. 18, p. 184105, 8 maio 2005.
- CHRISTENSEN, A. S.; KUBAŘ, T.; CUI, Q.; ELSTNER, M. Semiempirical Quantum Mechanical Methods for Noncovalent Interactions for Chemical and Biochemical Applications. **Chemical Reviews**, v. 116, n. 9, p. 5301–5337, 11 maio 2016.
- COLLINS, M. A.; BETTENS, R. P. A. Energy-Based Molecular Fragmentation Methods. **Chemical Reviews**, v. 115, n. 12, p. 5607–5642, 24 jun. 2015.
- CONEV, A.; RIGO, M. M.; DEVAURS, D.; FONSECA, A. F.; KALAVADWALA, H.; DE FREITAS, M. V.; CLEMENTI, C.; ZANATTA, G.; ANTUNES, D. A.; KAVRAKI, L. E. EnGens: a computational framework for generation and analysis of representative protein conformational ensembles. **Briefings in Bioinformatics**, v. 24, n. 4, 20 jul. 2023.
- COSSI, M.; REGA, N.; SCALMANI, G.; BARONE, V. Energies, structures, and electronic properties of molecules in solution with the C-PCM solvation model. **Journal of Computational Chemistry**, v. 24, n. 6, p. 669–681, 30 abr. 2003.
- DANTAS, D. S.; OLIVEIRA, J. I. N.; LIMA NETO, J. X.; DA COSTA, R. F.; BEZERRA, E. M.; FREIRE, V. N.; CAETANO, E. W. S.; FULCO, U. L.; ALBUQUERQUE, E. L. Quantum molecular modelling of ibuprofen bound to human serum albumin. **RSC Advances**, v. 5, n. 61, p. 49439–49450, 2015.
- DELLEY, B. An all-electron numerical method for solving the local density functional for polyatomic molecules. **The Journal of Chemical Physics**, v. 92, n. 1, p. 508–517, 1 jan. 1990.

- DELLEY, B. From molecules to solids with the DMol3 approach. **The Journal of Chemical Physics**, v. 113, n. 18, p. 7756–7764, 8 nov. 2000.
- DIRAC, P. A. M. Note on Exchange Phenomena in the Thomas Atom. **Mathematical Proceedings of the Cambridge Philosophical Society**, v. 26, n. 3, p. 376–385, 24 jul. 1930.
- DUARTE, H. Chemical reactivity indexes from density functional theory: Formalism and perspectives. **Quimica Nova**, v. 24, p. 501–508, jun. 2001.
- EHRlich, S.; MOELLMANN, J.; GRIMME, S. Dispersion-Corrected Density Functional Theory for Aromatic Interactions in Complex Systems. **Accounts of Chemical Research**, v. 46, n. 4, p. 916–926, 16 abr. 2013.
- FANALI, G.; DI MASI, A.; TREZZA, V.; MARINO, M.; FASANO, M.; ASCENZI, P. Human serum albumin: From bench to bedside. **Molecular Aspects of Medicine**, v. 33, n. 3, p. 209–290, jun. 2012.
- FERMI, E. Un metodo statistico per la determinazione di alcune proprietà dell'atomo. **Rend. Accad. Naz. Lincei**, v. 6 n° 12, p. 602–607, 1927.
- FOX, S. J.; DZIEDZIC, J.; FOX, T.; TAUTERMANN, C. S.; SKYLARIS, C.-K. Density functional theory calculations on entire proteins for free energies of binding: Application to a model polar binding site. **Proteins: Structure, Function, and Bioinformatics**, v. 82, n. 12, p. 3335–3346, dez. 2014.
- FRANÇA, V. L. B.; AMARAL, J. L.; MARTINS, Y. A.; CAETANO, E. W. S.; BRUNALDI, K.; FREIRE, V. N. Characterization of the binding interaction between atrazine and human serum albumin: Fluorescence spectroscopy, molecular dynamics and quantum biochemistry. **Chemico-Biological Interactions**, v. 366, p. 110130, out. 2022.
- G. SUDLOW; D. J. BIRKETT; D. N. WADE. Further Characterization of Specific Drug Binding Sites on Human Serum Albumin. **American Society for Pharmacology and Experimental Therapeutics**, v. 12, p. 1052–1061, 1976.
- GANESAN, A.; COOTE, M. L.; BARAKAT, K. Molecular dynamics-driven drug discovery: leaping forward with confidence. **Drug Discovery Today**, v. 22, n. 2, p. 249–269, fev. 2017.
- GELPI, J.; HOSPITAL, A.; GOÑI, R.; OROZCO, M. Molecular dynamics simulations: advances and applications. **Advances and Applications in Bioinformatics and Chemistry**, p. 37, nov. 2015.
- GHOSH, R.; KISHORE, N. Mechanistic physicochemical insights into glycation and drug binding by serum albumin: Implications in diabetic conditions. **Biochimie**, v. 193, p. 16–37, fev. 2022.
- GHUMAN, J.; ZUNSZAIN, P. A.; PETITPAS, I.; BHATTACHARYA, A. A.; OTAGIRI, M.; CURRY, S. Structural Basis of the Drug-binding Specificity of Human Serum Albumin. **Journal of Molecular Biology**, v. 353, n. 1, p. 38–52, out. 2005.
- GORDON, M. S.; FEDOROV, D. G.; PRUITT, S. R.; SLIPCHENKO, L. V. Fragmentation Methods: A Route to Accurate Calculations on Large Systems. **Chemical Reviews**, v. 112, n. 1, p. 632–672, 11 jan. 2012a.

- GORDON, M. S.; FEDOROV, D. G.; PRUITT, S. R.; SLIPCHENKO, L. V. Fragmentation Methods: A Route to Accurate Calculations on Large Systems. **Chemical Reviews**, v. 112, n. 1, p. 632–672, 11 jan. 2012b.
- GRINTER, S.; ZOU, X. Challenges, Applications, and Recent Advances of Protein-Ligand Docking in Structure-Based Drug Design. **Molecules**, v. 19, n. 7, p. 10150–10176, 11 jul. 2014.
- HAMILTON, J. A. NMR reveals molecular interactions and dynamics of fatty acid binding to albumin. **Biochimica et Biophysica Acta (BBA) - General Subjects**, v. 1830, n. 12, p. 5418–5426, dez. 2013.
- HAND, D. J.; MCLACHLAN, G. J.; BASFORD, K. E. Mixture Models: Inference and Applications to Clustering. **Applied Statistics**, v. 38, n. 2, p. 384, 1989.
- HE, X.; ZHANG, J. Z. H. A new method for direct calculation of total energy of protein. **The Journal of Chemical Physics**, v. 122, n. 3, 15 jan. 2005.
- HOHENBERG, P.; KOHN, W. Inhomogeneous Electron Gas. **Physical Review**, v. 136, n. 3B, p. B864–B871, 9 nov. 1964.
- HOLLINGSWORTH, S. A.; DROR, R. O. Molecular Dynamics Simulation for All. **Neuron**, v. 99, n. 6, p. 1129–1143, set. 2018.
- HU, X.; ZENG, Z.; ZHANG, J.; WU, D.; LI, H.; GENG, F. Molecular dynamics simulation of the interaction of food proteins with small molecules. **Food Chemistry**, v. 405, p. 134824, mar. 2023.
- JANA, A. K.; BATKULWAR, K. B.; KULKARNI, M. J.; SENGUPTA, N. Glycation induces conformational changes in the amyloid- $\beta$  peptide and enhances its aggregation propensity: molecular insights. **Physical Chemistry Chemical Physics**, v. 18, n. 46, p. 31446–31458, 2016.
- JAÑA, G. A.; MENDOZA, F.; OSORIO, M. I.; ALDERETE, J. B.; FERNANDES, P. A.; RAMOS, M. J.; JIMÉNEZ, V. A. A QM/MM approach on the structural and stereoelectronic factors governing glycosylation by GTF-SI from *Streptococcus mutans*. **Organic & Biomolecular Chemistry**, v. 16, n. 14, p. 2438–2447, 2018.
- JIANG, N.; MA, J.; JIANG, Y. Electrostatic field-adapted molecular fractionation with conjugated caps for energy calculations of charged biomolecules. **The Journal of Chemical Physics**, v. 124, n. 11, 21 mar. 2006.
- JONES, R. O. Density functional theory: Its origins, rise to prominence, and future. **Reviews of Modern Physics**, v. 87, n. 3, p. 897–923, 25 ago. 2015.
- JOSEPH, K. S.; HAGE, D. S. The effects of glycation on the binding of human serum albumin to warfarin and l-tryptophan. **Journal of Pharmaceutical and Biomedical Analysis**, v. 53, n. 3, p. 811–818, nov. 2010.
- KHALIFAH, R. G.; BAYNES, J. W.; HUDSON, B. G. Amadorins: Novel Post-Amadori Inhibitors of Advanced Glycation Reactions. **Biochemical and Biophysical Research Communications**, v. 257, n. 2, p. 251–258, abr. 1999.

- KITCHEN, D. B.; DECORNEZ, H.; FURR, J. R.; BAJORATH, J. Docking and scoring in virtual screening for drug discovery: methods and applications. **Nature Reviews Drug Discovery**, v. 3, n. 11, p. 935–949, 1 nov. 2004.
- KLAMT, A.; SCHÜÜRMAN, G. COSMO: a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. **J. Chem. Soc., Perkin Trans. 2**, n. 5, p. 799–805, 1993.
- KOHN, W.; SHAM, L. J. Self-Consistent Equations Including Exchange and Correlation Effects. **Physical Review**, v. 140, n. 4A, p. A1133–A1138, 15 nov. 1965.
- KRENZEL, E. S.; CHEN, Z.; HAMILTON, J. A. Correspondence of Fatty Acid and Drug Binding Sites on Human Serum Albumin: A Two-Dimensional Nuclear Magnetic Resonance Study. **Biochemistry**, v. 52, n. 9, p. 1559–1567, 5 mar. 2013.
- LEGGIO, C.; GALANTINI, L.; PAVEL, N. V. About the albumin structure in solution: cigar Expanded form versus heart Normal shape. **Physical Chemistry Chemical Physics**, v. 10, n. 45, p. 6741, 2008.
- LI, J.; ZHU, X.; YANG, C.; SHI, R. Characterization of the binding of angiotensin II receptor blockers to human serum albumin using docking and molecular dynamics simulation. **Journal of Molecular Modeling**, v. 16, n. 4, p. 789–798, 12 abr. 2010.
- LIU, J.; HERBERT, J. M. Pair–Pair Approximation to the Generalized Many-Body Expansion: An Alternative to the Four-Body Expansion for ab Initio Prediction of Protein Energetics via Molecular Fragmentation. **Journal of Chemical Theory and Computation**, v. 12, n. 2, p. 572–584, 9 fev. 2016.
- LIU, Y.; YANG, X.; GAN, J.; CHEN, S.; XIAO, Z.-X.; CAO, Y. CB-Dock2: improved protein–ligand blind docking by integrating cavity detection, docking and homologous template fitting. **Nucleic Acids Research**, v. 50, n. W1, p. W159–W164, 5 jul. 2022.
- LV, Y.; LIANG, Q.; LI, Y.; LIU, X.; ZHANG, D.; LI, X. Study of the binding mechanism between hydroxytyrosol and bovine serum albumin using multispectral and molecular docking. **Food Hydrocolloids**, v. 122, p. 107072, jan. 2022.
- MALIK, V. S.; HU, F. B. Fructose and Cardiometabolic Health. **Journal of the American College of Cardiology**, v. 66, n. 14, p. 1615–1624, out. 2015.
- MARDIROSSIAN, N.; HEAD-GORDON, M. Thirty years of density functional theory in computational chemistry: an overview and extensive assessment of 200 density functionals. **Molecular Physics**, v. 115, n. 19, p. 2315–2372, 2 out. 2017.
- MARTÍNEZ-ROSELL, G.; GIORGINO, T.; DE FABRITIIS, G. PlayMolecule ProteinPrepare: A Web Application for Protein Preparation for Molecular Dynamics Simulations. **Journal of Chemical Information and Modeling**, v. 57, n. 7, p. 1511–1516, 24 jul. 2017.
- MARUYAMA, Y.; IGARASHI, R.; USHIKU, Y.; MITSUTAKE, A. Analysis of Protein Folding Simulation with Moving Root Mean Square Deviation. **Journal of Chemical Information and Modeling**, v. 63, n. 5, p. 1529–1541, 13 mar. 2023.

MCKEEN, L. W. Introduction to the Properties of Plastic and Elastomer Films. *Em: Film Properties of Plastics and Elastomers*. [s.l.] Elsevier, 2012. p. 19–55.

MOEINPOUR, F.; MOHSENI-SHAHRI, F. S.; MALAEKEH-NIKOUEI, B.; NASSIRLI, H. Investigation into the interaction of losartan with human serum albumin and glycated human serum albumin by spectroscopic and molecular dynamics simulation techniques: A comparison study. **Chemico-Biological Interactions**, v. 257, p. 4–13, set. 2016.

MORAIS, P. A.; MAIA, F. F.; SOLIS-CALERO, C.; CAETANO, E. W. S.; FREIRE, V. N.; CARVALHO, H. F. The urokinase plasminogen activator binding to its receptor: a quantum biochemistry description within an in/homogeneous dielectric function framework with application to uPA–uPAR peptide inhibitors. **Physical Chemistry Chemical Physics**, v. 22, n. 6, p. 3570–3583, 2020.

MORGON, N. H.; COUTINHO, K. R. **Métodos de química teórica e modelagem molecular** Editora Livraria da Física, 2007.

NASIRI, R.; BAHRAMI, H.; ZAHEDI, M.; MOOSAVI-MOVAHEDI, A. A.; SATTARAHMADY, N. A Theoretical Elucidation of Glucose Interaction with HSA's Domains. **Journal of Biomolecular Structure and Dynamics**, v. 28, n. 2, p. 211–226, out. 2010.

NELSON H. MORGON; CUSTODIO R. Teoria do Funcional de Densidade. **Química Nova**, v. 18, p. 44–55, 1994.

NOCEDAL, J.; WRIGHT, S. Numerical Optimization. **ITM Web of Conferences ItmconfLibrary of Congress Cataloging-in-Publication**, v. Second, 1960.

PAN, X.-L.; LIU, W.; LIU, J.-Y. Mechanism of the Glycosylation Step Catalyzed by Human  $\alpha$ -Galactosidase: A QM/MM Metadynamics Study. **The Journal of Physical Chemistry B**, v. 117, n. 2, p. 484–489, 17 jan. 2013.

PEARLMAN, D. A.; CASE, D. A.; CALDWELL, J. W.; ROSS, W. S.; CHEATHAM, T. E.; DEBOLT, S.; FERGUSON, D.; SEIBEL, G.; KOLLMAN, P. AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules. **Computer Physics Communications**, v. 91, n. 1–3, p. 1–41, set. 1995.

PERDEW, J. P.; BURKE, K.; ERNZERHOF, M. Generalized Gradient Approximation Made Simple. **Physical Review Letters**, v. 77, n. 18, p. 3865–3868, 28 out. 1996.

PERDEW, J. P.; ZUNGER, A. Self-interaction correction to density-functional approximations for many-electron systems. **Physical Review B**, v. 23, n. 10, p. 5048–5079, 15 maio 1981.

PERUTZ, M. F. Electrostatic Effects in Proteins. **Science**, v. 201, n. 4362, p. 1187–1191, 29 set. 1978.

PETERS, T. Ligand Binding by Albumin. *Em: All About Albumin*. [s.l.] Elsevier, 1995. p. 76–132.

- PETITPAS, I.; GRÜNE, T.; BHATTACHARYA, A. A.; CURRY, S. Crystal structures of human serum albumin complexed with monounsaturated and polyunsaturated fatty acids. **Journal of Molecular Biology**, v. 314, n. 5, p. 955–960, dez. 2001.
- PONGPRAYOON, P.; GLEESON, M. P. Probing the binding site characteristics of HSA: A combined molecular dynamics and cheminformatics investigation. **Journal of Molecular Graphics and Modelling**, v. 54, p. 164–173, nov. 2014.
- PONGPRAYOON, P.; MORI, T. The critical role of dimer formation in monosaccharides binding to human serum albumin. **Physical Chemistry Chemical Physics**, v. 20, n. 5, p. 3249–3257, 2018.
- POTTER, E. D.; HEREK, J. L.; PEDERSEN, S.; LIU, Q.; ZEWAIL, A. H. Femtosecond laser control of a chemical reaction. **Nature**, v. 355, n. 6355, p. 66–68, jan. 1992.
- QI, X.; TESTER, R. F. Fructose, galactose and glucose – In health and disease. **Clinical Nutrition ESPEN**, v. 33, p. 18–28, out. 2019.
- RAPPE, A. K.; CASEWIT, C. J.; COLWELL, K. S.; GODDARD, W. A.; SKIFF, W. M. UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations. **Journal of the American Chemical Society**, v. 114, n. 25, p. 10024–10035, 1 dez. 1992.
- REYNOLDS, T. M. Chemistry of Nonenzymic Browning I. The Reaction between Aldoses and Amines. *Em: [s.l: s.n.]p. 1–52.*
- SAHA, A.; RAGHAVACHARI, K. Analysis of Different Fragmentation Strategies on a Variety of Large Peptides: Implementation of a Low Level of Theory in Fragment-Based Methods Can Be a Crucial Factor. **Journal of Chemical Theory and Computation**, v. 11, n. 5, p. 2012–2023, 12 maio 2015.
- SATTARAHMADY, N.; MOOSAVIMOVAHEDI, A.; AHMAD, F.; HAKIMELAHI, G.; HABIBIREZAEI, M.; SABOURY, A.; SHEIBANI, N. Formation of the molten globule-like state during prolonged glycation of human serum albumin. **Biochimica et Biophysica Acta (BBA) - General Subjects**, v. 1770, n. 6, p. 933–942, jun. 2007.
- SERGIO, L. M.; MARTINS, Y. A.; AMARAL, J. L.; FRANÇA, V. L. B.; DE FREITAS, C. F.; NETO, A. M.; HIOKA, N.; RAVANELLI, M. I.; MAREZE-COSTA, C.; CLAUDIO DA COSTA, S.; FREIRE, V. N.; BRUNALDI, K. Molecular insight on the binding of stevia glycosides to bovine serum albumin. **Chemico-Biological Interactions**, v. 344, p. 109526, ago. 2021.
- SHAM, L. J.; KOHN, W. One-Particle Properties of an Inhomogeneous Interacting Electron Gas. **Physical Review**, v. 145, n. 2, p. 561–567, 13 maio 1966.
- SHAMSI, A.; AHMED, A.; BANO, B. Probing the interaction of anticancer drug temsirolimus with human serum albumin: molecular docking and spectroscopic insight. **Journal of Biomolecular Structure and Dynamics**, v. 36, n. 6, p. 1479–1489, 26 abr. 2018.

SIMARD, J. R.; ZUNSZAIN, P. A.; HAMILTON, J. A.; CURRY, S. Location of High and Low Affinity Fatty Acid Binding Sites on Human Serum Albumin Revealed by NMR Drug-competition Analysis. **Journal of Molecular Biology**, v. 361, n. 2, p. 336–351, ago. 2006.

SITTIWANICHA, S.; JAPRUNG, D.; MORI, T.; PONGPRAYOON, P. Structural and Dynamic Alteration of Glycated Human Serum Albumin in Schiff Base and Amadori Adducts: A Molecular Simulation Study. **The Journal of Physical Chemistry B**, v. 127, n. 23, p. 5230–5240, 15 jun. 2023.

SMITH, G. R.; STERNBERG, M. J. E. Prediction of protein–protein interactions by docking methods. **Current Opinion in Structural Biology**, v. 12, n. 1, p. 28–35, fev. 2002.

SÖDERHJELM, P.; AQUILANTE, F.; RYDE, U. Calculation of Protein–Ligand Interaction Energies by a Fragmentation Approach Combining High-Level Quantum Chemistry with Classical Many-Body Effects. **The Journal of Physical Chemistry B**, v. 113, n. 32, p. 11085–11094, 13 ago. 2009.

SOUSA, B. L.; BARROSO-NETO, I. L.; OLIVEIRA, E. F.; FONSECA, E.; LIMA-NETO, P.; LADEIRA, L. O.; FREIRE, V. N. Explaining RANKL inhibition by OPG through quantum biochemistry computations and insights into peptide-design for the treatment of osteoporosis. **RSC Advances**, v. 6, n. 88, p. 84926–84942, 2016.

SOUSA, S. F.; FERNANDES, P. A.; RAMOS, M. J. General Performance of Density Functionals. **The Journal of Physical Chemistry A**, v. 111, n. 42, p. 10439–10452, 1 out. 2007.

SUDLOW, G.; BIRKETT, D. J.; WADE, D. N. The Characterization of Two Specific Drug Binding Sites on Human Serum Albumin. **Molecular Pharmacology**, v. 11, n. 6, p. 824, 1 nov. 1975. Disponível em: <<http://molpharm.aspetjournals.org/content/11/6/824.abstract>>.

SZKUDLAREK, A.; SUŁKOWSKA, A.; MACIĄŻEK-JURCZYK, M.; CHUDZIK, M.; RÓWNICKA-ZUBIK, J. Effects of non-enzymatic glycation in human serum albumin. Spectroscopic analysis. **Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy**, v. 152, p. 645–653, jan. 2016.

TAYEH, N.; RUNGASSAMY, T.; ALBANI, J. R. Fluorescence spectral resolution of tryptophan residues in bovine and human serum albumins. **Journal of Pharmaceutical and Biomedical Analysis**, v. 50, n. 2, p. 107–116, set. 2009.

THIEL, W. Semiempirical quantum–chemical methods. **WIREs Computational Molecular Science**, v. 4, n. 2, p. 145–157, 19 mar. 2014.

THOMAS, L. H. The calculation of atomic fields. **Mathematical Proceedings of the Cambridge Philosophical Society**, v. 23, n. 5, p. 542–548, 24 jan. 1927.

THORNALLEY, P. J.; LANGBORG, A.; MINHAS, H. S. Formation of glyoxal, methylglyoxal and 3-deoxyglucosone in the glycation of proteins by glucose. **Biochemical Journal**, v. 344, n. 1, p. 109–116, 15 nov. 1999.



TKATCHENKO, A.; SCHEFFLER, M. Accurate Molecular Van Der Waals Interactions from Ground-State Electron Density and Free-Atom Reference Data. **Physical Review Letters**, v. 102, n. 7, p. 073005, 20 fev. 2009.

TROZZI, F.; WANG, X.; TAO, P. UMAP as a Dimensionality Reduction Tool for Molecular Dynamics Simulations of Biomacromolecules: A Comparison Study. **The Journal of Physical Chemistry B**, v. 125, n. 19, p. 5022–5034, 20 maio 2021.

TRÜEB, B.; HOLENSTEIN, C. G.; FISCHER, R. W.; WINTERHALTER, K. H. Nonenzymatic glycosylation of proteins. A warning. **Journal of Biological Chemistry**, v. 255, n. 14, p. 6717–6720, jul. 1980.

VANOMMESLAEGHE, K.; HATCHER, E.; ACHARYA, C.; KUNDU, S.; ZHONG, S.; SHIM, J.; DARIAN, E.; GUVENCH, O.; LOPES, P.; VOROBYOV, I.; MACKERELL, A. D. CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. **Journal of Computational Chemistry**, p. NA-NA, 2009.

VOSKO, S. H.; WILK, L.; NUSAIR, M. Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis. **Canadian Journal of Physics**, v. 58, n. 8, p. 1200–1211, 1 ago. 1980.

WANG, W.; GAN, N.; SUN, Q.; WU, D.; GAN, R.; ZHANG, M.; TANG, P.; LI, H. Study on the interaction of ertugliflozin with human serum albumin in vitro by multispectroscopic methods, molecular docking, and molecular dynamics simulation. **Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy**, v. 219, p. 83–90, ago. 2019.

WANG, X.; LIU, J.; ZHANG, J. Z. H.; HE, X. Electrostatically Embedded Generalized Molecular Fractionation with Conjugate Caps Method for Full Quantum Mechanical Calculation of Protein Energy. **The Journal of Physical Chemistry A**, v. 117, n. 32, p. 7149–7161, 15 ago. 2013a.

WANG, Y.; YU, H.; SHI, X.; LUO, Z.; LIN, D.; HUANG, M. Structural Mechanism of Ring-opening Reaction of Glucose by Human Serum Albumin. **Journal of Biological Chemistry**, v. 288, n. 22, p. 15980–15987, maio 2013b.

WANI, T. A.; BAKHEIT, A. H.; AL-MAJED, A. A.; ALTWAIJRY, N.; BAQUAYSH, A.; ALJURAISSY, A.; ZARGAR, S. Binding and drug displacement study of colchicine and bovine serum albumin in presence of azithromycin using multispectroscopic techniques and molecular dynamic simulation. **Journal of Molecular Liquids**, v. 333, p. 115934, jul. 2021.

WARSHEL, A.; RUSSELL, S. T.; CHURG, A. K. Macroscopic models for studies of electrostatic interactions in proteins: limitations and applicability. **Proceedings of the National Academy of Sciences**, v. 81, n. 15, p. 4785–4789, ago. 1984.

WARSHEL, A.; SHARMA, P. K.; KATO, M.; PARSON, W. W. Modeling electrostatic effects in proteins. **Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics**, v. 1764, n. 11, p. 1647–1676, nov. 2006.

YANG, F.; BIAN, C.; ZHU, L.; ZHAO, G.; HUANG, Z.; HUANG, M. Effect of human serum albumin on drug metabolism: Structural evidence of esterase activity of human serum albumin. **Journal of Structural Biology**, v. 157, n. 2, p. 348–355, fev. 2007a.

YANG, F.; BIAN, C.; ZHU, L.; ZHAO, G.; HUANG, Z.; HUANG, M. Effect of human serum albumin on drug metabolism: Structural evidence of esterase activity of human serum albumin. **Journal of Structural Biology**, v. 157, n. 2, p. 348–355, fev. 2007b.

ZAMAN, A.; ARIF, Z.; MOINUDDIN; ALAM, K. Fructose-human serum albumin interaction undergoes numerous biophysical and biochemical changes before forming AGEs and aggregates. **International Journal of Biological Macromolecules**, v. 109, p. 896–906, abr. 2018.

ZHANG, D. W.; ZHANG, J. Z. H. Molecular fractionation with conjugate caps for full quantum mechanical calculation of protein–molecule interaction energy. **The Journal of Chemical Physics**, v. 119, n. 7, p. 3599–3605, 15 ago. 2003.

ZHANG, Y.; LEE, P.; LIANG, S.; ZHOU, Z.; WU, X.; YANG, F.; LIANG, H. Structural Basis of Non-Steroidal Anti-Inflammatory Drug Diclofenac Binding to Human Serum Albumin. **Chemical Biology & Drug Design**, v. 86, n. 5, p. 1178–1184, nov. 2015.

ZHAO, H.; CAFLISCH, A. Molecular dynamics in drug design. **European Journal of Medicinal Chemistry**, v. 91, p. 4–14, fev. 2015.

ZUNSZAIN, P. A.; GHUMAN, J.; KOMATSU, T.; TSUCHIDA, E.; CURRY, S. Crystal structural analysis of human serum albumin complexed with hemin and fatty acid. **BMC Structural Biology**, v. 3, n. 1, p. 6, 2003.

ZUNSZAIN, P. A.; GHUMAN, J.; MCDONAGH, A. F.; CURRY, S. Crystallographic Analysis of Human Serum Albumin Complexed with 4Z,15E-Bilirubin-IX $\alpha$ . **Journal of Molecular Biology**, v. 381, n. 2, p. 394–406, ago. 2008.

## APÊNDICE A – TABELAS

Tabela A1 - Resíduos, distância (Å), energia de interação (kcal.mol<sup>-1</sup>), átomo da HSA, átomo da FRCc (*cluster conf. 2*).

Resíduos	Distância (Å)	Energia de Interação (kcal.mol <sup>-1</sup> )	HSA Átomo	FRC Átomo
ALA291	2,07	-1,02	HB2	H9
ILE264	2,24	-2,24	HD12	H6
LEU238	2,34	-1,61	HD11	H5
PHE223	2,46	-0,93	HE2	O6
ILE290	2,51	-1,79	HG22	O2
LEU219	2,64	-2,05	HD21	H7
SER287	2,66	-2,16	HB3	H2
ARG257	2,70	-2,94	HD3	O1
ALA261	2,72	-1,45	HB3	H8
LEU260	2,79	-2,97	HD22	H5
LEU234	3,29	-1,05	HD22	H6
HIS242	4,32	-0,28	HE2	H11
LYS199	4,39	0,01	HZ3	H11
HIS288	4,57	-0,41	HA	H9
GLU153	4,93	-0,35	OE1	H10
GLU292	5,16	-0,26	H	H9
ARG218	5,25	-0,25	HE	H11
VAL241	5,26	-0,28	HG12	H3
TYR150	5,27	-0,26	HE1	H3
LYS286	5,33	-0,27	O	H1
ALA258	5,34	0,14	HA	O1
VAL235	5,67	-0,08	HA	H7
ARG222	5,73	-0,21	HD2	H4
LYS262	5,84	-0,20	H	H8
CYS289	5,92	-0,20	C	H9
PRO152	6,06	-0,15	HB3	H2
ASP237	6,21	-0,03	HB3	H6
ASP256	6,24	-0,11	O	H8
TRP214	6,28	-0,06	HZ3	H11
ASP259	6,32	0,08	C	H8
VAL293	6,37	-0,17	H	H9
TYR263	6,40	-0,10	HD2	H6
ILE271	6,43	-0,08	HG21	H12
LEU283	6,45	-0,08	O	H1
ALA215	6,50	-0,05	HB1	H7
LEU284	6,63	-0,10	O	H2
LEU275	6,75	-0,11	HD23	H12
PHE211	6,94	-0,02	HZ	O4
VAL231	7,24	-0,04	O	H7
THR239	7,31	-0,05	H	H7
LYS233	7,47	-0,09	O	H6
ALA254	7,48	-0,04	O	O1

<b>CYS265</b>	7,76	-0,07	H	H1
<b>CYS278</b>	7,78	-0,08	HB3	H1
<b>GLU285</b>	7,82	-0,03	O	H1
<b>GLN196</b>	7,89	-0,03	OE1	H11
<b>THR236</b>	8,06	-0,07	N	H7
<b>CYS253</b>	8,32	-0,05	O	O1
<b>GLU230</b>	8,32	0,11	O	H6
<b>PHE156</b>	8,35	-0,03	HB3	H2
<b>ASP255</b>	8,37	0,15	O	H8
<b>SER220</b>	8,61	-0,02	N	H7
<b>LYS195</b>	8,75	0,07	HZ3	H10
<b>LYS240</b>	8,77	-0,07	H	H5
<b>SER272</b>	8,84	-0,01	HB2	H12
<b>VAL216</b>	8,89	0,00	HA	H7
<b>CYS279</b>	8,92	-0,05	HA	H1
<b>PRO282</b>	9,23	0,01	O	H1
<b>ASN267</b>	9,28	-0,05	HB3	H6
<b>SER232</b>	9,33	0,02	O	H7
<b>CYS245</b>	9,53	-0,06	SG	H3
<b>GLN268</b>	9,87	-0,05	HA	H12

---

Tabela A2 - Resíduos, distância (Å), energia de interação (kcal.mol<sup>-1</sup>), átomo da HSA, átomo da FRC<sub>C</sub> (*cluster conf. 1*).

Resíduos	Distância (Å)	Energia de Interação (kcal.mol <sup>-1</sup> )	HSA Átomo	FRC Átomo
ARG257	2,03	-7,16	HH22	O3
ALA261	2,11	-1,03	HB2	H8
ALA291	2,14	-2,42	HA	H12
SER287	2,14	-4,06	O	H10
TYR150	2,26	-3,80	HH	O4
LEU238	2,29	-2,28	HD13	H5
HIS242	2,49	-1,64	HE2	O4
ILE290	2,54	-6,34	HG21	H9
LEU260	2,61	-2,39	HG	H1
LEU219	2,71	-1,15	HD22	H6
ILE264	3,24	-0,99	HD13	H1
LYS199	3,55	0,20	HZ1	H11
VAL241	3,67	-0,83	HG11	H3
PHE223	3,68	-0,66	HZ	O6
ARG222	3,88	-0,24	HG3	H12
HIS288	4,07	-0,89	HA	H10
ALA258	4,20	-1,04	HA	H8
LYS286	4,42	-0,52	O	H2
GLU153	4,84	0,10	OE1	O3
GLU292	4,87	-0,37	N	H12
VAL293	5,01	-0,24	HG23	H12
ASP256	5,19	-0,26	O	H8
LYS262	5,30	-0,25	H	H8
ARG218	5,30	-0,07	HE	H7
LEU234	5,44	-0,17	HB2	H6
TRP214	5,48	-0,11	HH2	H7
CYS289	5,52	-0,46	C	O2
LEU283	5,67	-0,09	HD11	H8
ASP259	5,67	-0,18	N	H8
GLN196	6,13	-0,06	OE1	H11
ASP237	6,37	-0,03	C	H5
PRO152	6,53	-0,15	HB3	O3
TYR263	6,68	-0,07	H	H1
PHE211	6,69	-0,05	HZ	H5
CYS278	6,74	-0,10	HB3	H2
LEU284	6,80	-0,05	O	H10
VAL235	6,81	-0,05	HA	H6
LEU275	6,82	-0,11	HD21	H9
CYS265	6,98	-0,06	H	H1
CYS253	6,99	-0,05	O	H8
THR239	6,99	-0,05	N	H5
ALA254	7,06	-0,05	O	H8
ALA215	7,27	-0,04	HB1	H6
GLU285	7,39	0,02	O	H10
ASP255	7,57	0,13	O	H8

<b>LYS195</b>	7,82	0,00	HB2	H11
<b>LYS240</b>	8,31	-0,09	H	H5
<b>VAL231</b>	8,37	-0,03	HG12	H6
<b>PHE156</b>	8,38	-0,03	HD1	H10
<b>CYS245</b>	8,42	-0,07	HB2	O4
<b>ILE271</b>	8,54	-0,03	HG21	H6
<b>SER220</b>	8,90	-0,01	N	H6
<b>GLN221</b>	8,94	0,00	C	H12
<b>SER192</b>	8,97	-0,01	HG	H11
<b>CYS279</b>	8,99	-0,05	SG	H2
<b>THR243</b>	9,04	-0,04	H	H5
<b>THR236</b>	9,08	-0,06	C	H5
<b>PRO282</b>	9,30	0,01	O	H2
<b>LYS233</b>	9,32	-0,02	HE3	H1
<b>LYS281</b>	9,38	-0,14	HD2	H2
<b>SER272</b>	9,43	-0,01	HB3	H12
<b>GLU294</b>	9,52	0,04	N	H12
<b>GLU266</b>	9,56	0,07	H	H1
<b>VAL216</b>	9,75	-0,01	N	H6
<b>GLU252</b>	9,82	0,10	O	H8
<b>VAL343</b>	9,84	0,00	HG21	H7

---

Tabela A3 - Resíduos, distância (Å), energia de interação (kcal.mol<sup>-1</sup>), átomo da HSA, átomo da GLC<sub>C</sub> (*cluster conf. 2*).

<b>Resíduos</b>	<b>Distância (Å)</b>	<b>Energia de Interação (kcal.mol<sup>-1</sup>)</b>	<b>HSA Átomo</b>	<b>FRC Átomo</b>
LEU219	2,13	-1,64	HD21	H9
LEU238	2,13	-3,08	HD13	H5
VAL293	2,27	-2,76	HG11	H1
ARG222	2,43	-0,28	HE	H8
ILE264	2,43	-0,97	HD11	H10
ALA291	2,50	-2,38	HA	H7
ILE290	2,62	-3,63	O	H4
LYS199	2,66	-0,69	HZ2	O6
PHE223	2,76	-1,53	HZ	O2
LEU260	3,88	-0,76	HD22	H11
ARG257	3,98	-0,50	HH22	H6
HIS242	4,17	-0,44	HE2	H5
LEU234	4,31	-0,50	HB2	H10
ARG218	4,56	-0,69	HG3	H8
GLU153	4,59	-0,33	OE1	H12
GLU292	4,66	-0,50	N	H7
LEU275	4,76	-0,26	HD22	H2
SER287	5,49	-0,18	O	H4
ILE271	5,68	-0,16	HG22	O2
TYR150	5,85	-0,26	HE2	H6
VAL241	5,86	-0,11	HG11	H11
ALA261	5,95	-0,10	HA	O3
CYS289	6,10	-0,33	O	H1
GLN196	6,27	-0,06	OE1	H12
GLU294	6,45	-0,21	O	H8
VAL235	6,61	-0,09	HA	H3
HIS288	6,72	-0,13	O	H4
LYS195	6,78	0,06	HZ2	H12
TRP214	6,88	-0,06	HH2	O6
SER272	7,10	-0,04	HB3	H2
SER192	7,14	-0,01	HG	H12
ASN295	7,17	-0,14	OD1	H8
TYR263	7,24	-0,05	HD2	H10
PHE211	7,27	-0,02	HZ	H5
ALA215	7,32	-0,05	O	H9
SER220	7,44	-0,05	N	H9
VAL231	7,60	-0,04	HG23	H9
THR239	7,60	-0,05	H	H3
ASP237	7,66	0,09	HB3	H10
PRO152	7,67	-0,07	HB3	H6
LYS286	7,75	-0,12	O	H4
LYS274	7,81	-0,03	HB3	H2
GLN221	8,11	-0,06	C	H9
ALA258	8,20	-0,01	N	H11
ASP256	8,30	0,04	O	H11

<b>CYS278</b>	8,34	-0,06	SG	H2
<b>VAL216</b>	8,47	-0,01	HA	H9
<b>LYS233</b>	8,48	-0,07	O	H10
<b>CYS265</b>	8,62	-0,04	H	O3
<b>VAL343</b>	8,91	-0,01	HG22	H8
<b>LYS262</b>	8,92	0,00	H	H11
<b>ASP259</b>	8,92	0,07	C	H11
<b>ASP296</b>	9,05	0,08	H	H9
<b>ALA226</b>	9,12	-0,01	HB3	H9
<b>ALA217</b>	9,25	-0,01	O	H9
<b>THR236</b>	9,26	-0,04	N	H3
<b>ASN267</b>	9,28	-0,05	HB3	H10
<b>GLN268</b>	9,54	-0,06	HA	H10
<b>GLU230</b>	9,54	0,10	O	H10
<b>PRO224</b>	9,56	-0,01	HD2	H9
<b>CYS253</b>	9,58	-0,04	O	H11
<b>CYS245</b>	9,72	-0,06	HB2	H6
<b>ALA254</b>	9,76	-0,04	HA	H11
<b>LEU283</b>	9,78	-0,01	HD12	H11
<b>GLU285</b>	9,91	-0,03	O	H4

---



Tabela A4 - Resíduos, distância (Å), energia de interação (kcal.mol<sup>-1</sup>), átomo da HSA, átomo da GLC<sub>D</sub> (conf. final).

<b>Resíduos</b>	<b>Distância (Å)</b>	<b>Energia de Interação (kcal.mol<sup>-1</sup>)</b>	<b>HSA Átomo</b>	<b>FRC Átomo</b>
<b>LEU219</b>	<b>2,00</b>	<b>-1,54</b>	<b>HD22</b>	<b>H11</b>
<b>ARG257</b>	2,01	-8,27	HE	O3
<b>ILE264</b>	2,02	-0,65	HD1	H1
<b>TYR150</b>	2,17	-3,24	HH	O4
<b>PHE223</b>	2,25	-1,13	HE2	H5
<b>ILE290</b>	2,29	-3,42	HG22	H1
<b>LEU260</b>	2,40	-3,61	HD22	H6
<b>ALA291</b>	2,49	-3,61	HA	O5
<b>LEU238</b>	2,57	-3,25	HD11	H9
<b>ALA261</b>	2,95	-0,67	HB3	H2
<b>HIS242</b>	3,02	-0,10	HE2	H3
<b>LYS199</b>	3,55	0,19	HZ3	H10
<b>LEU234</b>	3,65	-0,87	HB2	H12
<b>ARG218</b>	4,04	-0,27	HH22	H4
<b>SER287</b>	4,08	-0,71	O	H7
<b>ARG222</b>	4,62	0,01	HE	H4
<b>VAL241</b>	4,63	-0,30	HG12	H9
<b>VAL293</b>	5,03	-0,20	HG11	H5
<b>HIS288</b>	5,42	-0,24	HA	H7
<b>GLU292</b>	5,49	-0,53	N	H4
<b>LEU275</b>	5,52	-0,09	HD22	H5
<b>ILE271</b>	5,55	-0,10	HG21	H5
<b>TRP214</b>	5,75	-0,12	HZ3	H10
<b>PHE211</b>	5,87	-0,06	HZ	H10
<b>GLU153</b>	5,92	0,09	HG2	H3
<b>GLN196</b>	6,12	-0,17	OE1	H3
<b>ALA258</b>	6,13	-0,01	HA	H2
<b>PRO152</b>	6,13	-0,13	HB1	O4
<b>VAL235</b>	6,32	-0,12	HA	H12
<b>LYS262</b>	6,61	-0,22	HN	H2
<b>ASP237</b>	6,66	-0,01	HB1	H6
<b>ALA215</b>	6,76	-0,09	O	H11
<b>LYS286</b>	6,76	-0,24	O	H2
<b>CYS289</b>	6,91	-0,15	C	H7
<b>LEU283</b>	6,95	-0,02	HD13	H2
<b>TYR263</b>	6,96	-0,16	HD2	H1
<b>VAL231</b>	7,04	-0,05	HG11	H11
<b>ASP259</b>	7,12	0,07	C	H2
<b>CYS265</b>	7,20	-0,06	HN	H1
<b>ASP256</b>	7,33	-0,03	O	H2
<b>THR239</b>	7,60	-0,08	N	H9
<b>SER272</b>	7,70	0,00	HB1	H5
<b>SER220</b>	7,81	-0,02	N	H11
<b>CYS245</b>	8,03	-0,02	HB2	H3
<b>LYS233</b>	8,04	-0,06	O	H12

<b>CYS253</b>	8,27	0,00	SG	H3
<b>LEU284</b>	8,31	0,00	O	H7
<b>VAL216</b>	8,43	-0,03	HA	H11
<b>LYS195</b>	8,47	-0,02	HD1	H4
<b>CYS278</b>	8,48	-0,03	HB1	H1
<b>GLU230</b>	8,53	0,05	O	H12
<b>THR236</b>	8,56	-0,04	HN	H12
<b>ALA254</b>	8,64	0,01	O	O3
<b>GLU294</b>	8,65	-0,04	O	H5
<b>SER192</b>	8,85	-0,03	HG1	H3
<b>ASN295</b>	8,97	-0,03	OD1	H5
<b>GLN221</b>	9,00	0,00	C	H5
<b>ALA151</b>	9,09	-0,04	HN	H3
<b>GLU285</b>	9,09	0,08	O	H7
<b>LYS240</b>	9,15	-0,15	HN	H9
<b>ALA226</b>	9,22	-0,01	HB2	H5
<b>GLN268</b>	9,36	-0,01	HA	H5
<b>THR243</b>	9,36	-0,03	HN	H9
<b>ALA217</b>	9,38	0,00	C	H11
<b>SER232</b>	9,53	-0,01	HA	H12
<b>GLU266</b>	9,59	0,08	HN	H1
<b>ASP296</b>	9,72	0,01	HN	H5
<b>LYS274</b>	9,76	-0,02	HB1	H5
<b>ASN267</b>	9,78	-0,03	HB1	H1
<b>PHE156</b>	9,79	-0,02	HD2	O4
<b>CYS279</b>	9,93	-0,02	SG	H1

---

Tabela A5 - Resíduos, distância (Å), energia de interação (kcal.mol<sup>-1</sup>), átomo da HSA, átomo da FRO<sub>C</sub> (*cluster conf.* 0).

Resíduos	Distância (Å)	Energia de Interação (kcal.mol <sup>-1</sup> )	HSA Átomo	FRC Átomo
LYS195	1,99	-1,43	HE2	H9
LYS199	2,28	-1,27	HE3	H12
ASP451	2,33	-3,05	HB2	H1
ARG218	2,46	-1,92	HG2	H7
TYR452	2,50	-2,80	HB3	H8
VAL455	2,71	-1,74	HG21	H2
TRP214	3,23	-2,07	HZ2	H6
GLU292	3,93	0,71	OE2	H11
CYS448	3,99	-0,30	O	H1
VAL343	4,20	-0,48	HG22	H10
LYS436	4,37	-0,21	HD3	H8
LEU198	4,67	-0,48	HD22	H6
PRO447	5,08	-0,17	O	O2
ALA215	5,21	-0,18	HA	H12
LEU238	5,24	-0,12	HD22	H12
HIS242	5,25	-0,11	HE2	H12
LEU453	5,32	-0,26	H	H1
ALA449	5,43	-0,10	O	H1
ALA191	5,56	-0,08	HB3	H9
ALA291	5,58	-0,12	O	H11
LEU219	5,62	-0,18	HD12	H7
GLU450	5,64	0,06	C	H1
SER454	5,86	-0,12	HB2	H2
PHE211	6,12	-0,07	HZ	H12
SER342	6,14	-0,10	HG	H1
ARG222	6,26	-0,25	NH2	H4
ALA194	6,36	-0,05	HB2	H9
VAL456	6,40	-0,09	H	H2
VAL344	6,42	-0,09	HG23	H10
GLN196	6,62	-0,13	N	H5
SER192	6,73	-0,06	HA	H9
GLN221	6,94	-0,05	OE1	H4
CYS437	7,44	-0,01	SG	H8
VAL433	7,61	-0,01	HG21	H8
LEU347	7,64	-0,02	HD13	H10
GLU153	7,67	0,17	OE1	H11
ALA217	7,70	0,07	HB3	H10
VAL235	7,77	-0,01	HG21	H12
TYR150	7,87	-0,03	HE2	H11
MET446	7,90	-0,03	O	H1
ARG257	7,94	-0,19	HH21	H11
TYR341	8,06	-0,07	O	O2
VAL293	8,30	-0,05	H	H11
LEU481	8,40	-0,01	HD23	H10
VAL216	8,43	-0,02	N	H12

<b>SER202</b>	8,52	-0,01	HG	H6
<b>HIS288</b>	8,57	-0,01	HD1	H11
<b>ASN458</b>	8,71	-0,02	HB2	H2
<b>GLN459</b>	8,71	-0,02	HE21	H9
<b>GLU188</b>	8,75	0,22	OE2	H3
<b>LYS190</b>	8,77	-0,02	O	H9
<b>LEU457</b>	8,78	-0,01	H	H2
<b>ARG197</b>	8,80	-0,11	H	H5
<b>CYS200</b>	8,85	-0,01	H	H12
<b>SER193</b>	8,94	-0,01	N	H9
<b>LYS212</b>	8,95	-0,01	HA	H12
<b>SER220</b>	9,15	-0,01	H	H7
<b>THR239</b>	9,17	0,01	HA	H12
<b>LEU345</b>	9,34	-0,01	H	H1
<b>ASN295</b>	9,35	-0,12	HD22	O2
<b>ILE290</b>	9,37	-0,01	O	H11
<b>LEU260</b>	9,43	-0,04	HD21	H12
<b>VAL241</b>	9,45	-0,01	HB	H12
<b>PRO339</b>	9,46	-0,04	O	O2
<b>LYS432</b>	9,47	-0,03	HB3	H8
<b>LEU346</b>	9,50	-0,01	HD13	H10
<b>HIS440</b>	9,60	0,02	HE2	H8
<b>LEU203</b>	9,64	-0,01	HD13	H12
<b>ARG445</b>	9,74	0,03	O	H1
<b>ALA213</b>	9,76	-0,01	O	H10
<b>ASP187</b>	9,99	0,11	O	H9

---

Tabela A6 - Resíduos, distância (Å), energia de interação (kcal.mol<sup>-1</sup>), átomo da HSA, átomo da FRO<sub>D</sub> (conf. final).

<b>Resíduos</b>	<b>Distância (Å)</b>	<b>Energia de Interação (kcal.mol<sup>-1</sup>)</b>	<b>HSA Átomo</b>	<b>FRC Átomo</b>
<b>ARG257</b>	1,98	-5,23	HE	O4
<b>LEU238</b>	2,12	-2,28	HD11	H10
<b>ARG222</b>	2,17	-2,66	HB2	H11
<b>VAL293</b>	2,24	-0,22	HG13	H12
<b>ILE290</b>	2,26	-8,12	O	H5
<b>LEU219</b>	2,39	-1,79	HD23	H1
<b>ALA291</b>	2,45	-4,95	HA	O2
<b>LEU260</b>	2,46	-1,89	HD23	H10
<b>ILE264</b>	2,93	-1,10	HD3	O5
<b>TYR150</b>	3,07	-0,67	HE2	O4
<b>PHE223</b>	3,15	-2,08	HE2	O5
<b>HIS242</b>	3,47	-0,33	HE2	H4
<b>GLU292</b>	3,78	-0,39	N	H5
<b>ALA261</b>	4,37	-0,35	HB2	H9
<b>LEU275</b>	4,76	-0,18	HD21	H12
<b>ARG218</b>	4,78	-0,45	HB2	H3
<b>SER287</b>	4,80	-0,26	O	O2
<b>VAL241</b>	4,95	-0,27	HG13	O4
<b>LEU234</b>	5,00	-0,32	HD11	H1
<b>GLU153</b>	5,36	-0,20	HG1	H2
<b>CYS289</b>	5,37	-0,20	O	H5
<b>ILE271</b>	5,65	-0,13	HG23	O5
<b>HIS288</b>	5,70	-0,18	HA	H2
<b>LYS199</b>	5,73	-0,06	HE2	H4
<b>PRO152</b>	5,81	-0,10	HG2	O4
<b>TRP214</b>	6,34	-0,11	HH2	H7
<b>VAL235</b>	6,52	-0,06	HA	H1
<b>SER272</b>	6,63	-0,02	HB1	H12
<b>GLN221</b>	6,66	-0,01	C	H11
<b>GLU294</b>	6,69	-0,24	N	H5
<b>ALA258</b>	6,73	-0,04	N	H9
<b>ALA215</b>	6,93	-0,08	O	H3
<b>PRO224</b>	7,31	-0,02	HD2	H11
<b>ASP256</b>	7,35	-0,01	O	H9
<b>ASP237</b>	7,39	0,04	C	H10
<b>SER220</b>	7,46	-0,05	N	H11
<b>GLN196</b>	7,51	-0,03	OE1	H4
<b>LYS274</b>	7,58	-0,04	HB1	H12
<b>CYS253</b>	7,64	-0,01	O	O4
<b>THR239</b>	7,64	-0,04	N	H10
<b>ALA254</b>	7,65	0,00	HA	O4
<b>PHE211</b>	7,65	-0,02	HZ	H4
<b>CYS278</b>	7,77	-0,04	SG	H12
<b>LYS262</b>	7,83	-0,13	HN	H9
<b>LYS286</b>	7,88	-0,15	O	H6

<b>VAL231</b>	7,90	-0,02	HG11	H1
<b>TYR263</b>	8,15	-0,11	HD2	H1
<b>ASP259</b>	8,18	0,04	C	H9
<b>CYS245</b>	8,35	-0,02	HB2	H4
<b>ASN295</b>	8,38	-0,04	HA	H11
<b>CYS265</b>	8,42	-0,03	HN	O5
<b>ALA217</b>	8,51	-0,02	C	H3
<b>ALA226</b>	8,78	-0,01	HB3	H11
<b>ALA151</b>	8,79	-0,03	HN	H4
<b>LEU284</b>	9,08	-0,01	O	H2
<b>VAL216</b>	9,13	-0,02	HA	H1
<b>ASP296</b>	9,17	0,02	HN	H11
<b>GLN268</b>	9,20	-0,01	HA	O5
<b>THR236</b>	9,27	-0,03	N	H1
<b>LYS233</b>	9,38	-0,10	C	H1
<b>LYS240</b>	9,39	-0,13	C	H10
<b>LEU283</b>	9,40	0,00	HD12	H9
<b>GLU230</b>	9,43	0,09	O	H1
<b>SER192</b>	9,55	-0,01	HG1	H2
<b>THR243</b>	9,55	-0,02	HN	H4
<b>PHE156</b>	9,59	-0,01	HB2	H2
<b>PRO339</b>	9,70	-0,01	HB2	H11
<b>CYS279</b>	9,72	-0,01	SG	H12
<b>ASP255</b>	9,73	0,11	O	H9
<b>GLU285</b>	9,81	0,02	O	O2
<b>PHE149</b>	9,94	-0,01	HE2	H2
<b>LEU154</b>	9,95	-0,03	HN	O4
<b>VAL343</b>	10,00	0,00	HG12	H3

---

Tabela A7 - Resíduos, distância (Å), energia de interação (kcal.mol<sup>-1</sup>), átomo da HSA, átomo da GLO<sub>D</sub> (*cluster conf. 0*).

Resíduos	Distância (Å)	Energia de Interação (kcal.mol <sup>-1</sup> )	HSA Átomo	FRC Átomo
TYR150	2,15	-1,66	HD2	H14
ALA291	2,17	-1,77	HB1	H11
TRP214	2,39	-2,28	HH2	H4
ARG218	2,53	-1,88	HG2	H1
GLU153	2,62	-1,71	HG3	H8
HIS242	2,73	-1,77	HE2	H12
LEU238	3,11	-1,64	HD23	H2
ARG222	3,20	-1,30	HH12	H3
LEU219	3,20	-0,78	HD21	H2
SER192	3,40	-0,15	HG	H14
GLN196	3,45	-0,22	OE1	H14
LYS199	3,48	-0,78	HE2	O4
ARG257	3,59	-0,85	HH22	H8
PHE211	5,18	-0,19	HZ	O4
LYS195	5,27	-0,20	HD3	H7
ILE290	5,29	-0,36	HG22	O3
ALA215	5,34	-0,13	HA	H2
HIS288	5,37	-0,12	HA	H13
GLU292	5,39	-0,33	OE2	H13
GLU188	5,65	-0,27	OE2	H9
SER287	5,75	-0,08	O	H11
LEU234	5,90	-0,08	HD13	H2
PRO152	5,96	-0,11	HD3	H14
LEU260	6,07	-0,12	HD21	H5
VAL293	6,22	-0,12	HG12	H3
VAL343	6,52	-0,04	HG22	H1
ALA151	6,63	-0,07	N	H14
PHE149	6,65	0,03	HA	H14
VAL241	6,71	-0,12	HG12	H6
PHE223	6,83	-0,07	HZ	H3
ARG160	7,07	0,21	HH21	H9
LEU154	7,20	-0,06	H	H14
VAL235	7,22	-0,02	HG22	H2
SER193	7,43	-0,03	N	H14
PHE156	7,44	-0,02	HD2	H8
ALA191	7,45	0,02	O	H14
GLY189	7,59	-0,03	HA2	H9
TYR148	7,63	-0,02	O	H14
ALA217	8,03	0,02	C	H1
CYS245	8,04	-0,08	HB3	H14
ILE264	8,19	-0,03	HG13	O3
CYS289	8,24	-0,04	C	H11
VAL216	8,33	-0,01	N	H2
SER220	8,43	-0,02	H	H1
ASP451	8,52	0,13	OD1	H1

<b>THR239</b>	8,75	-0,02	N	H5
<b>GLN221</b>	8,77	-0,02	HB3	H1
<b>ALA261</b>	8,81	0,01	HB2	O3
<b>LEU284</b>	8,97	-0,04	O	H8
<b>LYS286</b>	8,99	-0,08	O	H11
<b>ARG197</b>	9,00	0,04	H	H14
<b>PHE157</b>	9,15	-0,01	HD1	H9
<b>CYS253</b>	9,19	-0,03	SG	H14
<b>LYS212</b>	9,29	-0,01	HA	H2
<b>ASP237</b>	9,37	0,08	C	H5
<b>ALA194</b>	9,42	0,02	N	H14
<b>CYS246</b>	9,44	-0,01	SG	H12
<b>ASN295</b>	9,49	-0,10	OD1	H1
<b>CYS200</b>	9,51	0,00	H	H12
<b>LEU198</b>	9,59	-0,03	HB2	H12
<b>LEU155</b>	9,60	-0,01	H	H8
<b>GLN29</b>	9,63	0,01	HE22	H14
<b>LEU203</b>	9,63	-0,01	HD12	H12
<b>VAL231</b>	9,78	-0,01	HG12	H2
<b>LEU275</b>	9,90	-0,03	HD23	H3

---



Tabela A8 - Resíduos, distância (Å), energia de interação (kcal.mol<sup>-1</sup>), átomo da HSA, átomo da GLO<sub>D</sub> (*cluster conf. 2*).

<b>Resíduos</b>	<b>Distância (Å)</b>	<b>Energia de Interação (kcal.mol<sup>-1</sup>)</b>	<b>HSA Átomo</b>	<b>FRC Átomo</b>
<b>ARG257</b>	1,81	-6,29	HH22	O5
<b>SER287</b>	1,87	-5,72	O	H10
<b>ALA291</b>	2,24	-3,64	HB2	H10
<b>LEU238</b>	2,38	-3,49	HD13	H13
<b>ALA261</b>	2,69	-1,64	HB2	H3
<b>LEU260</b>	2,79	-2,53	HD23	H5
<b>TYR150</b>	2,86	-0,87	HH	H13
<b>ILE290</b>	2,88	-2,19	HB	H4
<b>ILE264</b>	2,89	-1,21	HD11	H3
<b>HIS242</b>	3,31	-0,12	HE2	H13
<b>VAL241</b>	3,60	-0,67	HG11	H11
<b>LEU219</b>	3,76	-0,58	HD21	H12
<b>HIS288</b>	3,95	-1,76	N	H10
<b>ARG218</b>	4,38	-0,15	HE	H14
<b>TRP214</b>	4,59	-0,19	HH2	H14
<b>LYS199</b>	4,69	-0,05	HE2	H14
<b>ARG222</b>	4,71	-0,07	HH12	H9
<b>PHE223</b>	4,72	-0,20	HE2	H12
<b>GLU153</b>	4,85	-0,32	OE2	H7
<b>LYS286</b>	4,93	-0,26	O	H2
<b>LEU283</b>	5,28	-0,15	HD13	H2
<b>ALA258</b>	5,30	-0,02	N	H1
<b>GLU292</b>	5,44	-0,46	H	H10
<b>LEU234</b>	5,51	-0,23	HD22	H5
<b>CYS289</b>	5,54	-0,08	N	H10
<b>LYS262</b>	5,97	-0,16	N	H3
<b>ASP256</b>	6,13	-0,06	O	H1
<b>GLN196</b>	6,16	-0,08	OE1	H13
<b>VAL293</b>	6,37	-0,15	H	H9
<b>LEU284</b>	6,43	-0,08	O	H10
<b>CYS265</b>	6,59	-0,08	H	H3
<b>ALA215</b>	6,84	-0,06	HA	H14
<b>GLU285</b>	6,87	-0,04	O	H10
<b>ALA254</b>	6,91	-0,03	O	H1
<b>ASP259</b>	6,91	0,09	C	H3
<b>ASP237</b>	6,94	0,00	O	O3
<b>LEU275</b>	7,01	-0,10	HD23	H4
<b>TYR263</b>	7,05	-0,05	H	H3
<b>PRO152</b>	7,15	-0,12	HB3	H11
<b>PHE211</b>	7,19	-0,02	HZ	H14
<b>CYS253</b>	7,24	-0,02	O	O3
<b>CYS278</b>	7,36	-0,07	HB3	H2
<b>VAL235</b>	7,40	-0,02	HA	H6
<b>THR239</b>	7,72	-0,05	N	H6
<b>LYS195</b>	7,75	0,10	HD3	H8

<b>ASP255</b>	8,46	0,18	O	H1
<b>CYS279</b>	8,50	-0,05	HA	H2
<b>PHE156</b>	8,65	-0,03	HB3	H10
<b>PRO282</b>	8,73	0,00	O	H2
<b>VAL343</b>	8,81	0,00	HG21	H14
<b>CYS245</b>	8,85	-0,06	HB2	H13
<b>ILE271</b>	8,88	-0,02	HG21	H12
<b>LYS233</b>	9,10	-0,02	O	H5
<b>LYS240</b>	9,12	-0,11	C	H11
<b>GLU266</b>	9,30	0,08	H	H3
<b>SER220</b>	9,48	-0,01	N	H12
<b>LYS281</b>	9,52	-0,08	HB3	H2
<b>VAL231</b>	9,56	-0,02	HG12	H12
<b>THR236</b>	9,57	-0,05	N	H6
<b>ASP451</b>	9,67	0,01	OD1	H14
<b>LEU14</b>	9,69	-0,01	HD21	H1
<b>SER192</b>	9,78	-0,01	OG	H8
<b>VAL216</b>	9,87	-0,02	N	H14
<b>THR243</b>	9,88	-0,04	H	H11

---

Tabela A9 - Resíduos, distância (Å), energia de interação (kcal.mol<sup>-1</sup>), átomo de HSA, átomo da GLO<sub>mc</sub> (*cluster conf. 1*).

<b>Resíduos</b>	<b>Distância (Å)</b>	<b>Energia de Interação (kcal.mol<sup>-1</sup>)</b>	<b>HSA Átomo</b>	<b>FRC Átomo</b>
ALA291	2,29	-0,09	HB3	H13
LYS199	2,32	0,25	HE2	H12
LEU238	2,40	-3,11	HD11	H6
LEU219	2,45	-1,77	HD22	H3
ARG218	2,48	-0,20	HG3	H1
TYR150	2,79	-2,46	HE2	H9
ARG222	2,85	-2,55	HH12	H3
HIS242	3,07	-1,24	HE2	O4
TRP214	3,36	-1,21	HH2	H4
GLU292	3,66	0,11	OE1	H13
ARG257	3,86	-0,80	HH22	H13
SER192	3,96	-0,27	HG	H14
GLN196	3,99	-0,40	OE1	H8
GLU153	4,34	-0,18	HG3	O6
LEU260	4,79	-0,41	HD23	O3
ALA215	4,93	-0,22	HA	H2
PHE211	4,95	-0,24	HZ	O4
LYS195	5,01	-0,07	HD3	H14
ILE290	5,22	-0,27	O	H11
VAL241	5,54	-0,18	HG12	H6
LEU234	5,69	-0,12	HD23	H3
HIS288	5,83	-0,12	HA	H13
PHE223	5,95	-0,21	HE2	H3
SER287	6,04	-0,03	O	H13
VAL293	6,11	-0,20	HG13	H3
PRO152	6,42	-0,13	HG2	H9
VAL343	6,72	-0,02	HG23	H1
VAL235	6,82	-0,04	HG21	H2
GLN221	6,99	-0,04	HB3	H1
ILE264	7,08	-0,06	HG13	O3
PHE149	7,15	-0,08	HD1	H14
ASN295	7,32	-0,14	OD1	H1
ALA191	7,39	-0,02	O	H14
SER220	7,40	-0,02	H	H2
VAL216	7,41	-0,02	N	H2
ALA217	7,70	0,01	C	H1
VAL231	7,80	-0,02	HG11	H2
SER193	7,83	-0,02	N	H14
CYS245	7,85	-0,08	SG	H9
THR239	7,91	-0,05	N	H5
GLU188	7,91	0,05	HG3	H14
ALA151	7,96	-0,05	H	H9
CYS253	8,00	-0,04	SG	H9
CYS289	8,22	-0,05	N	H13
GLU294	8,40	0,05	HA	H10

<b>ASP451</b>	8,42	-0,06	OD1	H10
<b>ILE271</b>	8,48	-0,02	HD13	H3
<b>ASP237</b>	8,51	0,11	C	H6
<b>ALA261</b>	8,59	-0,01	HA	H11
<b>LEU154</b>	8,75	-0,04	H	O6
<b>ARG160</b>	8,81	0,07	HH22	O6
<b>GLY189</b>	8,90	-0,02	HA2	H14
<b>LEU198</b>	9,11	-0,02	HG	H12
<b>ALA194</b>	9,11	0,01	C	H14
<b>TYR148</b>	9,14	-0,01	O	H14
<b>ARG197</b>	9,16	0,05	H	H14
<b>LEU275</b>	9,17	-0,05	HD23	H11
<b>LYS212</b>	9,21	0,00	HA	H2
<b>PHE156</b>	9,27	-0,01	HD2	H13
<b>CYS200</b>	9,33	0,00	H	H8
<b>LYS286</b>	9,58	-0,10	O	H13
<b>ASP256</b>	9,61	0,06	O	H6
<b>VAL455</b>	9,63	0,00	HG22	H12
<b>LEU203</b>	9,68	-0,01	HD12	O4
<b>ALA254</b>	9,82	-0,02	HA	H9
<b>CYS246</b>	9,84	-0,01	SG	H8
<b>TYR263</b>	9,86	-0,01	HD2	O3
<b>THR243</b>	9,95	-0,04	H	H6

---

Tabela A10 - Resíduos, distância (Å), energia de interação (kcal.mol<sup>-1</sup>), átomo da HSA, átomo da GLO<sub>mC</sub> (conf. final).

Resíduos	Distância (Å)	Energia de Interação (kcal.mol <sup>-1</sup> )	HSA Átomo	FRC Átomo
ARG257	1,93	-9,75	O	H1
HIS242	2,10	-2,77	HE1	O5
LEU238	2,22	-2,21	HD11	H2
LYS199	2,29	-0,88	HZ3	H5
ALA291	2,32	-3,36	HB1	H10
LEU260	2,48	-3,41	HD12	H3
ILE290	2,75	-2,36	HB	H8
SER287	2,92	-2,09	HG1	H9
ILE264	2,98	-1,10	HD3	H6
ALA261	3,02	-0,85	HB1	H7
VAL241	3,48	-0,76	HG12	H3
ALA258	3,85	-1,83	N	H1
GLU153	4,12	-0,96	OE2	H14
LEU234	4,50	-0,28	HD11	H2
HIS288	4,74	-0,58	HA	H9
LEU219	4,87	-0,34	HD22	O2
ASP256	4,89	0,09	O	H1
ARG222	5,11	-0,27	HH11	H4
GLN196	5,20	-0,21	OE1	H5
LYS286	5,30	-0,20	O	H7
GLU292	5,61	-0,52	HG1	H4
ASP259	5,70	-0,27	N	H1
ARG218	5,93	-0,23	HH22	H4
TYR150	5,93	-0,18	HE2	H14
LYS262	5,93	-0,07	HN	H7
ALA254	5,96	-0,11	O	H1
PHE211	6,12	-0,06	HZ	O5
CYS245	6,13	-0,12	HB2	H13
TYR263	6,21	-0,11	HD2	H6
CYS289	6,27	-0,12	C	H8
TRP214	6,31	-0,09	HZ3	O5
VAL293	6,45	-0,14	HG12	H8
PHE223	6,46	-0,07	HZ	H8
CYS253	6,64	-0,09	O	H1
LEU283	6,75	-0,09	HD12	H7
ASP255	6,85	-0,07	O	H1
LYS195	6,89	-0,05	HB2	H5
LEU275	6,90	-0,03	HD23	H8
LEU284	7,09	-0,08	O	H9
VAL235	7,14	-0,04	HA	H2
CYS265	7,15	-0,04	HN	H7
ASP237	7,21	-0,08	C	H2
THR239	7,37	-0,07	HA	H13
CYS278	7,54	-0,04	HB1	H8
SER192	7,84	-0,01	HB2	H5

<b>GLU285</b>	7,91	-0,06	O	H9
<b>ALA215</b>	7,98	-0,04	HB2	H12
<b>THR243</b>	8,05	-0,03	HN	H13
<b>PRO152</b>	8,08	-0,04	HB1	H14
<b>ILE271</b>	8,41	-0,01	HD2	H6
<b>CYS246</b>	8,43	-0,02	SG	H13
<b>CYS200</b>	8,45	-0,02	HN	O5
<b>LYS240</b>	8,47	-0,10	C	O3
<b>GLU244</b>	8,64	0,04	HB1	H13
<b>PHE149</b>	8,67	-0,01	HE2	H5
<b>CYS279</b>	8,90	-0,01	SG	H7
<b>PHE156</b>	8,99	-0,01	HE2	H9
<b>LEU203</b>	9,41	-0,02	HD22	H13
<b>THR236</b>	9,51	-0,01	N	H2
<b>ARG197</b>	9,61	-0,06	N	H5
<b>GLU266</b>	9,64	0,03	HN	H7
<b>PRO282</b>	9,72	-0,01	O	H7
<b>LYS233</b>	9,90	-0,07	O	H2
<b>SER193</b>	9,95	0,00	HA	H5

---

## APÊNDICE B – DOCKING: CB-DOCK2

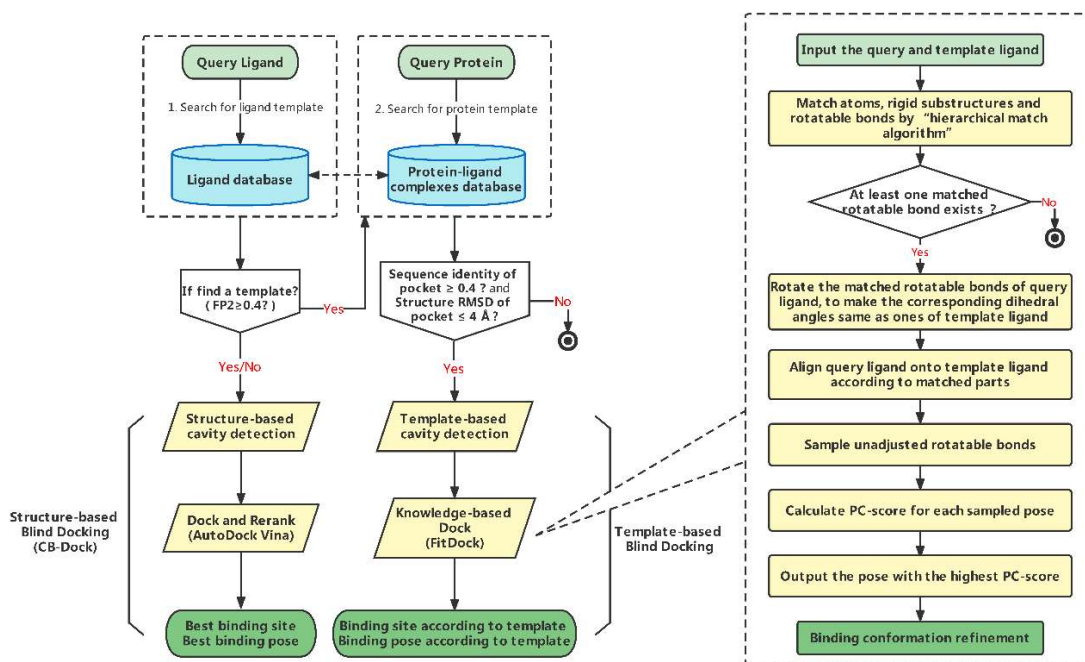
### B. Docking: CB-Dock2

#### B1 Como funciona?

O CB-Dock2 é uma versão melhorada da ferramenta de *docking* cego proteína-ligante que herda o procedimento de detecção de cavidades baseado na curvatura e o procedimento de *docking* molecular baseado no *AutoDock Vina* do servidor CB-Dock. Sobre esta base, foi integrado um procedimento de *docking* cego baseado em modelos homólogos, considerando que a informação sobre as cavidades nos modelos pode fornecer referências valiosas para a previsão de locais de ligação na presença de modelos homólogos. O fluxo de trabalho completo do CB-Dock2 é apresentado a seguir. Para a proteína e o ligante submetidos pelo utilizador, o CB-Dock2 vai buscar à base de dados de complexos proteína-ligante armazenada no servidor dos ligantes modelos com elevada semelhança topológica ( $FP2 \geq 0,4$ ). Se estiver presente, será calculada a semelhança entre a proteína de consulta e as proteínas complexadas com os ligantes modelos selecionados. Os complexos com mais de 40% de identidade de sequência (RMSD de cavidade  $\leq 4\text{\AA}$ ) no local de ligação do ligante modelo serão retidos para a subsequente detecção de cavidade baseada em modelo e *docking* molecular. O método de *docking* utilizado neste procedimento é o *FitDock*, um método desenvolvido que ajusta a conformação inicial ao modelo dado utilizando uma abordagem de alinhamento hierárquico de múltiplos recursos, explora subsequentemente as conformações possíveis e, finalmente, produz poses de *docking* refinada.

Depois de submeter os ficheiros necessários, um *script perl* processa estes ficheiros automaticamente, (Figura B1). É utilizado a versão mais recente (1.2.0) do *AutoDock Vina* para a acoplagem cega independente de modelo. O pipeline de *docking* cego baseado em modelos utiliza a base de dados *BioLip* (versão de 2021.09.15) como base de dados de modelos.

Figura B1 - Fluxo de trabalho do CB-Dock2. *Docking* cego com integração de modelos homólogos



Fonte: Adaptado de Yang Liu, et al. (2022), (LIU *et al.*, 2022), <https://cadd.labshare.cn/cb-dock2/index.php>.

## B2 Submeter trabalhos

Para satisfazer as diferentes exigências, é disponibilizado dois conjuntos de fluxos de tarefas com objetivos diferentes, nomeadamente "*Search Cavities*" e "*Auto Blind Docking*". O primeiro irá realizar a deteção de cavidades de acordo com a proteína e o ligante submetidos, incluindo a deteção de cavidades com base na estrutura e a deteção de cavidades com base no modelo (realizada no caso de existirem modelos homólogos). O último utilizará automaticamente os dois programas acima referidos para efetuar a deteção de cavidades e, em seguida, o docking com base nas cavidades detectadas. Os métodos e notas de submissão são os seguintes:

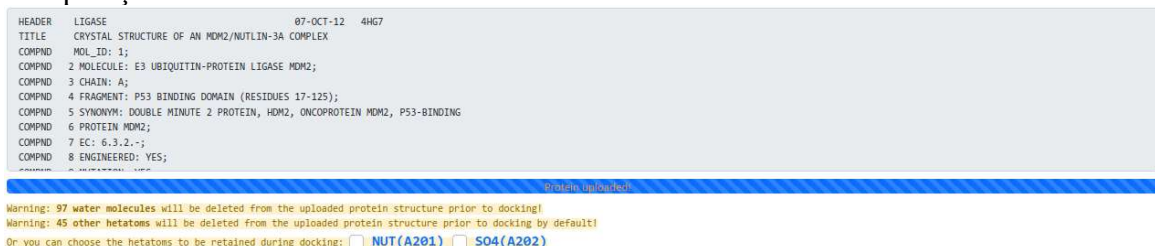
### B2.1 Proteína de consulta

A proteína alvo é carregada em formato PDB. Por predefinição, é removido todos os heteroátomos (incluindo pequenas moléculas, íons e água) da proteína antes do *docking* (denominam-se heteroátomos os elementos químicos diferentes do carbono que conseguem estabelecer ligação química, ao mesmo tempo, com dois carbonos, exemplos de heteroátomos



são o oxigênio - O, enxofre - S, nitrogênio -N e fósforo - P). Para otimizar a preparação da sua proteína para o *docking* molecular no CB-Dock2, é crucial prestar atenção às informações de aviso fornecidas na área de submissão de proteínas (Figura B2). Essas mensagens de aviso podem indicar a presença de heteroátomos na estrutura da proteína submetida.

Figura B2 - Área de submissão de proteínas no CB-Dock2. Avisos importantes para Otimização da Preparação



Fonte: Adaptado de Yang Liu, et al. (2022), (LIU *et al.*, 2022), <https://cadd.labshare.cn/cb-dock2/index.php>.

## B2.2 Consultar ligante

Para iniciar o processo de *docking* molecular no CB-Dock2, você tem a flexibilidade de carregar o seu ligante em diversos formatos de arquivo: MOL2, MOL, SDF ou PDB.

Se optar por usar os formatos MOL ou SDF, o sistema automaticamente converte esses arquivos para o formato PDB utilizando o *Open Babel*, uma ferramenta robusta e amplamente reconhecida para a conversão de formatos químicos. Isso permite que você use arquivos que talvez não estejam originalmente em PDB, simplificando o processo de preparação do ligante.

Além disso, se preferir, você pode desenhar o ligante diretamente no nosso servidor Web. O *JSME Molecule Editor*, uma ferramenta integrada e intuitiva, está disponível para você criar e editar estruturas moleculares de forma manual. Essa opção é especialmente útil se você deseja modificar rapidamente o ligante ou criar uma nova estrutura a partir do zero.

Uma dica crucial para otimizar o desempenho do *AutoDock Vina* durante o *docking* molecular é prestar atenção ao tamanho do arquivo do ligante. Para manter a precisão e a eficiência, recomenda-se que o tamanho do arquivo seja inferior a 15 KB. Isso ocorre porque à medida que o número de átomos e ligações rotativas no ligante aumenta, a complexidade da tarefa de *docking* também cresce. Como resultado, um arquivo maior pode reduzir a eficiência e a precisão do *AutoDock Vina*. Portanto, ao preparar seu ligante, é vantajoso manter a estrutura o mais simplificada possível, sem comprometer as características essenciais necessárias para a

interação com a proteína alvo. Isso ajudará a assegurar que o CB-Dock2 possa realizar o *docking* de maneira eficaz, oferecendo resultados precisos e confiáveis.

### **B2.3 Número de cavidades para detecção**

O CB-Dock2 permite a detecção de cavidades na proteína para prever potenciais locais de ligação. Por padrão, o método de detecção independente do modelo identifica até cinco cavidades principais. No entanto, essa configuração pode ser personalizada de acordo com suas necessidades específicas. Para isso, basta acessar a opção "*More parameters*" no sistema, onde você pode ajustar o número de cavidades a serem detectadas.

O processo de *docking* molecular é realizado para cada uma dessas cavidades detectadas. Após a execução, as cavidades são classificadas com base na melhor pontuação *Vina*, uma métrica que avalia a qualidade do *docking* do ligante na cavidade. Essa classificação ajuda a identificar os locais de ligação mais promissores para o ligante na estrutura da proteína.

É importante ter em mente que, embora seja possível aumentar o número de cavidades a serem consideradas, definir esse número para mais de 20 pode trazer desvantagens. Um número muito elevado de cavidades aumenta significativamente o tempo de execução do programa e pode também elevar a taxa de falsos positivos. Isso ocorre porque mais cavidades implicam em mais tentativas de *docking*, algumas das quais podem não ser relevantes ou biologicamente significativas. Por isso, recomenda-se manter o número de cavidades a um nível que balanceie a profundidade da análise e a eficiência do processamento.

Ao ajustar o número de cavidades, você otimiza o processo para obter resultados precisos e eficientes, focando nos locais de ligação mais relevantes para sua pesquisa. Essa flexibilidade é crucial para personalizar a análise de acordo com as particularidades da proteína em estudo e as hipóteses de interação que você deseja testar.

O número de cavidades detectadas pelo método independente do modelo, e o valor por defeito é 5. Os utilizadores podem clicar em "Mais parâmetros" para personalizar este valor. O programa efetua o *docking* em cada cavidade e classifica as cavidades de acordo com a melhor pontuação *Vina*. O número de cavidades deve ser inferior a 20, porque um número superior aumentará o tempo de execução e a taxa de falsos positivos.

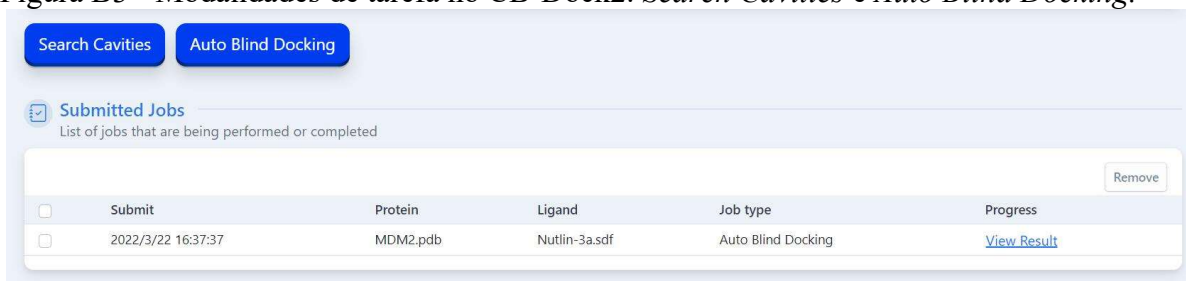
## B2.4 *Tempalte* - uso de modelos e gerenciamento de tarefas no CB-Dock2

O CB-Dock2 oferece uma funcionalidade opcional de carregar complexos proteína-ligante para o docking cego baseado em modelos. Este parâmetro opcional, denominado "*Template*", permite aos usuários aproveitar estruturas pré-existentes de complexos para guiar o processo de docking, aumentando a precisão da previsão dos locais de ligação.

Uma dica importante ao utilizar complexos proteína-ligante como modelos é considerar a natureza biológica da montagem. Se o complexo que você deseja carregar for uma montagem biológica completa, pode ser vantajoso dividi-lo em cadeias individuais antes de carregar (opção "*Split Chains*"). Isso facilita a manipulação do modelo e pode melhorar a eficiência do *docking*.

O CB-Dock2 também oferece duas principais modalidades de tarefa: "*Search Cavities*" e "*Auto Blind Docking*". Os usuários podem submeter uma dessas tarefas de cada vez. Assim que a tarefa é concluída, você pode clicar em "*View Result*" na lista de tarefas para acessar a página de resultados, onde os dados gerados podem ser visualizados ou baixados (Figura B3).

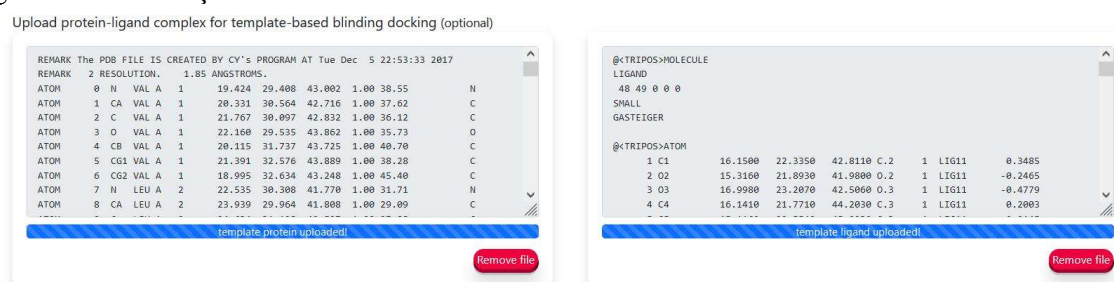
Figura B3 - Modalidades de tarefa no CB-Dock2. *Search Cavities* e *Auto Blind Docking*.



Fonte: Adaptado de Yang Liu, et al. (2022), (LIU *et al.*, 2022), <https://cadd.labshare.cn/cb-dock2/index.php>.

Se desejar, você pode remover os resultados do servidor clicando no botão "*Remover*" na lista de tarefas (Figura B4). Além disso, se você fornecer um endereço de e-mail ao configurar a tarefa, os resultados serão enviados diretamente para sua caixa de entrada, facilitando o acompanhamento e a análise.

Figura B4 - Remoção de resultados no CB-Dock2. Gerenciando tarefas.



Fonte: Adaptado de Yang Liu, et al. (2022), (LIU *et al.*, 2022), <https://cadd.labshare.cn/cb-dock2/index.php>.

Os resultados das tarefas que não forem removidos manualmente são automaticamente excluídos do servidor após um dia. Portanto, é essencial baixar e salvar os resultados das suas tarefas prontamente para evitar a perda de dados valiosos. Essa política de exclusão automática ajuda a manter o servidor limpo e eficiente, garantindo que você sempre tenha espaço disponível para novas análises.

Em resumo, o CB-Dock2 oferece flexibilidade e eficiência na análise de *docking* molecular, desde a utilização de complexos como modelos até a gestão dos resultados das tarefas. Aproveitar essas funcionalidades pode otimizar significativamente suas pesquisas e a gestão de dados no processo de docking.

## B3 Visualizar resultados

O CB-Dock2 oferece diversas ferramentas de visualização na página de resultados para ajudar os usuários a analisar os resultados das tarefas de "*Search Cavities*" e "*Auto BlindDock*". Esta seção facilita a exploração detalhada dos dados calculados, permitindo também o download dos resultados para análises adicionais.

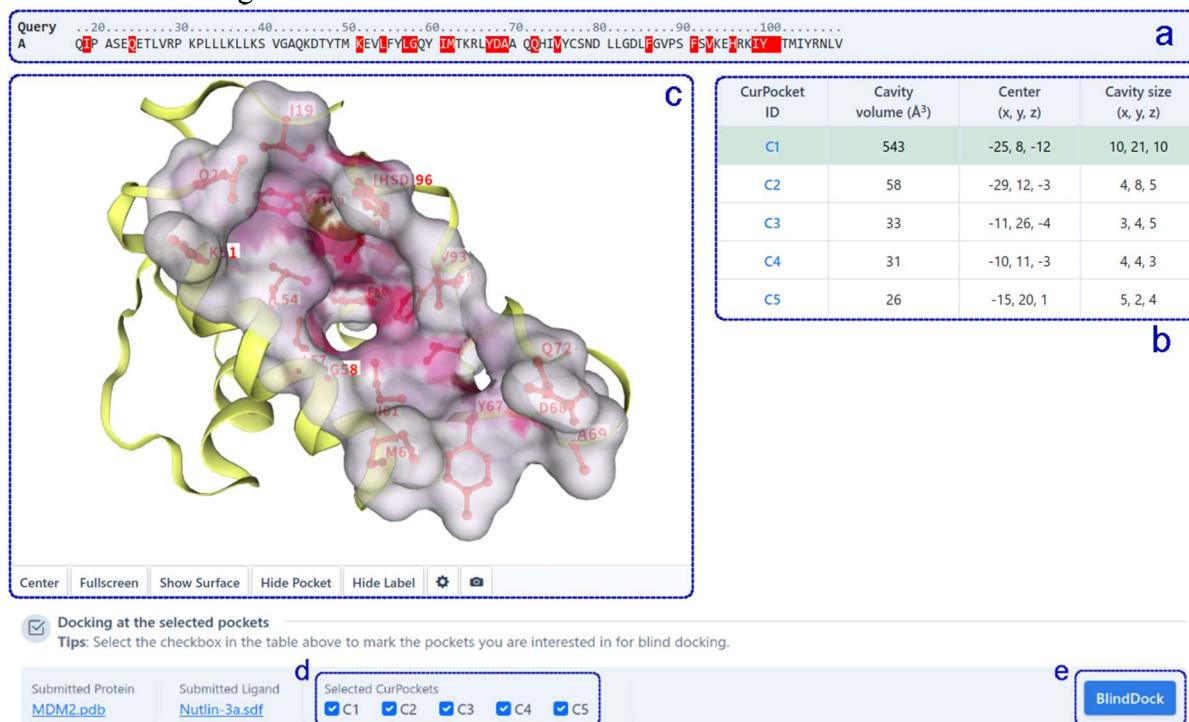
### B3.1 Resultados da pesquisa de cavidades

A interface de resultados da detecção de cavidades está dividida em duas partes, de acordo com o método de detecção: Detecção de cavidades com base na estrutura e detecção de cavidades com base no modelo (realizada no caso de existirem modelos homólogos).

Na seção de detecção de cavidades com base na estrutura, a interface exibe a sequência completa (a) da proteína de consulta, destacando os resíduos presentes em cada cavidade identificada (b). Para explorar mais detalhes, os usuários podem clicar diretamente nos resíduos

listados na sequência para visualizá-los em um modelo 3D interativo (c). As cavidades detectadas são classificadas por volume, proporcionando aos usuários a capacidade de selecionar diferentes cavidades na tabela para uma análise detalhada da estrutura em vista 3D (Figura B5).

Figura B5 - Detecção e análise de cavidades com base na estrutura no CB-Dock2. Exploração estrutural e docking molecular.



Fonte: Adaptado de Yang Liu, et al. (2022), (LIU *et al.*, 2022), <https://cadd.labshare.cn/cb-dock2/index.php>.

Adicionalmente, os usuários podem verificar as cavidades de interesse (d) para o *docking* molecular utilizando o *AutoDock Vina* (e). Isso permite explorar os potenciais modos de ligação do ligante de interesse nas cavidades selecionadas, fornecendo insights cruciais sobre as interações moleculares específicas que estão sendo investigadas (Figura B5).

Na seção de detecção de cavidades com base em modelos, a interface exibe o alinhamento das sequências (a) da proteína de consulta e do modelo, destacando os resíduos dos sítios detectados com base no modelo (b). Os usuários têm a opção de clicar nos resíduos listados na sequência para visualizá-los em um modelo 3D interativo (c) (Figura B6).

Os sítios detectados são inicialmente ordenados pela similaridade FP2 entre o modelo e a proteína de consulta. Em caso de empate na FP2, a ordenação é determinada pela identidade

dos sítios. Sítios provenientes de modelos diferentes, mas compartilhando o mesmo "FitPocket ID" na tabela (b), possuem pelo menos 50% de resíduos do sítio sobrepostos (Figura B6).

Além disso, o servidor realiza automaticamente uma análise de consenso entre os sítios detectadas com base no modelo e aquelas detectadas com base na estrutura. Para explorar os resultados do *docking* molecular, os usuários podem clicar em "Template-base Docking" (d) para visualizar os modos de ligação previstos para o ligante de interesse nos sítios identificados (Figura B6).

Figura B6 - Detecção e análise de cavidades com base em modelos no CB-Dock2. Exploração estrutural e docking molecular.

The screenshot displays the CB-Dock2 interface. At the top, it shows a sequence alignment for protein 4hg7 (A) and a query sequence (A). Below this is a 3D ribbon representation of the protein structure with several pockets highlighted in blue and labeled with residue IDs: HSD196, L54, G58, M62, Y61, HSD173, Q77, and Y68. To the right of the structure is a table of detected FitPockets. Below the structure is a control bar with options like 'Center', 'Fullscreen', 'Show Surface', 'Hide Pocket', 'Show Template(Pro)', 'Show Template(Lig)', and 'Hide Label'. At the bottom, there is a section for 'Docking with homologous protein complex guidance' with fields for 'Submitted Protein' (MDM2.pdb), 'Submitted Ligand' (Nutlin-3a.sdf), and 'Detected FitPockets' (1). A 'Template-based Docking' button is visible on the right.

FitPocket ID	Template ID	Template Protein	Template Ligand	FP2	Pocket Identity	Pocket RMSD	Consensus
F1	t1	4hg7	NA	1.0	1.0	0.0	C1 (65%); C2 (10%)
F1	t2	4j3e	NA	1.0	0.75	0.82	C1 (63%); C2 (11%)
F1	t3	2n14	NA	0.96	0.53	2.27	C1 (42%)
F1	t4	2n06	NA	0.95	0.6	1.72	C1 (57%)
F1	t5	2n0w	NA	0.88	0.55	1.51	C1 (53%); C2 (13%)

Fonte: Adaptado de Yang Liu, et al. (2022), (LIU *et al.*, 2022), <https://cadd.labshare.cn/cb-dock2/index.php>.

### B3.2 Resultados do *Auto BlindDock*

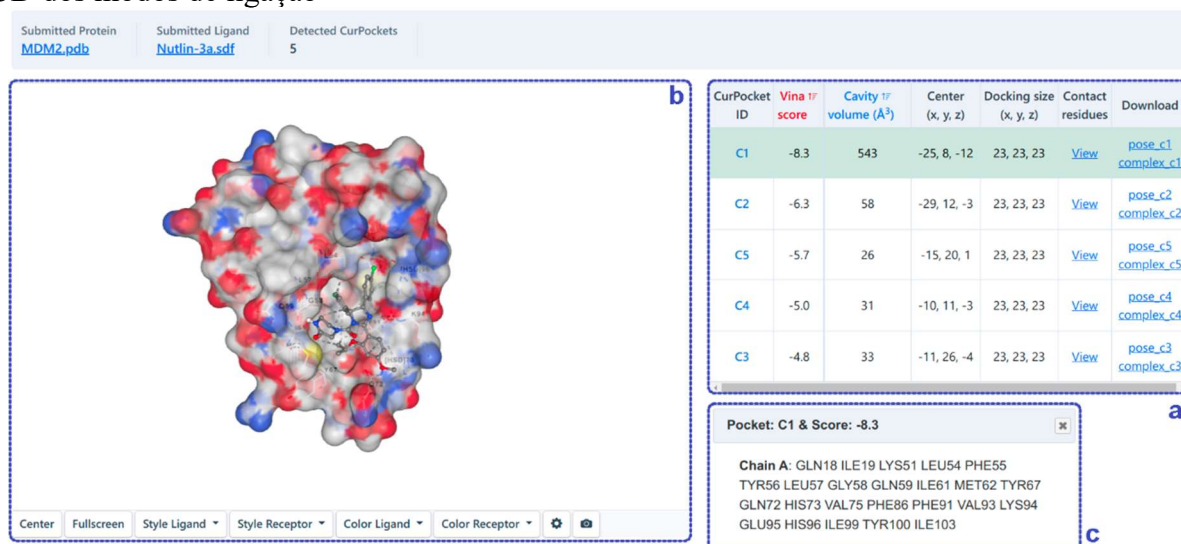
A interface de resultados de *docking* cego automático exhibe os resultados obtidos com base em dois princípios distintos de *docking* cego.

Para o *docking* cego baseada na estrutura, a acoplagem utilizando o *AutoDock Vina* é realizada em cada cavidade identificada. Os resultados são então classificados com base na pontuação *Vina* (kcal/mol), indicando os potenciais locais de ligação para o ligante de interesse. Os usuários têm a opção de visualizar detalhes adicionais clicando no botão "View" na tabela



(a), onde podem examinar os resíduos de contato (c) associados ao melhor modo de ligação em cada sítio. Isso permite uma análise mais profunda, integrando informações estruturais através de uma visualização 3D interativa (b) (Figura B7).

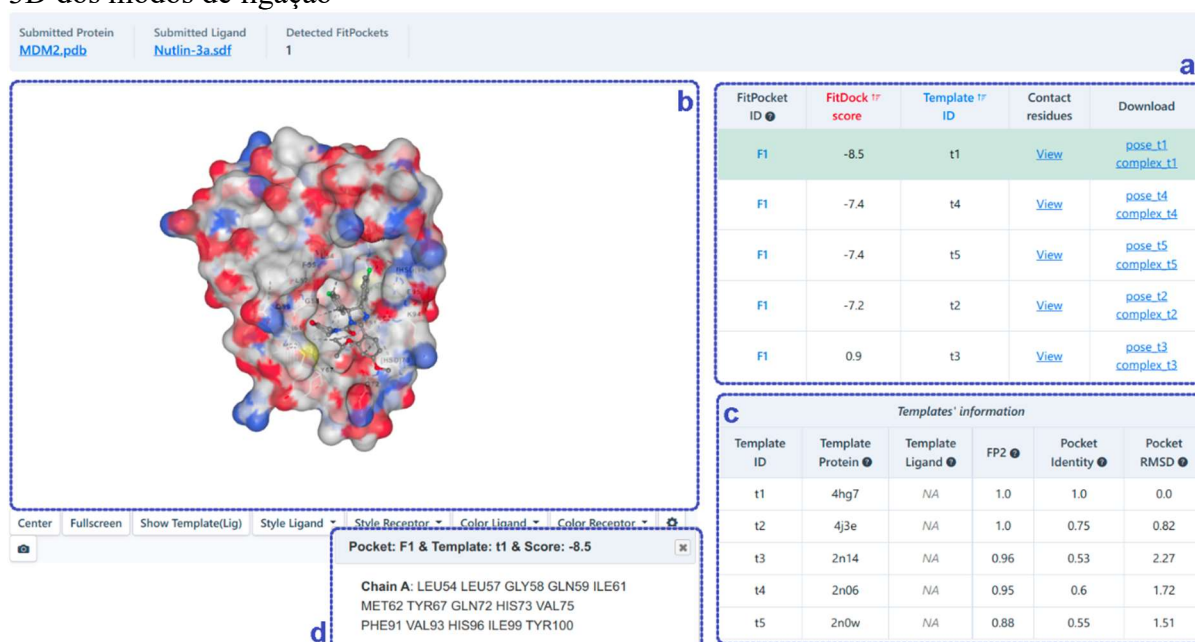
Figura B7 - *Docking* cego baseado na estrutura no CB-Dock2. Análise estrutural e visualização 3D dos modos de ligação



Fonte: Adaptado de Yang Liu, et al. (2022), (LIU *et al.*, 2022), <https://cadd.labshare.cn/cb-dock2/index.php>.

No docking cego baseada em modelos (como mostrado abaixo), utiliza-se um método de docking desenvolvido internamente em cada cavidade identificada. Da mesma forma, os potenciais locais de ligação do ligante são classificados conforme a pontuação de docking (kcal/mol). Modos de ligação com o mesmo "*FitPocket ID*" pertencem a diferentes conformações de ligação na mesma cavidade. Os usuários podem avaliar a melhor conformação de ligação em um local específico com base na pontuação de docking, nos modelos disponíveis e nas informações de contato. Os *templates* (c) e detalhes de contato (d) são acessíveis através da tabela (a) e da visualização 3D (b), clicando nos botões correspondentes (Figura B8).

Figura B8 - *Docking* cego baseado em modelos no CB-Dock2. Análise estrutural e visualização 3D dos modos de ligação



Fonte: Adaptado de Yang Liu, et al. (2022), (LIU *et al.*, 2022), <https://cadd.labshare.cn/cb-dock2/index.php>.

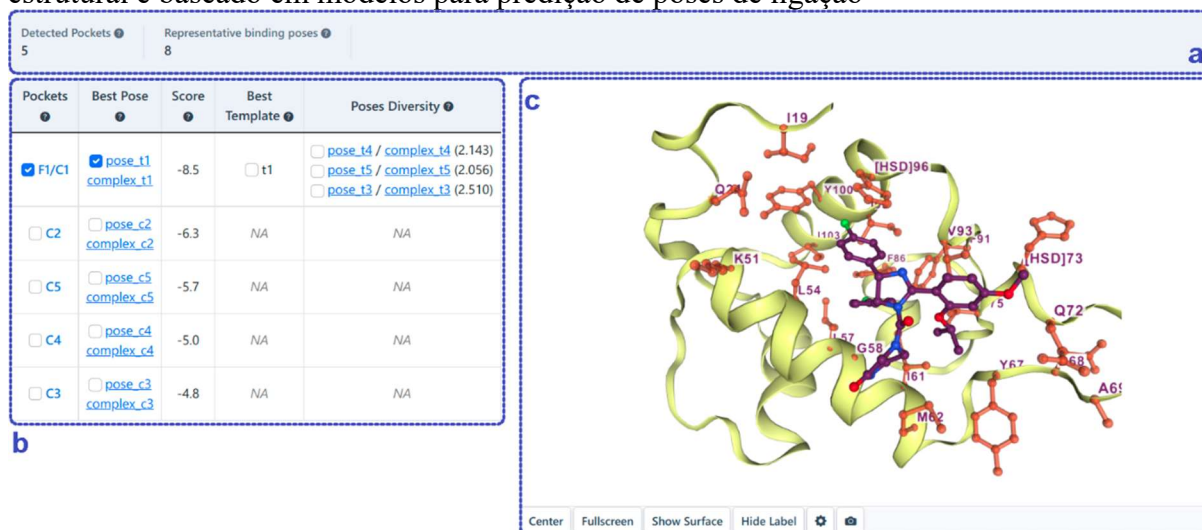
A análise integrativa combinou os resultados dos dois métodos de *docking* cego mencionados anteriormente (Figura B9). As cavidades identificadas pelos dois métodos são agrupadas com base na sobreposição dos resíduos da cavidade (a, b). Tanto os métodos de *docking* cego baseados na estrutura quanto no modelo podem detectar o mesmo local de ligação e prever as poses de ligação do ligante de interesse utilizando *AutoDock Vina* e *FitDock*, respectivamente. Nesse contexto, determinar a pose de ligação mais precisa pode ser uma consideração crucial para os usuários.

Quando os modelos estão disponíveis, o *Root-Mean-Square Deviation* (RMSD) entre as melhores poses obtidas pelos dois métodos de *docking* cego é calculado. Se o RMSD for inferior a 2 Å, a pose com a pontuação mais baixa é selecionada como a melhor previsão; caso contrário, a pose gerada pelo *docking* cego baseado em modelos é preferida como a melhor previsão (a, b). Na ausência de modelos, a pose obtida pelo *docking* cego baseado na estrutura é escolhida como a melhor previsão (a, b).

Os usuários podem visualizar previamente as cavidades integradas e a melhor pose de ligação para cada cavidade utilizando o *NGL-viewer* (c).



Figura B9 - Análise integrativa de docking cego no CB-Dock2. Combinação de métodos estrutural e baseado em modelos para predição de poses de ligação



Fonte: Adaptado de Yang Liu, et al. (2022), (LIU *et al.*, 2022),  
<https://cadd.labshare.cn/cb-dock2/index.php>.

## APÊNDICE C – DINÂMICA MOLECULAR

### C1. Formalismo da Dinâmica Molecular (DM)

O formalismo da dinâmica molecular (DM) é exposto nesta seção, onde uma descrição concisa das equações de movimento utilizadas na DM e apresentamos os métodos para integrar essas equações. Mais elucidações podem ser encontradas nas referências citadas. A simulação clássica de DM serve como uma abordagem de modelagem computacional baseada no movimento de partículas individuais dentro de um sistema especificado. Essas partículas aderem à equação de movimento de Newton, permitindo a determinação da posição e da velocidade de cada partícula em cada etapa da simulação. A equação de Newton para uma partícula solitária é descrita pela equação C1:

$$F_i(t) = m_i \alpha_i \quad \text{Eq. C1}$$

em que  $F_i$  representa a força atuando sobre uma partícula específica do sistema em um determinado momento  $t$ , com  $m_i$  e  $\alpha_i$  denotando a massa e a aceleração da partícula  $i$ , respectivamente.

O potencial de interação entre os átomos no sistema é calculado com base nos parâmetros que constituem o campo de força. Esse potencial é uma função da energia potencial total do sistema,  $V(r)$ , que compreende termos atribuídos a átomos ligados (como alongamento de ligações, deformação angular e torção) e termos atribuídos a átomos não ligados (incluindo interações de van der Waals e Coulomb). Um campo de força típico é caracterizado pela equação C2, a seguir:

$$V(r) = \sum V_{lig.}(r) + \sum V_{angular}(\theta) + \sum V_{diédro}(\phi) + \sum V_{vdw} + \sum V_{elet.} \quad \text{Eq. C2}$$

Ao utilizar aproximações de potencial harmônico, as expressões relativas ao alongamento da ligação e à deformação angular são apresentadas nas equações subsequentes,

$$\sum V_{lig.}(r) = \sum_{lig.} \frac{1}{2} K_r (r - r_0)^2 \quad \text{Eq. C3}$$

$$\sum V_{angular}(\theta) = \sum_{angular} \frac{1}{2} K_\theta (\theta - \theta_0)^2 \quad \text{Eq. C4}$$

onde  $r$  e  $\theta$  correspondem aos comprimentos e ângulos da ligação em torno das posições de equilíbrio,  $r_0$  e  $\theta_0$ . As constantes  $K_r$  e  $K_\theta$  representam as forças restauradoras do sistema. Uma representação comum para o potencial de torção é fornecida na equação C5,

$$\sum V_{diédrico}(\phi) = \frac{V_n}{2} (1 + \cos(n\phi - \gamma)) \quad \text{Eq. C5}$$

onde  $V_n$ ,  $n$ ,  $\phi$  e  $\gamma$  correspondem sequencialmente à barreira de energia para torção, ao número máximo ou mínimo de energia torcional, ao ângulo diédrico e ao ângulo de fase. As aproximações harmônicas também podem ser integradas em alguns campos de força para definir o potencial de torção com precisão.

As interações entre moléculas, envolvendo átomos não ligados, são delineadas pelos termos de van der Waals e eletrostáticos, normalmente representados pelos potenciais de Lennard-Jones e Coulomb, conforme mostrado nas equações C6 e C7:

$$V_{vaw} = 4\varepsilon_{ij} \left[ \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left( \frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] \quad \text{Eq. C6}$$

$$V_{elet.} = \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}} \quad \text{Eq. C7}$$

Na equação C6,  $\varepsilon_{ij}$ ,  $\sigma_{ij}$  e  $r_{ij}$  correspondem à profundidade do potencial, largura do poço de potencial e à separação entre os átomos  $i$  e  $j$ , respectivamente. Enquanto as partes repulsivas e atrativas são representadas pelo primeiro e segundo termos, respectivamente. A equação C7 envolve as cargas parciais dos átomos  $i$  e  $j$  denotadas por  $q_i$  e  $q_j$ , com  $\varepsilon_0$  representando a permissividade do meio e  $r_{ij}$  tendo a mesma definição da equação C6.

Inúmeros campos de força surgiram nos últimos anos, sendo os mais proeminentes para proteínas CHARMM, GROMOS, AMBER, OPLS, entre outros.

Uma vez definido o campo de força, as forças que atuam em cada átomo são calculadas avaliando a primeira derivada da energia potencial em relação às posições dos átomos, conforme indicado na equação C8 abaixo:

$$F_i(t) = -\frac{\partial V(r_i)}{\partial r_i} \quad \text{Eq. C8}$$

Ao utilizar a equação C8, a aceleração da partícula pode ser encontrada, por meio da integração das equações de movimento, as velocidades e as novas posições dos átomos podem ser determinadas. Posteriormente, a energia potencial e cinética do sistema são avaliadas com base nessas posições e velocidades atualizadas. Por meio da aplicação iterativa dessa metodologia, uma série de posições e velocidades evolui ao longo do tempo, formando uma trajetória. Existem vários algoritmos para a integração de equações de movimento, com o algoritmo Verlet descrito em DM sendo um método comumente empregado em DM, conforme apresentado na equação C9:

$$r(t + \Delta t) = r(t) + v(t)\Delta t + \frac{1}{2} \alpha(t)\Delta t^2 + \dots \quad \text{Eq. C9}$$

A equação C9 é obtida considerando as posições da partícula no tempo  $t$ , " $r(t)$ " e a posição da partícula no tempo  $t + \Delta t$ , " $r(t + \Delta t)$ " são obtidos por meio da utilização da expansão da série de Taylor nas direções para frente e para trás, conforme demonstrado nas equações C10 e C11, respectivamente:

$$r(t + \Delta t) = r(t) + v(t)\Delta t + \frac{1}{2} \alpha(t)\Delta t^2 + \dots \quad \text{Eq. C10}$$

$$r(t - \Delta t) = r(t) - v(t)\Delta t + \frac{1}{2} \alpha(t)\Delta t^2 - \dots \quad \text{Eq. C11}$$

A equação C9 é obtida por meio da soma das equações C10 e C11, seguida pelo isolamento de  $r(t + \Delta t)$ .

## APÊNDICE D – TEORIA DO FUNCIONAL DA DENSIDADE (DFT)

### D1. Teoria do Funcional da Densidade (DFT)

A Teoria do Funcional da Densidade (DFT) serve como um pilar fundamental no campo dos cálculos de estruturas eletrônicas para sólidos e moléculas. Atualmente, uma grande variedade de problemas abordados anteriormente usando técnicas *ab initio* Hartree-Fock e pós-Hartree-Fock são resolvidos de forma eficaz por meio da utilização do DFT. Essa teoria se mostra particularmente vantajosa para sistemas moderados a grandes, com mais de vinte átomos, pois produz resultados precisos com despesas computacionais significativamente menores em comparação com metodologias correlacionadas convencionais, como teoria de perturbação e agrupamento acoplado. A progressão de algoritmos de integração numérica eficientes junto com funcionais de troca e correlação desempenha um papel fundamental no avanço da teoria e na facilitação da aplicação do DFT em vários sistemas. Além disso, a incorporação do DFT em pacotes de software comerciais e de código aberto aumentou significativamente sua acessibilidade (MORGON; COUTINHO, 2007).

De acordo com os princípios do DFT, a energia de um sistema eletrônico está ligada à densidade de probabilidade eletrônica  $\rho$ . Dentro de um sistema composto por  $n$  elétrons,  $\rho(r)$  representa a densidade eletrônica cumulativa em um ponto espacial específico  $r$ , com a energia  $E$  sendo um funcional da densidade eletrônica  $E(\rho)$  (SOUSA; FERNANDES; RAMOS, 2007). A vantagem distinta de empregar DFT sobre metodologias baseadas em funções de onda está em sua carga computacional reduzida (NELSON H. MORGON; CUSTODIO R., 1994; SOUSA; FERNANDES; RAMOS, 2007). Por exemplo, em um sistema de  $n$ -elétrons, a função de onda de cada elétron exige a consideração de três coordenadas espaciais, além de uma quarta dimensão que contabiliza o *spin* (se aplicável), totalizando  $4n$  coordenadas. Por outro lado, a densidade eletrônica depende apenas das três coordenadas espaciais, independentemente da contagem de elétrons dentro do sistema. Conseqüentemente, os sistemas que empregam metodologias de função de onda experimentam uma complexidade crescente com o aumento do número de elétrons, enquanto a densidade eletrônica retém um número consistente de variáveis (SOUSA; FERNANDES; RAMOS, 2007). É importante destacar que o DFT, por meio de sua estrutura formal, elucida com sucesso os principais princípios químicos que foram previamente explorados empiricamente, como potencial químico, dureza de Pearson, eletronegatividade, entre outros (MORGON; COUTINHO, 2007).

A utilização da densidade eletrônica  $\rho(r)$  para explicar um sistema eletrônico teve origem na pesquisa conduzida por Drude (1900), que empregou a teoria cinética dos gases para analisar um metal como um gás eletrônico homogêneo. O objetivo de Drude era formular uma teoria sobre a condução térmica e elétrica. Desse ponto em diante, uma variedade de metodologias foi estabelecida por pesquisadores como Sommerfeld, Thomas, Fermi e Dirac (DUARTE, 2001; MORGON; COUTINHO, 2007).

Sousa, Fernandes e Ramos (2007) afirmam em seu estudo que o conceito de densidade funcional surgiu no final de 1927, ainda que implicitamente, nos trabalhos de Fermi e Thomas (FERMI, 1927; THOMAS, 1927). Esses estudiosos introduziram a noção de expressar a energia do sistema eletrônico em função da densidade eletrônica total. Em relação ao modelo Thomas-Fermi, Duarte e Rocha (2007) afirmam que os autores utilizaram o raciocínio estatístico para aproximar a distribuição de um gás eletrônico e elaborar o funcionamento da energia. Posteriormente, o modelo Thomas-Fermi sofreu modificações com a incorporação da energia de troca para um gás eletrônico por Dirac em 1930 (modelo Thomas-Fermi-Dirac) (DIRAC, 1930). Slater, em 1951, aplicou uma abordagem semelhante à do método Slater-Hartree-Fock ao modelo Thomas-Fermi. Esse método foi percebido como uma aproximação da técnica de Hartree-Fock, mas atualmente é reconhecido como o precursor do DFT (SOUSA; FERNANDES; RAMOS, 2007).

Sousa, Fernandes e Ramos (2007) sugerem que os métodos de Fermi, Thomas e Slater-Hartree-Fock facilitaram a associação de energia e outras características do sistema com a densidade eletrônica, embora com certas restrições (HOHENBERG; KOHN, 1964). Foi somente em 1964 que Hohenberg e Kohn demonstraram que a energia do estado fundamental de um sistema eletrônico não degenerado e sua propriedade eletrônica correspondente são determinadas exclusivamente pela densidade eletrônica. Eles destacaram que a densidade eletrônica determina o potencial externo dos elétrons e a contagem de elétrons, influenciando posteriormente o hamiltoniano e, conseqüentemente, a energia (MORGON; COUTINHO, 2007; SOUSA; FERNANDES; RAMOS, 2007). Os princípios da teoria introduzida por Hohenberg e Kohn foram concebidos para estados não degenerados de um gás eletrônico não homogêneo (NELSON H. MORGON; CUSTODIO R., 1994).

Embora o teorema de Hohenberg-Kohn tenha validado a presença de uma energia de ligação funcional à densidade eletrônica de um sistema, a natureza exata desse funcional permaneceu desconhecida. Desde então, vários funcionais foram desenvolvidos para esse propósito. Kohn (1965) e Sham (1966) refinaram o modelo Hohenberg-Kohn formulando

equações auto consistentes e incorporando efeitos de troca e correlação (KOHN; SHAM, 1965; SHAM; KOHN, 1966).

### D1.1 O Formalismo de Kohn-Sham

O referencial teórico estabelecido por Kohn e Sham foi crucial nos estágios iniciais da Teoria do Funcional da Densidade (DFT). Um obstáculo importante encontrado foi a identificação de um funcional capaz de capturar com precisão a energia cinética do sistema. Kohn e Sham propuseram um método para resolver esse problema dividindo o funcional de energia cinética em dois componentes. Um componente pode ser calculado com precisão sob a suposição de elétrons não interagentes, enquanto a outra parte serve como uma pequena correção para incorporar as interações elétron-elétron. Dentro dessa estrutura, a energia eletrônica de um sistema composto por  $n$  elétrons e  $N$  núcleos pode ser definida por meio da equação D1.

$$E(r) = -\frac{1}{2} \sum_{i=1}^n \int \Psi_i^*(r_1) \nabla_i^2 \Psi_i(r_1) dr_1 - \sum_{X=1}^N \int \frac{Z_X}{r_{Xi}} \rho(r_1) dr_1 + \frac{1}{2} \iint \frac{\rho(r_1)\rho(r_2)}{r_{12}} dr_1 dr_2 + E^{XC}(\rho) \quad \text{Eq. D1}$$

A equação D1 introduz os orbitais de Kohn-Sham  $\Psi_i = (i = 1, 2, \dots, n)$ , com o termo inicial correspondendo à energia cinética eletrônica na ausência de interações. O termo subsequente explica as interações entre núcleos e elétrons, seguido pela representação da repulsão coulombiana entre as distribuições de carga em  $r_1$  e  $r_2$ . O termo final, o funcional de troca e correlação, desempenha um papel crítico no ajuste da energia cinética para refletir as interações elétron-elétron, que são negligenciadas no primeiro termo, juntamente com as correções não clássicas da energia de repulsão elétron-elétron. A descrição precisa dos sistemas depende da capacidade da teoria DFT de enfrentar os desafios impostos pelo funcional de troca e correlação.

A densidade eletrônica do estado fundamental  $\rho(r)$  na posição  $r$  pode ser descrita por um conjunto de orbitais de um elétron, conhecidos como orbitais de Kohn-Sham (Equação D2). Esses orbitais específicos são obtidos por meio da solução da equação de Kohn-Sham (Equação D3) usando o princípio variacional para determinar a energia eletrônica  $E(\rho)$ , considerando a densidade fornecida na equação D2. A importância dos orbitais de Kohn-Sham está em sua capacidade de facilitar o cálculo da densidade eletrônica com base na Equação D2 (SOUSA; FERNANDES; RAMOS, 2007).

$$\rho(r) = \sum_{i=1}^n |\Psi_i(r)|^2 \quad \text{Eq. D2}$$

$$\hat{h}_i \Psi_i(r_1) = \varepsilon_i \Psi_i(r_1) \quad \text{Eq. D3}$$

Dentro da estrutura da equação de Kohn-Sham (Equação D3), o símbolo  $\hat{h}_i$  denota o hamiltoniano e  $\varepsilon_i$  representa a energia ligada ao orbital de Kohn-Sham. O Hamiltoniano de Kohn-Sham  $\hat{h}_i$  é formulado conforme mostrado na equação D4, incorporando  $V^{XC}$  como o funcional derivado da energia de troca e correlação ( $E^{XC}$  na Equação D1), conforme definido na equação D5. Se o valor de  $E^{XC}$  for conhecido, a determinação de  $V^{XC}$  será acelerada.

A energia de troca e correlação é normalmente segregada em dois componentes: energia de troca decorrente de interações elétron-elétron de spin idêntico e energia de correlação decorrente de interações entre elétrons de spins opostos (equação D6). Consequentemente, cada segmento de energia dá origem à função eletrônica de troca e correlação de densidade correspondente, conforme descrito na equação D5. Notavelmente, os funcionais de troca e correlação representam uma fonte significativa de imprecisões nos cálculos da Teoria do Funcional da Densidade (DFT) (SOUSA; FERNANDES; RAMOS, 2007).

$$\hat{h}_i = -\frac{1}{2} \nabla_i^2 - \sum_{X=1}^N \frac{Z_X}{r_{Xi}} + \int \frac{\rho(r_2)}{r_{12}} dr_2 + V^{XC}(r_1) \quad \text{Eq. D4}$$

$$V^{XC}(\rho) = \frac{\delta E^{XC}(\rho)}{\delta r} \quad \text{Eq. D5}$$

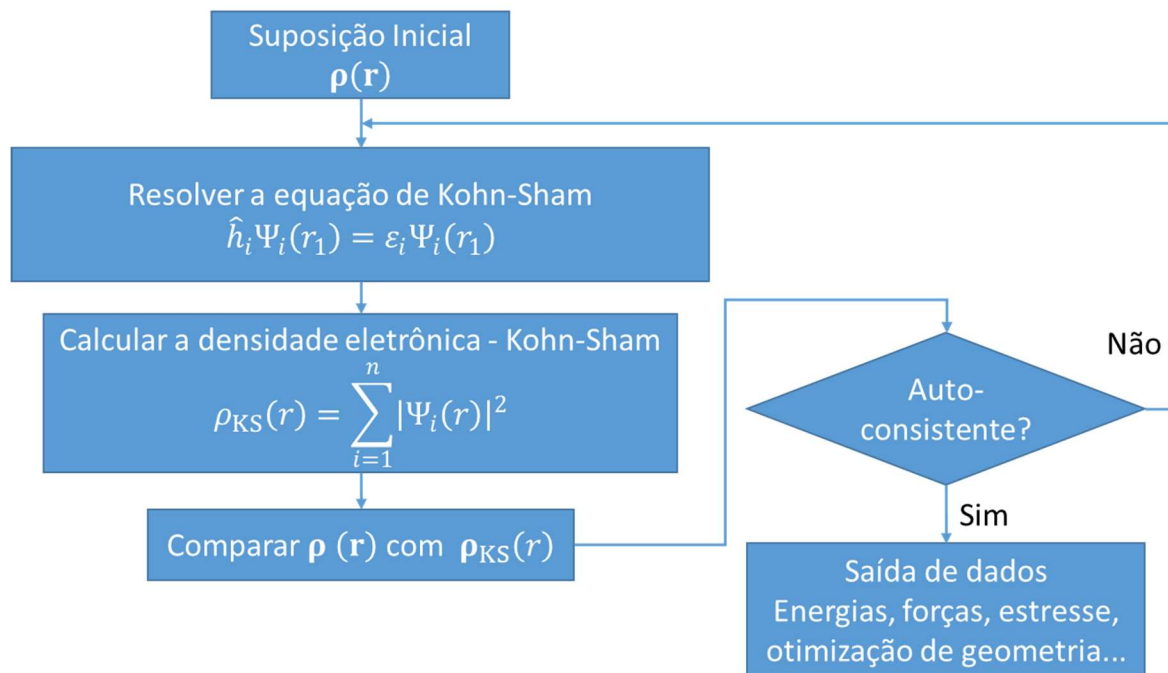
$$E^{XC}(\rho) = E^X(\rho) + E^C(\rho) \quad \text{Eq. D6}$$

Ao abordar a equação de Kohn-Sham (equação D3), é prática comum fazer uma suposição sobre a densidade eletrônica inicial  $\rho$ . Normalmente, isso envolve a utilização da densidade alcançada pela combinação das densidades dos átomos individuais que compõem o sistema em questão. Posteriormente, um funcional aproximado denotado como  $V^{XC}$ , que captura a relação entre a energia de troca, a correlação e a densidade eletrônica, é selecionado e mantido como uma constante na computação preliminar (teste inicial). A solução para a equação de Kohn-Sham é então determinada, produzindo uma série de orbitais de Kohn-Sham que são posteriormente empregados para calcular uma densidade mais precisa utilizando a equação D2. Essa densidade eletrônica atualizada é inserida na equação de Kohn-Sham (tentativa subsequente), iniciando uma repetição do procedimento mencionado acima que culmina na criação de uma nova densidade. Esse processo iterativo continua até que a densidade eletrônica e a energia de troca, junto com a correlação, estejam em conformidade com o critério



de convergência predeterminado. A convergência é realizada quando o desvio no valor calculado de um determinado parâmetro (densidade, energia de troca e correlação) entre as duas últimas etapas atinge o limite especificado. Finalmente, a energia eletrônica é derivável aplicando a equação D1 (SOUSA; FERNANDES; RAMOS, 2007), conforme a representação esquemática do loop autoconsistente para obtenção da solução de Kohn-Sham apresentado na figura D1 abaixo.

Figura D1 - Fluxograma do Processo Autoconsistente. Solução das Equações de Kohn-Sham em DFT.



## D1.2 Funcionais de Troca e Correlação

Os funcionais de troca  $V^X$  e correlação  $V^E$  desempenham papéis essenciais na Teoria do Funcional da Densidade (DFT), permitindo a aproximação da energia de um sistema em função da densidade eletrônica. Esses funcionais podem ser construídos considerando diferentes níveis de detalhe na densidade eletrônica. Na Aproximação da Densidade Local (LDA - *Local Density Approximation*), o funcional é determinado com base na densidade eletrônica  $\rho(r)$  apenas no ponto específico  $r$ . Essa abordagem assume que a densidade eletrônica varia lentamente no espaço, permitindo a utilização de informações locais para estimar a energia de troca e correlação.

No entanto, para sistemas onde a densidade eletrônica varia mais rapidamente, a Aproximação do Gradiente Generalizado (GGA - *Generalized Gradient Approximation*) se mostra mais precisa. Essa aproximação incorpora não apenas a densidade eletrônica local  $\rho(\mathbf{r})$ , mas também o gradiente da densidade  $\nabla\rho(\mathbf{r})$ . Ao incluir essa variação espacial da densidade, a GGA proporciona uma descrição mais detalhada das interações de troca e correlação em sistemas complexos.

Portanto, a escolha entre a LDA e a GGA depende da natureza do sistema estudado e do grau de precisão requerido para modelar suas propriedades eletrônicas.

### D1.2.1 Aproximação da Densidade Local (LDA - *Local Density Approximation*)

A Aproximação de Densidade Local (LDA) considera a função de troca e correlação ao ver a energia de troca e correlação em um ponto específico no espaço como exclusivamente dependente da densidade naquele ponto, que se alinha com o modelo de gás eletrônico homogêneo. A proposta de Dirac de 1930 da energia de troca  $E^X$  baseada na densidade local (Eq. D7), aplicada ao modelo Thomas-Fermi (FERMI, 1927; THOMAS, 1927), foi posteriormente desenvolvida por Slater para abranger a dependência funcional do spin (Eq. D8), resultando na Aproximação da Densidade de Spin Local (*Local Spin Density Approximation* - LSDA).

$$E_{LDA}^{X,Dirac}(\rho) = -C_X \int \rho^{4/3}(\mathbf{r}) d\mathbf{r}, \quad \text{onde } C_X = -\frac{3}{4} \left(\frac{3}{\pi}\right)^{1/3} \quad \text{Eq. D7}$$

$$E_{LDA}^X(\rho) = -2^{1/3} C_X \int (\rho_\alpha^{4/3} + \rho_\beta^{4/3})(\mathbf{r}) d\mathbf{r} \quad \text{Eq. D8}$$

Os símbolos  $\alpha$  e  $\beta$  na equação D8 representam o movimento do spin do elétron, com o funcional LSDA sendo equivalente ao funcional LDA em sistemas de camada fechada onde elétrons e spin são idênticos. A energia de correlação  $E^C$  é normalmente obtida junto com a energia de troca devido ao desafio de isolá-la, usando valores calculados para várias densidades no gás eletrônico homogêneo e fórmulas de interpolação como a desenvolvida por Vosko, Wilk e Nussair (VWN) e a fórmula de Perdew (PL) (PERDEW; ZUNGER, 1981; VOSKO; WILK; NUSAIR, 1980). Embora o LDA funcional frequentemente subestime as energias atômicas e de ionização do estado fundamental e superestime as energias de ligação, ele funciona melhor em sistemas maiores e onde a densidade eletrônica permanece relativamente constante. Apesar dos erros do LDA, resultados satisfatórios ainda são alcançáveis para sistemas com densidade

eletrônica flutuante, pois os erros tendem a se compensar - com a subestimação da energia de troca balanceada pela superestimação da energia de correlação (SOUSA; FERNANDES; RAMOS, 2007).

### D1.2.2 Aproximação do Gradiente Generalizado (GGA - *Generalized Gradient Approximation*)

A prosperidade da metodologia LDA resultou no surgimento de aproximações funcionais para troca e correlação, conhecidas como Aproximações de Gradiente Generalizado (GGA), que visam incorporar a variabilidade na densidade eletrônica. Esses GGAs aumentam a precisão da aproximação do LDA em vários cenários moleculares. A parametrização de um funcional GGA é notavelmente mais complexa do que a de um funcional LDA, abrangendo termos relacionados a gradientes eletrônicos de densidade. As formas comumente utilizadas para troca de energia em funcionais GGA incluem aquelas propostas por Becke (B88), Perdew e Wang (PW91) e Perdew-Burke-Enzerhof (PBE). Por outro lado, a representação matemática da energia de correlação é mais complexa, com seu impacto na energia normalmente inferior ao da energia de troca. Perdew e Wang, junto com Perdew-Burke-Enzerhof, criaram expressões amplamente adotadas para energia de correlação, enquanto os químicos frequentemente optam pelo funcional de correlação de Lee-Yang-Parr (LYP).

Os funcionais híbridos amalgamam energias de troca Hartree-Fock explícitas com uma troca DFT “puro” e um funcional de correlação, exibindo maior precisão nos cálculos da lacuna de energia em comparação com as aproximações de LDA e GGA. Os funcionais híbridos são predominantes na química quântica, com Becke propondo uma fusão da energia de troca de Hartree-Fock com as energias de troca e correlação de LDA e GGA, resultando em duas parametrizações: B3PW91 e B3LYP. A energia de troca e correlação em um funcional híbrido B3C é expressa da seguinte forma:

$$\varepsilon_{XC} = \varepsilon_{XC}^{LDA} + a_0(\varepsilon_X^{HF} - \varepsilon_X^{Funcional\ LDA\ ou\ GGA}) + a_X \varepsilon_X^{Becke} + a_C \varepsilon_C \quad \text{Eq. D8}$$

A denominação B3C se origina da inclusão da energia de troca de Becke (B), três (3) parâmetros empiricamente calibrados ( $a_0, a_X, a_C$ ) e um funcional de correlação DFT (C). A resolução computacional das equações de Kohn-Sham, semelhante aos métodos Hartree e Hartree-Fock, envolve a utilização de bases para representar os orbitais de Kohn-Sham. Pseudopotenciais ou potenciais efetivos são comumente empregados para representar os elétrons mais internos dos átomos, imitando os atributos de espalhamento do núcleo atômico

percebidos pelos elétrons de valência, diminuindo assim a carga computacional dos cálculos. Devido à forma precisa desconhecida do funcional de troca e correlação, bem como à utilização eficaz de funcionais híbridos calibrados empiricamente, há um debate contínuo sobre se as implementações da teoria do funcional da densidade devem ser consideradas estritamente como cálculos do primeiro princípio.

**ANEXO A – ARTIGO PUBLICADO**

É apresentado a seguir o artigo publicado pelo autor relacionado a tese.

Hadad A, França VLB, Crisostomo MW, Brunaldi K, Carvalho HF, Freire VN. **Unveiling fructose and glucose binding to human serum albumin: fluorescence measurements and docking, molecular dynamics and quantum biochemistry computations.** J Biomol Struct Dyn. 2024 Jan 30:1-21. doi: 10.1080/07391102.2024.2310211.

# Unveiling fructose and glucose binding to human serum albumin: fluorescence measurements and docking, molecular dynamics and quantum biochemistry computations


André Hadad, Victor L. B. França, Marcos William Crisostomo, Kellen Brunaldi, Hernandes F. Carvalho & Valder N. Freire

To cite this article: André Hadad, Victor L. B. França, Marcos William Crisostomo, Kellen Brunaldi, Hernandes F. Carvalho & Valder N. Freire (30 Jan 2024): Unveiling fructose and glucose binding to human serum albumin: fluorescence measurements and docking, molecular dynamics and quantum biochemistry computations, Journal of Biomolecular Structure and Dynamics, DOI: [10.1080/07391102.2024.2310211](https://doi.org/10.1080/07391102.2024.2310211)

To link to this article: <https://doi.org/10.1080/07391102.2024.2310211>

 View supplementary material [↗](#)


---

 Published online: 30 Jan 2024.

---

 Submit your article to this journal [↗](#)

---

 View related articles [↗](#)


---

 View Crossmark data [↗](#)

---



# Unveiling fructose and glucose binding to human serum albumin: fluorescence measurements and docking, molecular dynamics and quantum biochemistry computations

André Hadad<sup>a</sup> , Victor L. B. França<sup>a,b</sup>, Marcos William Crisostomo<sup>c</sup>, Kellen Brunaldi<sup>c</sup>, Hernandes F. Carvalho<sup>d</sup> and Valder N. Freire<sup>a</sup>

<sup>a</sup>Department of Physics, Federal University of Ceará, Fortaleza, Ceará, Brazil; <sup>b</sup>Department of Physiology and Pharmacology, Faculty of Medicine, Federal University of Ceará, Fortaleza, Ceará, Brazil; <sup>c</sup>Department of Physiological Sciences, State University of Maringá, Maringá, Paraná, Brazil; <sup>d</sup>Department of Structural and Functional Biology, Institute of Biology, State University of Campinas, Campinas, São Paulo, Brazil

Communicated by Ramaswamy H. Sarma

## ABSTRACT

This research examines the interaction between human serum albumin (HSA) and various sugar forms ( $\beta$ -D-fructofuranose (FRC),  $\alpha$ -D-glucopyranose (GLC), Keto-D-fructose (FRO), Aldehydo-D-glucose (GLO), and modified Aldehydo-D-glucose (GLOm)) using fluorescent spectroscopy, molecular docking simulations, molecular dynamics, protein conformational clusters (EnGens), molecular fractionation with conjugate caps (MFCC) and quantum biochemistry analysis. We analyze molecular and quantum aspects, uncovering interaction energies between sugar atoms and amino acids. Total interaction energy considers protein fragmentation, energetic decomposition, and interaction energy from a bottom-up perspective. Molecular dynamics reveal that unmodified Aldehydo-D-glucose (GLO) escapes HSA binding sites, explaining gradual glycation. We pioneer studying HSA's binding mechanism with glucose and fructose in a 1:1 ratio using long molecular dynamics simulations. Results suggest the transitional GLOm form has a higher Sudlow I site propensity than unmodified glucose, crucial for K195 glycation. FRO and GLOm interaction tendencies move toward a deeper FA7 cavity, near its center. This approach effectively elucidates small molecule binding mechanisms, consistent with previous experimental results.

## ARTICLE HISTORY

Received 29 October 2023  
Accepted 19 January 2024



## 1. Introduction


Human serum albumin (HSA), which constitutes up to 60% of the total serum proteins, fulfills a pivotal role in facilitating the transportation of both endogenous and exogenous ligands through reversible non-covalent interactions (Carter & Ho, 1994; Peters, 1995). The significance of HSA lies in its remarkable binding capacity, which allows it to interact with various drugs such as warfarin, ibuprofen, and indomethacin (Anguizola et al., 2013; Petitpas et al., 2001). It is important to note that changes in the pathological or physiological environment can potentially induce alterations in the conformation of a protein, consequently affecting the binding efficiency of ligands and leading to functional impairment of HSA (Awang et al., 2017; Pongprayoon & Gleeson, 2014). Among the multiple impairments found in HSA, the presence of glycation has emerged as a significant factor, as outlined by Jana et al. in 2016.

With a remarkable arrangement of 585 amino acids distributed across three homologous helical domains (I-III), each of which comprises two subdomains (A and B), HSA serves as an indispensable component in various crucial processes

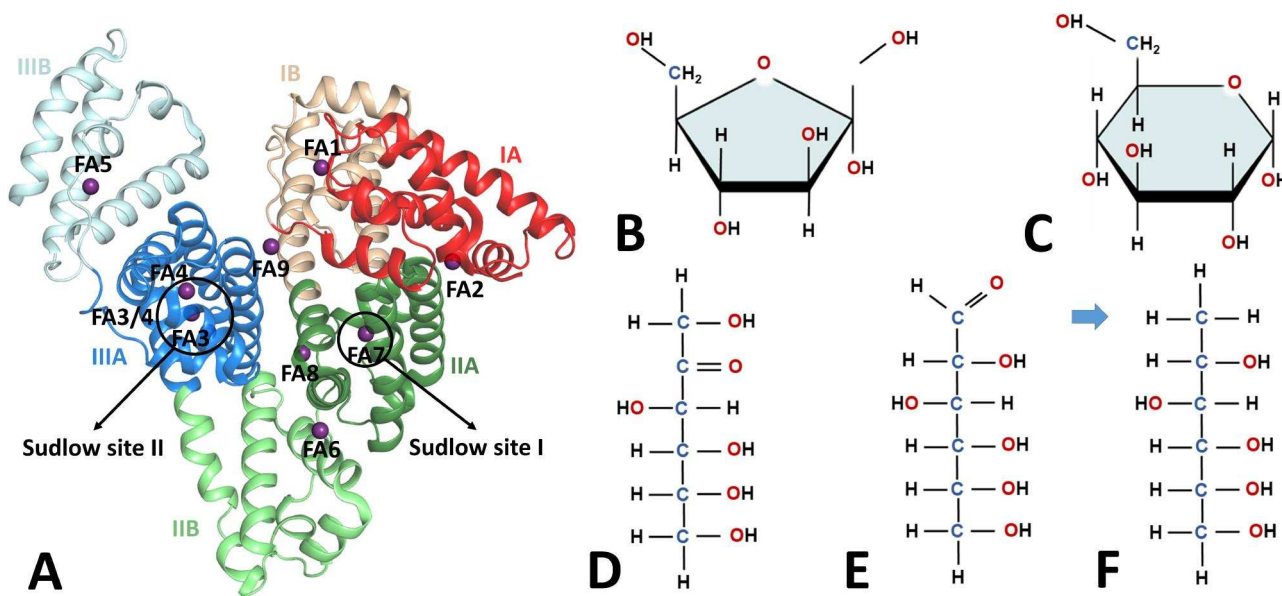
such as the absorption, distribution, metabolism, and excretion of small molecules. The structural organization of HSA is composed of nine distinct binding sites within its domains, specifically designed to accommodate long and medium chain fatty acids (FA). It is worth noting that the primary drug binding sites are situated in the hydrophobic pockets of subdomains II-A (commonly referred to as Sudlow site I or FA7) and III-A (known as Sudlow site II or FA3) (Ascenzi et al., 2015; Fanali et al., 2012), as depicted in Figure 1A.

Carbohydrates, which enjoy widespread recognition as fructose (FRC/FRO) and glucose (GLC/GLO) (Figure 1B–E), are highly regarded for their role as the principal physiological sources of carbon, a crucial component that assumes a pivotal position in an extensive array of physiological and biochemical processes that are inherent to the existence of living organisms. Glucose, an aldohexose, garners significant attention as the predominant sugar circulating in the bloodstream, while fructose, existing as a ketohexose isomer of glucose, showcases its distinctive characteristics. It is worthy of note that the open-chain manifestations of D-fructose and D-glucose, commonly denoted as FRO and GLO (Figure correspondingly, encompass a diverse range of structural configurations that

**CONTACT** André Hadad  andrehadad@gmail.com  Department of Physics, Federal University of Ceará, Bloco 929, Labnitio Laboratory, 60440-900, Fortaleza, Ceará, Brazil

 Supplemental data for this article can be accessed online at <https://doi.org/10.1080/07391102.2024.2310211>.

© 2024 Informa UK Limited, trading as Taylor & Francis Group



**Figure 1.** Structures of human serum albumin (HSA), fructose and glucose. (A) The previously determined overall structure of HSA with its subdomains identified by different colors. Chemical structure for (B)  $\beta$ -D-Fructofuranose (FRC), (C)  $\alpha$ -D-Glucopyranose (GLC) (D) Keto-D-Fructose (FRO), (E) Aldehyde-D-Glucose (GLO) and (F) modified Aldehyde-D-Glucose (GLOm).

encompass the keto-D-fructose and aldehyde-D-glucose forms as their respective representations (Qi & Tester, 2019). Interestingly, preliminary investigations conducted by our studies provided indications of the limited binding efficiency between open glucose and Human Serum Albumin (HSA), prompting a thorough exploration of their interaction. As a result, this investigation has ultimately culminated in a structural modification of the glucose molecule, which is visually portrayed in Figure 1F.

FRC/FRO and GLC/GLO are efficiently assimilated and metabolized by various tissues, particularly the liver, and are implicated in conditions related to monosaccharide metabolism, while an elevated intake of carbohydrates is linked to a range of ailments (Asp, 1994). Of particular note, the consumption of beverages rich in fructose is correlated with an escalated susceptibility to obesity, type 2 diabetes, and cardiovascular disease (Malik & Hu, 2015). Furthermore, research has emphasized the impact of heightened levels of glucose in the plasma on the binding capacity of HSA (Chaves et al., 2021).

The intricate interplay between reducing sugars and proteins is unraveled by chemical modifications, specifically non-enzymatic glycation. This post-translational modification, which results from a condensation reaction between the carbonyl group of a reducing sugar and the amino group of a protein, yields advanced glycation end products (Bai et al., 2012; Khalifah et al., 1999; Reynolds, 1963; Thornalley et al., 1999). Glycated proteins, such as HSA, exhibit functional discrepancies compared to their non-glycated counterparts (Brownlee, 1991) and contribute to the long-term tissue damage associated with the pathophysiology of diabetes (Trüeb et al., 1980).

In this comprehensive scenario, advances in the determination of molecular structure, facilitated by the use of X-ray crystallography (Bhattacharya et al., 2000; Ghuman et al., 2005; Petitpas et al., 2001; Zunszain et al., 2003, 2008) and NMR spectroscopy (Hamilton, 2013; Krenzel et al., 2013;

Simard et al., 2006), is complemented by the application of computational biology techniques that allow the study of protein-drug complexes beyond the scope of molecular interactions (Grinter & Zou, 2014; Kitchen et al., 2004). These techniques include classical molecular dynamics (Ganesan et al., 2017; Smith & Sternberg, 2002; Zhao & Caflisch, 2015), molecular mechanics (Pearlman et al., 1995; Rappe et al., 1992), and semiempirical quantum methods (Christensen et al., 2016; Thiel, 2014). Quantum biochemistry calculations using density functional theory (DFT) (Antony & Grimme, 2006, 2012; Becke, 1992; Ehrlich et al., 2013; Fox et al., 2014; Jones, 2015; Mardirossian & Head-Gordon, 2017) combined with molecular fractionation strategies (Collins & Bettens, 2015; Gordon et al., 2012; Jiang et al., 2006; J. Liu & Herbert, 2016; Saha & Raghavachari, 2015; Söderhjelm et al., 2009; Wang et al., 2013; Zhang & Zhang, 2003) are accurate for estimating interaction energies.

The three-dimensional crystal structures of HSA bound to fructose or glucose provided information on the binding of the closed-chain (GLC/FRC) and open-chain (GLO/FRO) forms at Sudlow site I (FA7). GLC was located at the bottom of the pocket, while GLO was covalently linked to K195 (Wang et al., 2013). K195 was the major site for glucose and HSA reactions (Nasiri et al., 2010). No covalent bond between fructose and HSA was reported (Wang et al., 2013). Molecular dynamics simulations have further delineated the dynamics of these interactions, highlighting the different degrees of domain motion that affect drug/solute binding affinity. While previous theoretical investigations have conducted comprehensive examinations of the interaction between HSA and diverse saccharides employing a 1:2 stoichiometric ratio (HSA simultaneously associated with the sugar's closed and open configurations), we opted to concentrate our computational endeavors on a 1:1 stoichiometric ratio (HSA attached to either the closed or open configuration of the sugar). This decision is motivated by the ring-opening mechanism of



these sugars, which requires one sugar molecule to initially occupy the lower part of Sudlow site I, while a second sugar molecule, positioned at the entrance of Sudlow site I, receives a proton from K199, facilitating the subsequent opening of the ring. Furthermore, previous theoretical/structural studies have demonstrated the escape of sugars from HSA pockets (Awang et al., 2021; Pongprayoon & Mori, 2018), highlighting the importance of a 1:1 inspection.

Studying the influence of different forms of glucose and fructose before the glycation reactions manifest themselves plays a crucial role in understanding the early stages of this process and how these steps affect protein dynamics. Investigation of these early stages can provide significant insight into the fundamental processes and possible interventions. Theoretical studies dedicated to the effects of different manifestations of glucose and fructose have the potential to contribute significantly to the development of biomarkers. This progress is particularly relevant in diseases such as diabetes, providing tools for early detection and effective monitoring of disease progression. Moreover, a thorough comprehension of the impacts of these sugars may hold significant significance in the advancement of specialized medications that aim to manipulate protein-ligand interactions. In our preliminary inquiry, we embarked upon an examination of the non-covalent interplay between FRO, FRC, GLO, and GLC and human serum albumin (HSA). Remarkably, our computational molecular dynamics simulations did not reveal a significant coupling of GLO to HSA. This observation led to an innovative approach: modification of the GLO molecule. By eliminating the double bond with oxygen and replacing it with hydrogen, we created a modified version called GLOm. This structural change targeted the known covalent bond between GLO and HSA, specifically the carbon site previously occupied by the oxygen double bond. Remarkably, upon implementation of this modification, GLOm showed a pronounced interaction with HSA, demonstrating the potential importance of this structural adaptation in modulating the binding dynamics between glucose derivatives and HSA. This inquiry will be carried out by means of a comprehensive and all-encompassing approach that combines the utilization of fluorescence spectroscopy measurements along with the integration of molecular dynamics simulations, protein conformational clusters (EnGens), molecular fractionation with conjugate caps (MFCC) and quantum biochemistry analysis.

## 2. Material and methods

### 2.1. Fluorescence spectroscopy

Fluorescence measurements were performed with a Perkin Elmer LS45 Fluorescence Spectrometer (Perkin-Elmer, UK) equipped with a 1 cm quartz cell. HSA (free of fatty acids, Sigma) was excited ( $\lambda_{exc}$ ) at 280 nm and emission recorded between 300 and 500 nm with the excitation and emission slit widths set to 10 nm each. Fluorescence spectra were obtained at a fixed concentration of HSA (10  $\mu$ M in 2.5 mL of phosphate buffered saline (PBS) solution, pH 7.4) titrated with 10  $\mu$ L aliquot from a D-glucose or D-fructose in PBS

buffer (pH 7.4). The final concentration of monosaccharides was 0.2, 0.4, 0.6, 1 and 1.4 mM. All measurements were performed in triplicate at 25 °C. HSA concentration was determined spectrophotometrically using the molar extinction coefficient of 35,700 M<sup>-1</sup> cm<sup>-1</sup>.

### 2.2. Acquiring and preparing the structures of HSA, FRC, FRO, GLC, GLO and GLOm

The three-dimensional structures of the FRC/FRO:HSA and GLC/GLO:HSA complexes were obtained from the Protein Data Bank with the identifiers 4IW1 (2.56 Å resolution) and 4IW2 (2.41 Å resolution), respectively (Wang et al., 2013). The removal of water and small ligands from the crystal structure of HSA was performed using Discovery Studio 2016 from Dassault Systems BIOVIA. To calculate the protonation states of the HSA residues at pH 7.4 and to complete their hydrogens, we used the ProteinPrepare web server as described by Martínez-Rosell et al. (2017). In addition, the three-dimensional structures of GLC, GLO, GLOm, FRC, and FRO were subjected to charge analysis at pH 7.4 using MarvinSketch 18.24 (Marvin Beans Suite, ChemAxon), followed by geometry optimization in Discovery Studio 2016.

As mentioned above, our computational molecular dynamics simulations did not reveal a significant coupling of GLO to HSA. Considering that the formation of the GLO:K195 covalent bond results in the release of an oxygen atom bound to carbon 1 of linear glucose (Nasiri et al., 2010), we decided to exclude the double bond (C1=O) originating from the pyranose opening and replace it with hydrogen atoms. This approach was implemented to assess whether this pre-glycation modification would result in a more stable complex compared to the unmodified version GLO.

### 2.3. Molecular docking

We decided to perform molecular dynamics simulations of human serum albumin (HSA) in duplicate to explore different perspectives on the interaction of the protein with sugars. However, to ensure a comprehensive approach, we decided to avoid using the same ligand as a starting point in both simulations. Therefore, in the first simulation we decided to use the ligand derived from crystallographic data to represent a more realistic interaction with HSA based on experimental structures. In the second simulation, we decided to use the result of molecular docking as the ligand, providing a computational approach that takes into account the complementarity and affinity between sugars and HSA. This strategic choice aims to provide a more comprehensive insight into the molecular interactions involved, thus justifying the molecular docking calculations as a crucial step in enriching our understanding of the molecular dynamics of HSA in conjunction with different sugars.

In this study, we implemented the CB-Dock2 algorithm to perform automated blind protein-ligand docking, following the four steps described in the literature (Liu et al., 2022). The first phase consisted of data input, where we provided the PDB files of the protein of interest and the ligand, either from

crystallographic data. CB-Dock2 then performed data processing, generated a model of the protein binding cavity, and created initial poses for the ligand. In the third step, cavity detection and docking were performed, with CB-Dock2 starting the search for known complexes and using two parallel pipelines, structure-based and template-based blind docking, to calculate the interaction energy between each pose of the ligand and the binding cavity. Finally, the fourth step involved visualization and analysis of the results, using the tools provided by CB-Dock2 to explore the structure of the protein, analyze the interaction energy between the molecules, and examine the pose of the ligand in the binding cavity of the protein.

## 2.4. Molecular dynamics (MD)

GROMACS 2021.2 was used to merge the ligands and the HSA in the water box and to perform energy minimization, NVT, NPT, and MD simulations. The topologies of the HSA were determined using the CHARMM36 force field (Best et al., 2012). The parameters of the atomic crystallographic coordinates and the molecular docking positions with the lowest RMSD values were provided by CgenFF (Vanommeslaeghe et al., 2009) and GROMACS 2021.2. The complexes constructed are shown in Table 1.

based on the previously prepared HSA structures. These complexes were separately inserted into different water boxes (based on the TIP3 model) with an ionic concentration of 0.15 M ( $\text{Na}^+$  and  $\text{Cl}^-$ ). The generated complexes were energetically minimized by a steepest descent algorithm, and these simulations were stopped when the maximum force of the systems was less than  $1,000 \text{ kJ.mol}^{-1}.\text{nm}^{-1}$ . Next, 5,000 steps of NVT simulations were performed to equilibrate the temperature of the systems to approximately 300 K using a velocity rescaling thermostat, V-rescale thermostat. Then 5,000 steps of NPT simulations based on the Berendsen-isotropic pressure couple were performed as a continuation of NVT, using the Berendsen barostat in conjunction with the V-rescale thermostat to maintain the system temperature and pressure during the NPT equilibration phase (Berendsen et al., 1984). The time integration of the NVT and NPT simulations was performed using a Leap-Frog algorithm with a time step of 2 fs, enabled by the combination of the LINCS algorithm and hydrogen mass repartitioning. The preference for the Leap-Frog algorithm over others is motivated by its remarkable stability in energy conservation, simplicity of implementation, and effectiveness in simulating complex systems. The Verlet cutoff was used to calculate the unbonded interactions, with a distance cutoff of 1.0 nm for Van Der Waals and real-space Coulomb interactions. At the same time, periodic boundary conditions and the particle mesh Ewald scheme with default parameters were used to treat the long-range electrostatic part of the Coulomb interactions.

Then, 100 ns MD simulations were performed for each complex.

The Root-Mean-Square Deviation (RMSD) of the atomic positions of HSA, FRC/FRO, and GLC/GLO/GLOm were also calculated for all molecular dynamics conformations in their computational samples using the `gmx rms` GROMACS module. The RMSD results were the average between two MD simulations (Crystal and Redocking). The final atomic coordinates obtained by this procedure were then used for all interaction calculations based on non-bonded interactions. VMD 1.9.3 was used to perform a detailed analysis of the mobility of the HSA subdomain and the hydrogen bonds formed between sugars and albumin during the simulations. The donor-acceptor distance ( $3.0 \text{ \AA}$ ) and the angle cutoff ( $20^\circ$ ) were the parameters used as criteria for the calculation of hydrogen bonds.

The results obtained showed an interesting behavior of GLO with respect to HSA, where the molecule had no significant binding capacity with the protein in a 1:1 ratio. However, a specific modification in the molecular structure of glucose, involving the elimination of the double bond with oxygen O1 and the saturation of the valences with hydrogen, led to a remarkable change in the interaction scenario, allowing the modified molecule GLOm to form effective bonds with HSA.

The study used the Discovery Studio Receptor-Ligand Interactions package to obtain 2D plots of the interactions between HSA molecules and residues in a static perspective. Based on a medium quality ligand and preserving hydrogen interactions, the types of interactions at a distance of  $4 \text{ \AA}$  from the molecule were explored. The criteria for defining hydrogen bonds involved measuring distances and angles between specific atoms, with a maximum D-A distance of  $3.4 \text{ \AA}$ . The angular relationships were limited to between  $90$  and  $180$  degrees for different combinations of atoms. Hydrogen bonds with water were also considered when the water molecule formed a bridge between different molecules that met the criteria.

## 2.5. Generation of representative protein conformational clusters

The EnGens (short for ensemble generation) method for generating representative protein conformations from molecular dynamics is divided into three main steps: importing the molecular dynamics results, the results are imported in the form of trajectory files containing the spatial coordinates of the protein atoms at each time instant; filtering the molecular dynamics results, removing incomplete trajectories or those with errors; and reconstructing the conformational structures, which is performed using the averaging algorithm, which calculates the average of the spatial coordinates of the atoms at each time instant (Conev et al., 2023).

**Table 1.** Complexes of sugars complexed with HSA used in simulations.

	FRC	FRO	GLC	GLO	GLOm
Crystallographic (C)	FRC <sub>c</sub> :HSA	FRO <sub>c</sub> :HSA	GLC <sub>c</sub> :HSA	GLO <sub>c</sub> :HSA	GLOm <sub>c</sub> :HSA
Docking (D)	FRC <sub>D</sub> :HSA	FRO <sub>D</sub> :HSA	GLC <sub>D</sub> :HSA	GLO <sub>D</sub> :HSA	GLOm <sub>D</sub> :HSA

After reconstructing the conformational structures, the Uniform Manifold Approximation and Projection (UMAP) algorithm is used to reduce the dimensionality of the structures. UMAP is a non-linear algorithm that preserves the structural relationships between the conformational structures (Trozzi et al., 2021). By reducing the dimensionality, the conformational structures can be visualized in a low-dimensional space, which facilitates the analysis and interpretation of the results.

To optimize computational efficiency, in this particular case we chose to study three representative conformations, i.e. the UMAP algorithm identifies three groups of conformational structures that are significantly different from each other. These groups of conformational structures are then submitted to the Gaussian Mixture Models (GMM) algorithm (Hand et al., 1989), which is a probabilistic machine learning algorithm that assumes the data come from a mixture of normal distributions. The GMM algorithm is then used to refine the initial grouping performed by UMAP. This combination of algorithms makes it possible to identify conformations that are representative of the protein's flexibility. Once the results were obtained, Molecular Fractionation with Conjugate Caps (MFCC) calculations and quantum biochemical analyses were performed on the three representative conformations provided by EnGens, as well as on the final conformation found in molecular dynamics.

## 2.6. Molecular fractionation with conjugate caps (MFCC) and quantum biochemistry

Based on the results obtained, the analysis continued with Molecular Fractionation with Conjugate Caps (MFCC) calculations of the three representative conformations generated by EnGens, in addition to the study of the final conformation identified by molecular dynamics. This method allowed the evaluation of the interaction energies between the most relevant molecules of the studied complexes using density functional theory (DFT) (Jones, 2015).

The MFCC approach allows a deeper study of many residues in a protein with reduced computational time while maintaining accuracy (Chen et al., 2005; He & Zhang, 2005; Zhang & Zhang, 2003). To limit the number of residues to be analyzed without missing critical interactions, non-covalent interactions within 10 Å of the sugar molecule and water molecules within 2.5 Å of each HSA residue and sugar were considered. It is worth noting that due to the high computational cost, only in the final conformation of the molecular dynamics considered water molecules, while. Additionally, the MFCC and DFT calculations performed for representative conformations provided by EnGens did not take solvation effects into account. This approach optimizes the computational efficiency of the study and enables calculations to be performed within a reasonable time using the available computational resources.

The interaction energy between the  $i$ -th HSA residue of interest ( $R_i$ ) and FRC/FRO/GLC/GLO/GLOm was calculated as follows. First, four auxiliary system were defined: (S1)  $R_i$  plus its respective conjugate caps  $C_{i-1}$  and  $C_{i+1}$  (the caps are

attached to the amine and carboxyl moieties of  $R_i$  and reflect the electronic environment of its neighborhood); (S2)  $R_i$  plus its conjugate caps only; (S3) the conjugate caps of  $R_i$  and FRC/FRO/GLC/GLO/GLOm; (S4) the conjugate caps of  $R_i$  only. For each system, the total energy  $E_{S_i}$  was evaluated using the DFT formalism. All dangling bonds were passivated with hydrogen atoms, and the caps were formed from the immediately adjacent residues of  $R_i$  in the protein chain,  $C_{i\pm 1} = R_{i\pm 1}$ . The interaction energy between FRC/FRO/GLC/GLO/GLOm and each amino acid residue was given by:

$$E(\text{FRC/FRO/GLC/GLO/GLOm} - R_i) = E_{S1} - E_{S2} - E_{S3} + E_{S4} \quad (\text{Eq 1})$$

DFT calculations were carried out to investigate the sub-components of Eq.1 in a detailed way (Delley, 1990, 2000). A Double Numerical plus Polarization (DNP) basis set was chosen to expand the Kohn-Sham orbitals of all electrons. The Perdew-Burke-Ernzerhof (PBE) exchange correlation function within the Generalized Gradient Approximation (GGA) (Perdew et al., 1996) and the Tkatchenko-Scheffler dispersion correction scheme (GGA + TS) (Tkatchenko & Scheffler, 2009) were chosen to evaluate the electronic energies. The orbital cutoff radius was set to 4 Å, and the self-consistent field (SCF) convergence threshold was calibrated to  $10^{-6}$  Ha. Water solvation was implemented through the COSMO implicit solvation model (Klamt & Schüürmann, 1993) with a dielectric constant of 40.

Finally, using the above mentioned basis, corrections, orbital cut-off, and SCF adjustments, electrostatic potential calculations (isosurfaces) were performed for both the initial and final sugar conformations in all HSA::Sugar complexes.

## 3. Results and discussion

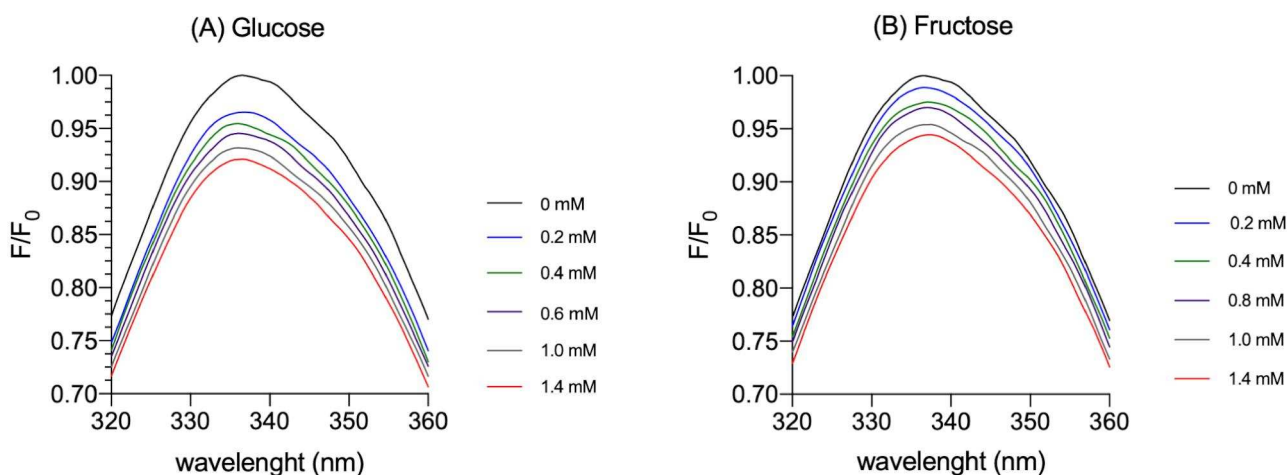
### 3.1. Fluorescence quenching

GLC and FRC caused a concentration-dependent quenching of HSA intrinsic fluorescence, with a slightly higher effect of GLC compared to FRC (Figure 2). These results clearly indicate binding of both monosaccharides to HSA. Although HSA presents 40 aromatic residues (31 phenylalanine, 18 tyrosine, and only one tryptophan residue) (Zhang et al., 2015) its intrinsic fluorescence is mainly attributed to the TRP214 residue found at the surface of subdomain II (site FA8) (Lv et al., 2022; Tayeh et al., 2009).

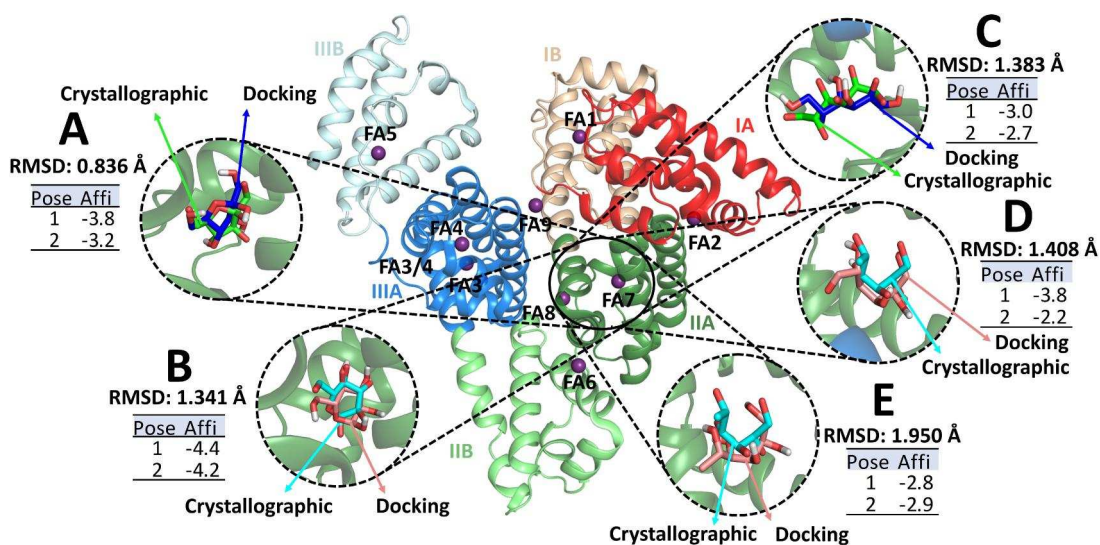
Spectra did not present any shift in the maximum emission peak of HSA (approximately at 340 nm) as well. Therefore, no change in the microenvironment hydrophobicity around TRP214 occurred upon the binding of GLC and FRC to HAS, consistent with previous studies (Sergio et al., 2021).

Our fluorescence quenching results are in accordance with a previous study that reported a weak and spontaneous binding of GLC and FRC to HSA (Chaves et al., 2021). The authors also showed that the interaction of GLC and FRC with HSA occurred *via* a ground-state association (static fluorescence quenching mechanism) and were mostly ruled by hydrogen bonding, Van Der Waals forces, as well as hydrophobic effects related to the release of hydration shell upon ligand binding





**Figure 2.** Fluorescence emission spectra. HSA titrated with glucose (A) and fructose (B). 10  $\mu$ M HSA in PBS buffer (pH 7.4). Concentration of glucose and fructose: 0.2, 0.4, 0.6, 1 and 1.4 mM. Excitation at 280 nm. Excitation and emission slits of 10 nm. Temperature was 25  $^{\circ}$ C.



**Figure 3.** Molecular template-based blind docking, listing the two poses with the lowest binding energy (kcal.mol $^{-1}$ ). (A) FRC, (B) GLC, (C) FRO, (D) GLO and (E) GLOm.

highlighting the importance of investigate these interactions on molecular level (Chaves et al., 2021).

As will be better discussed in section 3.4.4, the highest affinity conformations according to quantum biochemistry were glucose derivatives. Although this difference is not statistically significant, we hypothesize that this pattern may be related to a slightly greater effect of glucose on HSA fluorescence compared to fructose.

### 3.2. Molecular docking

The binding of numerous small molecules to human and bovine serum albumins has already been studied using molecular docking and computational methods with promising results (Li et al., 2010; Lv et al., 2022; Sergio et al., 2021; Shamsi et al., 2018; Wang et al., 2019; Wani et al., 2021). Here, we use blind docking tests with CB-Dock2 to predict FRC/GLC/FRO/GLO/GLOm positions similar to previously reported crystallographic conformations (Wang et al., 2019). CB-Dock2 produced template-based blind docking results

based on models very close to the crystal bonding modes (FRC/FRO/GLC/GLO) - see Figure 3.

The RMSD between the template-based blind docking poses (top rank 1) and the crystal structures (Figure 3) was less than 2  $\text{\AA}$ . These CB-Dock2 results are in full agreement with the crystallographic study that revealed an occupancy near the bottom of the Sudlow I pocket by these sugar closed forms (Wang et al., 2013).

CB-Dock2 values indicated a higher binding affinity of GLC ( $-4.4$  and  $-4.2$  kcal.mol $^{-1}$ ) for Sudlow site I (FA7) than that of FRC ( $-3.8$  and  $-3.2$  kcal.mol $^{-1}$ ) (Figure 3A,B). These results are in contrast to a previous Sudlow site I-directed docking investigation, which suggested more favorable docking values for FRC:HSA (18.4) rather than GLC:HSA (17.2) (Chaves et al., 2021). In the case of FRO, the docking scores obtained at Sudlow site I were  $-3.0$  kcal.mol $^{-1}$  and  $-2.7$  kcal.mol $^{-1}$ . In addition, the RMSD between the docking positions and the crystal structures was 1.383  $\text{\AA}$  (Figure 3C).

For GLO and GLOm, CB-Dock2 also showed satisfactory results, with docking scores at Sudlow site I of  $-3.8$  and  $-2.2$  kcal.mol $^{-1}$  for GLO (Figure 3D) and  $-2.8$  and

$-2.9 \text{ kcal.mol}^{-1}$  for GLOm (Figure 3E). The docking of GLO and GLOm yielded RMSDs of 1.408 and 1.950 Å, respectively (Figure 3D–E). Although the GLO and GLOm molecules are slightly different from the GLO molecule in the crystal, we found it interesting to perform the RMSD for comparison purposes.

In conclusion, CB-Dock2 proved to be effective in describing and predicting the interactions between FRC, GLC, FRO, GLO, GLOm and HSA. The template-based blind docking results showed agreement with the crystal structures and favorable docking scores, indicating higher binding affinity. Therefore, CB-Dock2 proves to be a suitable approach for future investigations of the interactions between FRC, GLC, FRO, GLO, GLOm and HSA.

### 3.3. Molecular dynamics

Molecular dynamics (MD) has been widely applied to detailed studies of proteins complexed to a variety of small molecules (Gelpi et al., 2015; Hollingsworth & Dror, 2018; Hu et al., 2023), including the binding of monosaccharides to HSA (Awang et al., 2017, 2021; Nasiri et al., 2010; Pongprayoon & Mori, 2018). We have identified binding interaction sites by template-based blind docking (Figure 3). Crystallography and spectroscopic studies pointed to Sudlow Site I as the leading binding site for both fructose and glucose (Barzegar et al., 2007; Chaves et al., 2021; Szkudlarek et al., 2016; Wang et al., 2019; Zaman et al., 2018). Molecular dynamics simulations were carried out in duplicate using two different ligands. The first ligand was derived from crystallographic data, while the second ligand was obtained from molecular docking (top 1). Combining the results of these simulations provided a more complete picture of the interaction between the ligands and the protein. Performing simulations in duplicate is a good practice to increase the reliability of the results. Simulations can be influenced by random factors, such as the initial position of the molecules. To reduce the likelihood of results being influenced by these factors, two simulations with different initial conditions can be carried out. In MD simulations, the measured root mean square deviation (RMSD) of atomic coordinates in a molecule can be used as a temporal indicator of structural stability (Maruyama et al., 2023). All heavy atom RMSD values were calculated to evaluate the overall stability of HSA and sugar over the course of the simulations. HSA shows consistent stability in all simulations, with no significant differences in the RMSD variations of albumin bound to different sugars - see Figure 4. The  $\Delta\text{RMSD-HSA}$  found here is consistent with previous C-alpha RMSD measurements of HSA simultaneously bound to closed and linear sugar forms of fructose (glucose), which generally fit in the ranges of  $\sim 2\text{-}4 \text{ \AA}$  or  $\sim 3\text{-}5 \text{ \AA}$  (Nocedal & Wright, 1960; Pongprayoon & Mori, 2018). Compared to albumin bound to sugars in a 1:2 ratio, the overall structure of the albumin studied here (1:1) does not appear to show significant structural changes.

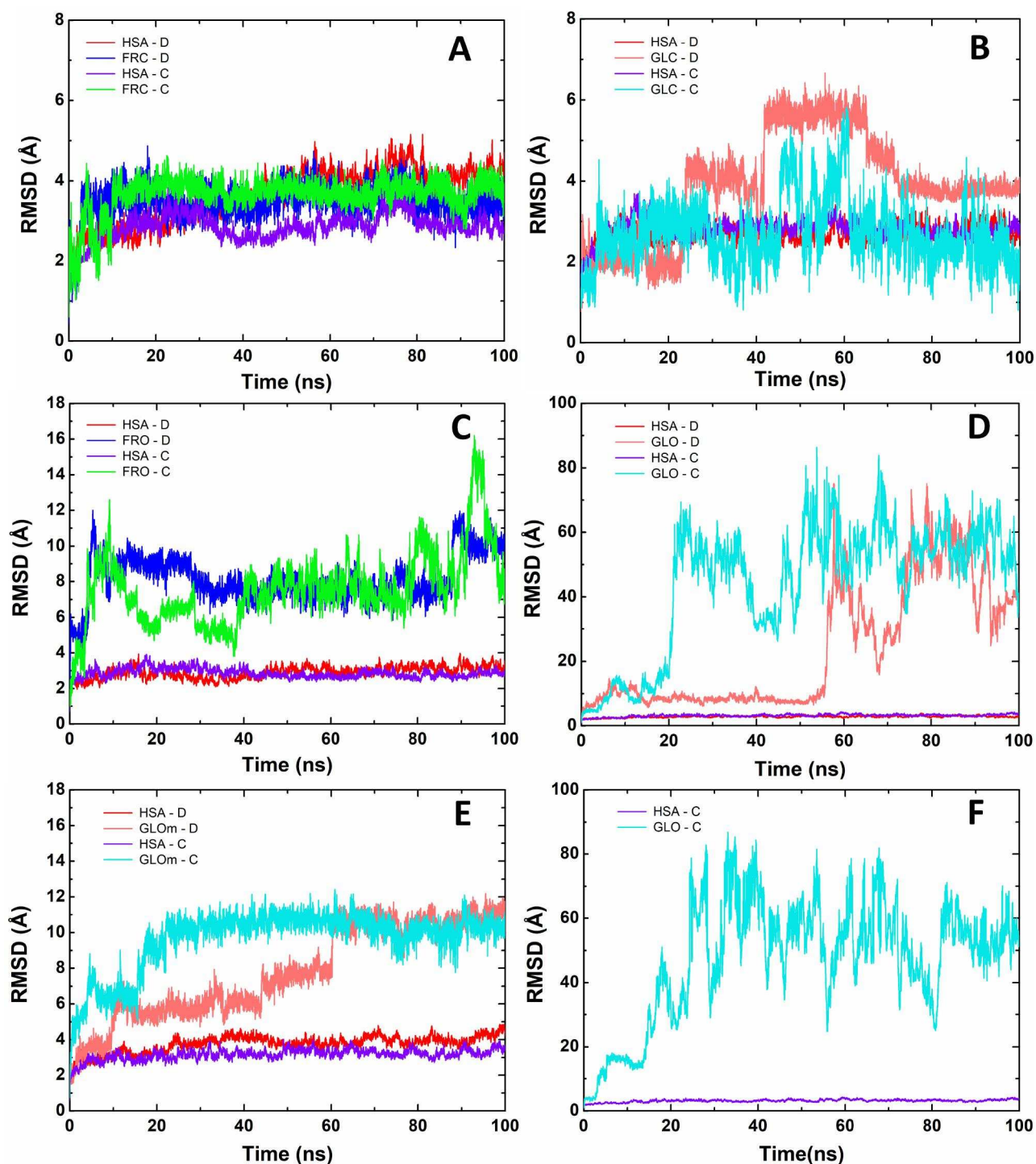
The RMSD values of the C-alpha atoms indicate that the binding of fructofuranose resulted in the highest degree of flexibility in subdomain IA (Figure S1). Long-chain sugars

appear to induce higher levels of structural variation in subdomains IB and IIIB (Figure S1). In addition, subdomains IIA and IIIA, which contain Sudlow site I and Sudlow site II, were the most rigid in all simulations. Considering previous computational studies investigating the nonbonded mode of monosaccharides with HSA, Awang et al. (2017) identified rigidity in Sudlow Site I and higher fluctuation in Sudlow Site II of pGHSA (pre-glycated HSA in a 1:2 ratio). Pongprayoon et al. (2018) identified low fluctuation levels in both subdomains IIA and IIIA. Awang et al. (2021) reported high flexibility in subdomain IA of HSA:FRC + FRO complexes and reduced rigidity in IIA of complex HSA:FRC-FRO number 1 (named FHSA\_1). Furthermore, all these measurements show that subdomain IIIB is highly flexible in the presence of sugars in their unbound modes. These results also suggest that the low rigidity found in subdomain IA of the HSA::FRC-FRO complex is due to the influence of the fructose closed form rather than its linear form.

We found that the different sugars did not result in significant differences in the proximity fluctuations between HSA subdomains IB and IIIA (represented by the K195-D451 distance), which generally varied between 6 Å to 12 Å (Figure S2A–D). The distances between V116 and D510 mainly varied between 8 Å and 34 Å, revealing the scissor-like motion between the upper albumin subdomains IB and IIIB of albumin (Figure S2E–H). Furthermore, the scissor-like motion observed in all 1 monosaccharide-bound HSA systems is similar and consistent with previous findings from 2 sugar-bound albumin studies (Nocedal & Wright, 1960; Sittivanichai et al., 2023; Sousa et al., 2016).

Although our fluorescence quenching results indicate that non-bonded sugars to HSA do not alter the polarity of the polar microenvironment of TRP214, we decided to examine the positional changes of TRP214 before and after simulations in the state of 1-sugar-bound HSA. As a result, we observed that both closed and linear sugar binding to HSA led to similar reorientations of TRP214 - see Figure S3. Compared to the structural reorientations of TRP214 observed in the simulations of glycosylated HSA over 500 ns, the modifications of TRP214 upon sugar binding to HSA in non-covalent mode and 1:1 molecular ratio showed remarkable similarity (Sittivanichai et al., 2023). These include the relocation of the indole ring in a parallel orientation to the alpha helix.

In contrast to albumin, the sugar ligands exhibited significant instability throughout the simulations, except for FRC, which showed an RMSD range similar to that of the protein. During the 100 ns of simulation, the total variation of RMSD for FRC (docking and crystallographic) ranged from 0.4 Å to 4.8 Å, for GLC (docking and crystallographic) from 0.7 Å to 6.6 Å, for FRO from 1.0 Å to 16.1 Å, for GLO from 1.0 Å to 86.3 Å, and for GLOm from 0.7 Å to 12.4 Å (Figure 4). In the last 50 ns, FRC<sub>D</sub> and FRC<sub>C</sub> showed  $\Delta\text{RMSD}$  of 2.4 and 1.7 Å, respectively (Figure 4A), GLC<sub>D</sub> and GLC<sub>C</sub> showed  $\Delta\text{RMSD}$  of 3.2 and 5.1 Å (Figure 4B), FRO<sub>D</sub> and FRO<sub>C</sub> showed  $\Delta\text{RMSD}$  of 6.1 Å and 10.8 Å (Figure 4C), GLO<sub>D</sub> and GLO<sub>C</sub> showed  $\Delta\text{RMSD}$  of 68.7 and 52.7 Å (Figure 4D), and GLOm<sub>D</sub> and GLOm<sub>C</sub> showed  $\Delta\text{RMSD}$  of 5.3 and 4.6 Å (Figure 4E). These results indicate that monosaccharides are structurally less



**Figure 4.** All heavy-atom Root-mean-square deviation (RMSD) as a function of time of the HSA:Monosaccharides complexes. FRC (A), GLC (B), FRO (C), GLO (D), GLOm (E) and repeated simulation GLO (F).

stable than HSA in all simulations, with GLC showing greater structural instability compared to FRC.

To confirm our results, we repeated the molecular dynamics simulation for GLO (Figure 4D) using the same conditions. This repetition was performed to address concerns about the drastically changed RMSDs (68.7 and 52.7 Å) in the GLO structure.

The results of the repeated simulation (Figure 4F) are consistent with the previous results. During the 100 ns of

simulation, the total RMSD of the GLO<sub>D</sub>/HSA complex varied from 1.0 Å to 86.8 Å. In the last 50 ns, the RMSD of the complex varied by 57.2 Å. These results confirm that GLO exhibits significant structural instability when it interacts with HSA. The RMSD (Figure 4D) and hydrogen bonding analysis indicate that GLO dissociates from HSA, suggesting that this conformation is not ideal for interaction with HSA, except if it is binding in a 1:2 ratio and have the effect of a pyranose as previously proposed (Wang et al., 2013). The Figures S4, S5 provides more



details on the binding and dissociation of GLO::HSA, the mobility of the domain and the displacement of TRP214. Interestingly, the GLOm with the excluded double bond and valences saturated with hydrogen exhibits greater conformational stability throughout the simulations. Since the GLOs escape from the HSA binding pockets, we only performed quantum biochemical calculations of the final conformation (100 ns) for this modified GLOm structure (Figure 4E). The hydrogen bonding behavior indicates that the complete exit occurred between 40-60 in both molecular simulations. These results confirm the mechanism previously proposed in the crystallographic study of HSA::monosaccharides, as this modified GLO is in the immediately pre-glycated state of linear glucose (Wang et al., 2013), suggesting that this intermediate has more affinity for Sudlow site I than its previous conformation with  $C1=O1$ . Although chemical reactions are generally fast processes (femtoseconds) (Potter et al., 1992), we believe that these data indicate that the loss of  $C1=O1$  is a critical step in the HSA::GLO glycation mechanism.

The MD results are supported by our fluorescence quenching data (Figure 2) and are consistent with a study by Chaves et al. (2021), who reported weak binding capacities of fructose and glucose to HSA (Chaves et al., 2021). For FRO and GLOm, the results showed greater structural instability compared to FRC and GLC. Structural analyses of sugar displacements near subdomain IIA clearly indicated that linear sugars exhibited greater flexibility than closed forms (see Figure S6). In addition, all sugars formed fewer than 5 hydrogen bonds with albumin (1:1), reinforcing the weak interactions between sugars and albumin - see Figure S7. Both FRC (~0.50 H-bonds) and FRO (~0.46) showed a slightly higher average number of strong contacts with HSA than GLC/GLOm (~0.35/~0.34). Given that human albumin is known to be promiscuous in binding a wide range of ligands (Fanali et al., 2012) and does not form strong interactions with fructose and glucose in a 1:1 molar ratio, our results are consistent with the well-documented fact that sugars are not primary ligands for this protein (Awang et al., 2017, 2021; Chaves et al., 2021).

The structural instability of GLO/FRO is consistent with previous theoretical reports (Awang et al., 2017, 2021), which suggest that linear sugars are more susceptible to solvation than their respective closed forms. This can be attributed to solvation effects and the absence of furanose/pyranose structures deep within Sudlow site I. Consequently, linear sugars were able to move from the entrance of Sudlow site I towards the upper right direction of the site (Figure S6C,G,H), except for FROD, which shifted towards subdomain IIIA (Figure S6D). In this regard, Pongpraynoo et al. (2018) reported that both forms of galactose were able to escape from the vicinity of Sudlow site I, an event not observed in HSA::GLC/GLO. Awang et al. (2021) detected the escape of galactose from HSA pockets and observed two distinct FRO behaviors around the IIA vicinity: a shift from the entrance of Sudlow Site I to its upper right, as well as wobbling around the same position. In addition, Awang et al. (2017) observed that GLC remained trapped within Sudlow Site I, while GLO showed a higher degree of fluctuation. Interestingly, the

closed form of glucose bound to glycated HSA was found to shift toward the gate of Sudlow Site I (Sittivanichai et al., 2023).

Consistently, the absence of the dimer formed between the linear and closed sugars did not result in the exit of the closed sugars. Interestingly, a combined analysis of the dynamic properties of monosaccharides bound to albumin at 1:1 and 1:2 ratios suggests that closed fructose is best trapped in Sudlow site I in the absence of FRO-FRC, as FRC escape was detected in the 2-fructose-bound HSA (Awang et al., 2021). Furthermore, this is consistent with the previously reported mechanism for GLC escape from the lower Sudlow site I within glycated HSA (Sittivanichai et al., 2023).

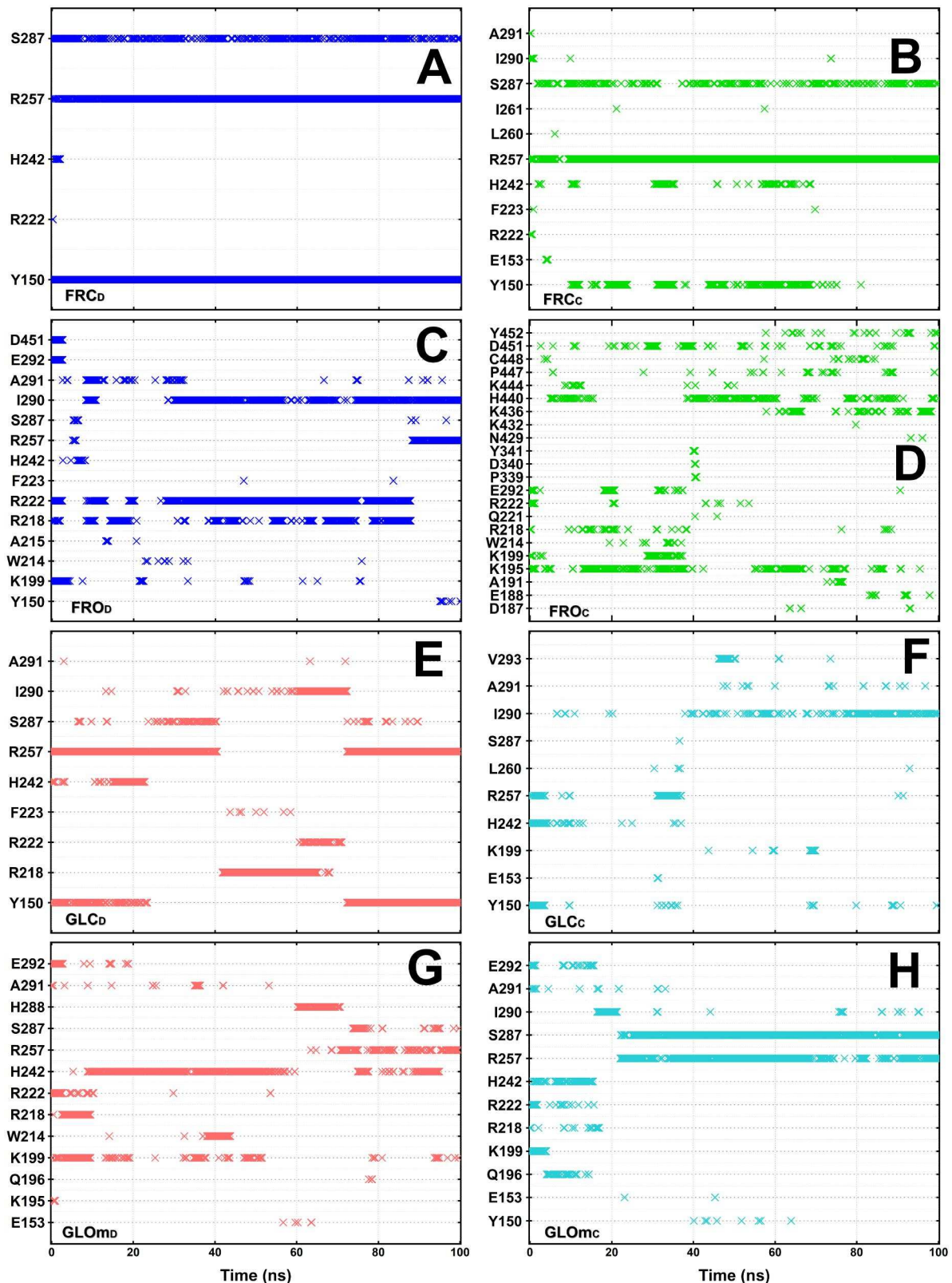
Furthermore, FRC exhibited the most linear binding at Sudlow site I, as represented by the permanent interactions with residues Y150, R222, and S287 (Figure 5). Initially, all linear forms except FRO<sub>C</sub> established strong contacts with residues R218, R222, and R292 located in the Sudlow site I gate, but these interactions were disrupted at ~5-20 ns (Figure 5). Subsequently, all linear forms showed a tendency to interact with deeper residues in Sudlow site I, such as Y150, R257, and S287 (Figure 5). This rapid change in hydrogen bonding pattern (~15-30 ns) also suggests that linear sugars have a higher affinity for Sudlow site I, as shown in Figure S6.

Since we found that the absence of pyranose at Sudlow site I led to a massive shift of GLOm to a more internal pocket of subdomain IIA (Figure 5 and Figure S6), our results seem to be in agreement with previous studies highlighting the critical role of pyranose for HSA glycation (Chaves et al., 2021; Wang et al., 2013). The peculiar behavior observed in FRO<sub>C</sub> suggests that linear fructose has greater flexibility than linear glucose at a temperature of 300 K, which may help explain the absence of the FRO<sub>C</sub>:K195 covalent bond in the previous crystallographic study (Wang et al., 2013).

The MD results are consistent with previous theoretical studies evaluating these biological complexes (Awang et al., 2021; Castro-Alvarez et al., 2017; Nasiri et al., 2010). For example, closed glucose appears to interact randomly with charged residues within Sudlow site I (Pongpraynoo & Mori, 2018), which is in good agreement with  $\Delta$ RMSD-GLO/GLOm (Figure 4) and the absence of permanent hydrogen bonding to HSA (Figure 5). For more details on the random occupancy of the closed and open forms, see Figures S5-S7. In addition, glycated HSA shows a tendency to retain the closed monosaccharides within the Sudlow site I, which was particularly observed for Amadori HSA (Sittivanichai et al., 2023). Therefore, our underlying hypothesis is that the random behavior of GLC near Sudlow Site I (Figure S6) is related to the different results obtained with glycated HSA samples used to investigate affinity changes associated with HSA glycation (Anguizola et al., 2013; Ghosh & Kishore, 2022; Joseph & Hage, 2010; Moeinpour et al., 2016).

### 3.4. Quantum biochemistry and FRC/FRO/GLC/GLO/GLOm atomic details of their binding to HSA

Chaves et al. (2021) have conducted observations which suggest that the interaction between glucose and fructose

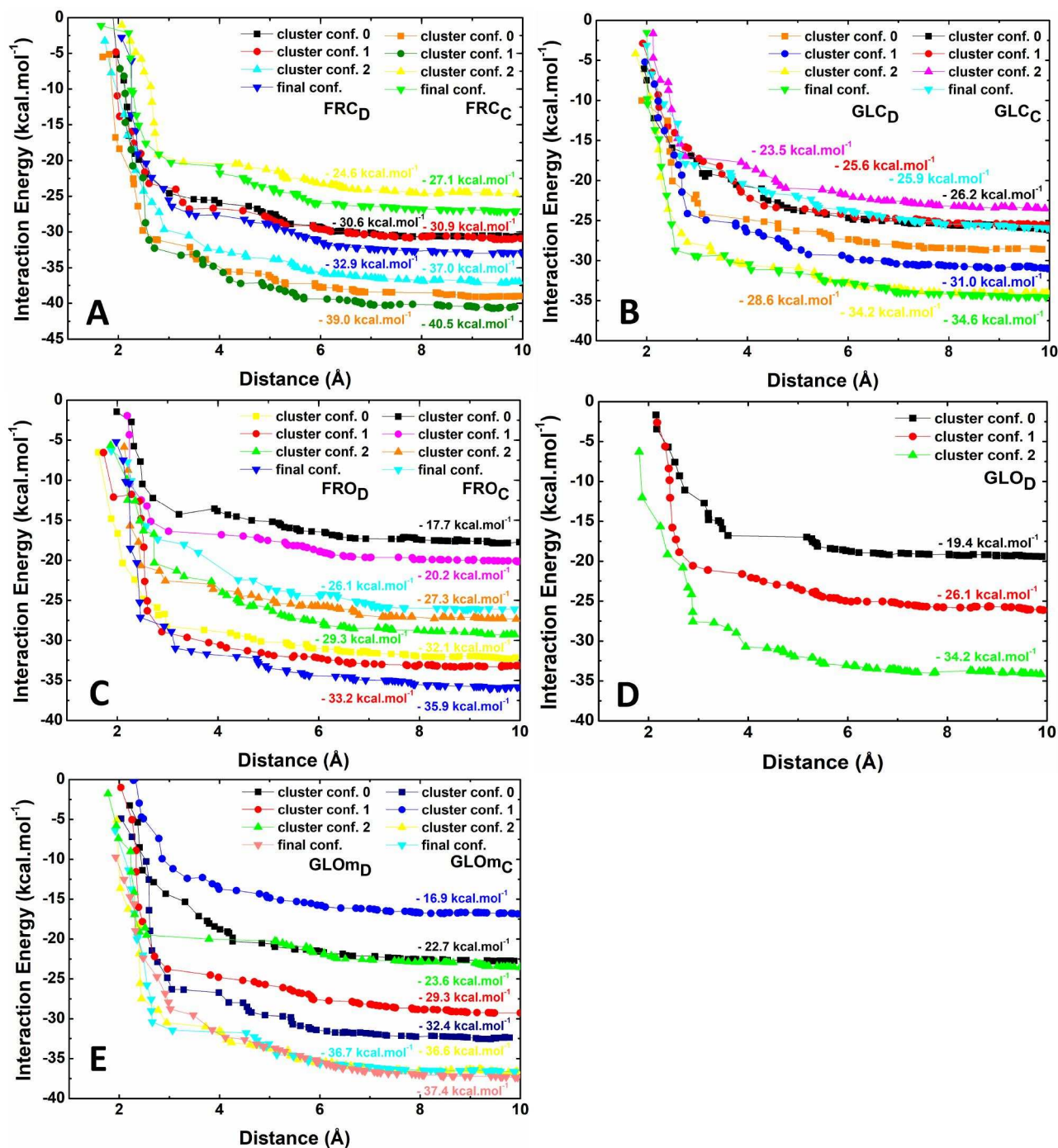


**Figure 5.** Detailed hydrogen bond analysis detected in the HSA::Monosaccharide systems over the simulation time. For (A) HSA::FRC<sub>D</sub>, (B) HSA::FRC<sub>C</sub>, (C) HSA::FRO<sub>D</sub>, (D) HSA::FRO<sub>C</sub>, (E) HSA::GLC<sub>D</sub>, (F) HSA::GLC<sub>C</sub>, (G) HSA::GLO<sub>D</sub>, (H) HSA::GLO<sub>mc</sub>. The y-axis represents the HSA residues, while the x-axis corresponds to the simulation time.

predominantly takes place *via* hydrophobic, electrostatic, and hydrogen bonding mechanisms. Given this scenario, we used DFT-based calculations in conjunction with the MFCC

scheme to perform a more comprehensive investigation of the non-bonded interactions between the monosaccharides and the HSA residues, considering each set of three



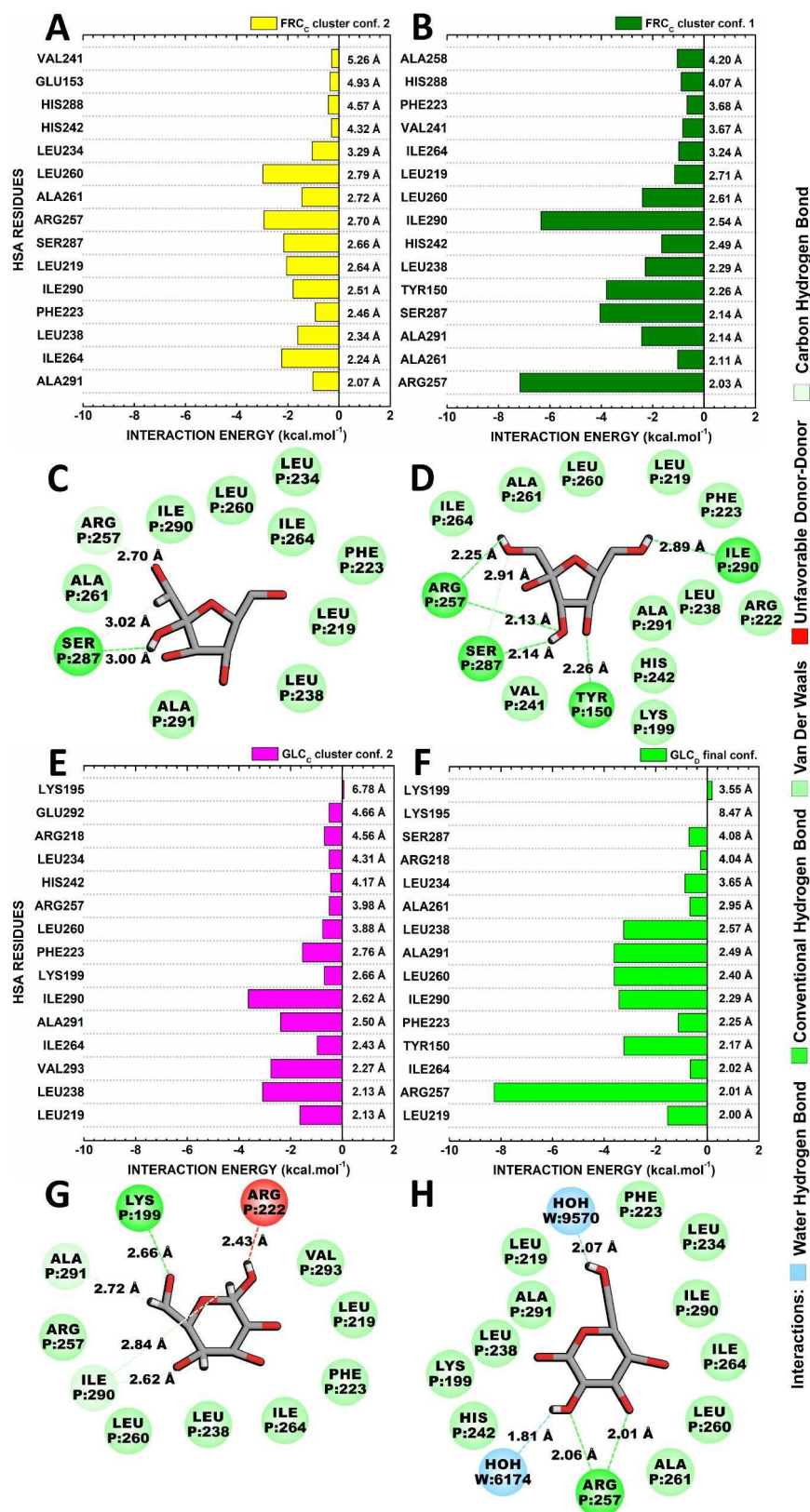


**Figure 6.** The total interaction energy between sugars and the HSA residues as a function of the binding site radius. (A) FRC<sub>D</sub>/FRC<sub>C</sub>; (B) GLC<sub>D</sub>/GLC<sub>C</sub>; (C) FRO<sub>D</sub>/FRO<sub>C</sub>; (D) GLO<sub>D</sub>; (E) GLO<sub>mD</sub>/GLO<sub>mC</sub>.

representative conformations generated by EnGens as well as the final conformation resultant from molecular dynamics.

Figures 6–9 provide detailed information on the interaction energy, the binding modes and the interactions between FRC/FRO/GLC/GLO/GLO<sub>m</sub> and HSA in the final frame of each of the MD simulations carried out here. The convergence of the interaction energy in the four samples occurred near the radius of 8 Å from the centre of FRC/FRO/GLC/GLO/GLO<sub>m</sub>, indicating that residues located further than 8 Å do not significantly contribute to the stability of these systems (Figure 6). Although a previous study performed

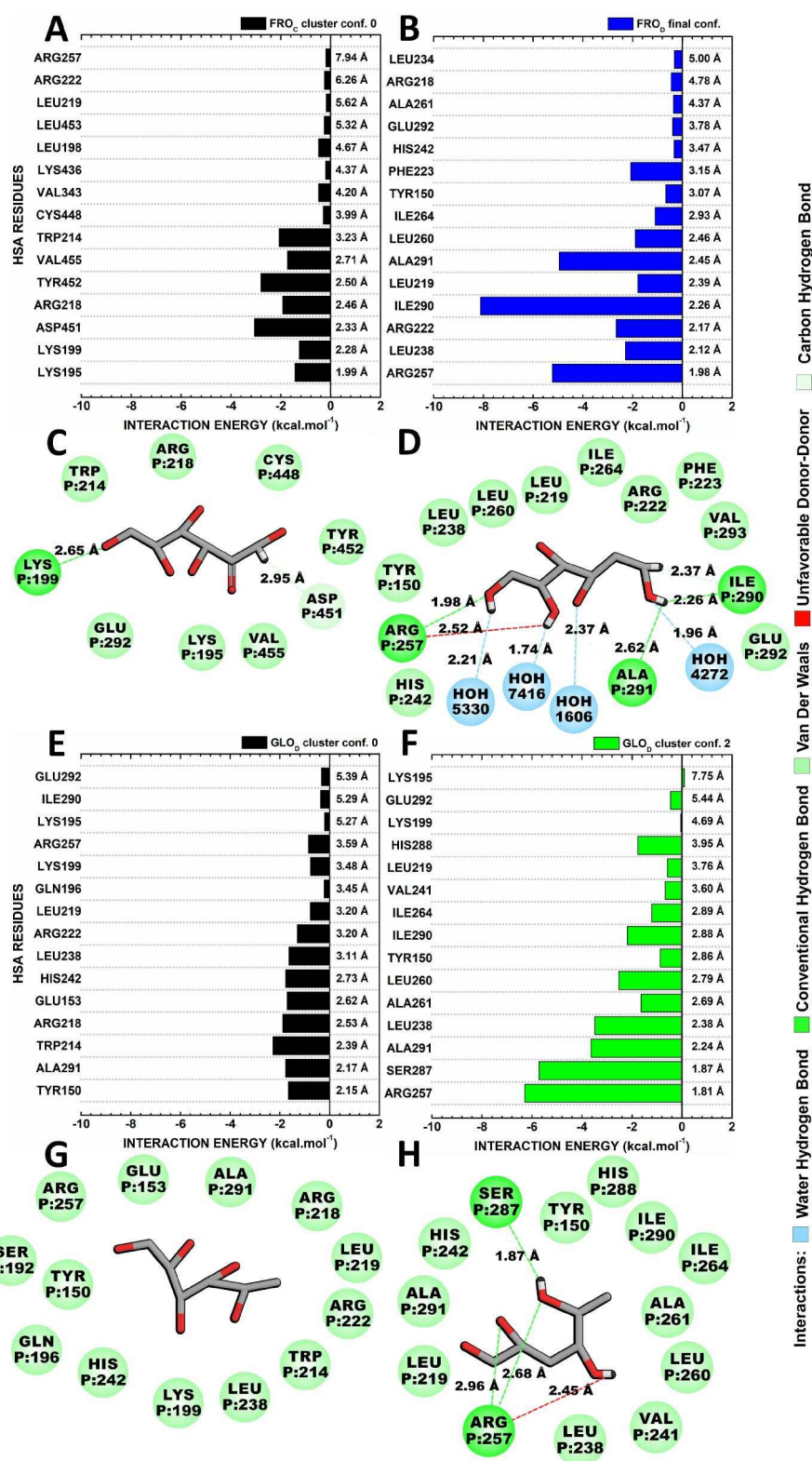
quantum calculations to estimate the energies of interactions between sugars and HAS (Awang et al., 2017), they limited their calculations to a smaller region of 4 Å, extending from the GLC/GLO centre towards the albumin residues. Quantum biochemistry revealed that residues beyond this radius are crucial for the capture of the sugars in the HSA. As mentioned above, we ran 8 MFCC/quantum biochemistry simulations for each simulation. For example, we performed 4 calculations of the FRC<sub>C</sub> complex and 4 of the FRC<sub>D</sub>. Of each these 4 simulations, 3 were generated by the EnGens method and 1 by the final



**Figure 7.** Quantum biochemistry results for the HSA::FRC and HSA::GLC complexes. BIRD panel shows the interaction energy of each residue with the regions of the molecule for the lowest and highest energy clusters, respectively. (A) FRC<sub>c</sub> cluster conf. 2; (B) FRC<sub>c</sub> cluster conf. 1; (E) GLC<sub>c</sub> cluster conf. 2; (F) GLC<sub>c</sub> final cluster. Non-covalent interactions within 10 Å of the molecule (C) FRC<sub>c</sub> cluster conf. 2; (D) FRC<sub>c</sub> cluster conf. 1; (G) GLC<sub>c</sub> cluster conf. 2; (H) GLC<sub>c</sub> final cluster.

conformation resulting from molecular dynamics. It is important to note that for the GLO complex only 3 results were obtained for the GLO<sub>D</sub>, all from the EnGens method. Furthermore, no results were obtained for the final

conformation of either GLO<sub>D</sub> or GLO<sub>C</sub>. This occurred because both molecules left the HSA binding site. The exit of these molecules from the binding site is shown in the RMSD [Figure 4](#).

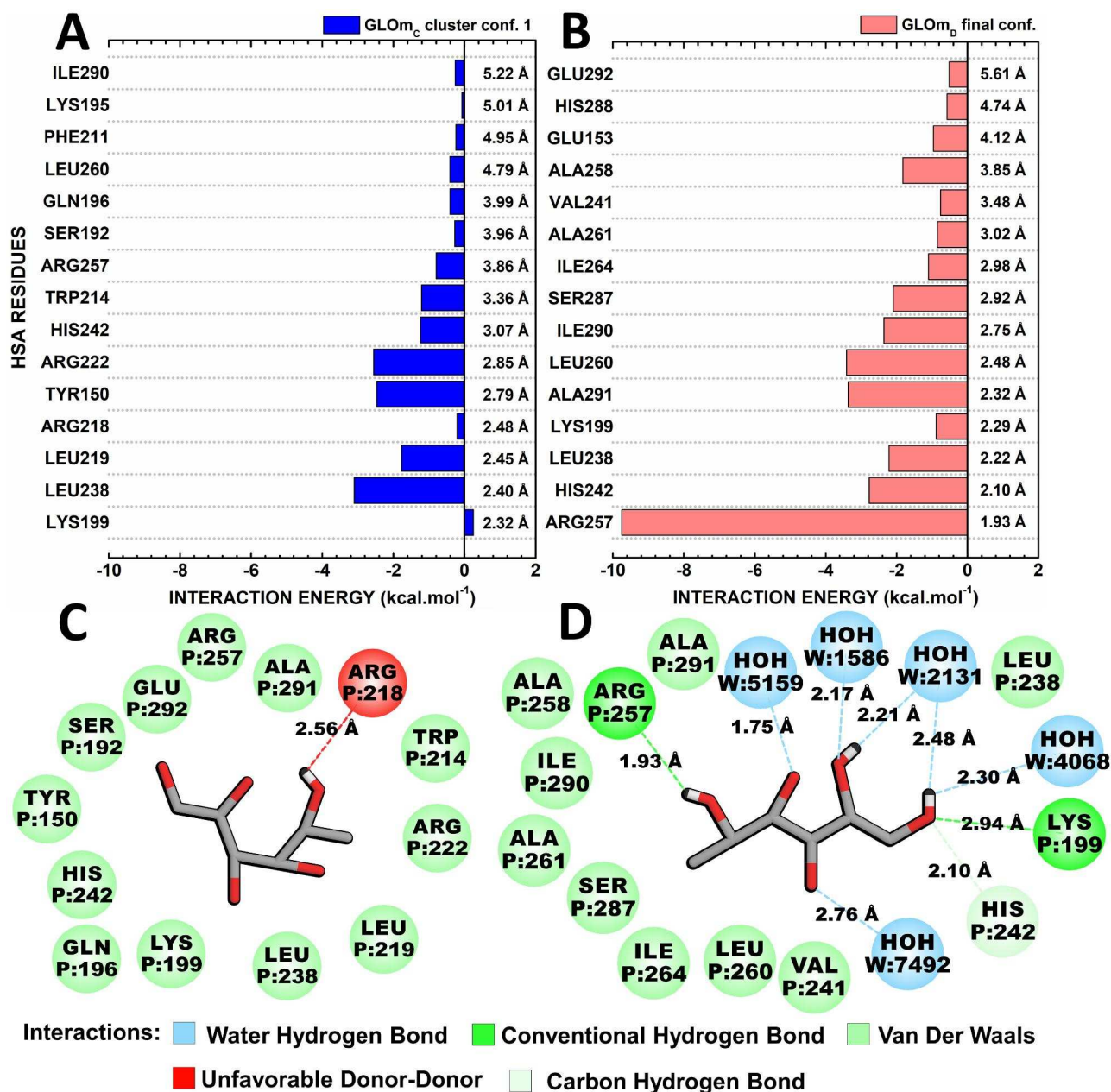


**Figure 8.** Quantum biochemistry results for the HSA::FRO and HSA::GLO complexes. BIRD panel shows the interaction energy of each residue with the regions of the molecule for the lowest and highest energy clusters, respectively. (A) FRO<sub>C</sub> cluster conf. 0; (B) FRO<sub>D</sub> final conf.; (E) GLO<sub>D</sub> cluster conf. 0; (F) GLO<sub>D</sub> cluster conf. 2. Non-covalent interactions within 10 Å of the molecule (C) FRO<sub>C</sub> cluster conf. 0; (D) FRO<sub>D</sub> final conf.; (G) GLO<sub>D</sub> cluster conf. 0; (H) GLO<sub>D</sub> cluster conf. 2.

The mean interaction energy between FRC and HSA reached a value of  $-32.8 \text{ kcal.mol}^{-1}$  with a standard deviation of  $5.29 \text{ kcal.mol}^{-1}$ . The coefficient of variation, which is 16.1%, indicates that the interaction energy has a relatively high degree of variability. Conversely, the average interaction

energy between GLC and HSA was  $-28.7 \text{ kcal.mol}^{-1}$  with a standard deviation of  $3.89 \text{ kcal.mol}^{-1}$ . The coefficient of variation of 13.6% suggests that the interaction energy is less variable compared to that of fructose (as shown in Table 2). These results indicate that FRC has a higher affinity for the





**Figure 9.** Quantum biochemistry results for the HSA:GLOm complex. BIRD panel shows the interaction energy of each residue with the regions of the molecule for the lowest and highest energy clusters, respectively. (A) GLOm<sub>c</sub> cluster conf. 1; (B) GLOm<sub>D</sub> final conf. Non-covalent interactions within 10 Å of the molecule (C) GLOm<sub>c</sub> cluster conf. 1; (D) GLOm<sub>D</sub> final conf.

**Table 2.** Interaction energy of all simulations, with mean value, standard deviation ( $\sigma$ ) and coefficient of variation ( $C_V$ ).

	Interaction energy kcal.mol <sup>-1</sup>								Mean kcal.mol <sup>-1</sup>	$\sigma$ kcal.mol <sup>-1</sup>	$C_V$
FRC	-24.6	-27.1	-30.6	-30.9	-32.9	-37.0	-39.0	-40.5	-32.8	5.29	16.1%
GLC	-23.5	-25.6	-25.9	-26.2	-28.6	-31.0	-34.2	-34.6	-28.7	3.89	13.6%
FRO	-17.7	-20.2	-26.1	-27.3	-29.3	-32.1	-33.2	-35.9	-27.7	5.90	21.3%
GLO	-19.4	-26.1	-34.2	—	—	—	—	—	-26.6	6.05	22.8%
GLOm	-16.9	-22.7	-23.6	-29.3	-32.4	-36.6	-36.7	-37.4	-29.5	7.18	24.4%

Sudlow I site than GLC. However, GLC shows a more stable variation than FRC. In other words, on average, the interaction energy values of GLC are closer to the mean than those of FRC. Interestingly, open and closed fructose showed more affinity for HSA than open and closed glucoses (Table 2). However, compared to the interaction energies obtained in our previous HSA quantum biochemistry studies (Dantas et al., 2015; França et al., 2022), closed and open

sugars appear to have weaker affinities for HSA. This is consistent with previous studies showing that closed sugars are not preferential ligands for HSA (Chaves et al., 2021; Pongprayoon & Mori, 2018).

The mean interaction energy for FRO was  $-27.7$  kcal.mol<sup>-1</sup>, as shown in Figure 6C and Table 2. The standard deviation for this value was calculated to be  $5.90$  kcal.mol<sup>-1</sup>, giving a coefficient of variation of 21.3%. This suggests that FRO

interacts consistently with HSA, although there is some variability in the results. The GLO system showed an average interaction energy of  $-26.6 \text{ kcal.mol}^{-1}$  with a standard deviation of  $6.05 \text{ kcal.mol}^{-1}$ . However, in some cases GLO did not interact with HSA (Figure 6D and Table 2). This may indicate that GLO has a higher tendency than FRO to detach from HSA during the interaction. GLOm had the highest average interaction energy value of  $-29.5 \text{ kcal.mol}^{-1}$ , together with the largest standard deviation of  $7.18 \text{ kcal.mol}^{-1}$ , resulting in a coefficient of variation of 24.4% (Figure 6E and Table 2). This indicates that GLOm has the strongest interaction with HSA, but also the greatest variation in results. Notably, free perturbation calculations have previously shown a greater affinity of HSA for glucose compared to fructose (Pongprayoon & Mori, 2018). Importantly, our results are consistent with previous studies (Awang et al., 2017; Wang et al., 2013), which reported a GLO:K195 covalent bond.

Moreover, quantum biochemistry suggests that isolated single sugar pyranoses do not play a crucial role in the slight difference between HSA:FRC and HSA:GLC affinities achieved through spectroscopic experiments that investigated monosaccharide binding to HSA (Chaves et al., 2021), which could be more consistently attributed both to the higher number of hydrogen bonding HSA::Fructose (Figure S7) and the permanent contacts established between FRC and HSA (Figure 5). However, the higher average affinity of HSA for long chain fructose than linear glucose seems not to be in line with the slightly stronger effect of glucose on HSA fluorescence compared to fructose (Figure 2).

The low binding affinity of both monosaccharides for HSA indicates a weak binding capacity of carbohydrates to albumin. Several factors contribute to this, such as the monosaccharides' small size, Sudlow Site I's large size, and the high solvation effect within the cavity. The factors mentioned facilitate easy access of monosaccharides to metabolic pathways (Yang et al., 2007), highlighting their primary function in providing energy, as previously reported by Chaves et al. (2021).

### 3.4.1. FRC:: HSA interactions inside Sudlow Site I

The previous crystallographic study revealed remarkable similarities between the FRC/GLC and FRO/GLO binding modes within Sudlow site I. This was evidenced by similar networks of interactions realised between FRC/GLC and the five HSA residues TYR150, ARG222, LEU238, ARG257 and ALA291, as well as the strong FRO/GLO contacts LYS195, LYS199, ARG218, ARG222 and ALA291 (Wang et al., 2013).

To better understand the nature of these interactions, in this section we analyse the lowest and highest affinities conformations of each complex shown in Figure 6. The lowest affinity conformation (FRC<sub>C</sub> cluster conf. 2), which include residues LEU260 ( $-2.97 \text{ kcal.mol}^{-1}$ ), ARG257 ( $-2.94 \text{ kcal.mol}^{-1}$ ) and ILE264 ( $-2.24 \text{ kcal.mol}^{-1}$ ), showed the highest interaction energies with FRC, suggesting that these residues may play a critical role in binding FRC to HSA (Figure 7A). In particular, residue SER287 ( $-2.16 \text{ kcal.mol}^{-1}$ ) formed a hydrogen bond and a carbon-hydrogen bond with FRC, while residue ARG257

formed a carbon-hydrogen bond with HSA. The other residues formed Van Der Waals interactions (Figure 7C).

On the other hand, the highest affinity conformation (FRC<sub>C</sub> cluster conf. 1), which include residues ARG257 ( $-7.16 \text{ kcal.mol}^{-1}$ ), ILE290 ( $-6.34 \text{ kcal.mol}^{-1}$ ) and SER287 ( $-4.06 \text{ kcal.mol}^{-1}$ ), showed the highest interaction energies with the FRC (Figure 7B). This suggests that these residues may have a particularly high affinity for the FRC, possibly due to specific charge or hydrophobic interactions. In this context, residue ARG257 formed two hydrogen bonds with FRC, while residues SER287, TYR150 and ILE290 ( $-3.80 \text{ kcal.mol}^{-1}$ ) each formed one hydrogen bond with FRC. The other residues depicted in Figure 7D formed Van Der Waals interactions.

It is noteworthy that several critical residues (13 of the 15 residues) such as LEU260, ARG257, ILE264, SER287, LEU219, ILE290, LEU238, ALA261, ALA291, PHE223, HIS288, HIS242 and VAL241 were identified in both the lowest and highest affinity conformations. This suggests that these residues may play a key role in modulating the interaction between FRC and HSA, irrespective of the energy conditions.

Quantum biochemistry also revealed that FRC interacts weakly with Sudlow site I. We found that the non-polar residues are mainly responsible for the HSA::FRC interaction energy. In addition, our quantum results identified new residues from subdomain IIA, such as ALA261, ILE264 and LEU219, as potentially essential for keeping FRC bound to HSA. Figure 7 and Tables S1, S2 provide further details of the calculated interactions.

Although Van Der Waals interactions mediated most of the HSA::FRC contacts, hydrogen bonds were also detected in the conformations. The low number of hydrogen bonds (Figure 7C,D) is consistent with the low interaction energy between HSA and FRC, the structural instability discussed above, and previous theoretical reports of the HSA::FRC complex (Awang et al., 2021). Furthermore, the residues surrounding the FRCRD in this study are similar to those discovered by crystallography (Wang et al., 2013) and slightly different from the pocket residues revealed by a previous theoretical approach (Awang et al., 2021).

### 3.4.2. GLC:: HSA interactions inside Sudlow Site I

The GLC results are consistent with the remarkable similarities previously reported between the FRC and GLC binding modes at the Sudlow site I. DFT calculations performed on the atomic coordinates of GLC revealed that 13 residues were shared among the top 15 residues with the highest affinity (Figure 7E-F).

Structural analysis and quantum biochemistry revealed that the lowest affinity conformation (GLC<sub>C</sub> cluster conf. 2), including residues ILE290 ( $-3.63 \text{ kcal.mol}^{-1}$ ), LEU238 ( $-3.08 \text{ kcal.mol}^{-1}$ ), VAL293 ( $-2.76 \text{ kcal.mol}^{-1}$ ) and LYS199 ( $-0.69 \text{ kcal.mol}^{-1}$ ), showed significant interactions with GLC (Figure 7E). In particular, LYS199 formed one hydrogen bond with GLC, ALA291 ( $-2.38 \text{ kcal.mol}^{-1}$ ) established one carbon-hydrogen bond with GLC and ILE290 formed two carbon-hydrogen bonds with GLC. The other residues formed Van Der Waals interactions (Figure 7G).

On the other hand, the highest affinity conformation (GLC<sub>D</sub> final conf. 2), which include residues ARG257 ( $-8.27 \text{ kcal.mol}^{-1}$ ), ALA291 ( $-3.61 \text{ kcal.mol}^{-1}$ ) and LEU260 ( $-3.61 \text{ kcal.mol}^{-1}$ ), showed significant interactions with the GLC (Figure 7F). In this context, residue ARG257 formed two hydrogen bonds with HSA and two hydrogen bonds between GLC and water molecules. The other residues formed Van Der Waals bonds (Figure 7H).

It is worth noting that several critical residues such as ALA291, ARG218, ARG257, ILE264, ILE290, LEU219, LEU234, LEU238, LEU260, LYS195, LYS199 and PHE223 were identified in both the lowest and highest affinity conformations; these residues may play a fundamental role in modulating the interaction between GLC and Sudlow site I of HSA. Furthermore, the residues forming the SAL and AZT subsites were expressively responsible for the non-binding interactions of HSA with GLC (Figure 7E–F) (Sudlow et al., 1976; Yang et al., 2007). Tables S3, S4 detail these and other weak interactions calculated within a radius of 10 Å from the centre of the GLC.

Consistent with previous quantum calculations between GLC and HSA (Awang et al., 2017), we also found a significant interaction energy at the GLC::ARG257 contact. Quantum biochemistry revealed additional residues within subdomain IIA, including LEU219, LEU260, ILE264 and ILE290, which may play a key role in retaining GLC trapped in the FA7 pocket. Furthermore, our computational results strongly suggest that the presence of a linear sugar at the entrance of Sudlow site I is critical for the glucose tuning revealed by crystallography and subsequently confirmed by molecular dynamics and interaction energy calculations (Awang et al., 2021; Nasiri et al., 2010; Wang et al., 2013).

The influence of LYS195 and LYS199 on glucose ring opening and subsequent covalent establishment with LYS195 was revealed by crystallographic studies of HSA::monosaccharide (Wang et al., 2013). Therefore, in order to gain more insight into these crucial residues, we believe it is essential to investigate them further. Despite the crucial role of LYS195 and LYS199 in HSA glycation, our quantum biochemistry results show that GLC has a weak and even repulsive affinity for these residues. Specifically, GLC binds weakly to LYS199 with interaction energies of  $-0.69 \text{ kcal.mol}^{-1}$  (GLC<sub>C</sub> cluster conf. 2) and  $0.19 \text{ kcal.mol}^{-1}$  (GLC<sub>D</sub> final conf. 2). Our results also show that the GLC was distant from LYS195 in both samples and had negligible interaction energies. Furthermore, our investigation revealed that the interaction between GLC and LYS199 is facilitated by Van Der Waals forces, as shown in Figure 7G,H. These results are not contradictory, as the glycation of HSA is a rather long process that takes days to weeks to complete (Barnaby et al., 2011; Sattarahmady et al., 2007).

Furthermore, the calculated electrostatic potential isosurfaces suggest that the final expressions (100 ns) of glucopyranose have the most extreme values of electronic occupancy (Figure S9, S10). This may indicate that GLC is more reactive than FRC and is more likely to be converted to its linear form (Wang et al., 2013). In addition, the weak electrostatic occupancy around C1 confirms this atom as an important anchor point for glycation (Wang et al., 2013).

### 3.4.3. FRO:: HSA interactions inside Sudlow site

The interaction of the FRO molecule with the residues of human serum albumin (HSA) was studied by analysing the lowest and highest affinity conformations. The results obtained reveal a number of significant interactions that contribute to the binding of FRO to HSA.

For FRO we observed similar results to those obtained for FRC in terms of interaction with HSA. The distribution of interaction energies showed significant differences between the two cases. The lowest affinity conformation (FROC cluster conf. 0) showed the highest interaction energies, with residues TRP214, ASP451 and TYR452 identified as the main contributors to these interactions, with interaction energies of  $-2.07 \text{ kcal.mol}^{-1}$ ,  $-3.05 \text{ kcal.mol}^{-1}$  and  $-2.80 \text{ kcal.mol}^{-1}$ , respectively (Figure 8A). In particular, residue LYS199 ( $-1.27 \text{ kcal.mol}^{-1}$ ) formed a hydrogen bond with FRO, while residue ASP451 formed a carbon-hydrogen bond with FRO. The other residues formed Van Der Waals interactions (Figure 8C).

On the other hand, the highest affinity conformation (FRO<sub>C</sub> cluster conf. 0) also showed significant interactions. The main participants in the process were determined to be ILE290 ( $-8.12 \text{ kcal.mol}^{-1}$ ), ARG257 ( $-5.23 \text{ kcal.mol}^{-1}$ ), and ALA291 ( $-4.95 \text{ kcal.mol}^{-1}$ ) residues (Figure 8B). Residue ARG257 formed a hydrogen bond and an unfavourable donor-donor bond with FRO. Moreover, the residues ALA291 and ILE290 established individual hydrogen bonds with FRO, while the residue ALA291 also created a carbon-hydrogen bond with FRO. Four hydrogen bonds were also observed between FRO and water molecules. The other residues formed van der Waals bonds (Figure 8D).

Analysis of the results reveals the presence of only five overlapping critical residues in the lowest and highest affinity conformation. These amino acid residues, namely ARG218, ARG222, ARG257 and LEU219, appear to have a critical function in the interaction between FRO and HSA, independent of the specific arrangement of the FRO molecule (Figure 8A,B). The presence of these overlapping critical residues suggests that they may be potential targets for modulation of the FRO-HSA interaction. We observed a notable effect of explicit water on interactions facilitated by polar residues, in contrast to interactions facilitated by non-polar residues. In particular, the quantum biochemistry and non-bonded interactions detected suggest that FRO is more susceptible to solvation effects than FRC (Figure 8D). This greater susceptibility by a linear sugar is consistent with previous theoretical findings involving HSA::Closed-Open complexes (Awang et al., 2017, 2021; Pongprayoon & Mori, 2018; Sergio et al., 2021). The residues localised around FRO<sub>D</sub> in this study were slightly different from those found by crystallography (ARG218, ARG222 and ALA291). However, the residues around FRO<sub>D</sub> were more similar to those found by FRC crystallography (TYR150, ARG222, LEU238, ARG257 and ALA291) (Wang et al., 2013). To obtain additional details on the computed interactions, refer to Tables S4, S5 within the supplemental materials.

### 3.4.4. GLO/GLOm:: HSA interactions inside Sudlow Site I

DFT calculations performed on the conformations with the lowest and highest affinity between HSA and GLO or GLOm



revealed that 9 and 8 residues were shared among the top 15 residues with the highest affinity (Figures 8E,F, 9A,B). Quantum biochemistry showed that the linear form of glucose has the highest binding affinity for Sudlow site I, despite the absence of a polar atom (due to the removal of C=O). This observation seems to confirm the dominant role of linear glucose in the establishment of glycation (Wang et al., 2013). Also, structural analysis (Figure 9) and interaction energy calculations (Figure S6) showed that, according to molecular dynamics, GLO<sub>m</sub> is retained at Sudlow site I. Furthermore, the residues forming the SAL and AZT subsites (Sudlow et al., 1976; Yang et al., 2007) were expressively responsible for the non-covalent interactions of HSA with GLO<sub>m</sub> (Figure 9).

As mentioned above, for the HSA::GLO complex only 3 results were obtained for GLO<sub>D</sub>, all from the EnGens method. In addition, no results were obtained for the final GLO<sub>D</sub> or GLO<sub>C</sub> conformation because these molecules were released from albumin as previously stated.

The lowest affinity conformation (GLO<sub>D</sub> cluster conf. 0) revealed the involvement of several residues, such as TRP214 ( $-2.28 \text{ kcal.mol}^{-1}$ ), ARG218 ( $-1.88 \text{ kcal.mol}^{-1}$ ) and ALA291 ( $-1.77 \text{ kcal.mol}^{-1}$ ), which interacted most attractively with the GLO molecule (Figure 8E). Interestingly, all these residues were involved in van der Waals interactions (Figure 8G).

The highest affinity conformation (GLO<sub>D</sub> cluster conf. 2) also involved several residues. Residues ARG257 ( $-6.29 \text{ kcal.mol}^{-1}$ ), SER287 ( $-5.72 \text{ kcal.mol}^{-1}$ ) and ALA291 ( $-3.64 \text{ kcal.mol}^{-1}$ ) interacted with the GLO molecule (Figure 8F). ARG257 was involved in different interaction types, forming two hydrogen bonds and one unfavourable donor-donor interaction with HSA and residue SER287 formed one hydrogen bond with GLO. The other residues formed van der Waals bonds (Figure 8H).

Notably, several critical residues such as ALA291, ARG257, GLU292, ILE290, LEU219, LEU238, LYS195, LYS199 and TYR150 were identified in both the lowest and highest affinity conformations of the GLO::HSA complex.

Although all linear sugars formed hydrogen bonds with LYS199 at different times in the simulation, GLO<sub>mD</sub> formed a hydrogen bond and a slightly attractive interaction energy ( $\sim -0.80 \text{ kcal.mol}^{-1}$ ) with this residue, just as in FRO. Notably, this interaction occurs precisely between the O6 atom of GLO<sub>m</sub> and the side chain of LYS199 (Figure 9D), suggesting that O6 may also play a key role in the pre-glycation step, in addition to the O5 indicated earlier in the crystallographic study. These data seem to suggest that O6 can be critical to stabilize the GLO<sub>m</sub> close to LYS195 and LYS199, which appears to be consistent with the covalent bond previously detected between linear glucose and LYS195 (Wang et al., 2013). The sites with the least amount of energy (GLO<sub>mC</sub> cluster conf. 1) exhibited the most substantial interaction energies. Among these interactions, the residues LEU238, ARG222, and TYR150 were identified as the primary contributors, with interaction energies of  $-3.11 \text{ kcal.mol}^{-1}$ ,  $-2.55 \text{ kcal.mol}^{-1}$ , and  $-2.46 \text{ kcal.mol}^{-1}$ , respectively. In particular, we observed that residue ARG222 formed unfavourable donor-donor bonds with GLO<sub>m</sub>. The other residues formed van der Waals interactions (Figure 9A-C).

On the other hand, the highest affinity conformation (GLO<sub>mC</sub> cluster conf. 1) including residues ARG257

( $-9.75 \text{ kcal.mol}^{-1}$ ), LEU260 ( $-3.41 \text{ kcal.mol}^{-1}$ ) and ALA291 ( $-3.36 \text{ kcal.mol}^{-1}$ ) were identified as the main contributors to these interactions. In particular, residues ARG257 and LYS199 formed a hydrogen bond with GLO<sub>m</sub> and residue HIS242 formed a carbon-hydrogen bond with GLO<sub>m</sub>. In addition, were detected six hydrogen bonds between HSA and water molecules (Figure 9B-D). The other residues formed van der Waals interactions. Figure 9 and Tables S9, S10 give full details of these and other weak interactions calculated within a radius of 10 Å from the centre of the GLO<sub>m</sub>.

Compared to the quantum detailing of the GLO<sub>m</sub>::HSA interaction using ONIOM calculations, we detected massive distinct contacts (Awang et al., 2017). Furthermore, we observed that these interactions are similar to those of closed-ring sugars - see Figures 7, 9. This behaviour reveals the fundamental role of pyranose in maintaining GLO<sub>m</sub> at the entrance of Sudlow site I in the preglycated state of albumin. However, our results cannot fully explain the role of pyranose/furanose in glycation. In view of this, future studies using a QM/MM approach coupled with molecular dynamics may provide a better understanding of the function of pyranoses in the mechanism of glucose ring opening and consequently in glycation (Jaña et al., 2018; Pan et al., 2013).

#### 4. Conclusion

This work reports the results of molecular and quantum mechanical analysis on  $\beta$ -D-Fructofuranose::HSA,  $\alpha$ -D-Glucopyranose::HSA, Keto-D-Fructose::HSA and modified-Aldehyde-D-Glucose::HSA complexes at the molecular and energetic levels by mapping and calculating the interaction energies between amino acid residues sugar atoms to assess the total interaction energy through representative conformational ensembles, protein fragmentation, energetic decomposition and a bottom-up interaction energy perspective. The unmodified Aldehyde-D-Glucose was also probed through molecular dynamics assays, which revealed its escape from HSA binding pockets and bulk. We believe this behavior helps explain why glycation is a very time-consuming event that can occur in a window of days to weeks. To the best of our knowledge, this is the first theoretical study to use concomitantly molecular dynamics and representative conformations based on the HSA::Monosaccharide aiming to evaluate the binding mechanism of HSA to glucose and fructose forms in a 1:1 ratio. The quantum biochemistry analysis revealed that the affinities of all HSA::Monosaccharide complexes exhibit high fluctuation. Therefore, it is necessary to scrutinize representative conformations to explore the binding between HSA and sugars in a reliable manner. In addition, this is clarified by the description of interaction types identified in these residues.

Interestingly, we showed that a possible reaction intermediate, namely GLO<sub>m</sub>, was more stable and propense to bind near Sudlow site I than unmodified glucose. We have shown that this transition conformation, just prior to the formation of the Glucose::K195 covalent bond, has a higher affinity for Sudlow site I than unmodified long-chain glucose, which left HSA. Future QM/MM studies may be useful in describing these intermediates and their propensity to

donate electrons around K195. In addition, FRO and GLOm presented a tendency to move from Sudlow Site I entrance to a deeper cavity near FA7 center. We consider that this movement behavior confirms the importance of Sudlow site occupancy for the adequate fit that results in the glycation of K195, as previously proposed by the authors of the crystallographic study. However, as previously stated in the previous 1:2 molecular dynamics studies, our classical simulations in combination with the quantum biochemistry approach also suggest that LYS195 and LYS199 cannot mediate strong non-bonded interactions with GLC, despite their essential role in HSA glycation. A consistent and similar pattern of TRP214 displacement was observed in all albumin samples, including those initially bound to unmodified linear glucose. Thus, the observed displacement of TRP214 does not appear to account for the modest differences in the fluorescence inhibition capacity of HSA by fructose compared to glucose.

Interestingly, the interaction energies of the HSA::monosaccharide complexes are less attractive compared to the interaction energies of HSA::atrazine-pesticide and HSA::ibuprofen-drug previously obtained using similar protocols. This indicates that this class of carbohydrate does not have a significant albumin binding capacity, confirming the availability of the monosaccharide for cell metabolism. In addition, we found that the most attractive interaction energy of fructose complexes and the higher average number of FRC/FRO hydrogen bonds to HSA than the ones established on GLC/GLOm::HSA may correlate with the binding affinity previously reported by spectroscopy (Chaves et al., 2021). Taken together, these findings suggest that the study of albumin as a carrier protein combining molecular and quantum mechanics can be a powerful tool to describe its binding mechanisms to a wide range of small molecules from both dynamic and energetic perspectives, as they are in full agreement with previous experimental reports.

## Disclosure statement

The authors declare that they have no conflicts of interest with the contents of this article.

## Funding

This work was supported by grants from the following Brazilian agencies: Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) (contract number 465699/2014-6); This study was also financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) Finance Code 001 (PROEX 88887.711696/2022-00); and Fundação Cearense de Apoio ao Desenvolvimento Científico e Tecnológico (FUNCAP).

## ORCID

André Hadad  <http://orcid.org/0009-0004-3396-0558>

## References

Anguizola, J., Matsuda, R., Barnaby, O. S., Hoy, K. S., Wa, C., DeBolt, E., Koke, M., & Hage, D. S. (2013). Review: Glycation of human serum

- albumin. *Clinica Chimica Acta; International Journal of Clinical Chemistry*, 425, 64–76. <https://doi.org/10.1016/j.cca.2013.07.013>
- Antony, J., & Grimme, S. (2006). Density functional theory including dispersion corrections for intermolecular interactions in a large benchmark set of biologically relevant molecules. *Physical Chemistry Chemical Physics: PCCP*, 8(45), 5287–5293. <https://doi.org/10.1039/b612585a>
- Antony, J., & Grimme, S. (2012). Fully ab initio protein-ligand interaction energies with dispersion corrected density functional theory. *Journal of Computational Chemistry*, 33(21), 1730–1739. <https://doi.org/10.1002/jcc.23004>
- Ascenzi, P., di Masi, A., Fanali, G., & Fasano, M. (2015). Heme-based catalytic properties of human serum albumin. *Cell Death Discovery*, 1(1), 15025. <https://doi.org/10.1038/cddiscovery.2015.25>
- Asp, N. (1994). Nutritional classification and analysis of food carbohydrates. *The American Journal of Clinical Nutrition*, 59(3 Suppl), 679S–681S. <https://doi.org/10.1093/ajcn/59.3.679Sa>
- Awang, T., Niramitranon, J., Japrun, D., Saparpakorn, P., & Pongprayoon, P. (2021). Investigating the binding affinities of fructose and galactose to human serum albumin: Simulation studies. *Molecular Simulation*, 47(9), 738–747. <https://doi.org/10.1080/08927022.2021.1922687>
- Awang, T., Wiriyatanakorn, N., Saparpakorn, P., Japrun, D., & Pongprayoon, P. (2017). Understanding the effects of two bound glucose in Sudlow site I on structure and function of human serum albumin: Theoretical studies. *Journal of Biomolecular Structure & Dynamics*, 35(4), 781–790. <https://doi.org/10.1080/07391102.2016.1160841>
- Bai, X., Wang, Z., Huang, C., Wang, Z., & Chi, L. (2012). Investigation of non-enzymatic glycosylation of human serum albumin using ion trap-time of flight mass spectrometry. *Molecules (Basel, Switzerland)*, 17(8), 8782–8794. <https://doi.org/10.3390/molecules17088782>
- Barnaby, O. S., Cerny, R. L., Clarke, W., & Hage, D. S. (2011). Comparison of modification sites formed on human serum albumin at various stages of glycation. *Clinica Chimica Acta; International Journal of Clinical Chemistry*, 412(3–4), 277–285. <https://doi.org/10.1016/j.cca.2010.10.018>
- Barzegar, A., Moosavi-Movahedi, A., Sattarahmady, N., Hosseinpour-Faizi, M., Aminbakhsh, M., Ahmad, F., Saboury, A., Ganjali, M., & Norouzi, P. (2007). Spectroscopic studies of the effects of glycation of human serum albumin on L-Trp binding. *Protein and Peptide Letters*, 14(1), 13–18. <https://doi.org/10.2174/092986607779117191>
- Becke, A. D. (1992). Density-functional thermochemistry. I. The effect of the exchange-only gradient correction. *The Journal of Chemical Physics*, 96(3), 2155–2160. <https://doi.org/10.1063/1.462066>
- Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, A., & Haak, J. R. (1984). Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics*, 81(8), 3684–3690. <https://doi.org/10.1063/1.448118>
- Best, R. B., Zhu, X., Shim, J., Lopes, P. E. M., Mittal, J., Feig, M., & MacKerell, A. D. (2012). Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone  $\phi$ ,  $\psi$  and side-chain  $\chi_1$  and  $\chi_2$  dihedral angles. *Journal of Chemical Theory and Computation*, 8(9), 3257–3273. <https://doi.org/10.1021/ct300400x>
- Bhattacharya, A. A., Grüne, T., & Curry, S. (2000). Crystallographic analysis reveals common modes of binding of medium and long-chain fatty acids to human serum albumin 1 Edited by R. Huber. *Journal of Molecular Biology*, 303(5), 721–732. <https://doi.org/10.1006/jmbi.2000.4158>
- Brownlee, M. (1991). Glycosylation products as toxic mediators of diabetic complications. *Annual Review of Medicine*, 42(1), 159–166. <https://doi.org/10.1146/annurev.me.42.020191.001111>
- Carter, D. C., & Ho, J. X. (1994). Structure of serum albumin. *Advances in Protein Chemistry*, 45, 153–203. [https://doi.org/10.1016/S0065-3233\(08\)60640-3](https://doi.org/10.1016/S0065-3233(08)60640-3)
- Castro-Alvarez, A., Costa, A., & Vilarrasa, J. (2017). The performance of several docking programs at reproducing protein-macrolide-like crystal structures. *Molecules (Basel, Switzerland)*, 22(1), 136. <https://doi.org/10.3390/molecules22010136>
- Chaves, O. A., Soares, M. A. G., & Campos de Oliveira, M. C. (2021). Monosaccharides interact weakly with human serum albumin. Insights for the functional perturbations on the binding capacity of albumin.



- Carbohydrate Research*, 501, 108274. <https://doi.org/10.1016/j.carres.2021.108274>
- Chen, X., Zhang, Y., & Zhang, J. Z. H. (2005). An efficient approach for *ab initio* energy calculation of biopolymers. *The Journal of Chemical Physics*, 122(18), 184105. <https://doi.org/10.1063/1.1897382>
- Christensen, A. S., Kubař, T., Cui, Q., & Elstner, M. (2016). Semiempirical quantum mechanical methods for noncovalent interactions for chemical and biochemical applications. *Chemical Reviews*, 116(9), 5301–5337. <https://doi.org/10.1021/acs.chemrev.5b00584>
- Collins, M. A., & Bettens, R. P. A. (2015). Energy-based molecular fragmentation methods. *Chemical Reviews*, 115(12), 5607–5642. <https://doi.org/10.1021/cr500455b>
- Conev, A., Rigo, M. M., Devaurs, D., Fonseca, A. F., Kalavadwala, H., de Freitas, M. V., Clementi, C., Zanatta, G., Antunes, D. A., & Kavradi, L. E. (2023). EnGens: A computational framework for generation and analysis of representative protein conformational ensembles. *Briefings in Bioinformatics*, 24(4), 1–11. <https://doi.org/10.1093/bib/bbad242>
- Dantas, D. S., Oliveira, J. I. N., Lima Neto, J. X., da Costa, R. F., Bezerra, E. M., Freire, V. N., Caetano, E. W. S., Fulco, U. L., & Albuquerque, E. L. (2015). Quantum molecular modelling of ibuprofen bound to human serum albumin. *RSC Advances*, 5(61), 49439–49450. <https://doi.org/10.1039/C5RA04395F>
- Delley, B. (1990). An all-electron numerical method for solving the local density functional for polyatomic molecules. *The Journal of Chemical Physics*, 92(1), 508–517. <https://doi.org/10.1063/1.458452>
- Delley, B. (2000). From molecules to solids with the DMol3 approach. *The Journal of Chemical Physics*, 113(18), 7756–7764. <https://doi.org/10.1063/1.1316015>
- Ehrlich, S., Moellmann, J., & Grimme, S. (2013). Dispersion-corrected density functional theory for aromatic interactions in complex systems. *Accounts of Chemical Research*, 46(4), 916–926. <https://doi.org/10.1021/ar3000844>
- Fanali, G., di Masi, A., Trezza, V., Marino, M., Fasano, M., & Ascenzi, P. (2012). Human serum albumin: From bench to bedside. *Molecular Aspects of Medicine*, 33(3), 209–290. <https://doi.org/10.1016/j.mam.2011.12.002>
- Fox, S. J., Dziejdz, J., Fox, T., Tautermann, C. S., & Skylaris, C.-K. (2014). Density functional theory calculations on entire proteins for free energies of binding: Application to a model polar binding site. *Proteins*, 82(12), 3335–3346. <https://doi.org/10.1002/prot.24686>
- França, V. L. B., Amaral, J. L., Martins, Y. A., Caetano, E. W. S., Brunaldi, K., & Freire, V. N. (2022). Characterization of the binding interaction between atrazine and human serum albumin: Fluorescence spectroscopy, molecular dynamics and quantum biochemistry. *Chemico-Biological Interactions*, 366, 110130. <https://doi.org/10.1016/j.cbi.2022.110130>
- Ganesan, A., Coote, M. L., & Barakat, K. (2017). Molecular dynamics-driven drug discovery: Leaping forward with confidence. *Drug Discovery Today*, 22(2), 249–269. <https://doi.org/10.1016/j.drudis.2016.11.001>
- Gelpi, J., Hospital, A., Goñi, J. R., Orozco, M. (2015). Molecular dynamics simulations: Advances and applications. *Advances and Applications in Bioinformatics and Chemistry: AABC*, 8, 37–47. <https://doi.org/10.2147/AABC.S70333>
- Ghosh, R., & Kishore, N. (2022). Mechanistic physicochemical insights into glycation and drug binding by serum albumin: Implications in diabetic conditions. *Biochimie*, 193, 16–37. <https://doi.org/10.1016/j.biochi.2021.10.008>
- Ghuman, J., Zunsain, P. A., Petitpas, I., Bhattacharya, A. A., Otagiri, M., & Curry, S. (2005). Structural basis of the drug-binding specificity of human serum albumin. *Journal of Molecular Biology*, 353(1), 38–52. <https://doi.org/10.1016/j.jmb.2005.07.075>
- Gordon, M. S., Fedorov, D. G., Pruitt, S. R., & Slipchenko, L. V. (2012). Fragmentation methods: A route to accurate calculations on large systems. *Chemical Reviews*, 112(1), 632–672. <https://doi.org/10.1021/cr200093j>
- Grinter, S., & Zou, X. (2014). Challenges, applications, and recent advances of protein-ligand docking in structure-based drug design. *Molecules (Basel, Switzerland)*, 19(7), 10150–10176. <https://doi.org/10.3390/molecules190710150>
- Hamilton, J. A. (2013). NMR reveals molecular interactions and dynamics of fatty acid binding to albumin. *Biochimica Et Biophysica Acta*, 1830(12), 5418–5426. <https://doi.org/10.1016/j.bbagen.2013.08.002>
- Hand, D. J., McLachlan, G. J., & Basford, K. E. (1989). Mixture models: Inference and applications to clustering. *Applied Statistics*, 38(2), 384. <https://doi.org/10.2307/2348072>
- He, X., & Zhang, J. Z. H. (2005). A new method for direct calculation of total energy of protein. *The Journal of Chemical Physics*, 122(3), 31103. <https://doi.org/10.1063/1.1849132>
- Hollingsworth, S. A., & Dror, R. O. (2018). Molecular dynamics simulation for all. *Neuron*, 99(6), 1129–1143. <https://doi.org/10.1016/j.neuron.2018.08.011>
- Hu, X., Zeng, Z., Zhang, J., Wu, D., Li, H., & Geng, F. (2023). Molecular dynamics simulation of the interaction of food proteins with small molecules. *Food Chemistry*, 405(Pt A), 134824. <https://doi.org/10.1016/j.foodchem.2022.134824>
- Jaña, G. A., Mendoza, F., Osorio, M. I., Alderete, J. B., Fernandes, P. A., Ramos, M. J., & Jiménez, V. A. (2018). A QM/MM approach on the structural and stereoelectronic factors governing glycosylation by GTF-SI from *Streptococcus mutans*. *Organic & Biomolecular Chemistry*, 16(14), 2438–2447. <https://doi.org/10.1039/C8OB00284C>
- Jiang, N., Ma, J., & Jiang, Y. (2006). Electrostatic field-adapted molecular fractionation with conjugated caps for energy calculations of charged biomolecules. *The Journal of Chemical Physics*, 124(11), 114112. <https://doi.org/10.1063/1.2178796>
- Jones, R. O. (2015). Density functional theory: Its origins, rise to prominence, and future. *Reviews of Modern Physics*, 87(3), 897–923. <https://doi.org/10.1103/RevModPhys.87.897>
- Joseph, K. S., & Hage, D. S. (2010). The effects of glycation on the binding of human serum albumin to warfarin and L-tryptophan. *Journal of Pharmaceutical and Biomedical Analysis*, 53(3), 811–818. <https://doi.org/10.1016/j.jpba.2010.04.035>
- Khalifah, R. G., Baynes, J. W., & Hudson, B. G. (1999). Amadorins: Novel post-amadori inhibitors of advanced glycation reactions. *Biochemical and Biophysical Research Communications*, 257(2), 251–258. <https://doi.org/10.1006/bbrc.1999.0371>
- Kitchen, D. B., Decornez, H., Furr, J. R., & Bajorath, J. (2004). Docking and scoring in virtual screening for drug discovery: Methods and applications. *Nature Reviews. Drug Discovery*, 3(11), 935–949. <https://doi.org/10.1038/nrd1549>
- Klamt, A., & Schüürmann, G. (1993). COSMO: A new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. *Journal of the Chemical Society, Perkin Transactions 2*, (5), 799–805. <https://doi.org/10.1039/P29930000799>
- Krenzel, E. S., Chen, Z., & Hamilton, J. A. (2013). Correspondence of fatty acid and drug binding sites on human serum albumin: A two-dimensional nuclear magnetic resonance study. *Biochemistry*, 52(9), 1559–1567. <https://doi.org/10.1021/bi301458b>
- Li, J., Zhu, X., Yang, C., & Shi, R. (2010). Characterization of the binding of angiotensin II receptor blockers to human serum albumin using docking and molecular dynamics simulation. *Journal of Molecular Modeling*, 16(4), 789–798. <https://doi.org/10.1007/s00894-009-0612-0>
- Liu, J., & Herbert, J. M. (2016). Pair-pair approximation to the generalized many-body expansion: An alternative to the four-body expansion for *ab initio* prediction of protein energetics via molecular fragmentation. *Journal of Chemical Theory and Computation*, 12(2), 572–584. <https://doi.org/10.1021/acs.jctc.5b00955>
- Liu, Y., Yang, X., Gan, J., Chen, S., Xiao, Z.-X., & Cao, Y. (2022). CB-Dock2: Improved protein–ligand blind docking by integrating cavity detection, docking and homologous template fitting. *Nucleic Acids Research*, 50(W1), W159–W164. <https://doi.org/10.1093/nar/gkac394>
- Lv, Y., Liang, Q., Li, Y., Liu, X., Zhang, D., & Li, X. (2022). Study of the binding mechanism between hydroxytyrosol and bovine serum albumin using multispectral and molecular docking. *Food Hydrocolloids*, 122, 107072. <https://doi.org/10.1016/j.foodhyd.2021.107072>
- Malik, V. S., & Hu, F. B. (2015). Fructose and cardiometabolic health. *Journal of the American College of Cardiology*, 66(14), 1615–1624. <https://doi.org/10.1016/j.jacc.2015.08.025>
- Mardirossian, N., & Head-Gordon, M. (2017). Thirty years of density functional theory in computational chemistry: An overview and extensive

- assessment of 200 density functionals. *Molecular Physics*, 115(19), 2315–2372. <https://doi.org/10.1080/00268976.2017.1333644>
- Martínez-Rosell, G., Giorgino, T., & De Fabritiis, G. (2017). PlayMolecule ProteinPrepare: A Web Application for Protein Preparation for Molecular Dynamics Simulations. *Journal of Chemical Information and Modeling*, 57(7), 1511–1516. <https://doi.org/10.1021/acs.jcim.7b00190>
- Maruyama, Y., Igarashi, R., Ushiku, Y., & Mitsutake, A. (2023). Analysis of protein folding simulation with moving root mean square deviation. *Journal of Chemical Information and Modeling*, 63(5), 1529–1541. <https://doi.org/10.1021/acs.jcim.2c01444>
- Moeinpour, F., Mohseni-Shahri, F. S., Malaekheh-Nikouei, B., & Nassirli, H. (2016). Investigation into the interaction of losartan with human serum albumin and glycosylated human serum albumin by spectroscopic and molecular dynamics simulation techniques: A comparison study. *Chemico-Biological Interactions*, 257, 4–13. <https://doi.org/10.1016/j.cbi.2016.07.025>
- Nasiri, R., Bahrami, H., Zahedi, M., Moosavi-Movahedi, A. A., & Sattarahmady, N. (2010). A theoretical elucidation of glucose interaction with HSA's domains. *Journal of Biomolecular Structure & Dynamics*, 28(2), 211–226. <https://doi.org/10.1080/07391102.2010.10507354>
- Nocedal, J., & Wright, S. (1960). Numerical optimization. ITM web of conferences itmconflibrary of congress cataloging-in-publication, second.
- Pan, X.-L., Liu, W., & Liu, J.-Y. (2013). Mechanism of the glycosylation step catalyzed by human  $\alpha$ -galactosidase: A QM/MM metadynamics study. *The Journal of Physical Chemistry. B*, 117(2), 484–489. <https://doi.org/10.1021/jp308747c>
- Pearlman, D. A., Case, D. A., Caldwell, J. W., Ross, W. S., Cheatham, T. E., DeBolt, S., Ferguson, D., Seibel, G., & Kollman, P. (1995). AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules. *Computer Physics Communications*, 91(1-3), 1–41. [https://doi.org/10.1016/0010-4655\(95\)00041-D](https://doi.org/10.1016/0010-4655(95)00041-D)
- Perdew, J. P., Burke, K., & Ernzerhof, M. (1996). Generalized gradient approximation made simple. *Physical Review Letters*, 77(18), 3865–3868. <https://doi.org/10.1103/PhysRevLett.77.3865>
- Peters, T. (1995). Ligand binding by albumin. In *All about albumin* (pp. 76–132). Elsevier. <https://doi.org/10.1016/B978-012552110-9/50005-2>
- Petitpas, I., Grüne, T., Bhattacharya, A. A., & Curry, S. (2001). Crystal structures of human serum albumin complexed with monounsaturated and polyunsaturated fatty acids. *Journal of Molecular Biology*, 314(5), 955–960. <https://doi.org/10.1006/jmbi.2000.5208>
- Pongprayoon, P., & Gleeson, M. P. (2014). Probing the binding site characteristics of HSA: A combined molecular dynamics and cheminformatics investigation. *Journal of Molecular Graphics & Modelling*, 54, 164–173. <https://doi.org/10.1016/j.jmgm.2014.10.007>
- Pongprayoon, P., & Mori, T. (2018). The critical role of dimer formation in monosaccharides binding to human serum albumin. *Physical Chemistry Chemical Physics: PCCP*, 20(5), 3249–3257. <https://doi.org/10.1039/C7CP06324E>
- Potter, E. D., Herek, J. L., Pedersen, S., Liu, Q., & Zewail, A. H. (1992). Femtosecond laser control of a chemical reaction. *Nature*, 355(6355), 66–68. <https://doi.org/10.1038/355066a0>
- Qi, X., & Tester, R. F. (2019). Fructose, galactose and glucose – In health and disease. *Clinical Nutrition ESPEN*, 33, 18–28. <https://doi.org/10.1016/j.clnesp.2019.07.004>
- Rappe, A. K., Casewit, C. J., Colwell, K. S., Goddard, W. A., & Skiff, W. M. (1992). UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations. *Journal of the American Chemical Society*, 114(25), 10024–10035. <https://doi.org/10.1021/ja00051a040>
- Reynolds, T. M. (1963). Chemistry of Nonenzymic Browning I. The Reaction between Aldoses and Amines. *Advances in Food Research*, 12, 1–52. [https://doi.org/10.1016/S0065-2628\(08\)60005-1](https://doi.org/10.1016/S0065-2628(08)60005-1)
- Saha, A., & Raghavachari, K. (2015). Analysis of different fragmentation strategies on a variety of large peptides: implementation of a low level of theory in fragment-based methods can be a crucial factor. *Journal of Chemical Theory and Computation*, 11(5), 2012–2023. <https://doi.org/10.1021/ct501045s>
- Sattarahmady, N., Moosavi-Movahedi, A. A., Ahmad, F., Hakimelahi, G. H., Habibi-Rezaei, M., Saboury, A. A., & Sheibani, N. (2007). Formation of the molten globule-like state during prolonged glycation of human serum albumin. *Biochimica Et Biophysica Acta*, 1770(6), 933–942. <https://doi.org/10.1016/j.bbagen.2007.02.001>
- Sergio, L. M., Martins, Y. A., Amaral, J. L., França, V. L. B., de Freitas, C. F., Neto, A. M., Hioka, N., Ravanelli, M. I., Mareze-Costa, C., Claudio da Costa, S., Freire, V. N., & Brunaldi, K. (2021). Molecular insight on the binding of stevia glycosides to bovine serum albumin. *Chemico-Biological Interactions*, 344, 109526. <https://doi.org/10.1016/j.cbi.2021.109526>
- Shamsi, A., Ahmed, A., & Bano, B. (2018). Probing the interaction of anti-cancer drug temsirolimus with human serum albumin: Molecular docking and spectroscopic insight. *Journal of Biomolecular Structure & Dynamics*, 36(6), 1479–1489. <https://doi.org/10.1080/07391102.2017.1326320>
- Simard, J. R., Zunszain, P. A., Hamilton, J. A., & Curry, S. (2006). Location of high and low affinity fatty acid binding sites on human serum albumin revealed by NMR drug-competition analysis. *Journal of Molecular Biology*, 361(2), 336–351. <https://doi.org/10.1016/j.jmb.2006.06.028>
- Sittiwanchai, S., Japrun, D., Mori, T., & Pongprayoon, P. (2023). Structural and dynamic alteration of glycosylated human serum albumin in schiff base and amadori adducts: A molecular simulation study. *The Journal of Physical Chemistry. B*, 127(23), 5230–5240. <https://doi.org/10.1021/acs.jpcc.3c02048>
- Smith, G. R., & Sternberg, M. J. E. (2002). Prediction of protein–protein interactions by docking methods. *Current Opinion in Structural Biology*, 12(1), 28–35. [https://doi.org/10.1016/S0959-440X\(02\)00285-3](https://doi.org/10.1016/S0959-440X(02)00285-3)
- Söderhjelm, P., Aquilante, F., & Ryde, U. (2009). Calculation of protein–ligand interaction energies by a fragmentation approach combining high-level quantum chemistry with classical many-body effects. *The Journal of Physical Chemistry. B*, 113(32), 11085–11094. <https://doi.org/10.1021/jp810551h>
- Sousa, B. L., Barroso-Neto, I. L., Oliveira, E. F., Fonseca, E., Lima-Neto, P., Ladeira, L. O., & Freire, V. N. (2016). Explaining RANKL inhibition by OPG through quantum biochemistry computations and insights into peptide-design for the treatment of osteoporosis. *RSC Advances*, 6(88), 84926–84942. <https://doi.org/10.1039/C6RA16712H>
- Sudlow, G., Birkett, D. J., & Wade, D. N. (1976). Further characterization of specific drug binding sites on human serum albumin. *American Society for Pharmacology and Experimental Therapeutics*, 12, 1052–1061.
- Szudlarek, A., Sułkowska, A., Maciążek-Jurczyk, M., Chudzik, M., & Równicka-Zubik, J. (2016). Effects of non-enzymatic glycation in human serum albumin. Spectroscopic analysis. *Spectrochimica Acta. Part A, Molecular and Biomolecular Spectroscopy*, 152, 645–653. <https://doi.org/10.1016/j.saa.2015.01.120>
- Tayeh, N., Rungassamy, T., & Albani, J. R. (2009). Fluorescence spectral resolution of tryptophan residues in bovine and human serum albumins. *Journal of Pharmaceutical and Biomedical Analysis*, 50(2), 107–116. <https://doi.org/10.1016/j.jpba.2009.03.015>
- Thiel, W. (2014). Semiempirical quantum–chemical methods. *WIREs Computational Molecular Science*, 4(2), 145–157. <https://doi.org/10.1002/wcms.1161>
- Thornalley, P. J., Langborg, A., & Minhas, H. S. (1999). Formation of glyoxal, methylglyoxal and 3-deoxyglucosone in the glycation of proteins by glucose. *Biochemical Journal*, 344(1), 109–116. <https://doi.org/10.1042/bj3440109>
- Tkatchenko, A., & Scheffler, M. (2009). Accurate molecular van der waals interactions from ground-state electron density and free-atom reference data. *Physical Review Letters*, 102(7), 073005. <https://doi.org/10.1103/PhysRevLett.102.073005>
- Trozzi, F., Wang, X., & Tao, P. (2021). UMAP as a dimensionality reduction tool for molecular dynamics simulations of biomacromolecules: A comparison study. *The Journal of Physical Chemistry. B*, 125(19), 5022–5034. <https://doi.org/10.1021/acs.jpcc.1c02081>
- Trüb, B., Hostenstein, C. G., Fischer, R. W., & Winterhalter, K. H. (1980). Nonenzymatic glycosylation of proteins. A warning. *The Journal of Biological Chemistry*, 255(14), 6717–6720. [https://doi.org/10.1016/S0021-9258\(18\)43630-7](https://doi.org/10.1016/S0021-9258(18)43630-7)

- Vanommeslaeghe, K., Hatcher, E., Acharya, C., Kundu, S., Zhong, S., Shim, J., Darian, E., Guvench, O., Lopes, P., Vorobyov, I., & Mackerell, A. D. (2009). CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *Journal of Computational Chemistry*, 31(4), 671–690. <https://doi.org/10.1002/jcc.21367>
- Wang, W., Gan, N., Sun, Q., Wu, D., Gan, R., Zhang, M., Tang, P., & Li, H. (2019). Study on the interaction of ertugliflozin with human serum albumin in vitro by multispectroscopic methods, molecular docking, and molecular dynamics simulation. *Spectrochimica Acta. Part A, Molecular and Biomolecular Spectroscopy*, 219, 83–90. <https://doi.org/10.1016/j.saa.2019.04.047>
- Wang, X., Liu, J., Zhang, J. Z. H., & He, X. (2013). Electrostatically embedded generalized molecular fractionation with conjugate caps method for full quantum mechanical calculation of protein energy. *The Journal of Physical Chemistry. A*, 117(32), 7149–7161. <https://doi.org/10.1021/jp400779t>
- Wang, Y., Yu, H., Shi, X., Luo, Z., Lin, D., & Huang, M. (2013). Structural mechanism of ring-opening reaction of glucose by human serum albumin. *The Journal of Biological Chemistry*, 288(22), 15980–15987. <https://doi.org/10.1074/jbc.M113.467027>
- Wani, T. A., Bakheit, A. H., Al-Majed, A. A., Altwaijry, N., Baquaysh, A., Aljuraissy, A., & Zargar, S. (2021). Binding and drug displacement study of colchicine and bovine serum albumin in presence of azithromycin using multispectroscopic techniques and molecular dynamic simulation. *Journal of Molecular Liquids*, 333, 115934. <https://doi.org/10.1016/j.molliq.2021.115934>
- Yang, F., Bian, C., Zhu, L., Zhao, G., Huang, Z., & Huang, M. (2007). Effect of human serum albumin on drug metabolism: Structural evidence of esterase activity of human serum albumin. *Journal of Structural Biology*, 157(2), 348–355. <https://doi.org/10.1016/j.jsb.2006.08.015>
- Zaman, A., Arif, Z., & Alam, K. (2018). Fructose-human serum albumin interaction undergoes numerous biophysical and biochemical changes before forming AGEs and aggregates. *International Journal of Biological Macromolecules*, 109, 896–906. <https://doi.org/10.1016/j.ijbiomac.2017.11.069>
- Zhang, D. W., & Zhang, J. Z. H. (2003). Molecular fractionation with conjugate caps for full quantum mechanical calculation of protein-molecule interaction energy. *The Journal of Chemical Physics*, 119(7), 3599–3605. <https://doi.org/10.1063/1.1591727>
- Zhang, Y., Lee, P., Liang, S., Zhou, Z., Wu, X., Yang, F., & Liang, H. (2015). Structural basis of non-steroidal anti-inflammatory drug diclofenac binding to human serum albumin. *Chemical Biology & Drug Design*, 86(5), 1178–1184. <https://doi.org/10.1111/cbdd.12583>
- Zhao, H., & Cafilisch, A. (2015). Molecular dynamics in drug design. *European Journal of Medicinal Chemistry*, 91, 4–14. <https://doi.org/10.1016/j.ejmech.2014.08.004>
- Zunzain, P. A., Ghuman, J., Komatsu, T., Tsuchida, E., & Curry, S. (2003). Crystal structural analysis of human serum albumin complexed with hemin and fatty acid. *BMC Structural Biology*, 3(1), 6. <https://doi.org/10.1186/1472-6807-3-6>
- Zunzain, P. A., Ghuman, J., McDonagh, A. F., & Curry, S. (2008). Crystallographic analysis of human serum albumin complexed with 4Z,15E-Bilirubin-IX $\alpha$ . *Journal of Molecular Biology*, 381(2), 394–406. <https://doi.org/10.1016/j.jmb.2008.06.016>