

CB-Dock2

Implementando a detecção de cavidades em proteínas guiada por docagem cega

Renato Massaharu Hassunuma
Aguinaldo Robinson de Souza



CB-Dock2

Implementando a detecção de cavidades em proteínas guiada por docagem cega

Renato Massaharu Hassunuma

Professor Titular do Curso de Biomedicina
Universidade Paulista – UNIP
Câmpus Bauru

Aguinaldo Robinson de Souza

Professor Adjunto Doutor
Departamento de Química
Faculdade de Ciências
Universidade Estadual Paulista “Júlio de Mesquita Filho” – UNESP,
Câmpus Bauru

1ª Edição / 2024
Bauru, SP

© Renato Massaharu Hassunuma.

Conselho Editorial

PROF. DR. MAURICIO IKEDA YOGUIM

Professor Bolsista do Departamento de Química da Universidade Estadual Paulista “Júlio de Mesquita Filho” – UNESP, Faculdade de Ciências – Câmpus Bauru

PROF. DR. VALDECIR FARIAS XIMENES

Professor Associado do Departamento de Química da Universidade Estadual Paulista “Júlio de Mesquita Filho” – UNESP, Faculdade de Ciências – Câmpus Bauru

Capa e Design

PROF. DR. RENATO MASSAHARU HASSUNUMA

Dados Internacionais de Catalogação na Publicação (CIP)
(BENITEZ Catalogação Ass. Editorial, MS, Brasil)

H284c Hassunuma, Renato Massaharu
1.ed. CB-Dock2: Implementando a detecção de cavidades em proteínas guiada por docagem cega [livro eletrônico] / Renato Massaharu Hassunuma, Aguinaldo Robinson de Souza. – 1. ed. – Bauru, SP : Canal 6, 2024.
PDF

ISBN 978-85-7917-652-4
DOI 10.52050/9788579176524

1. Bioquímica. 2. Bioinformática. 3. Docagem cega. 4. Proteínas.
I. Hassunuma, Renato Massaharu. II. Souza, Aguinaldo Robinson de.
III. Título.

04-2024/228

CDD 572

Índice para catálogo sistemático:
1. Bioquímica 572

Bibliotecária : Aline Grazielle Benitez CRB-1/3129

Agradecimentos

Agradecemos o **Prof. Dr. Mauricio Ikeda Yoguim** e o **Prof. Dr. Valdecir Farias Ximenes**, membros do Conselho Editorial deste livro, pelas suas valiosas contribuições.

*Prof. Dr. Renato Massaharu Hassunuma e
Prof. Dr. Aginaldo Robinson de Souza*



“A fúria do vendaval mostrou que o navio era seguro”.
(Edgar A. Poe. Manuscrito encontrado em uma garrafa, 1833)



Apresentação

Esta obra foi desenvolvida a partir de esforços conjuntos do **Prof. Dr. Renato Massaharu Hassunuma**, professor titular do Curso de Biomedicina da Universidade Paulista – UNIP, Câmpus Bauru, sob supervisão do **Prof. Sênior Aguinaldo Robinson de Souza**, professor titular da Universidade Estadual Paulista “Júlio de Mesquita Filho” – UNESP, Câmpus Bauru.

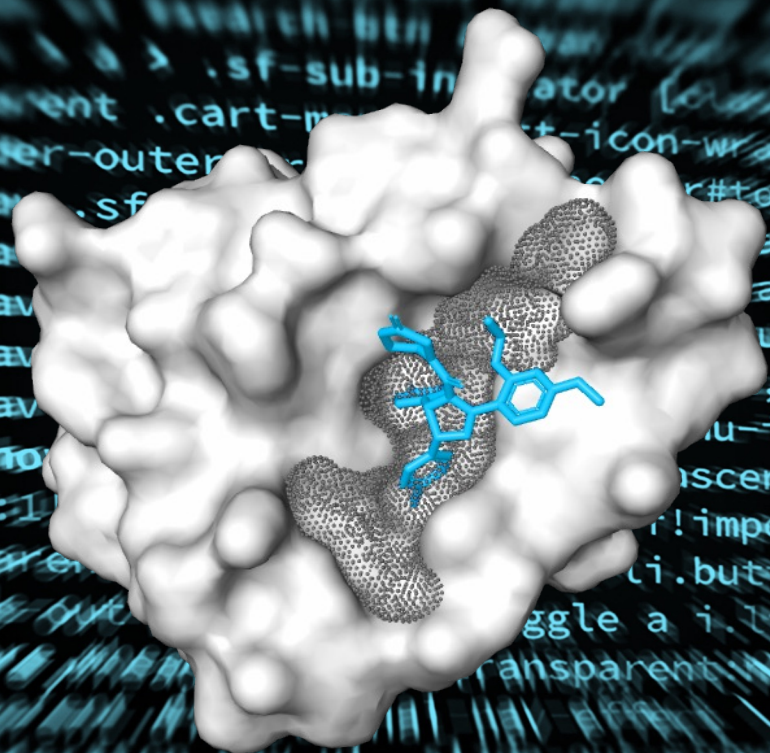


Sumário

Parte 1: Docagem molecular e o CB-Dock2®	10
1. Docagem molecular	11
2. Aplicações dos programas de docagem molecular	12
3. Docagem cega	13
4. CB-Dock2®	14
5. Funções do CB-Dock2®	15
6. Aplicações do CB-Dock2® em pesquisas	16
Parte 2: Modelos usados na docagem cega	18
1. Modelos de docagem cega	19
2. Albumina	20
3. Estrutura química da albumina	21
4. Ibuprofeno	22
5. Estrutura química do ibuprofeno	23
Parte 3: Arquivos de docagem cega	24
1. Arquivos para docagem cega	25
2. Arquivo PDB da proteína	26
3. Arquivos do ligante	27
Parte 4: Página inicial do CB-Dock2®	30

1. Página inicial do CB-Dock2®	31
2. Seção 1: Carregamento da proteína (<i>Upload protein</i>)	32
3. Seção 2: Carregamento do ligante (<i>Upload ligand</i>)	34
4. Seção 3: Mais parâmetros (<i>More parameters</i>)	37
Parte 5: Detecção de Cavidades do CB-Dock2®	41
1. Detecção de cavidades do CB-Dock2®	42
2. Seção 1: Detecção de cavidades baseada na estrutura (<i>Structure-based cavity detection</i>)	43
2.1. Subseção 1: Consulta A (<i>Query A</i>)	44
2.2. Subseção 2: Visualizador NGL® (<i>NGL®-viewer</i>)	45
2.3. Subseção 3: Quadro de informações das cavidades	48
3. Seção 2: Detecção de cavidades baseada em modelos (<i>Template-based cavity detection</i>)	49
3.1. Subseção 1: Sequências de resíduos de aminoácidos da proteína modelo e da proteína submetida – Consulta A (<i>Query A</i>)	50
3.2. Subseção 2: Visualizador NGL® (<i>NGL® -viewer</i>)	51
3.3. Subseção 3: Quadro de informações das cavidades	53
Parte 6: Docagem cega no CB-Dock2®	57
1. Docagem cega no CB-Dock2®	58
2. Seção 1: Cavidades detectadas e poses de ligação de ligante previstas (<i>Detected cavities and predicted ligand-binding poses</i>)	59
2.1. Subseção 1: Informações gerais	60
2.2. Subseção 2: Consulta A (<i>Query A</i>)	61
2.3. Subseção 3: Quadro de informações gerais	62
2.4. Subseção 4: Visualizador NGL® (<i>NGL®-viewer</i>)	64

3. Seção 2: Docagem cega baseada em estrutura (<i>Structure-based blind docking</i>)	66
3.1. Subseção 1: Informações gerais	68
3.2. Subseção 2: Visualizador NGL [®] (NGL [®] -viewer)	69
3.3. Subseção 3: Quadro de resultados	70
3.4. Interpretando os resultados de docagem cega baseada em estrutura (<i>Structure-based blind docking</i>)	71
4. Seção 3: Docagem cega baseada em modelo (<i>Template-based blind docking</i>)	72
4.1. Subseção 1: Informações gerais	74
4.2. Subseção 2: Visualizador NGL [®] (NGL [®] -viewer)	75
4.3. Subseção 3: Quadro de resultados	76
4.4. Subseção 4: Informações dos modelos	77
4.5. Interpretando os resultados de docagem cega baseada em modelo (<i>Template-based blind docking</i>)	78
4.6. Comparando os resultados de docagem cega baseada em estrutura (<i>Structure-based blind docking</i>) e docagem cega baseada em modelo (<i>Template-based blind docking</i>)	79
Créditos das figuras	80
Referências	87
Anexos	94
Anexo 1 – Abreviaturas de uma e três letras dos aminoácidos	95
Anexo 2 – Representação das interações químicas no visualizador NGL [®]	96
Anexo 3 – Esquemas de cores do visualizador NGL [®]	97



PARTE 1

Docagem molecular e o CB-Dock2[®]

1. Docagem molecular

Segundo o dicionário Collins, a **docagem molecular (DM)**, também conhecida como simulação de acoplamento molecular, ancoragem molecular ou *docking* molecular, corresponde à um método capaz de prever a orientação preferencial de uma molécula em relação à uma segunda, à qual está ligada de forma estável. O termo remete ao ato de atracar uma embarcação em uma doca (Figura 1) (Collins English Dictionary, 2023).

Na DM, é possível prever como ocorre a(s) interação(ões) de um ligante com uma proteína de estrutura tridimensional conhecida. Pode ser considerada uma ferramenta muito importante em Bioquímica Estrutural, especialmente no estudo de fármacos com biomoléculas (Morris, Lim-Wilby, 2008).

De forma geral, o processo de DM envolve a realização de duas etapas: a primeira é denominada pose, quando o programa prevê a conformação do ligante e de sua posição no sítio ativo; e a segunda é a avaliação da afinidade de ligação entre os componentes do complexo formado (Meng, Zhang, Mezei, Cui, 2011).

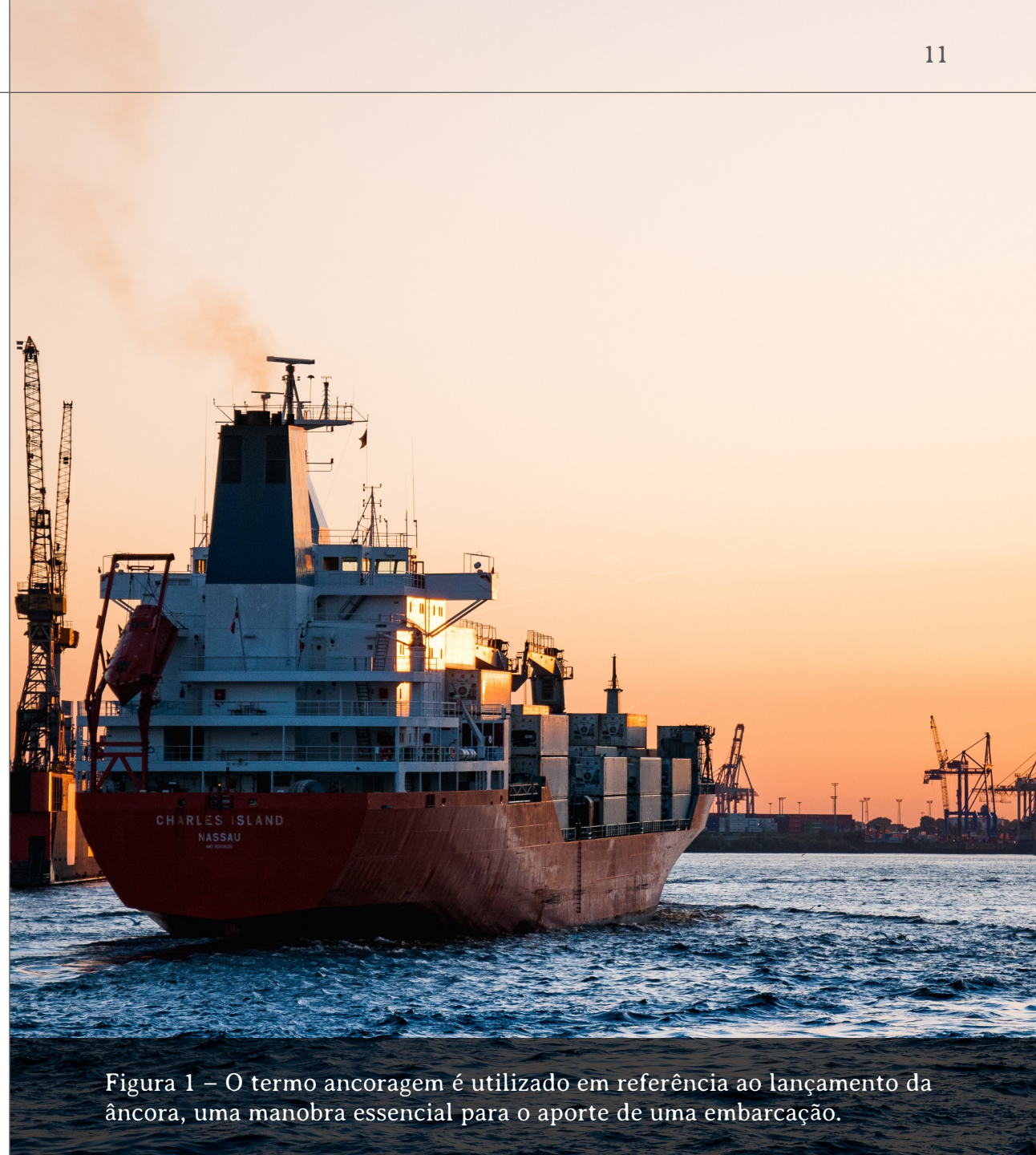


Figura 1 – O termo ancoragem é utilizado em referência ao lançamento da âncora, uma manobra essencial para o aporte de uma embarcação.

2. Aplicações dos programas de docagem molecular

Os métodos de DM são utilizados principalmente em pesquisas farmacêuticas, em especial na descoberta e desenvolvimento de novos fármacos (Figura 2), bem como em estudos de sistemas bioquímicos complexos (Ferreira, Dos Santos, Oliva, Andricopulo, 2015).

Em especial, o CB-Dock2[®] pode ser utilizado para:

1. Caracterizar a estrutura de proteínas e sugerir novos alvos terapêuticos em estudos biológicos e farmacêuticos;
2. Classificar fármacos de acordo com suas afinidades de ligação por determinados alvos;
3. Estimar e estudar possíveis interações entre ligantes e moléculas-alvo (Yang Cao Lab, 2021).

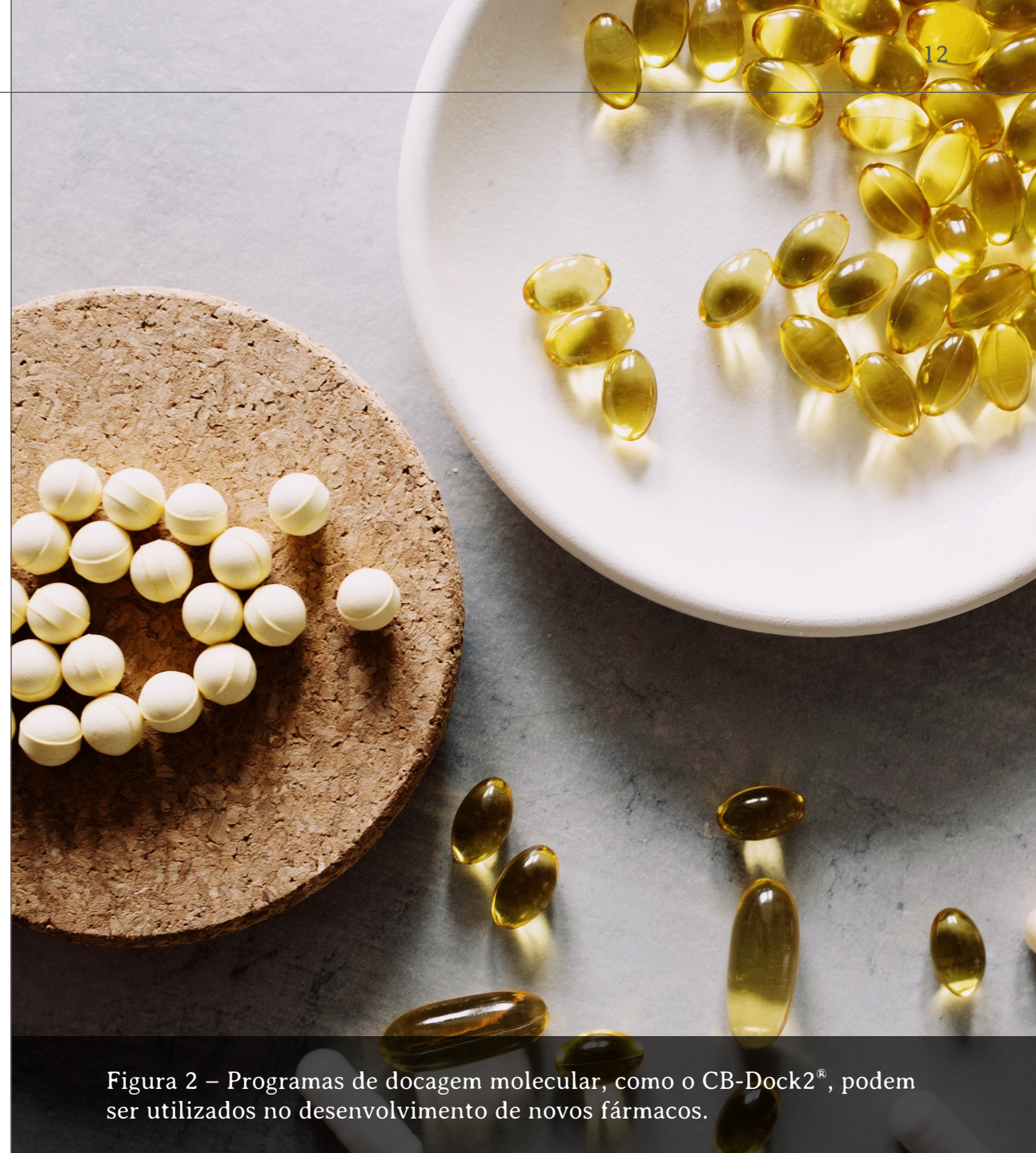


Figura 2 – Programas de docagem molecular, como o CB-Dock2[®], podem ser utilizados no desenvolvimento de novos fármacos.

3. Docagem cega

A **docagem cega (DC)** refere-se acoplar um determinado ligante a toda superfície de um proteína sem qualquer conhecimento prévio a respeito das cavidades-alvo, como quantidade e localização. Este processo envolve vários testes e cálculos para determinar uma posição mais favorável do complexo ligante-proteína (Hassam, Alhossary, Mu, Kwoh, 2017).

Vale ressaltar que a DC corresponde a uma abordagem inicial no processo de DM onde toda a superfície de uma proteína é investigada para localização de possíveis sítios de ligação. É um excelente recurso, por exemplo, para um mapeamento inicial de ligações de candidatos a medicamentos com proteínas (Figura 3) (Hetényi, van der Spoel, 2006).

A DC permite rastrear enormes quantidades de ligantes em um tempo relativamente curto e com precisão, sem necessidade de conhecer ou determinar um sítio ativo. Porém, por outro lado, são necessárias outras abordagens posteriores, usando outros programas de DM como, por exemplo, o AutoDock Vina[®], para obter mais informações sobre o complexo ligante-proteína (Hassam, Alhossary, Mu, Kwoh, 2017).



Figura 3 – Na DC, a interação de uma proteína com diferentes tipos de candidatos a medicamentos pode ser testada.

4. CB-Dock2®

O **CB-Dock**® (*Cavity-detection guided blind docking*) corresponde a um servidor *web* para docagem cega desenvolvido em 2019 pelo *Yang Cao Lab – Structural Bioinformatics Research Group* da Universidade de Sichuan na China. Está disponível gratuitamente no *link*: <http://clab.labshare.cn/cb-dock/php/index.php> (Figura 3) (Yang Cao Lab, 2021).

A confiabilidade do **CB-Dock**® foi verificada em uma pesquisa desenvolvida em 2020 por Liu, Grimm, Dai, Hou, Xiao e Cao, comparando os resultados obtidos pelo servidor com outros programas tradicionalmente utilizados em DM.

Em 2022, foi divulgado a segunda versão do servidor *web*, o **CB-Dock2**®, também disponível para acesso gratuito no *link*: <https://cadd.labshare.cn/cb-dock2/php/index.php>, que apresentou uma taxa de sucesso de DM de 16% superior à versão original e aumento de 4 a 5 vezes na velocidade de busca por cavidades (Liu, Yang, Gan, Chen, Xiao, Cao, 2022).



CB-Dock2 is an improved version of the **CB-Dock** server for protein-ligand blind docking, integrating cavity detection, docking and homologous template fitting.

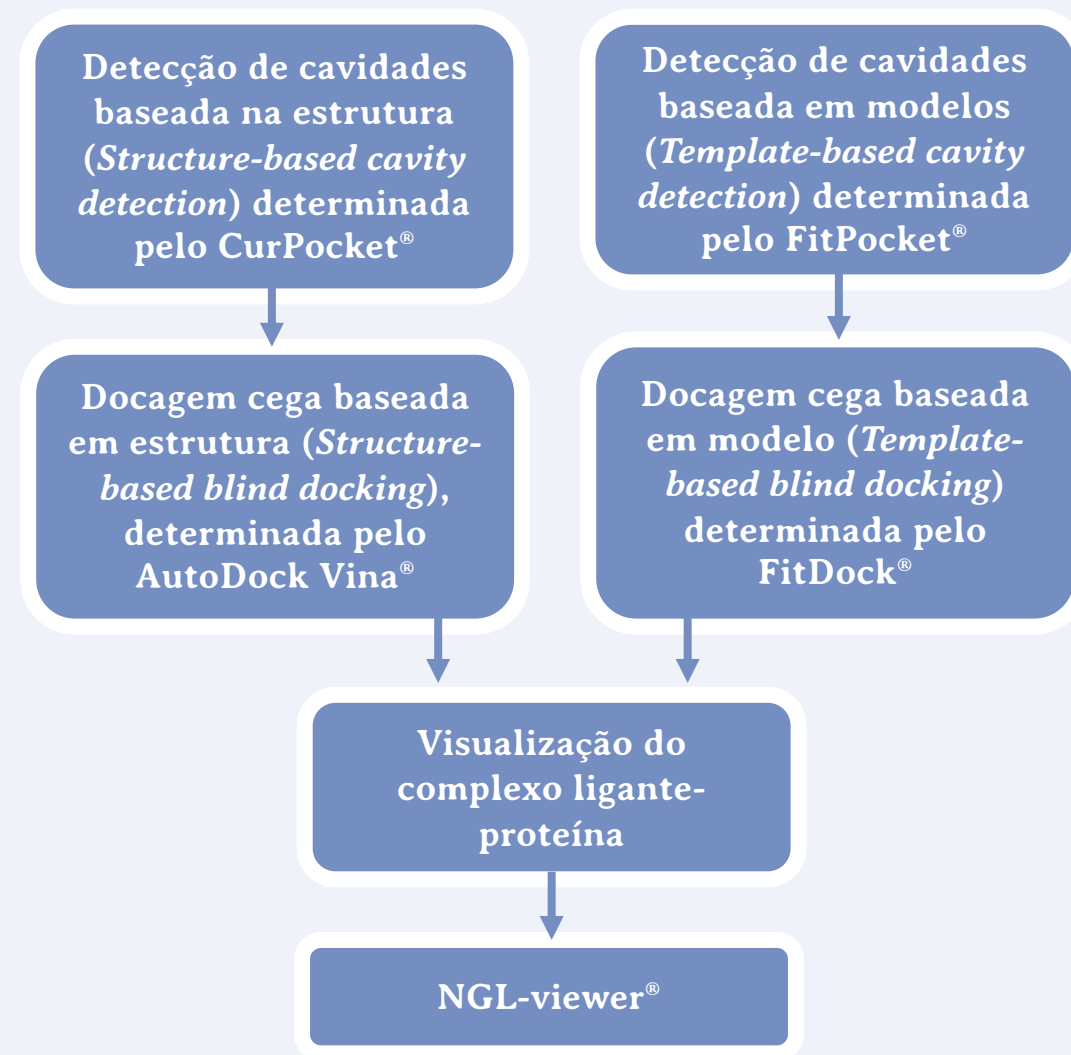
Given the three-dimensional (3D) structure of a protein and a ligand, we can predict their binding sites and affinity for computer-aided drug discovery.

Figura 3 – Servidor *web* CB-Dock2®.

5. Funções do CB-Dock2[®]

De forma resumida (Figura 4), o servidor CB-Dock2[®] realiza três atividades principais:

- **Detecção de cavidades baseada na estrutura (*Structure-based cavity detection*)**, onde a localização das cavidades é realizada pelo algoritmo **CurPocket[®]**. Os resultados obtidos são usados para **Docagem cega baseada em estrutura (*Structure-based blind docking*)** realizada pelo programa **AutoDock Vina[®]**. Nesta análise, as cavidades são denominadas **CurPockets**;
- **Detecção de cavidades baseada em modelos (*Template-based cavity detection*)**, onde a localização das cavidades é realizada pelo algoritmo **FitPocket[®]**. Os resultados são usados para **Docagem cega baseada em modelo (*Template-based blind docking*)** realizada pelo programa **FitDock[®]**. Neste caso, as cavidades são denominadas **FitPockets**;
- **Visualização do complexo ligante-proteína** por meio do visualizador **NGL[®] (NGL-viewer[®])** (Liu, Yang, Gan, Chen, Xiao, Cao, 2022).



Fonte: Autores, 2023.

Figura 4 – Fluxograma do funcionamento do CB-Dock2[®].

6. Aplicações do CB-Dock2[®] em pesquisas

Devido ao fato de o CB-Dock2[®] ter sido desenvolvido na *Sichuan University* na China, o mesmo tem sido muito utilizado na investigação de fitoterápicos conhecidos na Medicina Tradicional Chinesa (Figura 5), como:

- Raiz de *Huang Qi (Astragalus membranaceus)* (Tao, Huang, Wang, Kang, Wang, Xian, 2020);
- Composto *Shen Qi Wan* obtida a mistura de raízes e ervas, entre elas: raiz de *Huang Qi (Astragalus membranaceus)* (erva principal) e raiz da *Dang Shen (Codonopsis pilosula)* (Ye, Deng, Zhang, Liu, Liu, Xie, Wan, Huang, Zhang, Xi, Zhang, 2021);
- Preparação injetável *Xiyanping* obtida a partir do extrato da erva *Chuan Xin Lian (Andrographis paniculata)* (Xia, Zhang, Yang, Chen, Ye, Yan, 2022);
- Compostos derivados da videira Lei Gong Teng (*Tripterygium wilfordii*) (Wang, An, Yin, Li, Sun, Lan, Zhao, Li, 2021; Wang, Liu, Wu, Xu, Li, Feng, 2022);
- Óleo essencial da artemísia chinesa (*Artemisia argyi*) (Lu, Tang, Zhou, Lai, Liu, Duan, 2022);

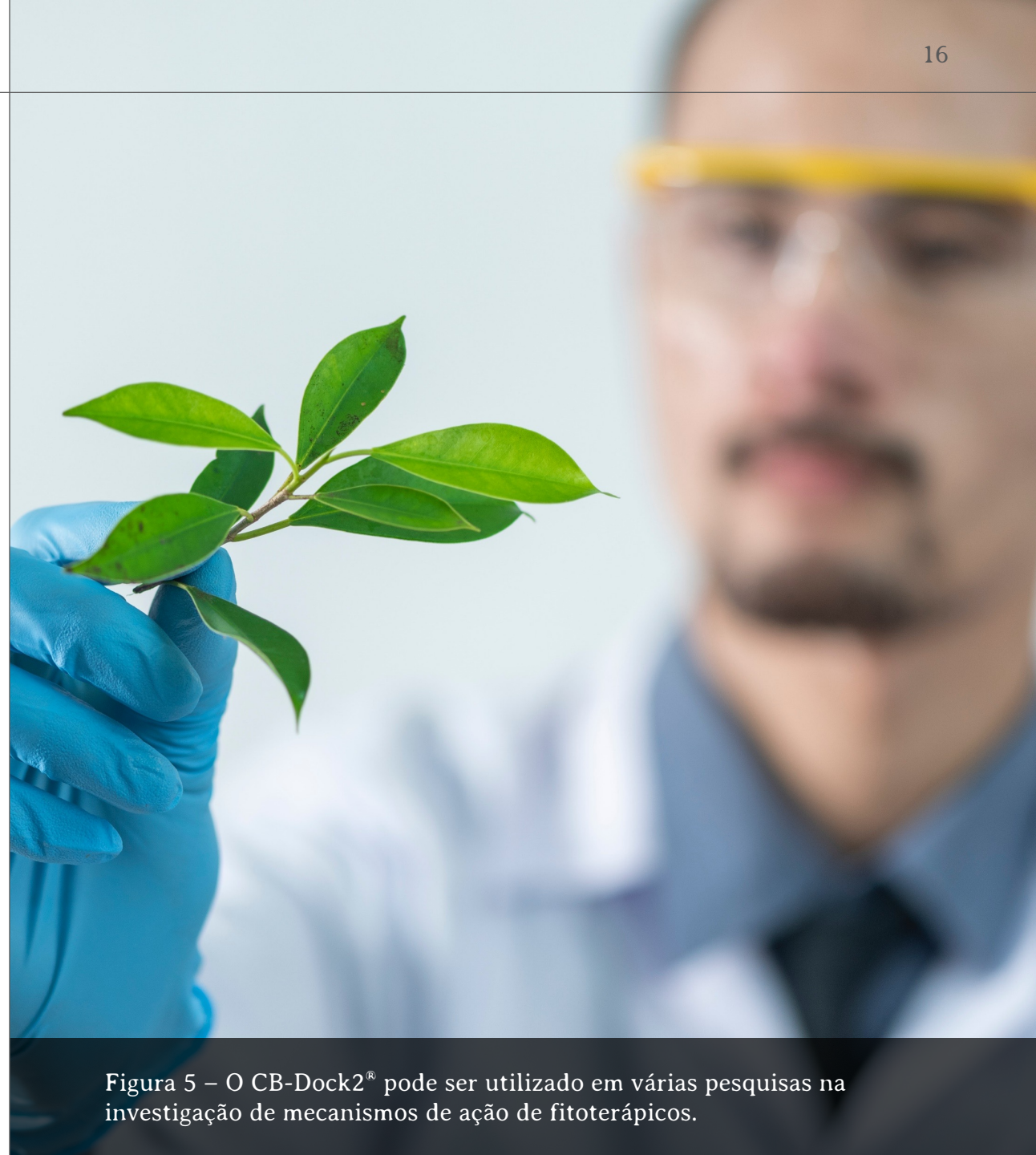


Figura 5 – O CB-Dock2[®] pode ser utilizado em várias pesquisas na investigação de mecanismos de ação de fitoterápicos.

- Composto *Shenbao Heji* obtido a partir da mistura de componentes como: raízes de *Huang Qi* (*Astragalus membranaceus*), *Bai Shao* (*Paeoniae alba*), Chuan Niu Xi (*Cyathulae officinalis*), folha de *Yin Yang Huo* (*Epimedii epimedium*), caule de *Ye Jiao Teng* (*Polygoni multiflori*), e *Mu Li* (concha de ostra) (Lv, Wang, Zhu, Jiang, 2022).

Além disso, foi utilizado em várias pesquisas com o objetivo de analisar os mecanismos de ação multialvo de diferentes compostos em doenças humanas:

- Patogenia dos compostos Amadori (Fru-Trp e Glc(α 1,4)-Fru-Trp) produzidos na reação de Maillard, na lesão pulmonar aguda e no diabetes (Bai, Li, Wang, Liang, He, Hu, Chen, 2021);
- Ação de flavonoides e fenóis na doença de Alzheimer (Zhang, Hu, Shao, Qin, Chen, Wang, Zhou, Cai, 2021);
- Mecanismos de ação de desoxiantocianidinas fermentadas na colite ulcerativa (Bai, Wang, Lan, Wu, Liang, Huang, Wu, Wang, Chen, 2022);
- Ação do pracinostat, um inibidor da enzima hidroxamato histona desacetilase (HDAC) que possui atividade antitumoral (Figura 6) (Ganai, 2021).



Figura 6 – O CB-Dock2[®] pode ser utilizado em pesquisa para investigar a ação de diferentes fármacos.



PARTE 2

Modelos usados na docagem cega

1. Modelos de docagem cega

Os programas de docagem cega fazem uso de arquivos referentes pelo menos uma proteína e seu ligante (Figura 7). Este ligante pode ser, por exemplo, um inibidor ou um fármaco que interage com a proteína estudada.

No processo de docagem cega, geralmente são pesquisadas proteínas e ligantes cuja ligação e sítios de ligação ainda não são conhecidos. Entretanto, como este livro tem uma finalidade didática e com o objetivo de facilitar o entendimento por novos usuários, serão utilizados, como modelo, a albumina sérica humana (HSA) e o ibuprofeno (IBP), cujas ligações já são conhecidas e bem estabelecidas na literatura.

Assim, serão apresentadas a seguir algumas informações gerais sobre estes modelos usados neste livro.

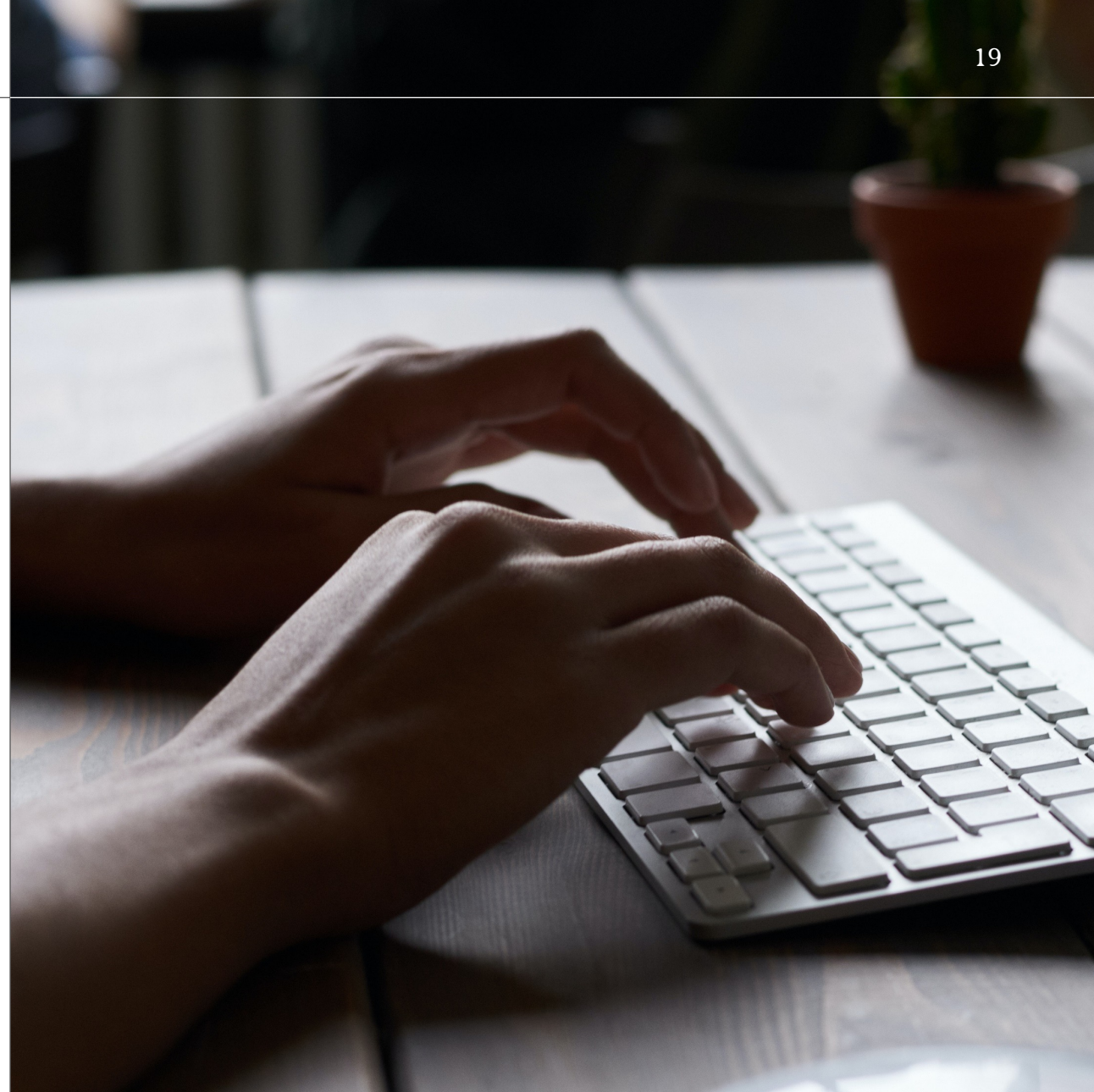


Figura 7 – Para usar um programa de docagem cega, o usuário deve escolher primeiramente a proteína e ligante a serem estudados.

2. Albumina

O termo albumina deriva da palavra *albus*, que em latim significa branco e refere-se à cor da clara cozida (Figura 8) (Fanali, di Masi, Trezza, Marino, Fasano, Ascenzi, 2012).

A albumina presente no soro humano (HSA) corresponde à proteína mais abundante do plasma sanguíneo, presente numa concentração de 3,5 a 5 g/dL, o que representa a metade do conteúdo total de proteínas. Cerca de 10 a 15 g de HSA são secretadas no sangue pelo fígado diariamente. Suas principais funções são:

- Manutenção da pressão osmótica plasmática;
- Transportador de ligantes endógenos (bilirrubina, íons, ácidos graxos);
- Transportador de ligantes exógenos (drogas como o IBP, varfarina, metotrexato, e muitos outros).

Em testes laboratoriais, pode ser utilizada como indicador de estado nutricional, função hepática, entre outros (Moman, Gupta, Varacallo, 2022).



Figura 8 – O termo albumina refere-se à cor branca da clara do ovo após cozimento.

3. Estrutura química da albumina

A HSA (Figura 9) apresenta 585 resíduos de aminoácidos, massa molecular de 66.438 Da (Mishra, Heath, 2021) e dimensões aproximadas de 80 x 80 x 30 Å (Fanali, di Masi, Trezza, Marino, Fasano, Ascenzi, 2012). É formada pelos domínios I (resíduos 1 a 195), II (resíduos 196 a 383) e III (resíduos 384 a 585), sendo que cada domínio é subdividido em 2 subunidades idênticas (A e B), que apresentam 4 e 6 alfa-hélices, respectivamente, as quais correspondem a 68% da sua estrutura secundária. Sua estrutura terciária é classificada como globular, com formato semelhante ao de um coração. Não apresenta fitas beta e possui 17 pares de pontes dissulfeto (Raoufinia, Mota, Keyhanvar, Safari, Shamekhi, Abdolalizadeh, 2016).

Apresenta 9 sítios ativos para ligação de ácidos graxos (FA): FA1 (ou sítio 3 de fármacos ou DS3), FA2-FA7 (ou sítio 1 de fármacos ou DS1), FA3-FA4 (ou sítio 2 de fármacos ou DS2), FA5 (no subdomínio IIIB), FA6 (entre os subdomínios IIA e IIB), FA8 e FA9 (ambos localizados no espaço entre os subdomínios IB e IIA de um lado e IIB, IIIA e IIIB do outro) (Ascenzi, di Masi, Fanali, Fasano, 2015; Czub, Stewart, Shabalin, Minor, 2022).



Legenda de cores		
Domínios	Cor	Amostra de cor
IA	Branco (<i>white</i>)	
IB	Amarelo-matiz (<i>yellowtint</i>)	
IIA	Rosa-matiz (<i>pinktint</i>)	
IIB	Azul-matiz (<i>bluetint</i>)	
IIIA	Verde-matiz (<i>greentint</i>)	
IIIB	Cinza (<i>grey</i>)	

Figura 9 – Estrutura secundária da albumina, com indicação dos subdomínios e sítios de ligação a ácidos graxos.

4. Ibuprofeno

O ibuprofeno (IBP) é um anti-inflamatório não esteroideal desenvolvido e patenteado em 1961 por Stewart Adams, farmacologista da *Research Department for the Boots Pure Drug Company Ltd.* Interessantemente, foi o próprio Adams que tomou a primeira dose do fármaco para ajudar a curar de uma ressaca. O IBP foi liberado para uso sob prescrição em 1969 no Reino Unido e em 1983, foi autorizada a venda sem prescrição médica devido seu perfil de segurança e tolerabilidade (Figura 10) (Halford, Lordkipanidzé, Watson, 2012).

Atualmente, é encontrado na forma de comprimidos, cápsulas e gotas para uso oral (AHFS, 2022), sendo prescrito como analgésico, antipirético e anti-inflamatório para o tratamento de osteoartrite, artrite idiopática juvenil e dor aguda em condições músculo-esqueléticas, além da artrite reumatoide (Halford, Lordkipanidzé, Watson, 2012).

Sua ação anti-inflamatória decorre da inibição não seletiva da enzima ciclo-oxigenase-2, produzida nos tecidos durante o processo inflamatório. Esta inibição impede a síntese de prostaglandinas, importantes mediadores da inflamação (Orlando, Lucido, Malkowski, 2015).

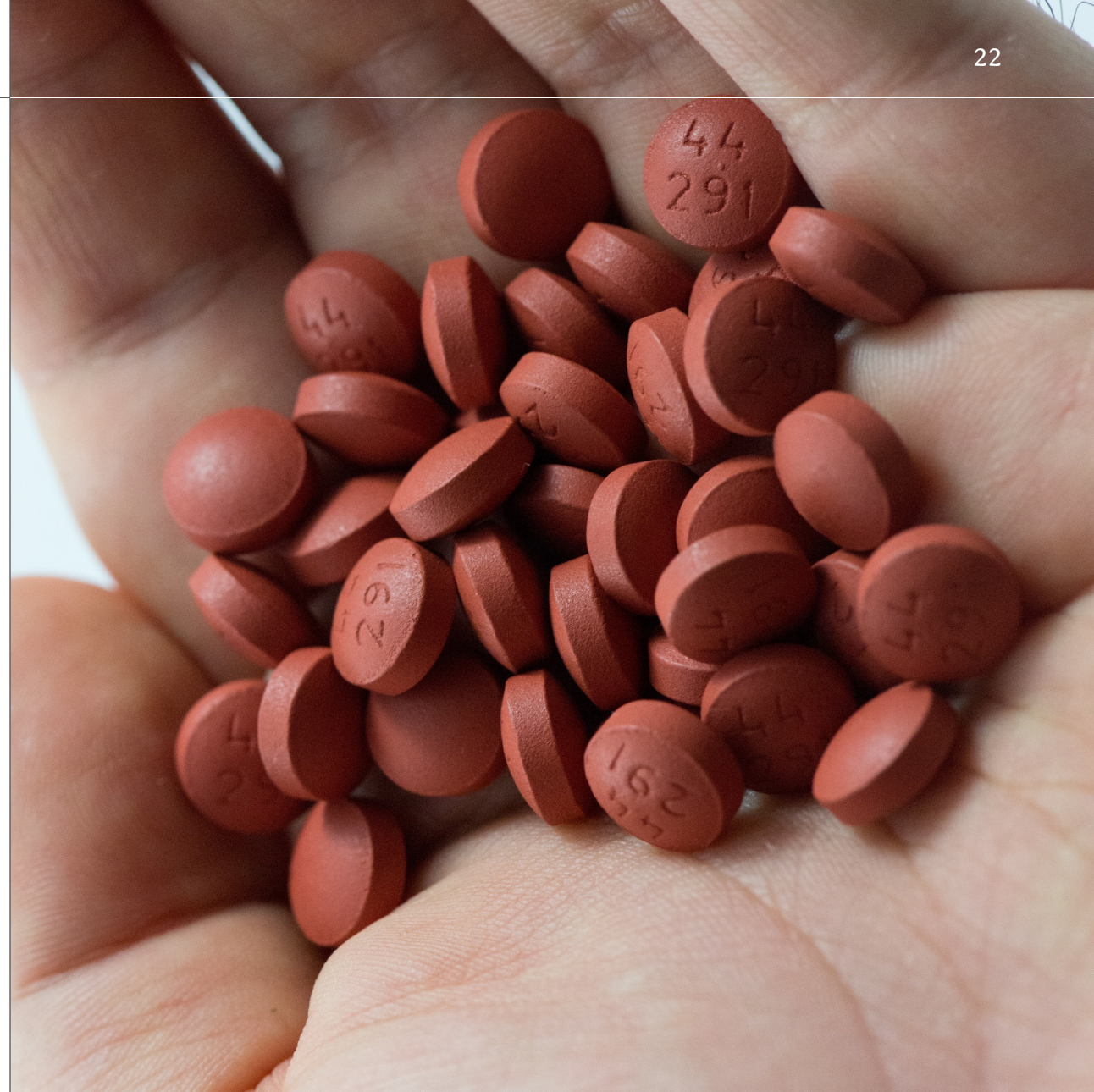


Figura 10 – O ibuprofeno pode ser comprado no formato de comprimidos, sem necessidade de prescrição médica.

5. Estrutura química do ibuprofeno

O IBP (Figura 11) possui fórmula química $C_{13}H_{18}O_2$ e seu nome IUPAC é ácido 2-(-4-isobutilfenil) propiônico. Assim, seu nome deriva de seus três grupos funcionais:

- Grupo **isobutil**: radical $(CH_3)_2-CH-CH_2-$, indicado em I nas Figuras 7a e 7b;
- Grupo **fenil**: radical derivado do anel aromático benzeno C_6H_5 , indicado em II nas Figuras 7a e 7b;
- Ácido **propiônico**: radical representado por $-CH-CH_2COOH$, indicado por III nas Figuras 7a e 7b (Halford, Lordkipanidzé, Watson, 2012).

Um fato interessante a ser destacado em relação à estrutura do IBP é que a ausência de ligações de hidrogênio e a presença do grupo CH entre o anel benzeno do grupo fenil e o grupo carboxilato (COOH) do ácido propiônico torna a estrutura da molécula mais flexível (Okulik, Jubert, 2006).

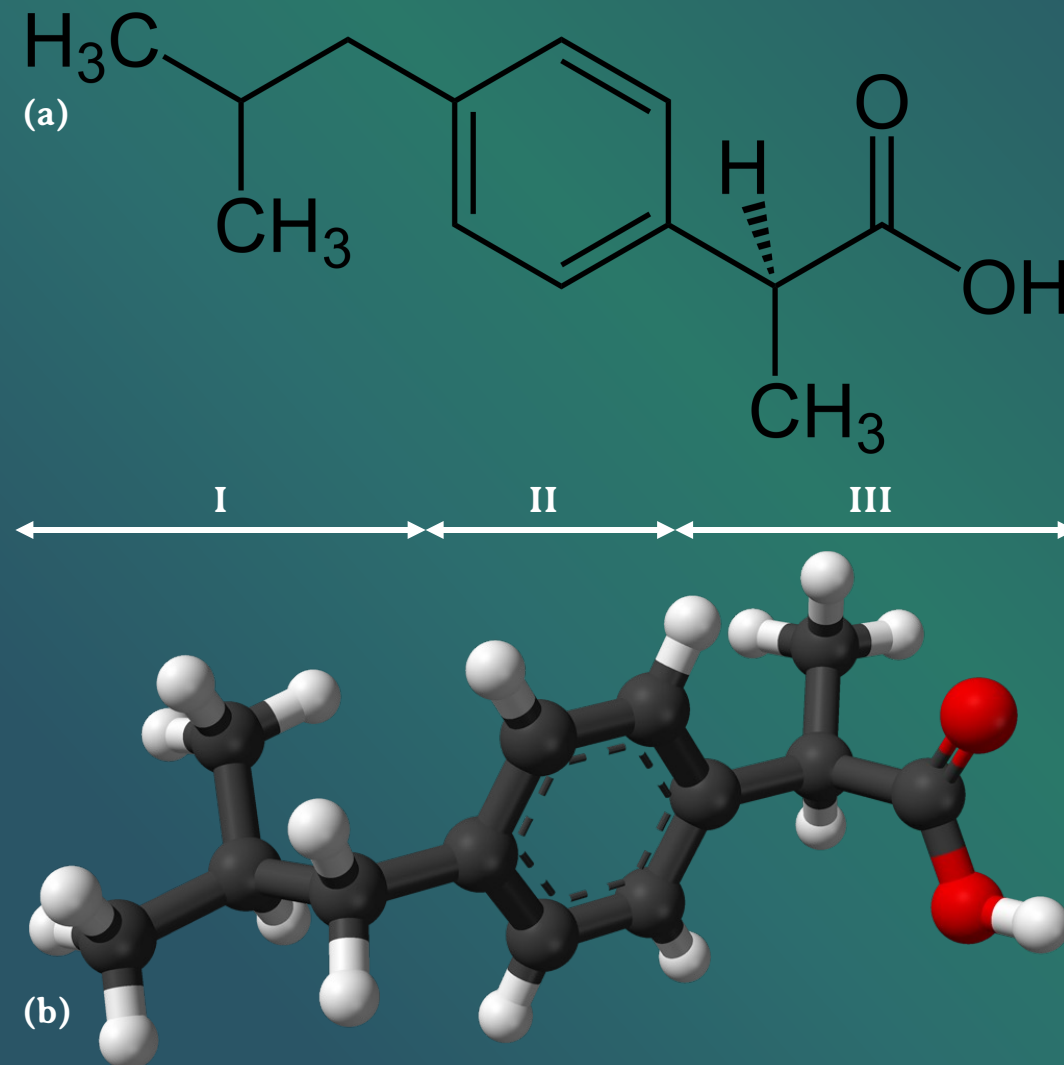
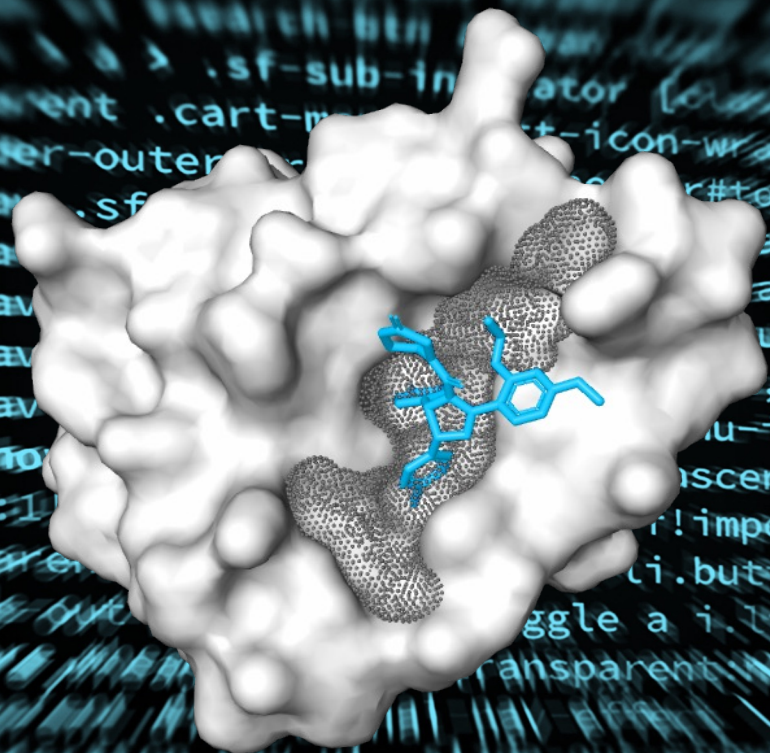


Figura 11 – Estrutura bioquímica do ibuprofeno representada no esquema de fórmula estrutural plana (a) e no modelo de bolas e varetas (b).



PARTE 3

Arquivos para docagem cega

1. Arquivos para docagem cega

Para realizar a docagem cega no CB-Dock2[®], são utilizados dois arquivos: um arquivo PDB referente à uma proteína e um arquivo MOL2, MOL, SDF ou PDB de um ligante, que podem ser obtidos gratuitamente na *internet* (Figura 12).

Nos próximos capítulos serão apresentadas algumas sugestões para obtenção destes arquivos.

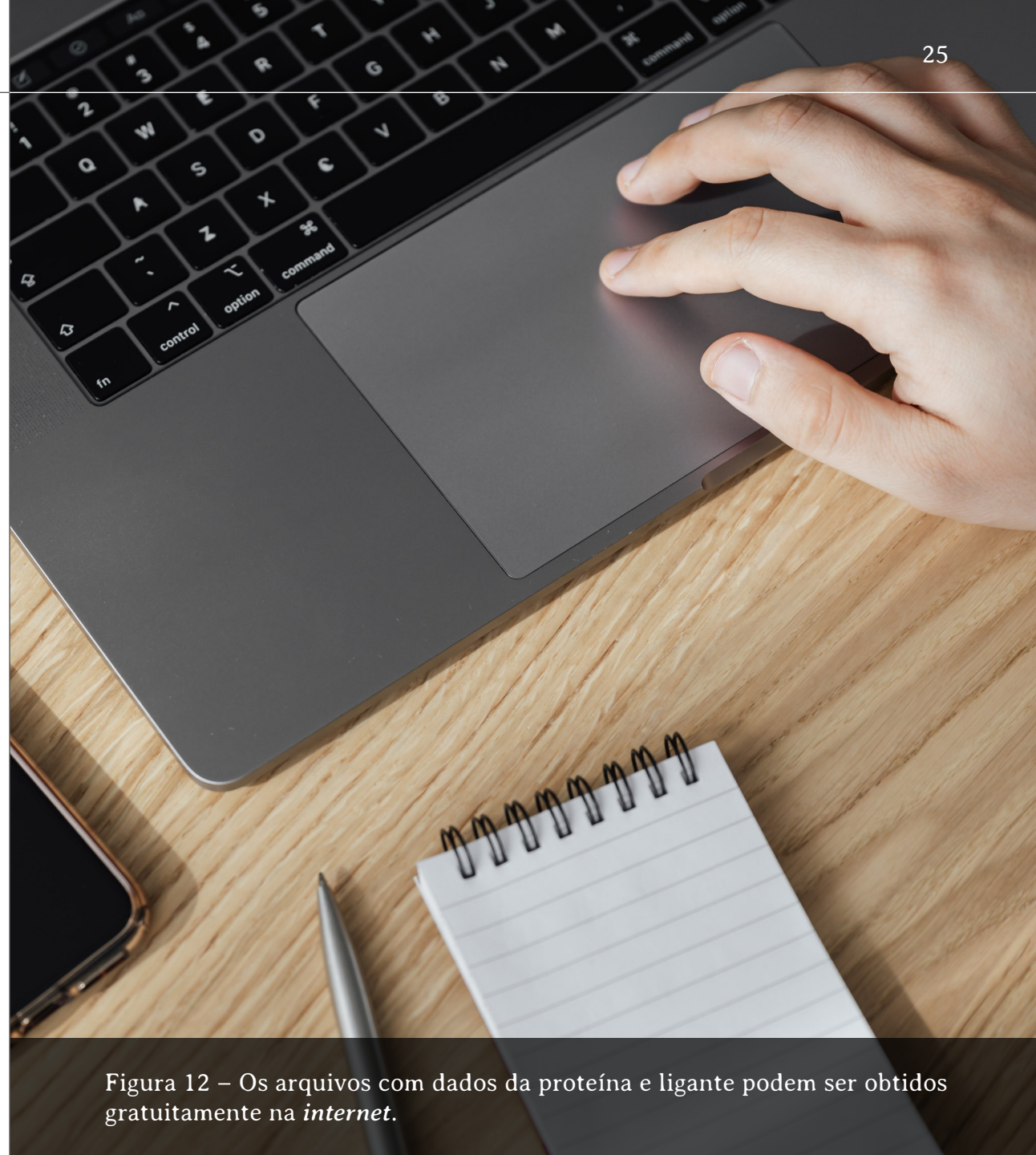
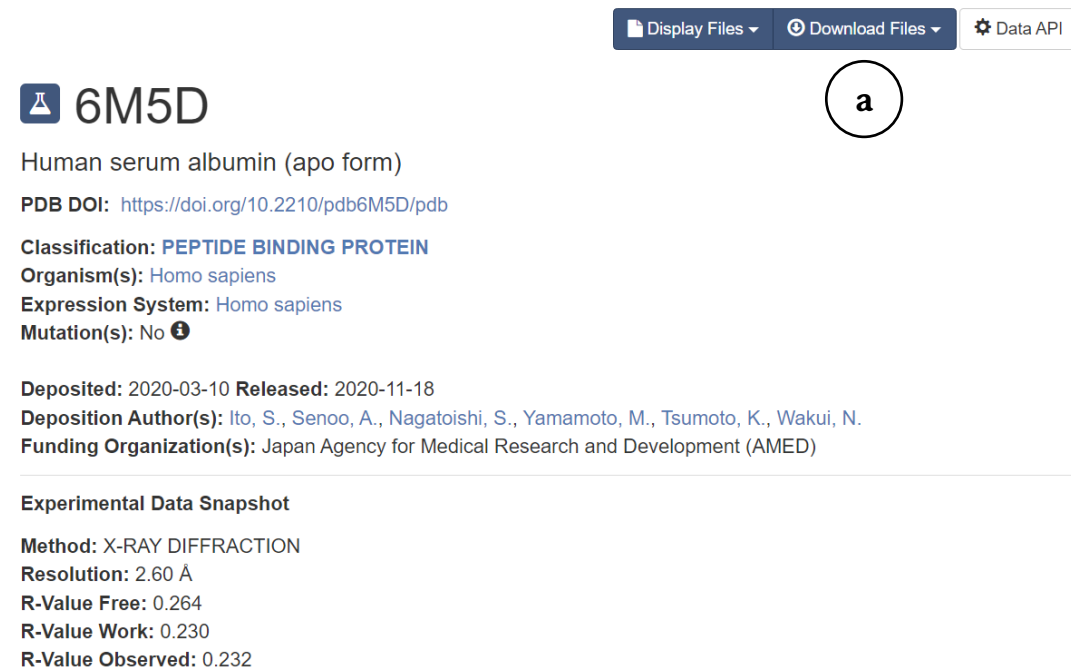


Figura 12 – Os arquivos com dados da proteína e ligante podem ser obtidos gratuitamente na *internet*.

2. Arquivo PDB da proteína

No processo de docagem cega no CB-Dock2, as informações referentes à proteína a ser testada são enviadas no formato de arquivo do tipo PDB. Estes arquivos podem ser obtidos gratuitamente no *site RCSB Protein Data Bank*, pelo *link*: <https://www.rcsb.org/>, o qual representa a principal e maior base mundial de dados deste tipo de arquivo.

No exemplo a ser utilizado neste livro, será realizada a redocagem da proteína HSA e do ibuprofeno, sendo utilizado o arquivo 6M5D.pdb (Ito, Senoo, Nagatoishi, Yamamoto, Tsumoto, Wakui, 2020), que corresponde à HSA na forma monomérica (Figura 13). O arquivo pode ser obtido gratuitamente no *link*: <https://www.rcsb.org/structure/6M5D>, clicando no botão *Download Files* (a) e selecionando a opção *PDB format*. O arquivo PDB será baixado em seu computador.



Display Files ▾ Download Files ▾ Data API

6M5D

Human serum albumin (apo form)

PDB DOI: <https://doi.org/10.2210/pdb6M5D/pdb>

Classification: **PEPTIDE BINDING PROTEIN**

Organism(s): [Homo sapiens](#)

Expression System: [Homo sapiens](#)

Mutation(s): No **i**

Deposited: 2020-03-10 Released: 2020-11-18

Deposition Author(s): [Ito, S.](#), [Senoo, A.](#), [Nagatoishi, S.](#), [Yamamoto, M.](#), [Tsumoto, K.](#), [Wakui, N.](#)

Funding Organization(s): [Japan Agency for Medical Research and Development \(AMED\)](#)

Experimental Data Snapshot

Method: X-RAY DIFFRACTION

Resolution: 2.60 Å

R-Value Free: 0.264

R-Value Work: 0.230

R-Value Observed: 0.232

Figura 13 – Obtendo do arquivo PDB de uma proteína.

3. Arquivos do ligante

Conforme mencionado anteriormente, no processo de docagem cega no CB-Dock2 é utilizado um arquivo referente ao ligante nos formatos: MOL2, MOL, SDF ou PDB.

Os arquivos nos formatos citados acima podem ser obtidos de diferentes fontes da *internet*, segue abaixo alguns exemplos:

- **Site:** *RCSB Protein Data Bank*

Link: www.rcsb.org

Orientações: caso o ligante a ser estudado esteja disponibilizado em um arquivo PDB da proteína a ser estudada ou de outra proteína, o arquivo do ligante pode ser baixado na Seção *Small molecules* (Pequenas moléculas) da página da proteína. Clicando em *Download Instance Coordinates* (Baixar coordenadas de instância), é possível obter os arquivos do ligante nos formatos SDF e MOL2. No caso do exemplo a ser usado no livro, a interação entre HSA e IBP pode ser observada por meio do arquivo PDB 2BXG (Ghuman, Zunszain, Petitpas, Bhattacharya, Curry, 2011) (Figura 14). Existem disponíveis os formatos SDF e MOL2 em diferentes versões criadas por diferentes autores (a).

2BXG

Human serum albumin complexed with ibuprofen

PDB DOI: [10.2210/pdb2BXG/pdb](https://doi.org/10.2210/pdb2BXG/pdb)

Classification: TRANSPORT PROTEIN

Organism(s): Homo sapiens

Expression System: Komagataella pastoris

Mutation(s): No

Deposited: 2005-07-26 **Released:** 2005-09-22

Deposition Author(s): Ghuman, J., Zunszain, P.A., Petitpas, I., Bhattacharya, A.A., Curry, S.

Experimental Data Snapshot

Method: X-RAY DIFFRACTION

Resolution: 2.70 Å

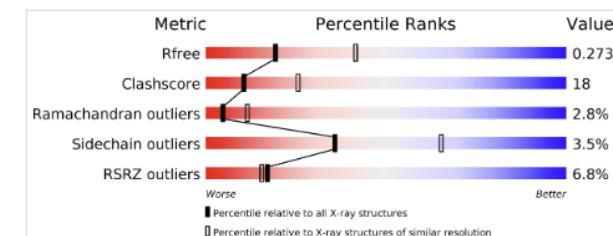
R-Value Free: 0.281

R-Value Work: 0.234

R-Value Observed: 0.234

wwPDB Validation

[3D Report](#) [Full Report](#)



Small Molecules

Ligands (1 Unique)

ID	Chains	Name / Formula / InChI Key	2D Diagram	3D Interactions
IBP Query on IBP	C [auth A], D [auth A], E [auth B], F [auth B]	IBUPROFEN C ₁₃ H ₁₉ O ₂ HEFNWWSXXWATRW-JTQLQIEISA-N		Ligand Interaction

Download Ideal Coordinates CCD File [Download Instance Coordinates](#)

- SDF format, chain C [auth A]
- SDF format, chain D [auth A]
- SDF format, chain E [auth B]
- SDF format, chain F [auth B]
- MOL2 format, chain C [auth A]
- MOL2 format, chain D [auth A]
- MOL2 format, chain E [auth B]
- MOL2 format, chain F [auth B]

Source	Binding Affinity
BindingDB: 2BXG	Ki: 1.00e+14 (nM) from 1 assay(s)

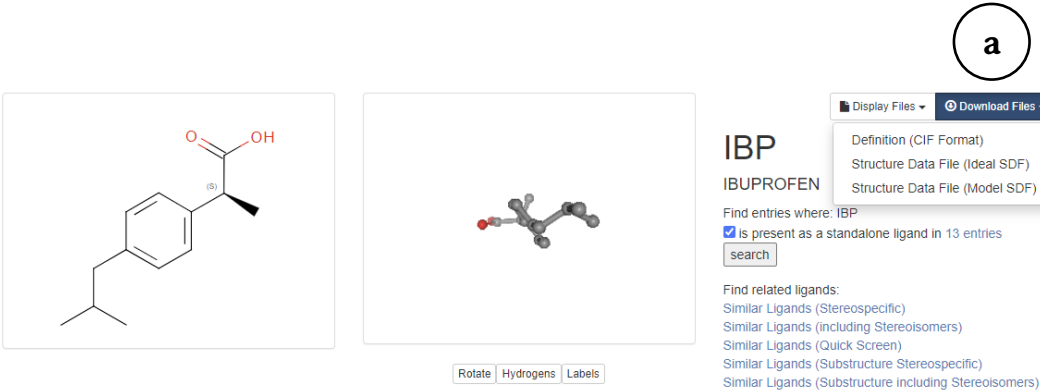
Figura 14 – Baixando o arquivo referente ao IBP no *site Protein Data Bank*.

- Site: *RCSB Protein Data Bank*

- Seção: *Ligand*

- Link: www.rcsb.org

Orientações: a Seção *Ligand* (Ligante) do site *Protein Data Bank* pode ser acessado clicando no código do ligante na Seção *Small Molecules* na página da proteína. No caso do exemplo a ser usado no livro, basta clicar em IBP. O usuário será redirecionado para a Seção *Ligand* (Figura 15). Ao clicar em Baixar arquivos (*Download Files*) (a), o usuário pode escolher entre as opções *Definition (CIF Format)*, *Structure Data File (Ideal SDF)* ou *Structure Data File (Model SDF)*. Lembrando que para o uso no CB-Dock2®, é possível utilizar apenas os arquivos no formato SDF, dentre as opções mencionadas.



Chemical Component Summary	
Name	IBUPROFEN
Synonyms	2-(4-ISOBUTYLPHENYL)PROPIONIC ACID
Identifiers	(2S)-2-[4-(2-methylpropyl)phenyl]propanoic acid
Formula	C ₁₃ H ₁₈ O ₂
Molecular Weight	206.28
Type	NON-POLYMER
Isomeric SMILES	CC(C)Cc1ccc(cc1)[C@H](C)C(=O)O
InChI	InChI=1S/C13H18O2/c1-9(2)8-11-4-6-12(7-5-11)10(3)13(14)15/h4-7,9-10H,8H2,1-3H3,(H,14,15)/t10-m/s1
InChIKey	HEFNWWSXXWATRW-JTQLQIEISA-N

Chemical Details	
Formal Charge	0
Atom Count	33
Chiral Atom Count	1
Bond Count	33
Aromatic Bond Count	6

Figura 15 – Baixando o arquivo referente ao IBP na Seção *Ligand*.

- **Site: PubChem**

Link: www.pubchem.ncbi.nlm.nih.gov

Orientações: Neste site, podem ser obtidos arquivos de ligantes nos formatos SDF, JSON, XML e ASNT. Entretanto, apenas o formato SDF é aceito no CB-Dock2[®], dentre as opções.

- **Site: Drugbank Online**

Link: www.go.drugbank.com

Orientações: Neste site, podem ser obtidos arquivos de ligantes nos formatos MOL, SDF, 3D-SDF, PDB, SMILES, InChI, sendo que apenas os formatos MOL e SDF podem ser usados no CB-Dock2[®].

Vale ressaltar que existem vários outros *sites* disponíveis que fornecem gratuitamente os arquivos com extensão MOL2, MOL, SDF ou PDB de ligantes. Estas páginas da *internet* podem ser facilmente encontradas realizando uma breve busca em qualquer navegador (Figura 16).

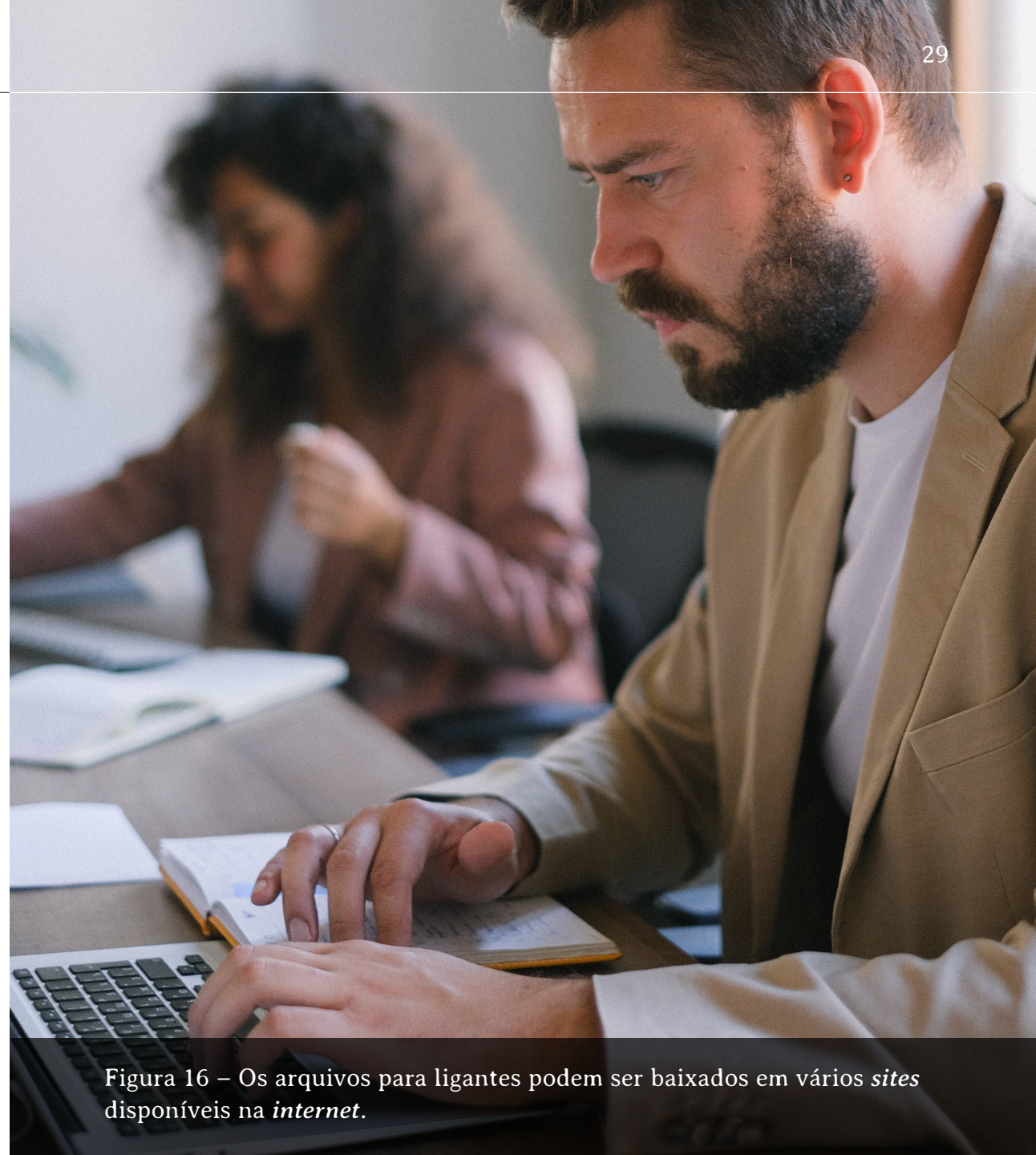
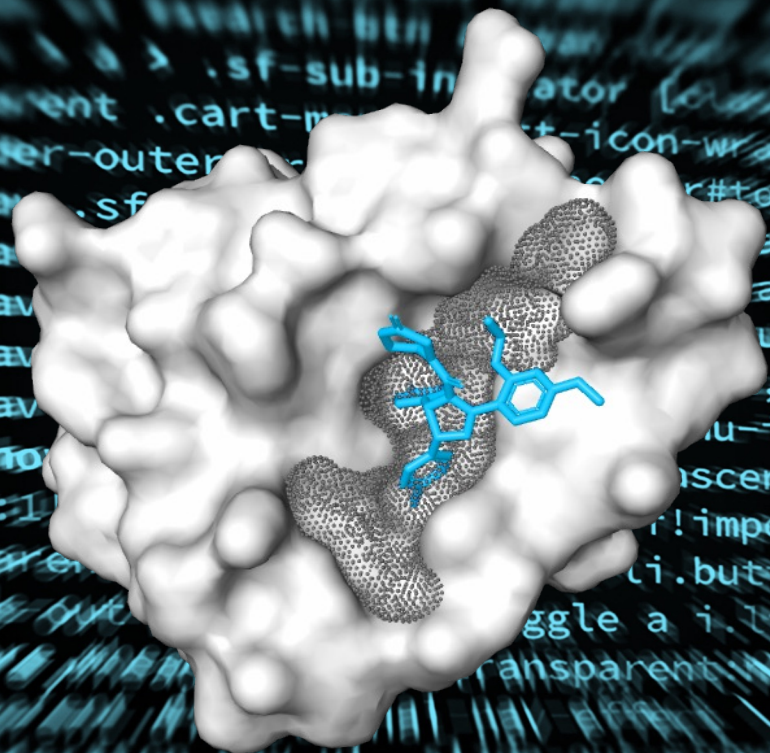


Figura 16 – Os arquivos para ligantes podem ser baixados em vários *sites* disponíveis na *internet*.



PARTE 4

Página inicial do CB-Dock2[®]

1. Página inicial do CB-Dock2®

Após obter os arquivos da proteína e ligante, o usuário deve realizar o carregamento (*Upload*) destes arquivos no servidor CB-Dock2® no *link*: <https://cadd.labshare.cn/cb-dock2/php/blinddock.php> (Figura 17). Esta primeira página do site CB-Dock2® é organizada em três seções:

- **Seção 1:** Carregamento da proteína (*Upload protein*) (a);
- **Seção 2:** Carregamento do ligante (*Upload ligand*) (b);
- **Seção 3:** Mais parâmetros (*More parameters*) (c);
- **Seção 4:** Trabalhos submetidos (*Submitted jobs*).

A **Seção 4: Trabalhos submetidos (*Submitted jobs*)** é exibida apenas após a solicitação de serviços no CB-Dock2®.

The screenshot displays the CB-Dock2 web interface. At the top, the logo 'CB-DOCK2' is shown in blue, with the tagline 'Cavity-detection guided Blind Docking' below it. A navigation bar contains buttons for 'Home', 'Dock', 'Example', 'Manual', and 'Contact'. The main content is divided into three sections, labeled 'a', 'b', and 'c' in circles on the left. Section 'a' is 'Upload Protein', with instructions to 'Select a protein file (pdb format)' and an example link '[Example: MDM2.pdb (PDB: 4HG7)]'. It features a dashed box with an upload icon and the text 'Choose or drop a protein file here'. Section 'b' is 'Upload Ligand', with instructions to 'Select a ligand file (mol2, mol, sdf, pdb format)' and an example link '[Example: Nutlin-3a.sdf]'. It includes 'Upload Local' and 'Draw Ligand' buttons, a dashed box with an upload icon, and the text 'Choose or drop a ligand file here'. Section 'c' is 'More parameters', with the text 'Other customizable parameters' and a text input field labeled 'Enter your email address to receive the results (optional)'. At the bottom, there are two buttons: 'Search Cavities' and 'Auto Blind Docking'.

Figura 17 – Página inicial do CB-Dock2®.

2. Seção 1: Carregamento da proteína (*Upload protein*)

Nesta seção é realizado o **carregamento da proteína** (*Upload protein*) (Figura 18), o usuário pode escolher o arquivo salvo no computador clicando com o botão esquerdo do mouse sobre o *link Escolha ou arraste um arquivo de proteína aqui* (*Choose or drop a protein file here*) ou arrastar o arquivo PDB até a área pontilhada.

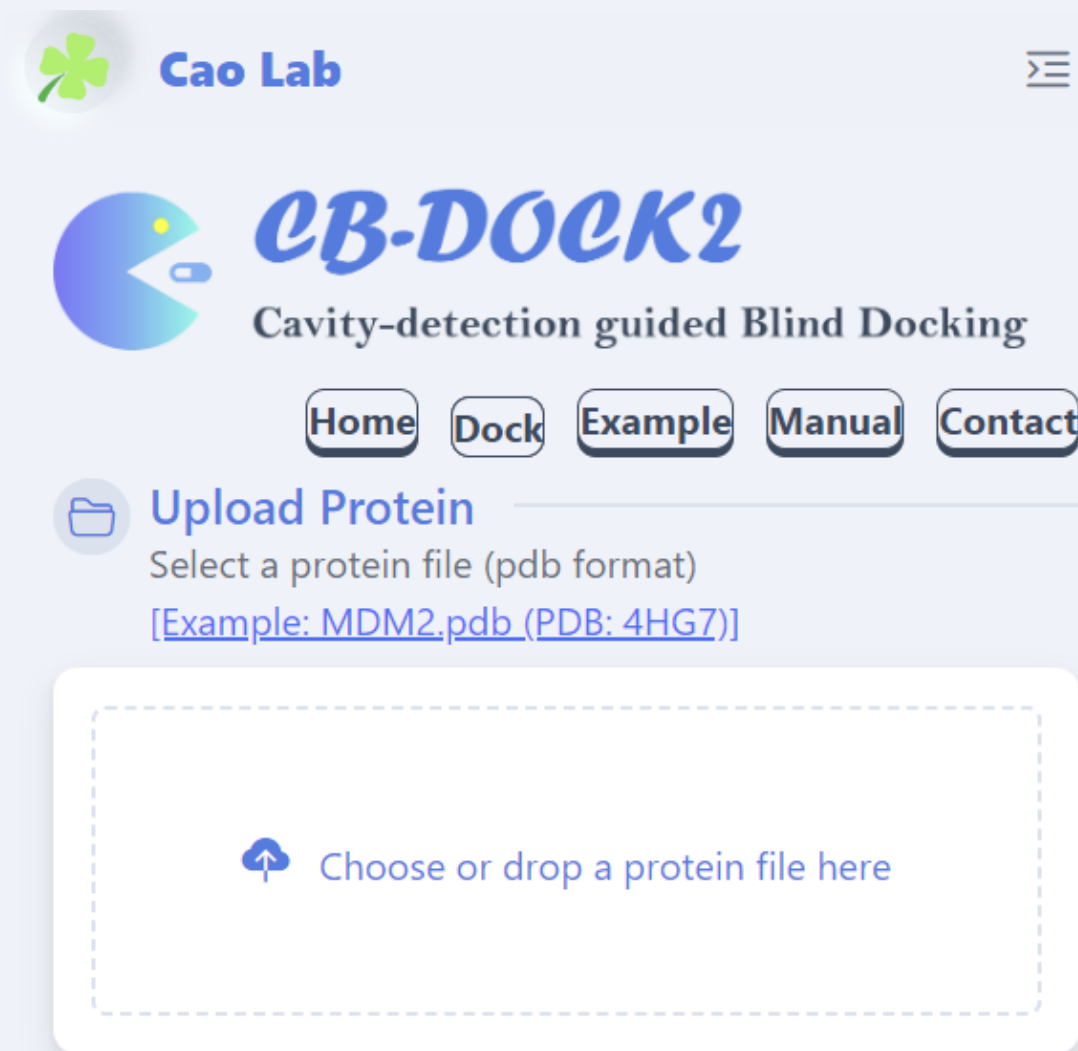


Figura 18 – Seção 1: Carregamento da proteína.

Após realizar o upload do arquivo PDB da proteína escolhida (Figura 19), o usuário poderá consultar o arquivo PDB no **quadro superior (a)** para obter eventuais informações.

A **barra azul (b)**, logo abaixo do quadro com o arquivo PDB, irá informar a atividade do programa e quando o *upload* for concluído com sucesso (*Protein uploaded!*).

Abaixo da barra, o programa irá emitir alguns **avisos (c)** aos usuários como por exemplo: deleção de moléculas de água, solventes e heteroátomos, eventuais problemas observados no arquivo analisado, presença de múltiplas cadeias (no caso de proteínas poliméricas), entre outros.

Após a análise do arquivo PDB e dos avisos emitidos pelo servidor, o usuário pode optar por **remover o arquivo (Remove file) (d)**, caso deseje.

Upload Protein
Select a protein file (pdb format)
[\[Example: MDM2.pdb \(PDB: 4HG7\)\]](#) [Submitted Protein: 6m5d.pdb]

a

```

HEADER  PEPTIDE BINDING PROTEIN                                10-MAR-
20      6M5D
TITLE   HUMAN SERUM ALBUMIN (APO FORM)
COMPND  MOL_ID: 1;
COMPND  2 MOLECULE: SERUM ALBUMIN;
COMPND  3 CHAIN: A;
COMPND  4 ENGINEERED: YES
SOURCE  MOL_ID: 1;
SOURCE  2 ORGANISM_SCIENTIFIC: HOMO SAPIENS;
SOURCE  3 ORGANISM_COMMON: HUMAN;

```

b

Protein uploaded!

Attention: 3 water molecules will be deleted from the uploaded protein structure prior to docking!

Attention: 5 other hetatoms will be deleted from the uploaded protein structure prior to docking by default!

Or you can choose the hetatoms to be retained during docking:

c

P04(A601)

Attention: The chains of the uploaded protein are broken at position A76(THR), A88(ALA), A89(ASP), A91(CYS), A101(CYS), A103(LEU)!

d **Remove file**

Figura 19 – Analisando o carregamento do arquivo da proteína.

3. Seção 2: Carregamento do ligante (*Upload ligand*)

O ligante a ser testado com a proteína deve ter o arquivo no formato MOL2, MOL, SDF ou PDB carregado no servidor on-line CB-Dock2[®] no *link*: <https://cadd.labshare.cn/cb-dock2/php/blinddock.php>, logo a seguir do carregamento da proteína.

Para realizar o **carregamento do ligante (*Upload ligand*)** (Figura 20), o usuário pode escolher o arquivo salvo no computador clicando com o botão esquerdo do mouse sobre o *link Escolha ou arraste um arquivo de proteína aqui (Choose or drop a protein file here)* ou arrastar o arquivo PDB até a área pontilhada.

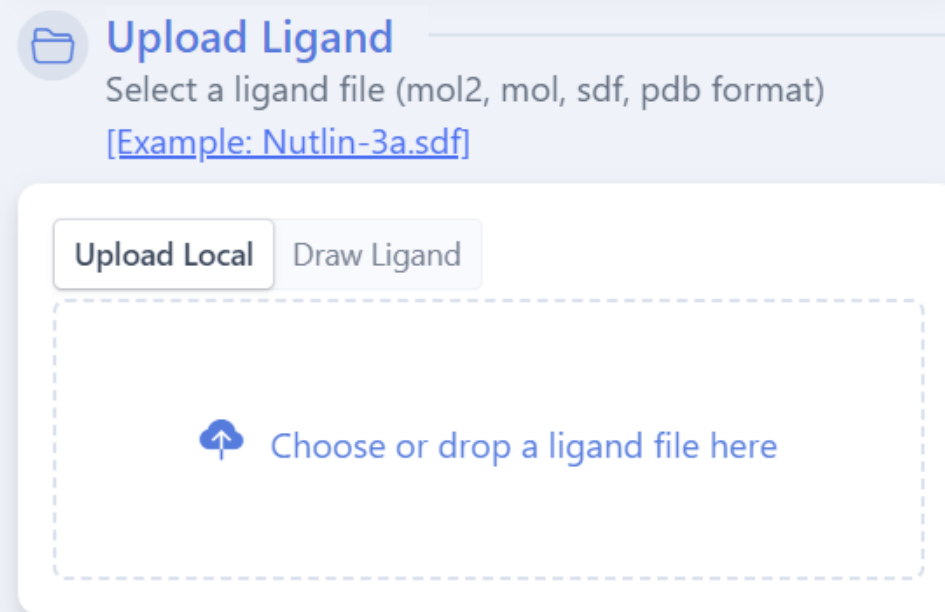


Figura 20 – Seção 2: Carregamento do ligante.

Após realizar o *upload* do ligante escolhido (Figura 21), o usuário poderá consultar informações sobre o arquivo no **quadro superior (a)**.

A **barra azul (b)**, logo abaixo do quadro, irá informar a atividade do programa e momento em que o *upload* for concluído com sucesso (*Ligand uploaded!*).

Após a análise do arquivo do ligante, o usuário pode optar por **remover o arquivo (Remove file) (c)**, caso deseje.

Upload Ligand
Select a ligand file (mol2, mol, sdf, pdb format)
[\[Example: Nutlin-3a.sdf\]](#) [Submitted Ligand: IBP_ideal.sdf]

Upload Local Draw Ligand

```
IBP
-OECHEM-06302316593D

33 33 0 1 0 0 0 0 0999 V2000
-0.3640 -0.1760 -3.4420 C 0 0 0 0 0 0 0 0
0 0 0 0
-0.9310 -0.3490 2.9700 C 0 0 0 0 0 0 0 0
0 0 0 0
0.0100 0.4730 3.8510 C 0 0 0 0 0 0 0 0
0 0 0 0
0.1550 0.2010 5.2070 C 0 0 0 0 0 0 0 0
```

Ligand uploaded!

Remove file

Figura 21 – Analisando o carregamento do arquivo do ligante.

Caso o ligante não esteja disponível nos formatos MOL2, MOL, SDF ou PDB, é possível **Desenhar o ligante (Draw ligand)** (Figura 22), clicando na aba indicada pela letra a.

O usuário pode usar os vários recursos de desenho disponíveis para desenhar o ligante na área de trabalho (b).

Após concluir o desenho do ligante, o usuário deve submeter o ligante ao provedor clicando no **botão Submeter ligante (Submit ligand)**.

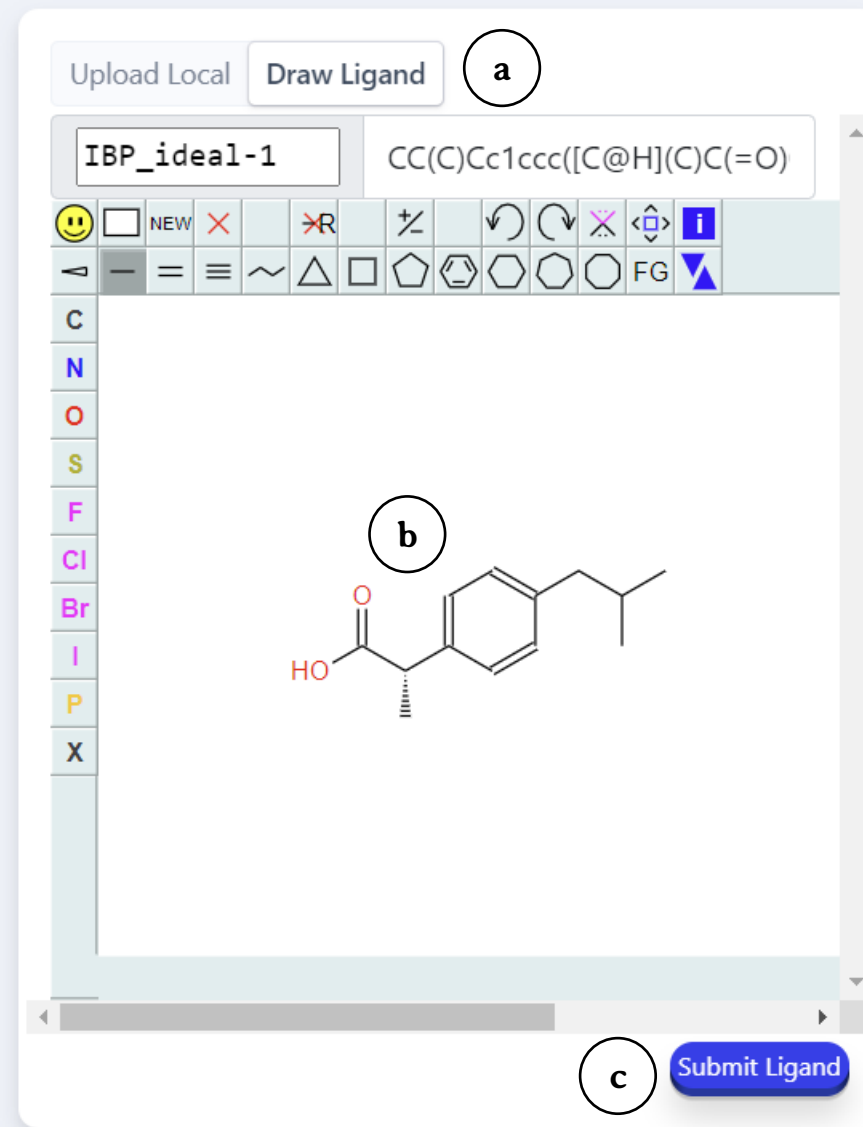


Figura 22 – Desenhando um ligante.

4. Seção 3: Mais parâmetros (*More parameters*)

O processo de docagem cega pode ser personalizado na seção **Mais parâmetros: outros parâmetros customizáveis (*More parameters: other customizable parameters*)** (Figura 23a). Para observar todos parâmetros personalizáveis, o usuário deve clicar na **dupla seta (b)** à direita do título da seção.

Pode ser escolhido o **número de cavidades a serem testadas na docagem cega (*Number of cavities for docking*) (c)**, sendo que o valor 5 é estabelecido como padrão.

É possível realizar também o **carregamento de uma proteína e um ligante bioquimicamente semelhantes ao serem testados como modelo na docagem cega (*Upload protein-ligand complex for template-based blind docking (optional)*) (d)**. O arquivo da proteína deve ser carregada em ***Upload the protein part of your complex here (pdb format)*** e o ligante em ***Upload the ligand part of your complex here (mol2, mol, sdf, pdb format)***.

O usuário pode receber os resultados por e-mail, digitando o endereço em ***Enter your email address to receive the results (optional)***.

a **More parameters** **b**
Other customizable parameters

c Number of cavities for docking
5 Default

d Upload protein-ligand complex for template-based blind docking (optional)
Upload the protein part of your complex here (pdb) Browse
Upload the ligand part of your complex here (mol2, mol, sdf, pdb format). Browse

e Enter your email address to receive the results (optional)

Search Cavities Auto Blind Docking

Figura 23 – Seção 3: Mais parâmetros.

Nesta mesma seção, o usuário pode solicitar dois serviços no CB-Dock2® (Figura 24):

- **Busca de cavidades (*Search cavities*):** onde o servidor localiza as cavidades presentes na proteína submetida;
- **Docagem cega automática (*Auto blind docking*):** onde o servidor realiza a docagem cega da proteína e ligante submetidos, testando as cavidades localizadas.

Os botões para solicitação da busca de cavidades e a docagem cega estão localizados na seção **Mais parâmetros (*More parameters*)** (Figura 26), apresentada anteriormente. Nesta seção, estão disponíveis o botão ***Search cavities*** (a) para busca de cavidades e o botão ***Auto blind docking*** (b) para docagem cega.

More parameters
Other customizable parameters

Number of cavities for docking

5 Default

Upload protein-ligand complex for template-based blind docking (optional)

Upload the protein part of your complex here (pdb) Browse

Upload the ligand part of your complex here (mol2, pdb format). Browse

Enter your email address to receive the results (optional)

a Search Cavities Auto Blind Docking b

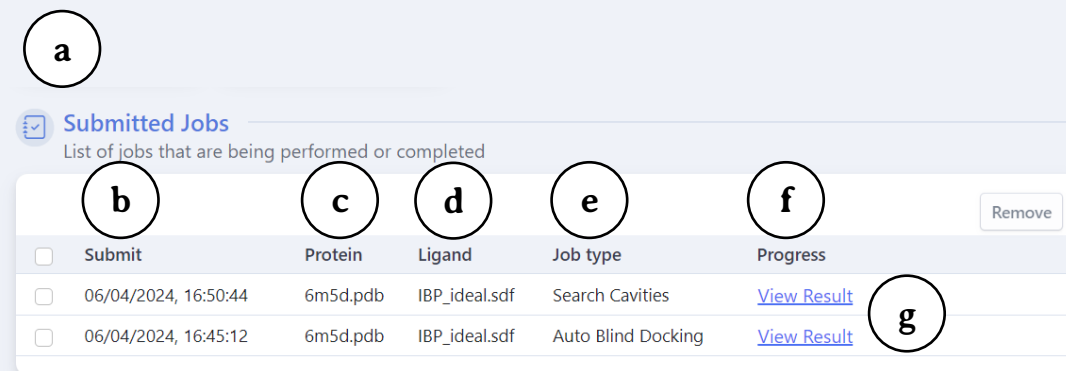
Figura 24 – Botões de solicitação da busca de cavidades e a docagem cega.

3. Seção 4: Trabalhos submetidos (*Submitted jobs*)

Após clicar no botão *Search cavities* ou no botão *Auto blind docking*, será aberta a seção **Trabalhos submetidos (*Submitted jobs*)** (Figura 25), que apresenta a **Lista de trabalhos realizados ou completos (*List of Jobs that are being performed ou completed*)** (a), a qual apresenta:

- **Submissão (*Submit*) (b)**: a data e horário de solicitação do serviço;
- **Proteína (*Protein*) (c)**: o código da proteína carregada;
- **Ligante (*Ligand*) (d)**: o código do ligante carregada;
- **Tipo de trabalho (*Job type*) (e)**: que pode ser a busca de cavidades (*Search cavities*) ou a docagem cega automática (*Auto blind docking*);
- **Progresso (*Progress*) (f)**: que apresenta os *links* para verificação dos resultados (*View result*).

Para observar os resultados de busca de cavidades (*Search cavities*) ou de docagem cega automática (*Auto blind docking*), o usuário deve clicar no *link View result* (g).



<input type="checkbox"/> Submit	Protein	Ligand	Job type	Progress	Remove
<input type="checkbox"/> 06/04/2024, 16:50:44	6m5d.pdb	IBP_ideal.sdf	Search Cavities	View Result	
<input type="checkbox"/> 06/04/2024, 16:45:12	6m5d.pdb	IBP_ideal.sdf	Auto Blind Docking	View Result	

Figura 25 – Seção 4: Verificação dos trabalhos submetidos.

Vale ressaltar que estes serviços são relativamente demorados, podendo variar de alguns segundos até um pouco mais de um minuto, sendo que o tempo aumenta de acordo com o tamanho da proteína. Caso leve muito tempo, é sugerido que o usuário atualize a página (Liu, Yang, Gan, Chen, Xiao, Cao, 2022) (Figura 26).

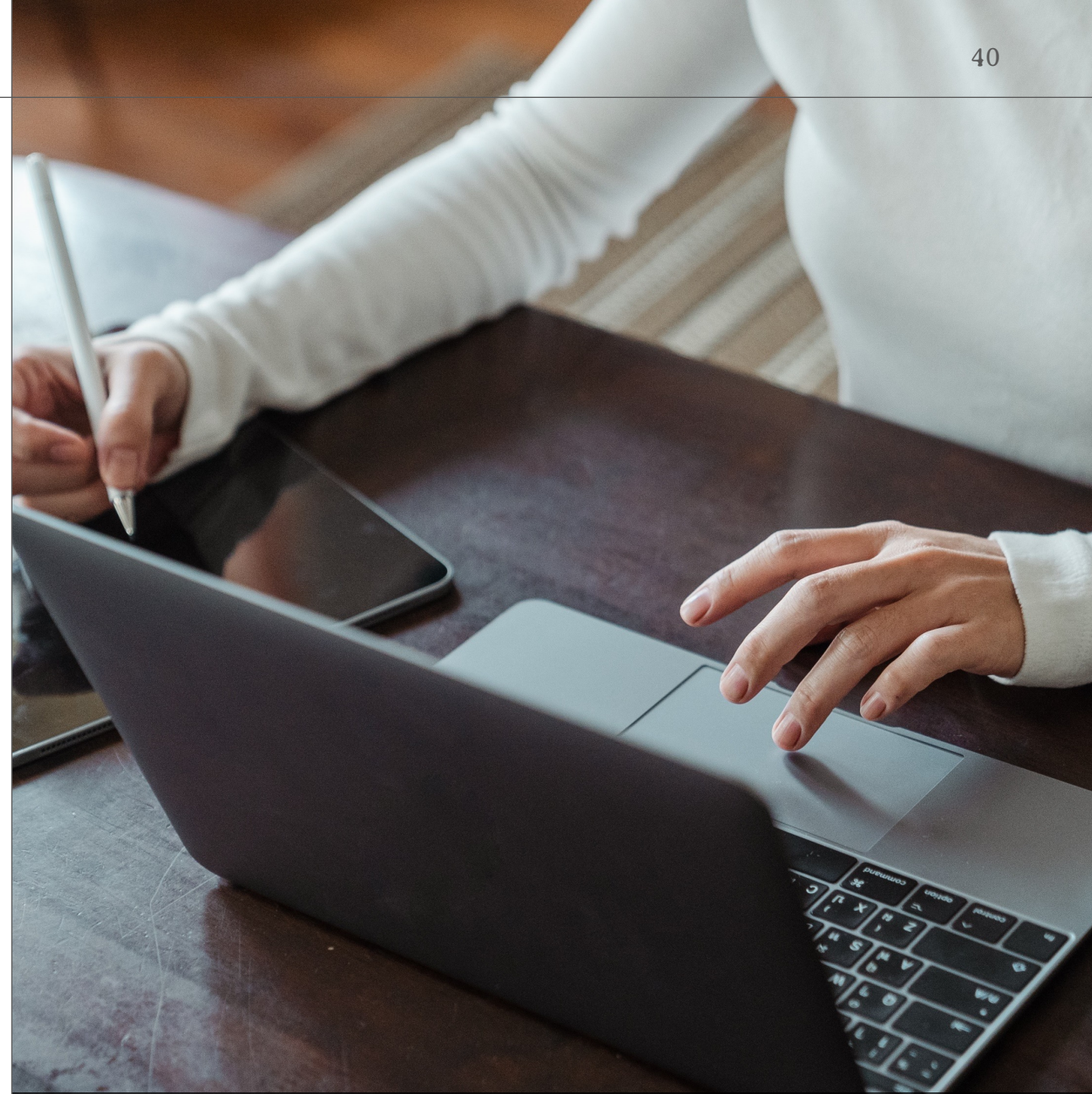
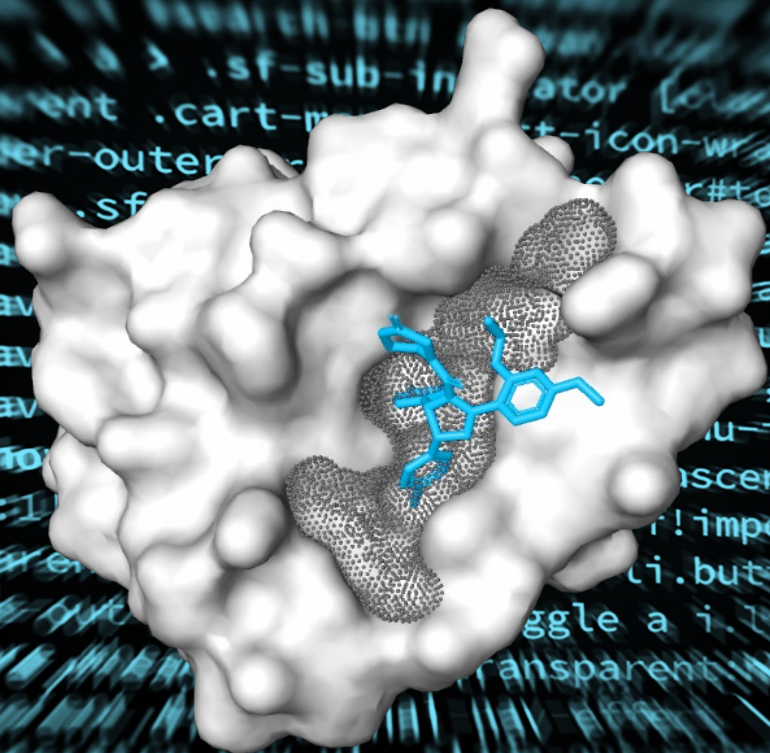


Figura 26 – Caso o serviço solicitado no CB-Dock2[®] demore mais do que o esperado, é sugerido que o usuário atualize a página.



PARTE 5

Detecção de Cavidades do CB-Dock2[®]

1. Detecção de cavidades do CB-Dock2®

Após clicar no *link: View Result* da opção **Busca de cavidades (Search cavities)** na Seção **Trabalhos submetidos (Submitted jobs)**, será aberta uma nova guia com a página de **Detecção de cavidades (Cavity detection)** (Figura 27), que é dividida nas seguintes seções:

- **Seção 1:** Detecção de cavidades baseada na estrutura (*Structure-based cavity detection*) (a);
- **Seção 2:** Detecção de cavidades baseada em modelos (*Template-based cavity detection*) (b).

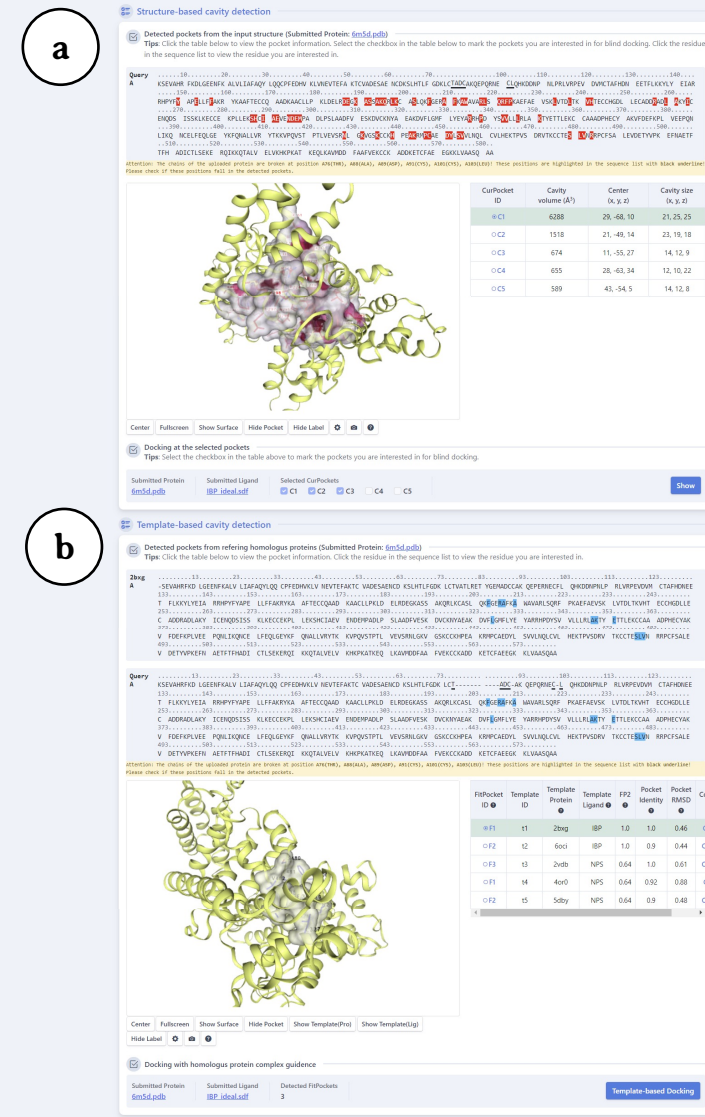


Figura 27 – Botão *Auto Blind Docking*.

2. Seção 1: Detecção de cavidades baseada na estrutura (*Structure-based cavity detection*)

A Detecção de cavidades baseada na estrutura (*Structure-based cavity detection*) (Figura 28) é realizada pelo algoritmo **CurPocket**[®], que analisa as curvaturas da superfície da proteína para prever os sítios de ligação, ou seja, as cavidades de uma determinada proteína, além de calcular os centros e os tamanhos delas (Liu, Grimm, Dai, Hou, Xiao, Cao, 2020; Liu, Yang, Gan, Chen, Xiao, Cao, 2022). Esta seção é organizada em quatro subseções:

- **Subseção 1:** Consulta A (*Query A*) (a);
- **Subseção 2:** Visualizador NGL[®] (*NGL[®]-viewer*) (b);
- **Subseção 3:** Quadro de informações das cavidades (c).

Structure-based cavity detection

Detected pockets from the input structure (Submitted Protein: [6m5d.pdb](#))
 Tips: Click the table below to view the pocket information. Select the checkbox in the table below to mark the pockets you are interested in for blind docking. Click the residue in the sequence list to view the residue you are interested in.

Query A
 KSEVAHR FKDLGEEFNK ALVLIQFAQY LQCCPFEDHV KLVNEVTEFA KTCVADESAA INCOKSLHTLFGDKLCTADCAKQEPORNE CLQHKDDNP NLPRLRPEV DVM
 CTAFHDI EETFLKLY EIARRHPYF APQLLFAKR YKAAFECCQ AADKAACLLP KLDELRLDEE ASAKOPLIC ASLQKGERA FAWAVARLS RRFPP
 KAFAE VSKVTDTK VTECCHGDL LECADDAD AKYICENQDS ISSKLEKCE KPLLEKSTAI AVENDEPA DLPSLADEV ESKDVCKNYA EAKDV
 FLGMF LYEYARHFD YSWLLRLA TYETLEKC CAAADPHECY AKVFDEFKPL VEEPQLIKQ NCELFEQLGE YKFNALLVR YTKKVPQVST PTLLEV
 SRLLGVGSCKK PEAKHPCAE YLWVLMQL CVLHEKTPVS DRVTKCCTE VWRPCFSA LEVDETYVPK EFMAETFTFH ADICTLSEKE RQIKKQT
 ALV ELVKHKPKAT KEQLKAVMDD FAAFVEKCK ADDKETFAE EGKLVAAASQ AA

Attention: The chains of the uploaded protein are broken at position A76(THR), A88(ALA), A89(ASP), A91(CYS), A101(CYS), A103(LEU)! These positions are highlighted in the sequence list with black underline! Please check if these positions fall in the detected pockets.

CurPocket ID	Cavity volume (Å ³)	Center (x, y, z)	Cavity size (x, y, z)
<input checked="" type="checkbox"/> C1	6288	29, -68, 10	21, 25, 25
<input type="checkbox"/> C2	1518	21, -49, 14	23, 19, 18
<input type="checkbox"/> C3	674	11, -55, 27	14, 12, 9
<input type="checkbox"/> C4	655	28, -63, 34	12, 10, 22
<input type="checkbox"/> C5	589	43, -54, 5	14, 12, 8

atom: [GLU]230:A:CA (RECEPTOR)

Center Fullscreen Show Surface Hide Pocket Hide Label

Figura 28 – Detecção de cavidades baseada na estrutura (*Structure-based cavity detection*)

2.1. Subseção 1: Consulta A (Query A)

Nesta primeira subseção, o usuário pode observar a sequência de resíduos de aminoácidos da proteína submetida. No exemplo apresentado na Figura 29, estão observados os resultados obtidos para a cavidade C1 da albumina sérica humana utilizando o arquivo 6M5D.pdb, sendo que os resíduos destacados em vermelho correspondem aos associados à cavidade C1. Caso o usuário selecione outras cavidades, novos resíduos serão destacados em cores diferentes.

```

Query .....10.....20.....30.....40.....50.....60...
A      KSEVAHR FKDLGEENFK ALVLIAFAQY LQQCPFEDHV KLVNEVTEFA KTCVADESAE NCD
      .....70.....100.....110.....120.....130...
      KSLHTLF GDKLCTADCAKQEPQRNE CLQHKDDNP NLPRLVRPEV DVMCTAFHDN EET
      .....140.....150.....160.....170.....180.....
      FLKKYLY EIARRHPYFY APEELLFFAKR YKAAFTECCQ AADKAACLLP KLDELRDEG
190.....200.....210.....220.....230.....240...
K ASSAKQRLKC ASLQKFGERA FKAWAVARLS QRFPKAEFAE VSKLVTDLTK VHT
      .....250.....260.....270.....280.....290.....
      ECCHGDL LECADDRADL AKYICENQDS ISSKLKECCE KPLLEKSHCI AEVNDEMP
300.....310.....320.....330.....340.....350...
      A DLPSLAADFV ESKDVCKNYA EAKDVFLGMF LYEYARRHPD YSVLLLRLA KTY
      .....360.....370.....380.....390.....400.....
      ETTLEKC CAAADPHECY AKVFDEFKPL VEEPQNLIKQ NCELFEQLGE YKFQNALLV
410.....420.....430.....440.....450.....460...
      R YTKKVPQVST PTLVEVSRNL GKVGSKCCKH PEAKRMPCAE DYLSVVLNQL CVL
      .....470.....480.....490.....500.....510.....
      HEKTPVS DRVTKCTES LVNRRPCFSA LEVDETYVPK EFNAETFTFH ADICTLSEK
520.....530.....540.....550.....560.....570...
      E RQIKKQTALV ELVKHKPKAT KEQLKAVMDD FAAFVEKCK ADDKETCFAE EGK
      .....580..
      KLVAASQ AA

```

Figura 29 – Subseção 1: Consulta A.

2.2. Subseção 2: Visualizador NGL® (NGL®-viewer)

Esta subseção utiliza o **visualizador NGL® (NGL-viewer®)** para exibir a proteína submetida e suas cavidades. Este programa possui vários **botões de controle** (Figura 30):

- **Centro (Center) (a)**: clicando neste botão, a molécula é reposicionada no centro da imagem;
- **Tela cheia (Full screen) (b)**: este botão exibe a proteína em tela cheia. Para retornar ao formato anterior, basta clicar na tecla ESC;
- **Exibir superfície (Show surface) (c)**: este botão exibe a superfície completa da proteína (não apenas da área da cavidade, como mostrado na Figura 24);
- **Esconder bolso (Hide pocket) (d)**: este botão permite deixar de exibir a superfície da parte da molécula ao redor da cavidade;
- **Esconder legendas (Hide label) (e)**: este botão remove as legendas de tipo e número de resíduo de aminoácido exibidas;

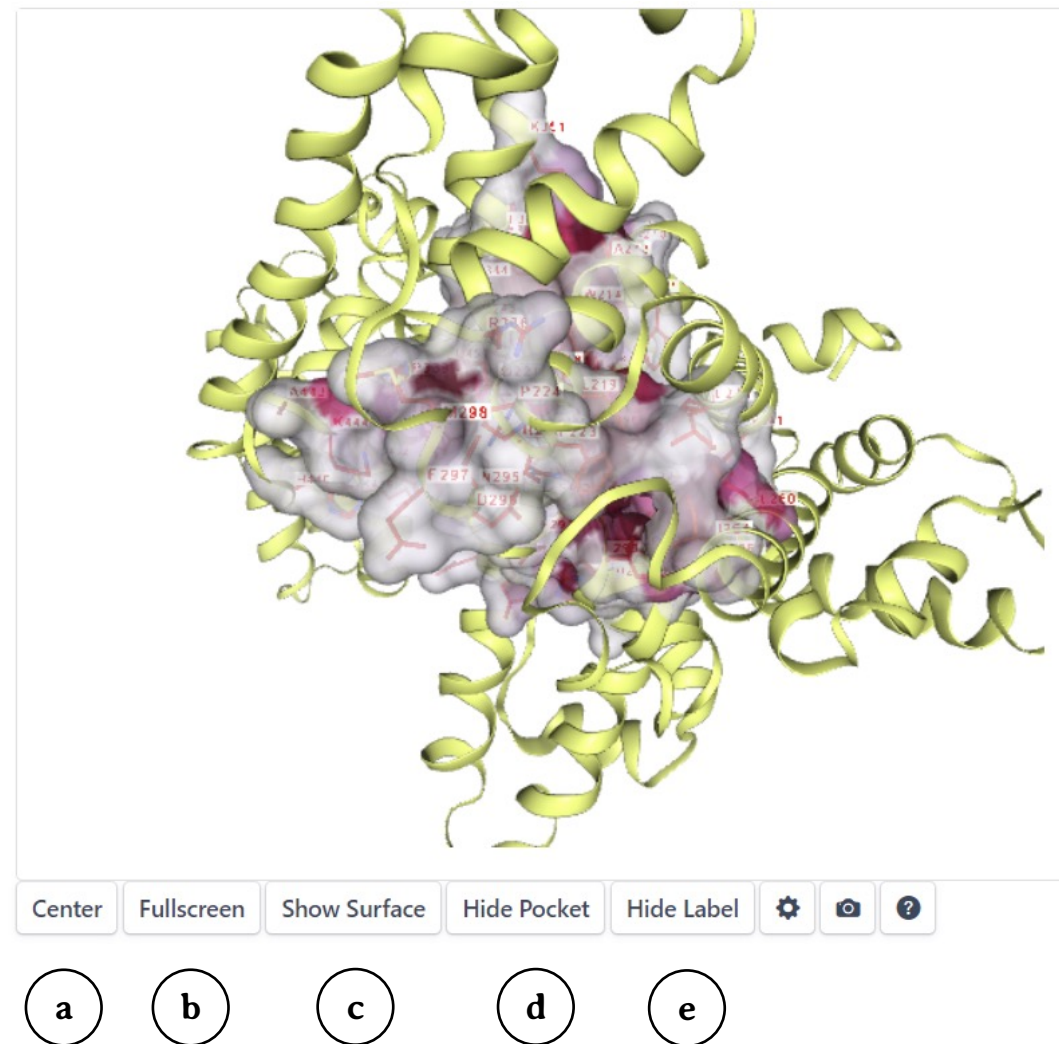


Figura 30 – Botões de controle do visualizador NGL®.

No NGL-viewer®, é observado em destaque o resíduo de aminoácido ou a cavidade selecionada. Assim, se o usuário clicar em um resíduo de aminoácido na seção Consulta A (*Query A*), a molécula será girada e o resíduo selecionado será exibido em destaque nesta visualização.

Por outro lado, se o usuário selecionar uma cavidade, o programa irá girar a molécula para mostrar a cavidade selecionada.

Na Figura 31, observe em destaque a metionina 298 da cavidade C1 (a), onde os resíduos de aminoácidos estão escritos na cor vermelha. Observe também a treonina 467 da cavidade C2 (b), sendo que nesta cavidade os resíduos são escritos em fonte de cor azul. Vale ressaltar que a translucidez observada na figura deve-se ao fato de a superfície da proteína ser exibida em ambos os casos.

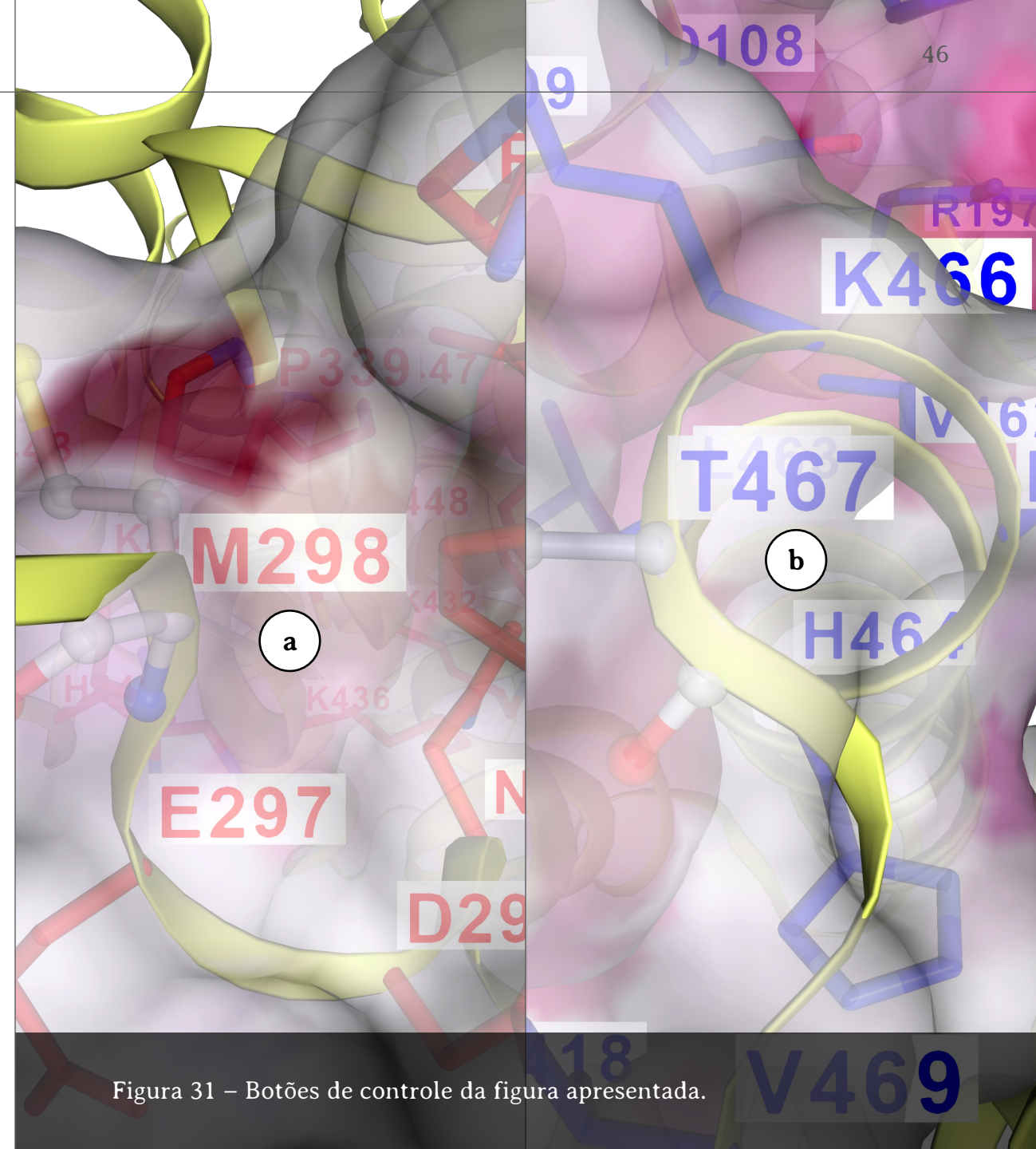


Figura 31 – Botões de controle da figura apresentada.

O visualizador NGL[®] apresenta também ícones de controle de exibição da figura apresentada (Figura 32). A figura apresenta a cavidade C1 da albumina sérica humana, onde é possível realizar os seguintes ajustes por meio dos ícones:

- **Configurações (Settings) (a):** este ícone abre a **janela Configurações (Settings) (b)**, onde o usuário pode controlar a opacidade (*Opacity*), corte da superfície próxima ao operador (*ClipNear*), corte da superfície distante (*ClipFar*), exibição do fundo (*Background*) e cor preta do fundo (*Dark stage*);
- **Salvar figura (Save picture) (c):** este ícone abre a **janela Salvar figura (Save Picture) (d)** e determina as formas de salvar a figura em relação ao tamanho (*Scale*), Antisserilhamento (*Antialias*), Aparar (*Trim*), Fundo transparente (*Transparent*) e nome da figura (*PicName*). Para salvar a imagem, o usuário deve clicar no botão Exporte (*Export*) e a figura será salva no formato PNG;
- **Manual (e):** este ícone direciona o usuário para a página do manual do CB-Dock2[®].

As representações das interações químicas e os esquemas de cores do visualizador NGL estão apresentados nos **Anexos 2 e 3**, respectivamente.

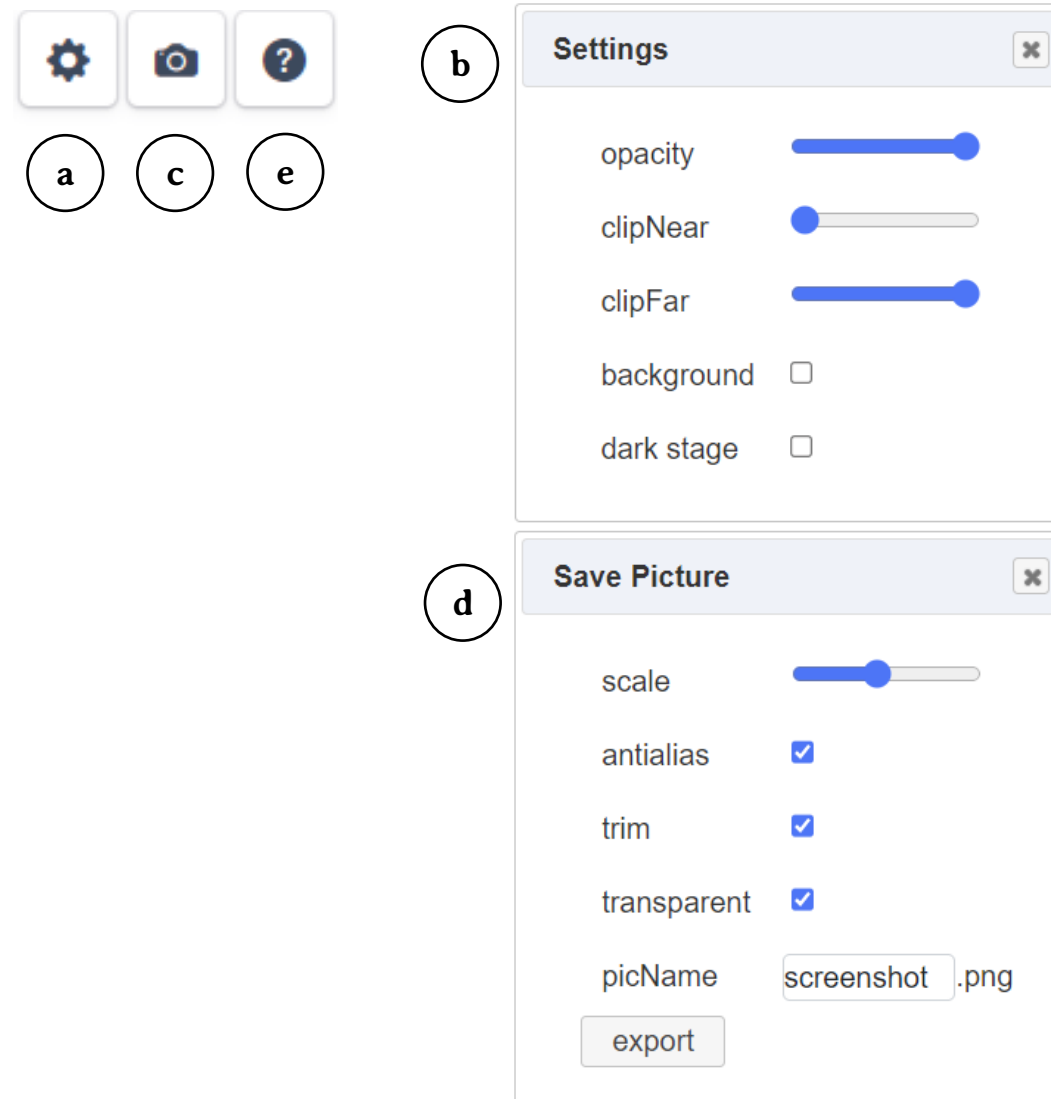


Figura 32 – Botões de controle da figura apresentada.

2.3. Subseção 3: Quadro de informações das cavidades

Na terceira seção, é apresentado o Quadro observado na Figura 33, com os resultados obtidos para as cinco cavidades da albumina sérica humana (6M5D.pdb). Vale ressaltar que, conforme apresentado anteriormente, o número de cavidades é determinado pelo usuário, porém é sugerido que este número seja preferencialmente abaixo de 20 para que não existam resultados falsos positivos (Liu, Yang, Gan, Chen, Xiao, Cao, 2022). Neste quadro são apresentados os seguintes resultados:

- **Identificação da cavidade (*CurPocket ID*) (a):** onde cada cavidade é indicada por um código. No caso da primeira cavidade, o código é C1, na segunda C2, e assim por diante;
- **Volume da cavidade em ângstrons cúbicos (*Cavity volume* Å³) (b):** indica o volume da cavidade;
- **Centro da cavidade (x,y,z) (*Center (x,y,z)*) (c):** indica a posição do centro da cavidade nos eixos x, y e z;
- **Tamanho da cavidade (*Cavity size (x,y,z)*) (d):** indica o diâmetro da cavidade nos eixos x, y e z em ângstrons.

(a)	(b)	(c)	(d)
CurPocket ID	Cavity volume (Å ³)	Center (x, y, z)	Cavity size (x, y, z)
⊙ C1	6288	29, -68, 10	21, 25, 25
○ C2	1518	21, -49, 14	23, 19, 18
○ C3	674	11, -55, 27	14, 12, 9
○ C4	655	28, -63, 34	12, 10, 22
○ C5	589	43, -54, 5	14, 12, 8

Figura 33 – Analisando as cavidades da proteína pela detecção de cavidades baseada na estrutura (*Structure-based cavity detection*).

3. Seção 2: Detecção de cavidades baseada em modelos (*Template-based cavity detection*)

A Detecção de cavidades baseada em modelos (*Template-based cavity detection*) (Figura 34) utiliza o algoritmo **FitPocket**[®] que realiza uma busca em bancos de dados para encontrar proteínas semelhantes à submetida. Assim, a detecção das cavidades da proteína é realizada usando como modelo outros arquivos PDB da mesma proteína de seres vivos da mesma espécie ou de espécies diferentes, ou ainda arquivos PDB de outras proteínas semelhantes à submetida (Liu, Yang, Gan, Chen, Xiao, Cao, 2022). Da mesma forma que o **CurPocket**[®], este algoritmo prevê os possíveis sítios de ligação (localizados nas cavidades da superfície da proteína submetida) e calcula os centros e os tamanhos delas (Liu, Grimm, Dai, Hou, Xiao, Cao, 2020). Esta seção é organizada em:

- **Subseção 1:** Sequências de resíduos de aminoácidos da proteína modelo e da proteína submetida – Consulta A (*Query A*) (a);
- **Subseção 2:** Visualizador NGL[®] (NGL[®]-viewer) (b);
- **Subseção 3:** Quadro de informações das cavidades (c).

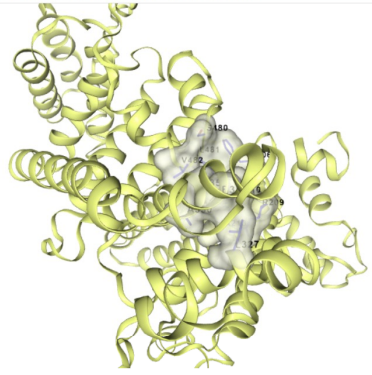
Template-based cavity detection

Detected pockets from referring homologous proteins (Submitted Protein: [6m5d.pdb](#))
 Tips: Click the table below to view the pocket information. Click the residue in the sequence list to view the residue you are interested in.

2bxg
 A13.....23.....33.....43.....53.....63.....73.....83.....93.....103.....113.....123.....
 -SEVAHRFKD LGENFKALV LIAFAQYLQ CPFDHVKLV NEVTEFAKT VADESAEND KSLHTLFGDK LCTVATRET YGEMADCCAK QEPERNECL QHKKDNNPLP RLVRPEVDVM CTAHFHNEE
 133.....143.....153.....163.....173.....183.....193.....203.....213.....223.....233.....243.....
 T FLKKLYEIA RRHPFYAPE LFFAFKRYKA AFTECCQAD KAACLLPKLD ELRDEGKASS AKQRLKCASL QKGEBAFKA IAVARLSQRF PKAEFAEVSK LVTDLTQVHT ECHGDLLE
 253.....263.....273.....283.....293.....303.....313.....323.....333.....343.....353.....363.....
 C AODRADLANY ICENQDSSSS KKECECKPL LEKSHCIAEV ENDEMPADLP SLADRVESK DVCKWYAEK DVFQGFILVE YARRPDPVSV VLLRLRLRY FTLEKCCAA ADPHEVAK
 373.....383.....393.....403.....413.....423.....433.....443.....453.....463.....473.....483.....
 V FDEFKPLVEE PQLIKQNC E LFEQLGEYK QNALLVRYTK KVPQVSTPTL VEVSRNLGKV GSKCKHPEA KRMPCAEDYL SVLNLQCLV HEKTPVSDRV TKCTESLVN RRPCFALE
 493.....503.....513.....523.....533.....543.....553.....563.....573.....
 V DETYVPKEFN AETFTFHADI CTLSEKERQI KKQATLVELV KHKPKATKEQ LKAVMDFAA FVEKCKADD KETCFAEEGK KLVAASQAA

Query
 A13.....23.....33.....43.....53.....63.....73.....83.....93.....103.....113.....123.....
 KSEVAHRFKD LGENFKALV LIAFAQYLQ CPFDHVKLV NEVTEFAKT VADESAEND KSLHTLFGDK LCT.....ADK AK QEPQRNEC L QHKKDNNPLP RLVRPEVDVM CTAHFHNEE
 133.....143.....153.....163.....173.....183.....193.....203.....213.....223.....233.....243.....
 T FLKKLYEIA RRHPFYAPE LFFAFKRYKA AFTECCQAD KAACLLPKLD ELRDEGKASS AKQRLKCASL QKGEBAFKA IAVARLSQRF PKAEFAEVSK LVTDLTQVHT ECHGDLLE
 253.....263.....273.....283.....293.....303.....313.....323.....333.....343.....353.....363.....
 C AODRADLANY ICENQDSSSS KKECECKPL LEKSHCIAEV ENDEMPADLP SLADRVESK DVCKWYAEK DVFQGFILVE YARRPDPVSV VLLRLRLRY FTLEKCCAA ADPHEVAK
 373.....383.....393.....403.....413.....423.....433.....443.....453.....463.....473.....483.....
 V FDEFKPLVEE PQLIKQNC E LFEQLGEYK QNALLVRYTK KVPQVSTPTL VEVSRNLGKV GSKCKHPEA KRMPCAEDYL SVLNLQCLV HEKTPVSDRV TKCTESLVN RRPCFALE
 493.....503.....513.....523.....533.....543.....553.....563.....573.....
 V DETYVPKEFN AETFTFHADI CTLSEKERQI KKQATLVELV KHKPKATKEQ LKAVMDFAA FVEKCKADD KETCFAEEGK KLVAASQAA

Attention: The chains of the uploaded protein are broken at position A76(HIS), A88(ALA), A89(ASP), A91(CYS), A101(CYS), A103(LEU)! These positions are highlighted in the sequence list with black underline!
 Please check if these positions fall in the detected pockets.



FitPocket ID	Template ID	Template Protein	Template Ligand	FP2	Pocket Identity	Pocket RMSD	Cc
F1	t1	2bxg	IBP	1.0	1.0	0.46	C
F2	t2	6oci	IBP	1.0	0.9	0.44	C
F3	t3	2vdb	NPS	0.64	1.0	0.61	C
F1	t4	4or0	NPS	0.64	0.92	0.88	C
F2	t5	5dby	NPS	0.64	0.9	0.48	C

Figura 34 – Detecção de cavidades baseada em modelos (*Template-based cavity detection*)

3.1. Subseção 1: Sequências de resíduos de aminoácidos da proteína modelo e da proteína submetida – Consulta A (Query A)

Nesta primeira subseção, são apresentadas as sequências de resíduos de aminoácidos da proteína modelo e da proteína submetida.

No exemplo da Figura 35, é apresentada a sequência de resíduos de aminoácidos da **proteína modelo (2bxg A)** (a). Assim, o CB-Dock2[®] realizou uma busca e selecionou como proteína modelo a do arquivo 2BXG.pdb, que refere-se à albumina sérica humana ligada ao ibuprofeno (Ghuman, Zunszain, Petitpas, Bhattacharya, Curry, 2011). Os resíduos destacados em azul referem-se aos associados à cavidade C1.

Logo abaixo, em **Consulta A (Query A)** (b), é apresentada a sequência de resíduos de aminoácidos da proteína submetida (no caso, a 6M5D.pdb), sendo que os resíduos destacados em azul referem-se aos associados à cavidade C1.

Detected pockets from referring homologous proteins (Submitted Protein: [6m5d.pdb](#))
 Tips: Click the table below to view the pocket information. Click the residue in the sequence list to view the residue you are interested in.

2bxg
 A
13.....23.....33.....43.....53.....63.....73.....83.....
 -SEVAHRFKD LGEENFKALV LIAFAQYLQQ CPFEDHVKLV NEVTEFAKTC VADESAENC D KSLHTLFGDK LCTVATLRET YGEMADC
 .93.....103.....113.....123.....133.....143.....153.....163.....
 CAK QEPERNECFL QHKDDNP NLP RLVRPEVDVM CTA FHDNEET FLKKYLYEIA RRHPYFYAPE L LFFAKRYKA AFTECCQA
 .173.....183.....193.....203.....213.....223.....233.....243.....
 AD KAACLLPKLD ELRDEGKASS AKQRLKASL QKGERAFKA WAVARLSQRF PKAEFAEVSK LVTDLTKVHT ECCHGDLL
 .253.....263.....273.....283.....293.....303.....313.....323.....
 EC ADDRDLAKY ICENQDSISS KLKECCEKPL LEKSHCIAEV ENDEMPADLP SLAADFVESK DVCKNYAEAK DVFGMFL
 .333.....343.....353.....363.....373.....383.....393.....403.....
 YE YARRHPDYSV VLLLR LAKTY ETTLEKCCAA ADPHECYAKV FDEFKPLVEE PQLIKQNC E LFEQLGEYKF QNALLVRY
 .413.....423.....433.....443.....453.....463.....473.....483.....
 TK KVPQVSTPTL VEVSRLGKV GSKCKHPEA KRMPCAEDYL SVVLNQLCVL HEKTPVSDRV TKCTE SLVN RRPCFSAL
 .493.....503.....513.....523.....533.....543.....553.....563.....
 EV DETYVPKEFN AETTFHADI CTLSEKERQI KKQ TALVELV KHKPKATKEQ LKAVMDDFAA FVEKCKKADD KETCF AEE
 .573.....
 GK KLVAASQAA

Query
 A
13.....23.....33.....43.....53.....63.....73.....
 KSEVAHRFKD LGEENFKALV LIAFAQYLQQ CPFEDHVKLV NEVTEFAKTC VADESAENC D KSLHTLFGDK LCT-----ADC
 .93.....103.....113.....123.....133.....143.....153.....163.....
 -AK QEPQRNEC-L QHKDDNP NLP RLVRPEVDVM CTA FHDNEET FLKKYLYEIA RRHPYFYAPE L LFFAKRYKA AFTECCQA
 .173.....183.....193.....203.....213.....223.....233.....243.....
 AD KAACLLPKLD ELRDEGKASS AKQRLKASL QKGERAFKA WAVARLSQRF PKAEFAEVSK LVTDLTKVHT ECCHGDLL
 .253.....263.....273.....283.....293.....303.....313.....323.....
 EC ADDRDLAKY ICENQDSISS KLKECCEKPL LEKSHCIAEV ENDEMPADLP SLAADFVESK DVCKNYAEAK DVFGMFL
 .333.....343.....353.....363.....373.....383.....393.....403.....
 YE YARRHPDYSV VLLLR LAKTY ETTLEKCCAA ADPHECYAKV FDEFKPLVEE PQLIKQNC E LFEQLGEYKF QNALLVRY
 .413.....423.....433.....443.....453.....463.....473.....483.....
 TK KVPQVSTPTL VEVSRLGKV GSKCKHPEA KRMPCAEDYL SVVLNQLCVL HEKTPVSDRV TKCTE SLVN RRPCFSAL
 .493.....503.....513.....523.....533.....543.....553.....563.....
 EV DETYVPKEFN AETTFHADI CTLSEKERQI KKQ TALVELV KHKPKATKEQ LKAVMDDFAA FVEKCKKADD KETCF AEE
 .573.....
 GK KLVAASQAA

Attention: The chains of the uploaded protein are broken at position A76(THR), A88(ALA), A89(ASP), A91(CYS), A101(CYS), A103(LEU)! These positions are highlighted in the sequence list with black underline! Please check if these positions fall in the detected pockets.

Figura 35 – Analisando as cavidades da proteína pela detecção de cavidades baseada em modelos (*Template-based cavity detection*).

3.2. Subseção 2: Visualizador NGL® (NGL® - viewer)

Esta subseção utiliza o **visualizador NGL® (NGL-viewer®)** (Figura 36) para exibir a proteína submetida e suas cavidades. Este programa possui vários **botões de controle**:

- **Centro (Center) (a)**: clicando neste botão, a molécula é reposicionada no centro da imagem;
- **Tela cheia (Full screen) (b)**: este botão exibe a proteína em tela cheia. Para retornar ao formato anterior, basta clicar na tecla ESC;
- **Exibir superfície (Show surface) (c)**: este botão exibe a superfície completa da proteína (não apenas da área da cavidade, como mostrado na Figura 24);
- **Esconder bolso (Hide pocket) (d)**: este botão permite deixar de exibir a superfície da parte da molécula ao redor da cavidade;
- **Exibir proteína modelo (Show template(Pro)) (e)**: exibe a proteína modelo;
- **Exibir ligante modelo (Show template(Lig)) (f)**: exibe o ligante modelo;

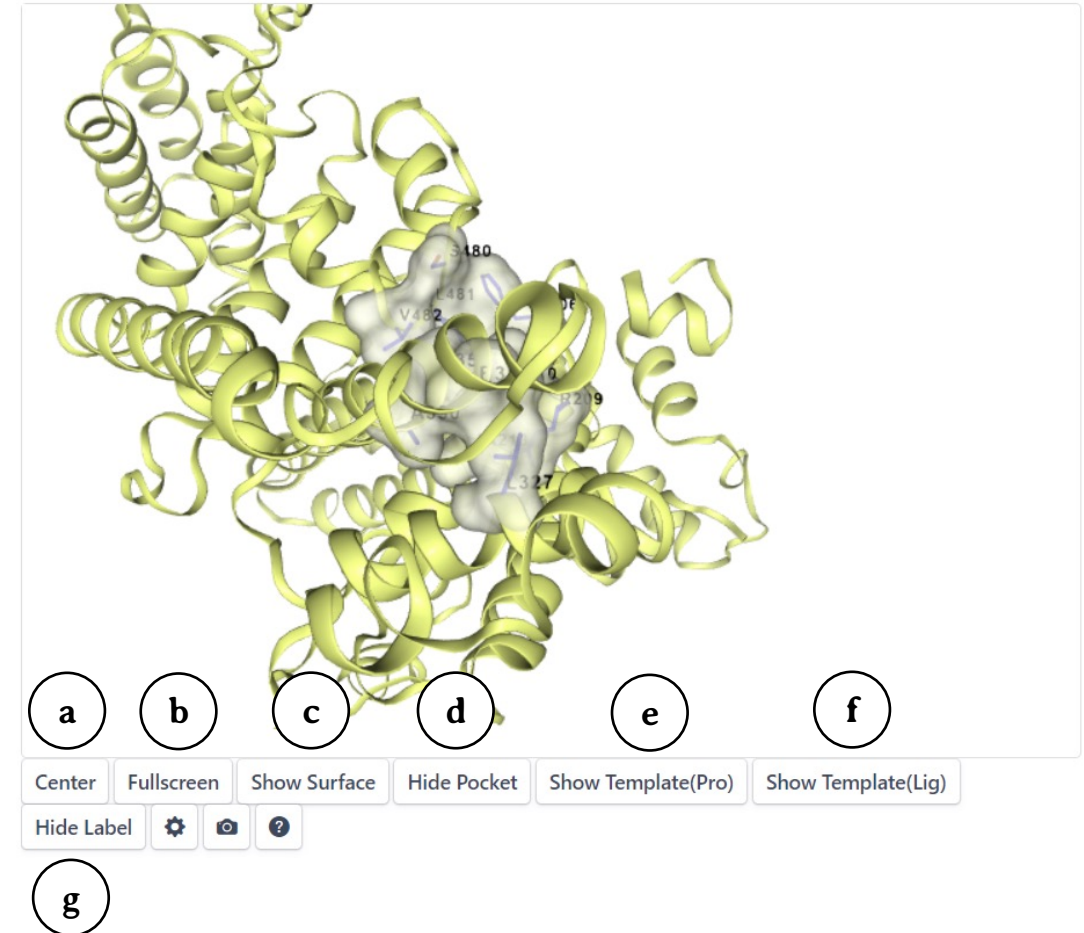


Figura 36 – Botões de controle da figura apresentada.

- **Esconder legendas (*Hide label*) (Figura 37a):** este botão remove as legendas do tipo e número de resíduo de aminoácido exibidas;
- **Botão de configurações (*Settings*) (b):** que abre a **janela Configurações (*Settings*) (c)**, onde o usuário pode controlar a opacidade (*Opacity*), corte da superfície próxima ao operador (*ClipNear*), corte da superfície distante (*ClipFar*), exibição do fundo (*Background*) e cor preta do fundo (*Dark stage*);
- **Botão para salvar figura (*Save picture*) (d):** que abre a **janela Salvar figura (*Save Picture*) (e)** e determina as formas de salvar a figura em relação ao tamanho (*Scale*), Antisserilhamento (*Antialias*), Aparar (*Trim*), Fundo transparente (*Transparent*) e nome da figura (*PicName*). Para salvar a imagem, o usuário deve clicar no botão Exporte (*Export*) e a figura será salva no formato PNG;
- **Botão de ajuda (f):** que direciona o usuário para a página do manual do CB-Dock2[®].



Figura 37 – Botões de controle da figura apresentada.

3.3. Subseção 3: Quadro de informações das cavidades

Nesta subseção (Figura 38), são apresentados os resultados obtidos pelo **FitPocket**[®], que busca proteínas que possuam estruturas e composição semelhantes para determinar a localização de possíveis cavidades de interação com o ligante. Vale ressaltar que o número de cavidades foi determinado pelo usuário, porém é sugerido que este número seja preferencialmente abaixo de 20 para que não existam resultados falsos positivos (Liu, Yang, Gan, Chen, Xiao, Cao, 2022).

Neste quadro são apresentados os seguintes resultados:

- **Código de identificação da cavidade ajustada (*FitPocket ID*) (a):** Neste caso, a cavidade é determinada por comparação com outro arquivo PDB;
- **Código do modelo (*Template ID*) (b);**
- **Proteína modelo (*Template protein*) (c):** indica o código PDB do arquivo usado como modelo;
- **Ligante modelo (*Template ligand*) (d):** indica o código do ligante usado como modelo;

a	b	c	d	FitPocket ID ?	Template ID	Template Protein ?	Template Ligand ?	FP2 ?	Pocket Identity ?	Pocket RMSD ?	Consensus ?
⦿ F1	t1	2bxg	IBP	1.0	1.0	0.46	C1 (10%)				
○ F2	t2	6oci	IBP	1.0	0.9	0.44	C5 (78%)				
○ F3	t3	2vdb	NPS	0.64	1.0	0.61	C4 (38%)				
○ F1	t4	4or0	NPS	0.64	0.92	0.88	C1 (11%)				
○ F2	t5	5dby	NPS	0.64	0.9	0.48	C5 (70%)				

Figura 38 – Analisando as cavidades da proteína pela detecção de cavidades baseada na estrutura (*Structure-based cavity detection*).

- **Similaridade da proteína modelo e do ligante submetido (FP2) (Figura 39a):** que corresponde à similaridade dos resíduos de aminoácidos do modelo proteína-ligante com a proteína-ligante submetidos. Assim, ao usar a detecção de cavidades baseada em modelos (*Template-based cavity detection*), o CB-Dock2® procura em banco de dados, arquivos de proteína-ligante idênticos ou similares aos submetidos, sendo escolhidos aqueles que possuam $FP2 \geq 0,4$. Vale a pena ressaltar a diferença entre **similaridade e identidade de sequência de aminoácidos**. Enquanto na identidade de sequência, os resíduos de aminoácidos devem ser iguais; na similaridade, os aminoácidos podem ser substituídos por outros com propriedades físico-químicas semelhantes (Liu, Yang, Gan, Chen, Xiao, Cao, 2022). No exemplo da Figura 30, as cavidades F1 e F2 apresentaram o valor de FP2 igual a 1, o que corresponde a 100% de similaridade. Isto se deve ao fato de as proteínas modelo e os ligantes serem similares aos submetidos. No caso da F1, o arquivo modelo é o 2BXG.pdb (Ghuman, Zunszain, Petitpas, Bhattacharya, Curry, 2011), que corresponde à albumina sérica ligada ao ibuprofeno e no caso da F2, o arquivo modelo é o 6OCI.pdb, (Sekula, Zielinski, Bujacz, Bujacz, 2020), que corresponde à albumina equina ligada ao ibuprofeno;

a

FitPocket ID ?	Template ID	Template Protein ?	Template Ligand ?	FP2 ?	Pocket Identity ?	Pocket RMSD ?	Consensus ?
⦿ F1	t1	2bxg	IBP	1.0	1.0	0.46	C1 (10%)
○ F2	t2	6oci	IBP	1.0	0.9	0.44	C5 (78%)
○ F3	t3	2vdb	NPS	0.64	1.0	0.61	C4 (38%)
○ F1	t4	4or0	NPS	0.64	0.92	0.88	C1 (11%)
○ F2	t5	5dby	NPS	0.64	0.9	0.48	C5 (70%)

Figura 39 – Similaridade da proteína modelo e do ligante submetido (FP2).

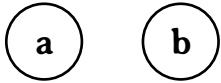
- **Identidade de sequência da proteína modelo e da proteína submetida na região da cavidade (*Pocket identity*) (Figura 40a):** conforme mencionado anteriormente, a identidade de sequência dos resíduos de aminoácidos das cavidades da proteína modelo e submetida corresponde à quantidade de resíduos idênticos. No caso da F1, o arquivo modelo é o 2BXG.pdb (Ghuman, Zunszain, Petitpas, Bhattacharya, Curry, 2011), que corresponde à albumina sérica ligada ao ibuprofeno. Por isso a identidade foi de 1.0 ou seja 100% dos resíduos de aminoácidos são idênticos, uma vez que os arquivos modelo e submetido apresentam a mesma sequência de aminoácidos na cavidade F1, uma vez que a proteína é a mesma, no caso a albumina sérica humana. No caso da F2, o arquivo modelo é o 6OCI.pdb, (Sekula, Zielinski, Bujacz, Bujacz, 2020), que corresponde à albumina equina ligada ao ibuprofeno e a identidade de sequência foi de 0,9, ou seja 90%, o que indica que a sequência de resíduos de aminoácidos da cavidade F1 da albumina sérica humana é 90% igual a do cavalo;

a

FitPocket ID ?	Template ID	Template Protein ?	Template Ligand ?	FP2 ?	Pocket Identity ?	Pocket RMSD ?	Consensus ?
⊙ F1	t1	2bxg	IBP	1.0	1.0	0.46	C1 (10%)
○ F2	t2	6oci	IBP	1.0	0.9	0.44	C5 (78%)
○ F3	t3	2vdb	NPS	0.64	1.0	0.61	C4 (38%)
○ F1	t4	4or0	NPS	0.64	0.92	0.88	C1 (11%)
○ F2	t5	5dby	NPS	0.64	0.9	0.48	C5 (70%)

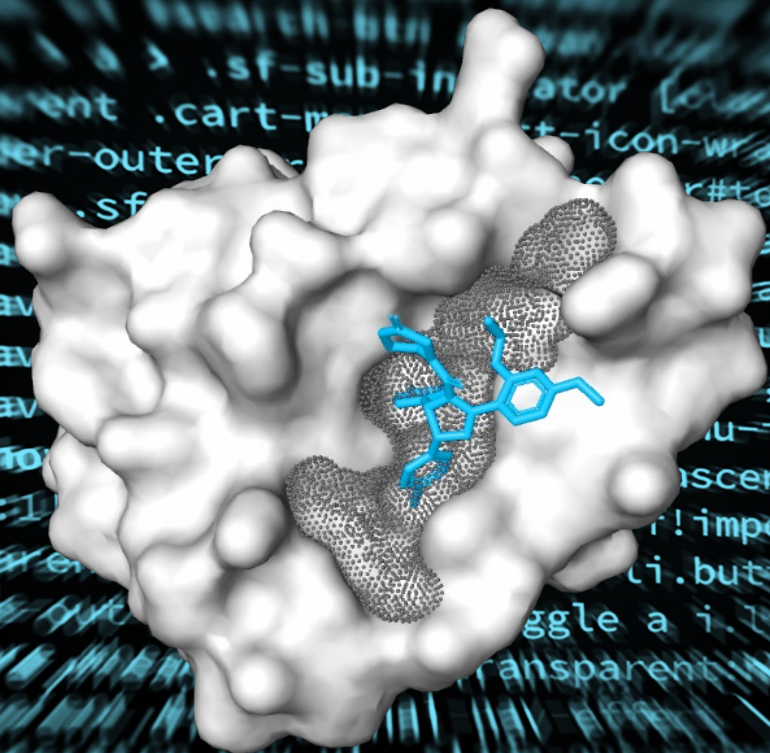
Figura 40 – Identidade de sequência da proteína modelo e da proteína submetida na região da cavidade (*Pocket identity*).

- **Raiz quadrada do desvio quadrático médio das proteínas modelo e submetida na região da cavidade (*Pocket RMSD*) (Figura 41a):** a raiz quadrada do desvio quadrático médio das posições atômicas (RMSD, do inglês *root-mean-square deviation of atomic positions*) corresponde à medida da distância média entre os átomos da cadeia principal de proteínas sobrepostas. Corresponde a uma medida de precisão, comparando a proteína modelo com a submetida. O valor do RMSD não apresenta valor negativo e quanto menor for o seu valor, mais próximos os átomos da proteína submetida estão da proteína modelo (Jewett, Huang, Ferrin, 2003);
- **Consenso de encaixe entre as cavidades estabelecidas pelo método de detecção de cavidades baseada na estrutura de acordo com a taxa de sobreposição dos resíduos da cavidade (*Consensus*) (Figura 41b):** apresenta a média de átomos da cadeia principal que estão posicionados no mesmo local, comparando os resíduos da proteína modelo e submetida. Quanto maior o valor, maior o número de átomos que apresentam a mesma posição (Blanes-Mira, Fernández-Aguado, de Andrés-López, Fernández-Carvajal, Ferrer-Montiel, Fernández-Ballester, 2022).



FitPocket ID ?	Template ID	Template Protein ?	Template Ligand ?	FP2 ?	Pocket Identity ?	Pocket RMSD ?	Consensus ?
⊙ F1	t1	2bxg	IBP	1.0	1.0	0.46	C1 (10%)
○ F2	t2	6oci	IBP	1.0	0.9	0.44	C5 (78%)
○ F3	t3	2vdb	NPS	0.64	1.0	0.61	C4 (38%)
○ F1	t4	4or0	NPS	0.64	0.92	0.88	C1 (11%)
○ F2	t5	5dby	NPS	0.64	0.9	0.48	C5 (70%)

Figura 41 – Raiz quadrada do desvio quadrático médio das proteínas modelo e submetida na região da cavidade (*Pocket RMSD*) e consenso de encaixe (*Consensus*)



PARTE 6

Docagem cega no CB-Dock2[®]

1. Docagem cega no CB-Dock2®

Após clicar no *link: View Result* da opção **Busca de cavidades (Search cavities)** na Seção **Trabalhos submetidos (Submitted jobs)**, será aberta uma nova guia que apresenta os resultados de docagem cega (Figura 42), a qual é dividida em três seções:

- **Seção 1:** Cavidades detectadas e poses de ligação de ligante previstas (*Detected cavities and predicted ligand-binding poses*) (a);
- **Seção 2:** Docagem cega baseada em estrutura (*Structure-based blind docking*) (b);
- **Seção 3:** Docagem cega baseada em modelo (*Template-based blind docking*) (c).

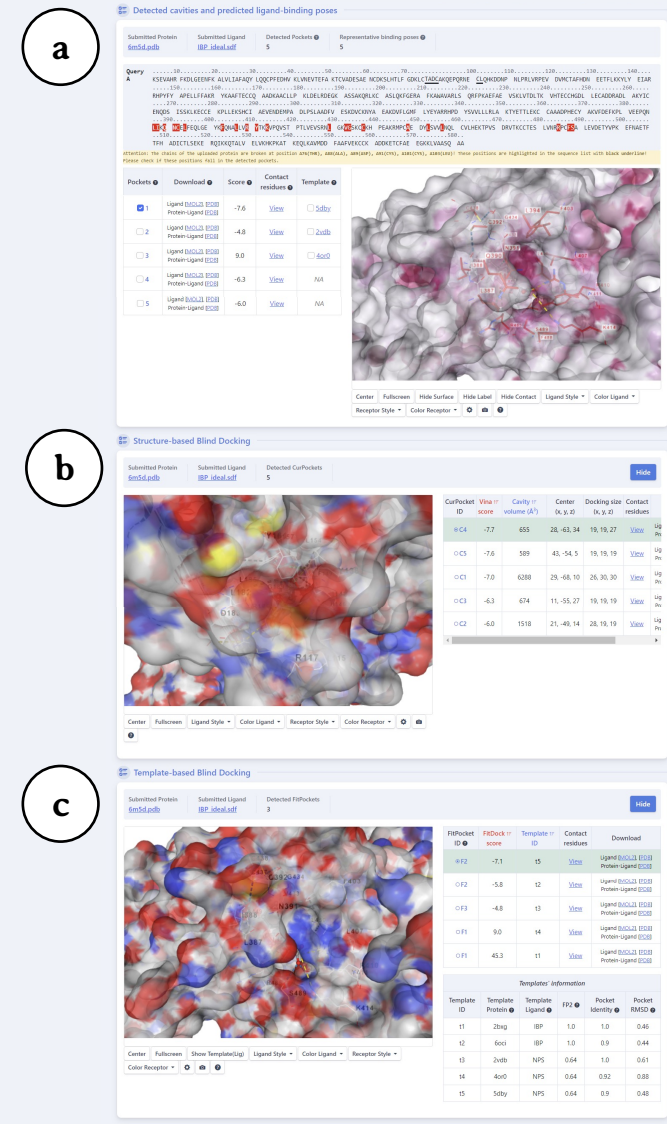


Figura 42 – Página com resultados de docagem cega.

2. Seção 1: Cavidades detectadas e poses de ligação de ligante previstas (*Detected cavities and predicted ligand-binding poses*)

A Seção Cavidades detectadas e poses de ligação de ligante previstas (*Detected cavities and predicted ligand-binding poses*) (Figura 43) é subdividida em quatro subseções:

- **Subseção 1:** Informações gerais (a);
- **Subseção 2:** Sequência de resíduos de aminoácidos da proteína analisada - Consulta A (*Query A*) (b);
- **Subseção 3:** Quadro de informações sobre as cavidades (c);
- **Subseção 4:** Visualizador NGL[®] (NGL[®]-viewer) (d).

Detected cavities and predicted ligand-binding poses

a

Submitted Protein	Submitted Ligand	Detected Pockets	Representative binding poses
6m5d.pdb	IBP_ideal.sdf	5	5

b

Query A

```

KSEVAHR FKDLGEEFK ALVLIFAAQY LQQCFEDHV KLVIEVTEFA KTCVADESAE NCCKSLHTLF GDKLCADCAKQEPORNE CLQHKDDNP NLPRLRPEV DVMCTAFHDI EETFKKLYL EIAH
RHPYFY APELLFAKR YKAAATECCQ AADKAACLLP KLDLRDEGK ASSAKRLKC ASLQKFGERA FKAMAVARLS QRFPKAEAE VSKLVDTLTK VHTECCHSGL LECADDRADL AKYIC
ENQDS ISSKKECE KPILLEKSHCI AEVENDEHPA DLPSLAADVF ESKDVCKNYA EAKDVFLGHF LVEYARRHPD YSVVLLRLA KYVETTLKCK CAAADPECV AKVFDEKPL VEEPQN
FEQLGGE YKQNA LQ TTKVPOVST PTLVEVSA GQSKKSH PEAKRMPCE DYSLVAGL CVLHEKTPVS DRVTXCCES LVMKPCA LEVDTYVPK EFNAETF
TFH ADICTLSEKE RQTKKQATLV ELVKHKPKAT KEQLKAVWDD FAAFVEKCKC ADDKETCFAE EGKLLVAASQ AA
  
```

Attention: The chains of the uploaded protein are broken at position A76(W), A88(ALA), A89(ASP), A91(CYS), A101(CYS), A103(LU) These positions are highlighted in the sequence list with black underline! Please check if these positions fall in the detected pockets.

c

Pockets	Download	Score	Contact residues	Template
<input checked="" type="checkbox"/> 1	Ligand [MOL] [PDB] Protein-Ligand [PDB]	-9.0	View	<input type="checkbox"/> 6oci
<input type="checkbox"/> 2	Ligand [MOL] [PDB] Protein-Ligand [PDB]	-4.8	View	<input type="checkbox"/> 2ydb
<input type="checkbox"/> 3	Ligand [MOL] [PDB] Protein-Ligand [PDB]	12.4	View	<input type="checkbox"/> 4o0
<input type="checkbox"/> 4	Ligand [MOL] [PDB] Protein-Ligand [PDB]	-6.3	View	NA
<input type="checkbox"/> 5	Ligand [MOL] [PDB] Protein-Ligand [PDB]	-6.0	View	NA

d

Center Fullscreen Hide Surface Hide Label Hide Contact Ligand Style Color Ligand Receptor Style Color Receptor

Figura 43 – Subseções da Seção Cavidades detectadas e poses de ligação de ligante previstas (*Detected cavities and predicted ligand-binding poses*).

2.1. Subseção 1: Informações gerais

Esta primeira subseção (Figura 44) apresenta as informações apresentadas a seguir:

- **Proteína submetida (*Submitted protein*) (a):** apresenta o código PDB do arquivo da proteína submetida;
- **Ligante submetido (*Submitted ligand*) (b):** apresenta o arquivo do ligante submetido;
- **Cavidades detectadas (*Detected pockets*) (c):** apresenta o número de cavidades detectadas, definida previamente pelo usuário. Cavidades que apresentem uma sobreposição de resíduos acima de 50% são consideradas as mesmas cavidades;
- **Poses de ligação representativas (*Representative binding poses*) (d):** poses com o valor de RMSD abaixo de 2 Å são consideradas as mesmas poses.

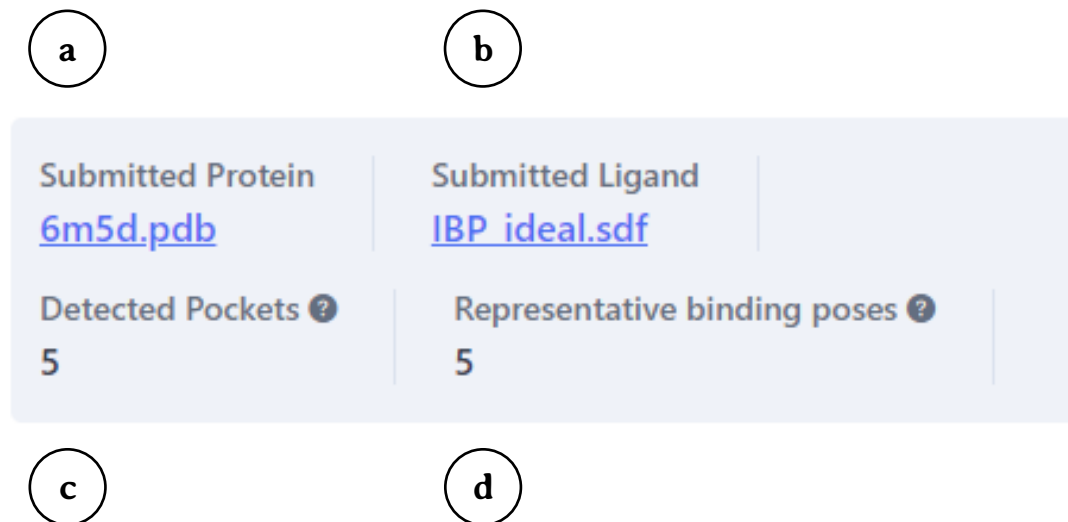


Figura 44 – Subseção 1: Informações gerais.

2.2. Subseção 2: Consulta A (Query A)

Esta segunda subseção (Figura 45) apresenta a sequência de resíduos de aminoácidos da proteína submetida, denominada **Consulta A (Query A)** (a). Os resíduos evidenciados na cor vermelha pertencem à cavidade C1 e o traço subscrito indica regiões “quebradas” da proteína. Esta seção apresenta também, no final, eventuais observações do CB-Dock2[®] sobre o arquivo submetido (b).

a

```

Query .....10.....20.....30.....40...
A      KSEVAHR FKDLGEENFK ALVLIAFAQY LQQCPFEDHV KLV
      .....50.....60.....70.....
      NEVTEFA KTCVADESAE NCDKSLHTLF GDKLCTADCAKQEP
      ...100.....110.....120.....130...
      QRNE CLQHKDDNP NLPRLVRPEV DVMCTAFHDN EET
      .....140.....150.....160.....
      FLKKYLY EIARRHPYFY APELLFFAKR YKAAFTECC
      170.....180.....190.....200....
      Q AADKAACLLP KLDELDEGK ASSAKQRLKC ASLQK
      ...210.....220.....230.....240.
      FGERA FKAWAVARLS QRFPKAEFAE VSKLVTDLTK V
      .....250.....260.....270.....
      HTECCHGDL LECADDRADL AKYICENQDS ISSKLKECC
      280.....290.....300.....310.....
      E KPLLEKSHCI AEVENDEMPA DLPSLAADFV ESKDV
      ...320.....330.....340.....350.
      CKNYA EAKDVFLGMF LYEYARRHPD YSVVLLRLA K
      .....360.....370.....380.....
      TYETTLEKC CAAADPHECY AKVFDEFKPL VEEPQNLIK
      390.....400.....410.....420....
      Q NCELFEQLGE YKQNALLVR YTKKVPQVST PTLVE
      ...430.....440.....450.....460.
      VSRNL GKVGSKCKH PEAKRMPCAE DYLSVVLNQL C
      .....470.....480.....490.....
      VLHEKTPVS DRVTKCCTES LVNRRPCFSA LEVDETYVP
      500.....510.....520.....530....
      K EFNAETTFH ADICTLSEKE RQIKQTALV ELVKH
      ...540.....550.....560.....570.
      KPKAT KEQLKAVMDD FAAFVEKCKC ADDKETCFAE E
      .....580..
      GKKLVAASQ AA
  
```

b

Attention: The chains of the uploaded protein are broken at position A76(THR), A88(ALA), A89(ASP), A91(CYS), A101(CYS), A103(LEU)! These positions are highlighted in the sequence list with black underline! Please check if these positions fall in the detected pockets.

Figura 45 – Subseção 2: Consulta A.

2.3. Subseção 3: Quadro de informações gerais

A terceira subseção (Figura 46) apresenta as seguintes informações:

- **Cavidades (Pockets) (a):** indica a numeração de cada cavidade da proteína;
- **Baixar (Download) (b):** permite que o usuário baixe um arquivo no formato MOL2 ou PDB do ligante (*Ligand [MOL2], [PDB]*) ou do complexo proteína-ligante no formato PDB (*Protein-Ligand [PDB]*) com a melhor posição prevista para cada cavidade;
- **Pontuação (Score) (c):** corresponde ao menor valor obtido a partir do cálculo do RMSD entre as melhores poses obtidas por cada método de docagem;
- **Resíduos de contato (Contact residues) (d):** indica os resíduos de contato em relação à melhor posição em cada cavidade. O link **View** abre uma janela que mostra a pontuação e os resíduos de contato de cada cavidade (e);
- **Modelo (Template) (f):** indica o arquivo PDB usado como modelo para detecção de cavidades. A abreviatura NA (*No Archive*) indica que não foram encontrados arquivos PDB para serem usados como modelo.

a	b	c	d	f
Pockets ?	Download ?	Score ?	Contact residues ?	Template ?
<input checked="" type="checkbox"/> 1	Ligand [MOL2] , [PDB] Protein-Ligand [PDB]	-9.0	View	<input type="checkbox"/> 6oci
<input type="checkbox"/> 2	Ligand [MOL2] , [PDB] Protein-Ligand [PDB]	-4.8	View	<input type="checkbox"/> 2vdb
<input type="checkbox"/> 3	Ligand [MOL2] , [PDB] Protein-Ligand [PDB]	12.4	View	<input type="checkbox"/> 4or0
<input type="checkbox"/> 4	Ligand [MOL2] , [PDB] Protein-Ligand [PDB]	-6.3	View	NA
<input type="checkbox"/> 5	Ligand [MOL2] , [PDB] Protein-Ligand [PDB]	-6.0	View	NA

e Pocket: 1 & Score: -9.0 ✕

Chain A: LEU387 ILE388 LYS389 GLN390 ASN391
 CYS392 PHE395 PHE403 LEU407 ARG410 TYR411
 LYS414 LEU430 VAL433 GLY434 SER435 CYS437
 CYS438 ALA449 LEU453 LEU457 ARG485 SER489

Figura 46 – Subseção 3: Quadro de informações gerais.

A partir dos arquivos obtidos na coluna **Baixar (Download)**, é possível observar o ligante e o complexo proteína ligante em programas de simulação computacional.

Na Figura 47, foram utilizados os arquivos de ligante e proteína-ligante para observar no programa computacional RasMol: o **ibuprofeno (a)** no modo *Sticks* e padrão de cores *CPK* e o **complexo ibuprofeno-albumina sérica humana com o ibuprofeno (b)** representado no modo *Spacefill* e padrão de cor *Chains* e a albumina sérica humana no modo *Cartoons* e padrão de cor *Chains*.

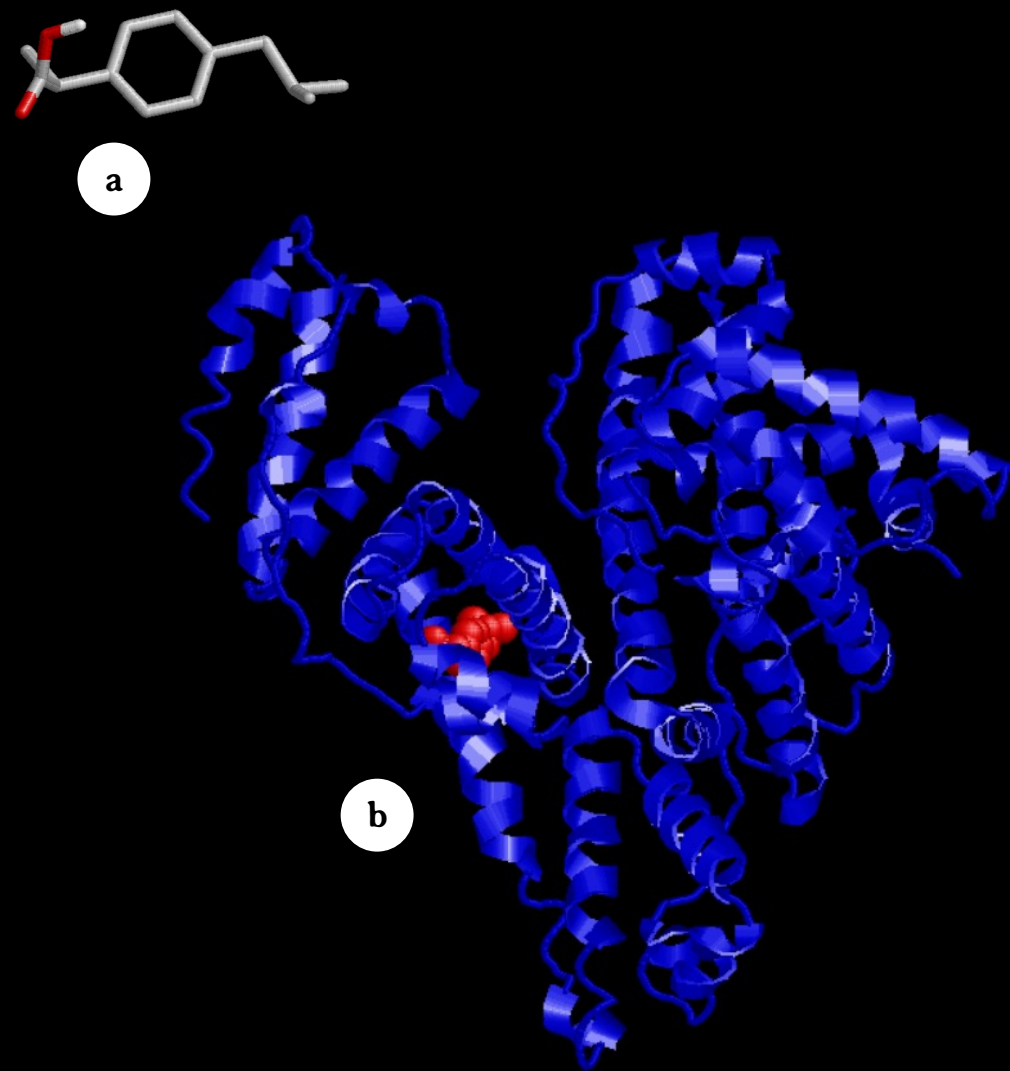
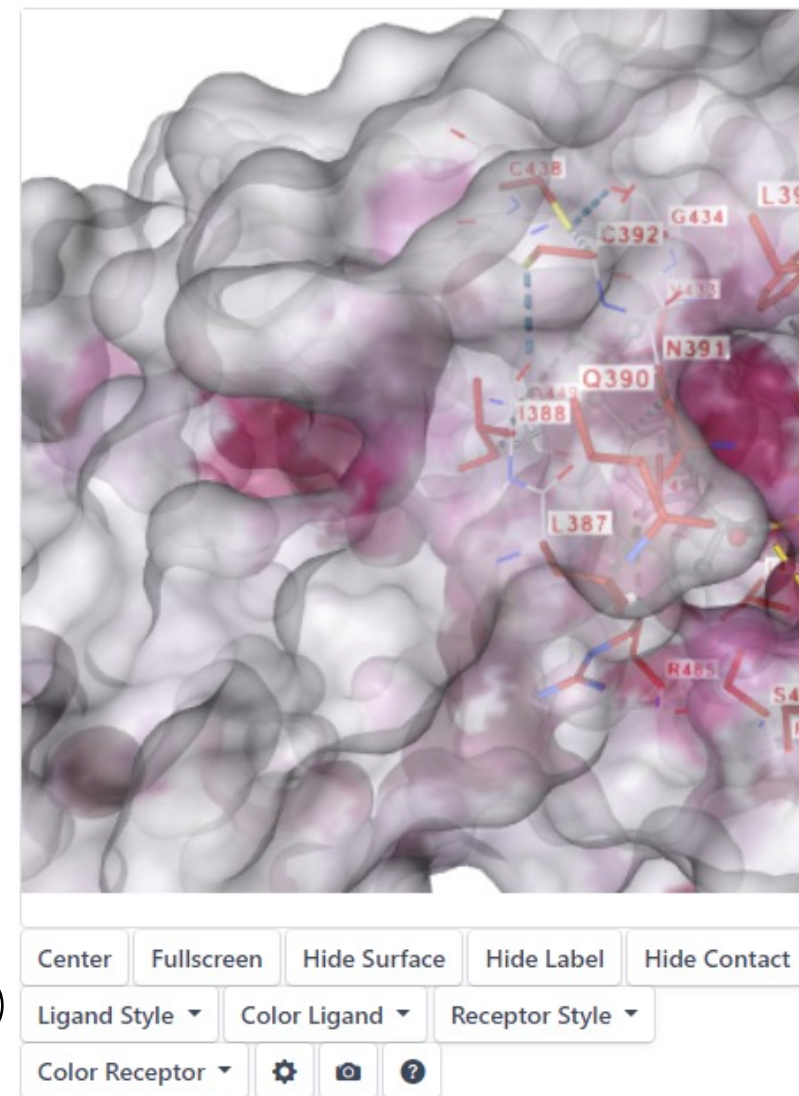


Figura 47 – Observação do ibuprofeno e do complexo ibuprofeno-albumina sérica humana no *software* RasMol.

2.4. Subseção 4: Visualizador NGL® (*NGL®-viewer*)

O visualizador NGL® (*NGL®-viewer*) (Figura 54) apresenta as seguintes funções:

- **Centro (*Center*);**
- **Tela cheia (*Fullscreen*);**
- **Esconder superfície (*Hide surface*);**
- **Esconder legendas (*Hide labels*);**
- **Esconder ligações químicas (*Hide contact*);**
- **Estilo do ligante (*Ligand style*):** exibe o ligante nos modos Bola e varetas (*Ball&Stick*), Preenchimento de espaço (*Spacefill*), Varetas (*Licorice*) e Superfície (*Surface*);
- **Cor do ligante (*Color ligand*):** por elemento químico (*by element*), por ligante (*by ligand*) ou monocromático (*uniform*);
- **Estilo do receptor (*Receptor style*):** Desenho (*Cartoon*), Cadeia principal (*Backbone*), Preenchimento de espaço (*Spacefill*), Varetas (*Licorice*), Superfície (*Surface*).



a

Figura 48 – Subseção 4: Visualizador NGL.

- **Cor do receptor (*Receptor color*):** por cadeia (*by chain*), por estrutura secundária (*by secondary structure*), por hidrofobicidade (*by hydrophobicity*), por elemento químico (*by element*), pelo fator-B (*by B-factor*) ou monocromático (*uniform*) (ver Anexo 3);
- **Botão de Configurações (*Settings*)**
- **Botão de Salvar figura (*Save picture*);**
- **Botão de Menu de ajuda (Figura 49a).**

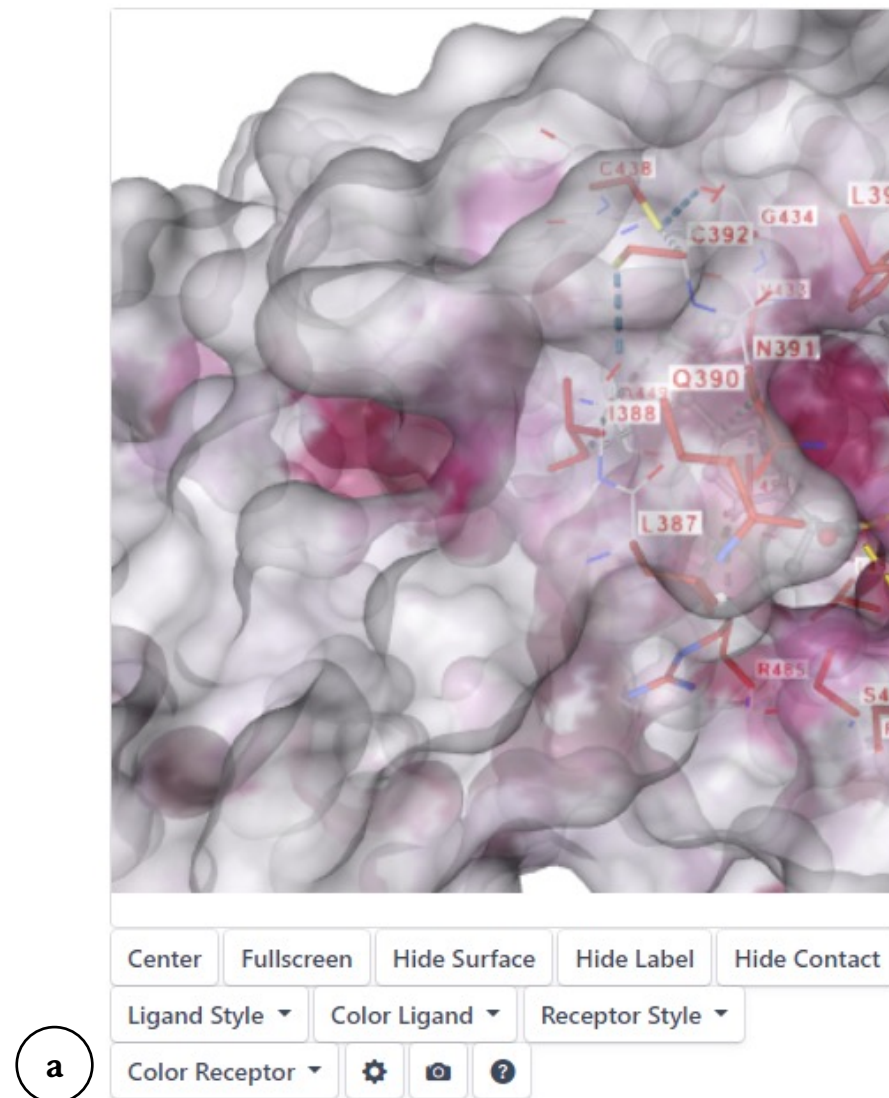


Figura 49 – Ícones do visualizador NGL.

3. Seção 2: Docagem cega baseada em estrutura (*Structure-based blind docking*)

A **Docagem cega baseada em estrutura** (*Structure-based blind docking*) utiliza o algoritmo **CurPocket**[®] para localização das cavidades da proteína e o programa **AutoDock Vina**[®] para obter os resultados de docagem cega (Figura 50).

O **CurPocket**[®] é um algoritmo que analisa as curvaturas da superfície da proteína para prever os sítios de ligação, ou seja, as cavidades de uma determinada proteína, além de calcular os centros e os tamanhos delas (Liu, Grimm, Dai, Hou, Xiao, Cao, 2020; Liu, Yang, Gan, Chen, Xiao, Cao, 2022).

O **AutoDock Vina**[®] realiza a Docagem cega baseada em estrutura (*Structure-based blind docking*) e utiliza três estratégias para estabelecer uma pontuação:

- **Campo de força:** baseia-se na força intermolecular de van der Waals e as interações eletrostáticas entre todos os átomos do ligante e da proteína usando um campo de força;
- **Empírico:** baseia-se na contagem do número de vários tipos de interações entre os dois parceiros de ligação;

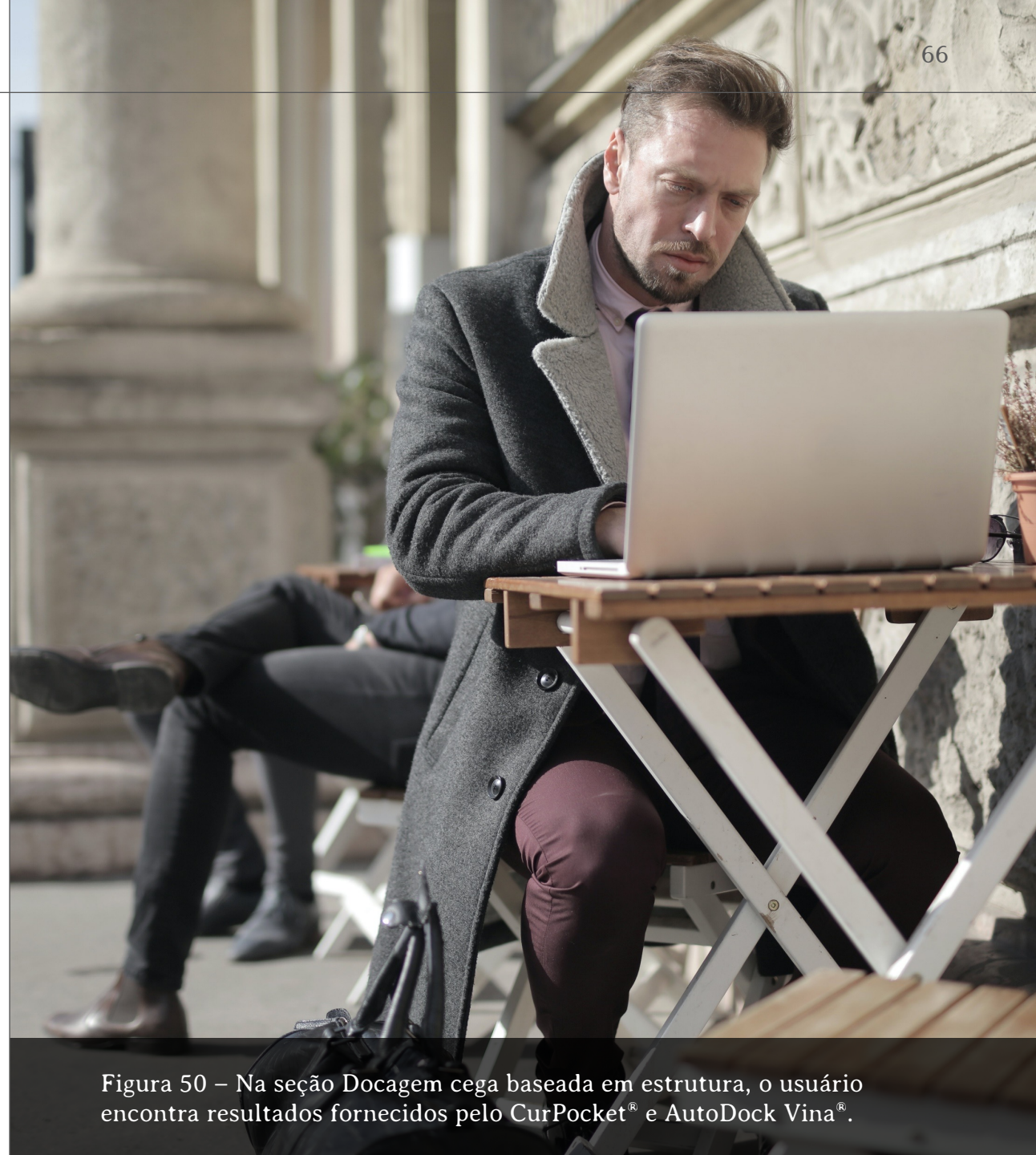


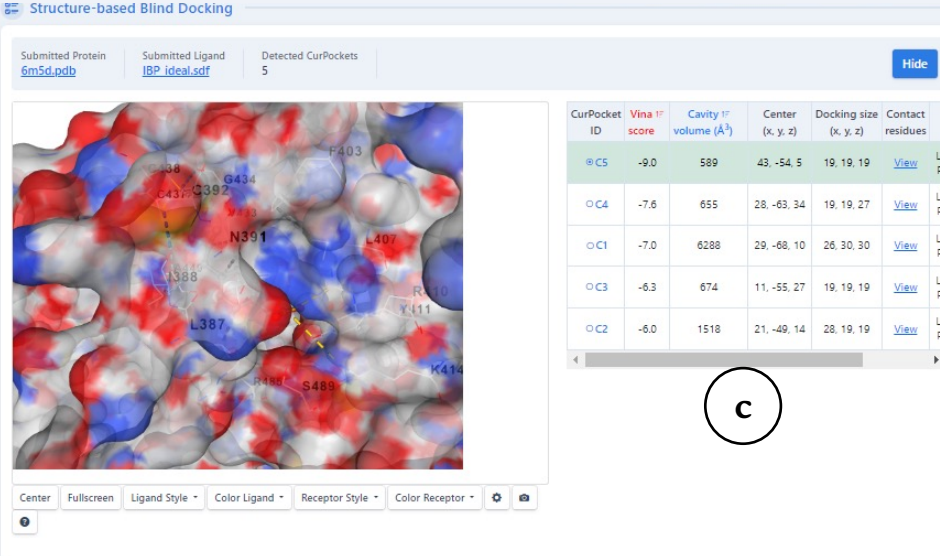
Figura 50 – Na seção Docagem cega baseada em estrutura, o usuário encontra resultados fornecidos pelo CurPocket[®] e AutoDock Vina[®].

- **Baseado em conhecimento:** é determinada a partir de observações estatísticas de contatos próximos intermoleculares disponíveis em bancos de dados (como *Protein Data Bank*) (Quiroga, Villarreal, 2016).

De forma resumida, o melhor resultado de docagem cega é estimado a partir do valor mais baixo indicado pelo AutoDock Vina[®] (kcal/mol), indicando a maior afinidade de ligação entre o ligante e a proteína (Xue, Liu, Russell, Li, Pan, Fu, Zhang, 2022) (Figura 51).

A **Seção 2: Docagem cega baseada em estrutura (*Structure-based blind docking*)** é subdividida em três subseções (Figura 51):

- **Subseção 1 (a):** Informações gerais;
- **Subseção 2 (b):** Visualizador NGL[®];
- **Subseção 3 (c):** Quadro de resultados.



The screenshot displays the 'Structure-based Blind Docking' interface. It includes a header with 'Submitted Protein' (6m5d.pdb), 'Submitted Ligand' (IBP_ideal.sdf), and 'Detected CurPockets' (5). A 'Hide' button is present in the top right. The main area is divided into three sections: (a) a 3D visualization of the protein-ligand complex, (b) a NGL viewer, and (c) a table of results.

CurPocket ID	Vina score	Cavity volume (Å ³)	Center (x, y, z)	Docking size (x, y, z)	Contact residues
OC5	-9.0	589	43, -54, 5	19, 19, 19	View Li Pr
OC4	-7.6	655	28, -63, 34	19, 19, 27	View Li Pr
OC1	-7.0	6288	29, -68, 10	26, 30, 30	View Li Pr
OC3	-6.3	674	11, -55, 27	19, 19, 19	View Li Pr
OC2	-6.0	1518	21, -49, 14	28, 19, 19	View Li Pr

Figura 51 – Subseções da Seção Docagem cega baseada em estrutura (*Structure-based blind docking*).

3.1. Subseção 1: Informações gerais

Esta primeira subseção (Figura 52) apresenta as informações apresentadas a seguir:

- **Proteína submetida (*Submitted protein*) (a)**: apresenta o código PDB do arquivo da proteína submetida;
- **Ligante submetido (*Submitted ligand*) (b)**: apresenta o arquivo do ligante submetido;
- **Cavidades detectadas (*Detected pockets*) (c)**: apresenta o número de cavidades detectadas, definida previamente pelo usuário. Cavidades que apresentem uma sobreposição de resíduos acima de 50% são consideradas as mesmas cavidades;
- **Botão ocultar (*Hide*) (d)**: oculta os resultados de Docagem cega baseada em estrutura (*Structure-based blind docking*).

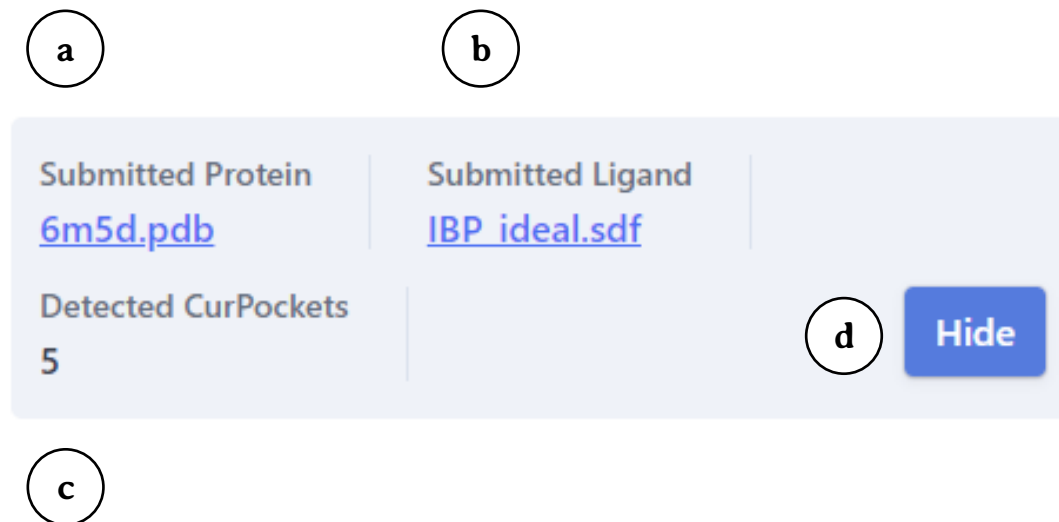


Figura 52 – Subseção 1: Informações gerais.

3.2. Subseção 2: Visualizador NGL® (NGL®-viewer)

O visualizador NGL® (Figura 53) apresenta os botões:

- **Centro (Center) e Tela cheia (Fullscreen);**
- **Estilo do ligante (Ligand style):** exibe o ligante nos modos Bola e varetas (*Ball&Stick*), Preenchimento de espaço (*Spacefill*), Varetas (*Licorice*) e Superfície (*Surface*);
- **Cor do ligante (Color ligand):** por elemento químico (*by element*), por ligante (*by ligand*) ou monocromático (*uniform*);
- **Estilo do receptor (Receptor style):** Desenho (*Cartoon*), Cadeia principal (*Backbone*), Preenchimento de espaço (*Spacefill*), Varetas (*Licorice*), Superfície (*Surface*) e Esconder o receptor (*Hide receptor*).
- **Cor do receptor (Receptor color):** por cadeia (*by chain*), por estrutura secundária (*by secondary structure*), por hidrofobicidade (*by hydrophobicity*), por elemento químico (*by element*), pelo fator-B (*by B-factor*) ou monocromático (*uniform*) (ver Anexo 3);
- **Botões de Configurações (Settings), Salvar figura (Save picture) e Menu de ajuda.**

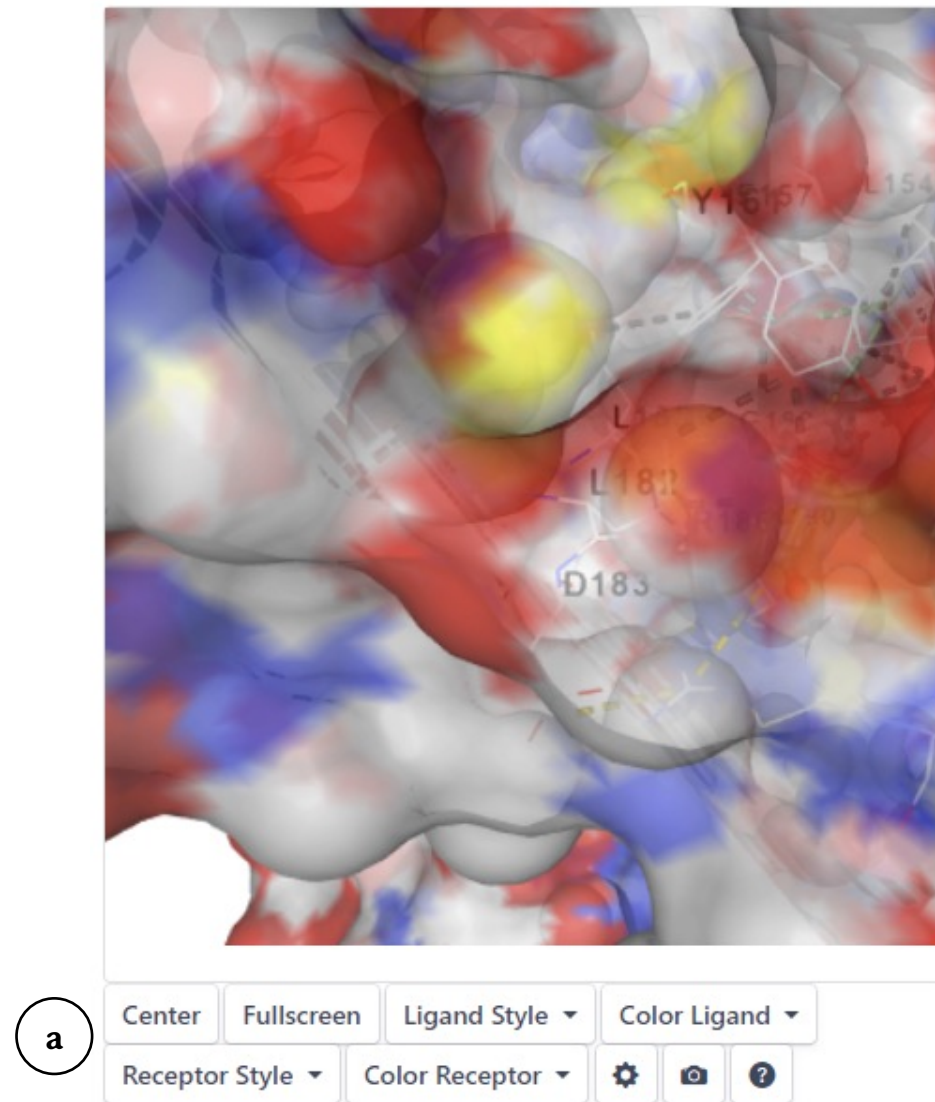


Figura 53 – Subseção 4: Visualizador NGL.

3.3. Subseção 3: Quadro de resultados

A terceira subseção (Figura 54) apresenta as seguintes informações:

- **Código de identificação da cavidade (*CurPocket ID*) (a):** indica a numeração da cavidade;
- **Pontuação Vina (*Vina score*) (b):** corresponde ao resultado de Docagem cega baseada em estrutura (*Structure-based blind docking*) obtido pelo AutoDock Vina®;
- **Volume da cavidade (*Cavity volume*) (c):** indica o tamanho da cavidade em angstroms (Å);
- **Centro (*Center*) (d):** indica a posição do centro da cavidade nos eixos x, y e z;
- **Tamanho do sítio de ligação (*Docking size*) (e):** o tamanho da cavidade nos eixos x, y e z;
- **Resíduos de contato (*Contact residues*) (f):** ao clicar no *link View* (g), exibe todos os resíduos da cavidade que interagem com o ligante;
- **Baixar (*Download*) (h):** permite baixar um arquivo no formato MOL2 ou PDB do ligante (*Ligand [MOL2], [PDB]*) ou do complexo proteína-ligante no formato PDB (*Protein-Ligand [PDB]*).

a	b	c	d	e	f	h
CurPocket ID	Vina score	Cavity volume (Å ³)	Center (x, y, z)	Docking size (x, y, z)	Contact residues	Download
OC5	-9.0	589	43, -54, 5	19, 19, 19	View	Ligand [MOL2] , [PDB] Protein-Ligand [PDB]
OC4	-7.6	655	28, -63, 34	19, 19, 27	View	Ligand [MOL2] , [PDB] Protein-Ligand [PDB]
OC1	-7.0	6288	29, -68, 10	26, 30, 30	View	Ligand [MOL2] , [PDB] Protein-Ligand [PDB]
OC3	-6.3	674	11, -55, 27	19, 19, 19	View	Ligand [MOL2] , [PDB] Protein-Ligand [PDB]
OC2	-6.0	1518	21, -49, 14	28, 19, 19	View	Ligand [MOL2] , [PDB] Protein-Ligand [PDB]

g

Pocket: F2 & Template: t2 & Score: -5.7

Chain A: LEU387 ILE388 ASN391 CYS392 PHE403
LEU407 ARG410 TYR411 LYS414 LEU430
VAL433 GLY434 CYS437 CYS438 ALA449 LEU453
ARG485 SER489

Figura 54 – Subseção 3: Quadro de resultados.

3.4. Interpretando os resultados de docagem cega baseada em estrutura (*Structure-based blind docking*)

A **Docagem cega baseada em estrutura (*Structure-based blind docking*)** corresponde a uma investigação inicial da interação entre um ligante e um substrato, onde o **CurPocket**[®] determina quais são as cavidades das proteínas que possuem maior chance de servirem como sítio de ligação (por padrão, geralmente as cinco cavidades mais prováveis) e o **AutoDock Vina**[®] apresenta um valor denominado **Vina score** que irá indicar qual cavidade da proteína apresenta a maior chance de interagir com o ligante.

No exemplo apresentado (Figura 55), a cavidade que apresenta a maior chance de interação com o ligante corresponde à C5, pois apresenta um menor valor para o **Vina score** (a). Entretanto, vale ressaltar que os resultados de **docagem cega baseada em estrutura** é apenas uma investigação inicial e que a proteína apresenta outros sítios de interação que devem ser buscados por meio de outros recursos.

a

CurPocket ID	Vina score	Cavity volume (Å ³)	Center (x, y, z)	Docking size (x, y, z)	Contact residues	Download
C5	-9.0	589	43, -54, 5	19, 19, 19	View	Ligand [MOL2] , [PDB] Protein-Ligand [PDB]
C4	-7.6	655	28, -63, 34	19, 19, 27	View	Ligand [MOL2] , [PDB] Protein-Ligand [PDB]
C1	-7.0	6288	29, -68, 10	26, 30, 30	View	Ligand [MOL2] , [PDB] Protein-Ligand [PDB]
C3	-6.3	674	11, -55, 27	19, 19, 19	View	Ligand [MOL2] , [PDB] Protein-Ligand [PDB]
C2	-6.0	1518	21, -49, 14	28, 19, 19	View	Ligand [MOL2] , [PDB] Protein-Ligand [PDB]

Figura 55 – Para análise de resultados da docagem cega baseada em estrutura, o usuário deve considerar os valores apresentados no *Vina Score*.

4. Seção 3: Docagem cega baseada em modelo (*Template-based blind docking*)

A **Docagem cega baseada em modelo** (*Template-based blind docking*) é realizada a partir da detecção de cavidades obtida pelo algoritmo **FitPocket**[®] e o programa **FitDock**[®] para obtenção dos escores de docagem cega (Figura 55). O **FitDock**[®] foi desenvolvido por Yang, Liu, Gan, Xiao, Cao em 2022, pesquisadores do Laboratório Yang Cao Lab da Universidade de Sichuan, também responsável pela programação do **CB-Dock2**[®]. O **FitDock**[®] explora as possíveis conformações do ligante no sítio ativo da proteína e gera poses de docagem refinadas, avaliando a afinidade do complexo ligante-proteína (Figura 8). Testes realizados pelos pesquisadores que desenvolveram este programa mostraram uma melhoria de 40% a 60% em termos de taxa de sucesso de encaixe e em velocidade mais rápida em relação aos métodos de encaixe populares, nos casos em que existem estruturas de moléculas modelo. O **FitDock** pode ser utilizado como uma ferramenta de *design* de medicamentos e de exploração de mecanismos moleculares. Este programa está disponível gratuitamente no [link: http://cao.labshare.cn/fitdock/](http://cao.labshare.cn/fitdock/).

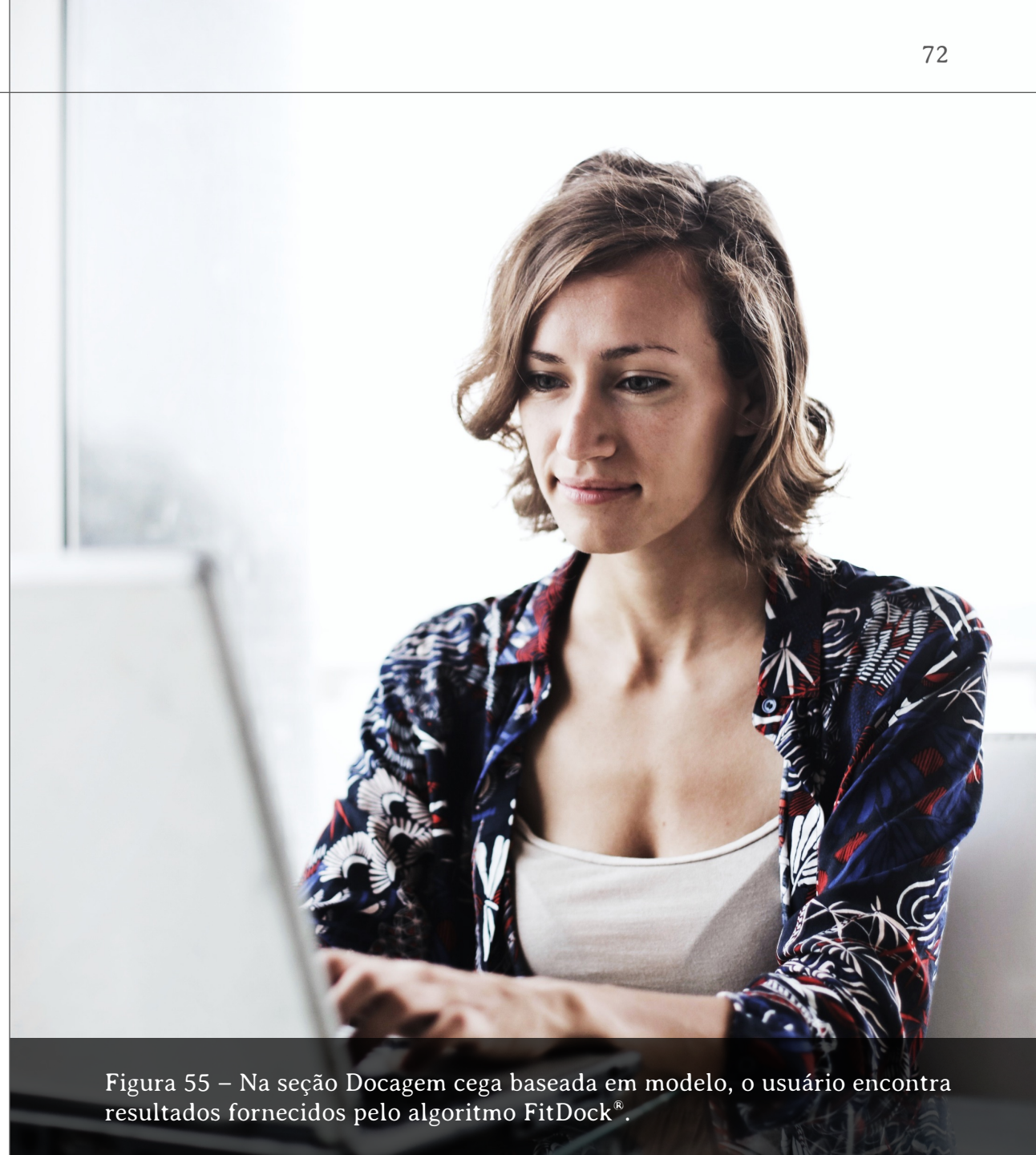


Figura 55 – Na seção Docagem cega baseada em modelo, o usuário encontra resultados fornecidos pelo algoritmo **FitDock**[®].

A Seção 3: Docagem cega baseada em modelo (*Template-based blind docking*) é subdividida em quatro subseções (Figura 56):

- **Subseção 1:** Informações gerais (a);
- **Subseção 2:** Visualizador NGL® (*NGL®-viewer*)(b);
- **Subseção 3:** Quadro de resultados (c);
- **Subseção 4:** Informações dos modelos (d).

a

Template-based Blind Docking

Submitted Protein: 6m5d.pdb | Submitted Ligand: IBP_ideal.sdf | Detected FitPockets: 3 | Hide

b

Center | Fullscreen | Show Template(Lig) | Ligand Style ▾ | Color Ligand ▾

Receptor Style ▾ | Color Receptor ▾ | ⚙️ | 🖨️ | 🔍

c

FitPocket ID	FitDock score	Template ID	Contact residues	Download
○ F2	-5.7	t2	View	Ligand [MOL3] [EDB] Protein-Ligand [EDB]
○ F2	-5.1	t5	View	Ligand [MOL3] [EDB] Protein-Ligand [EDB]
○ F3	-4.8	t3	View	Ligand [MOL3] [EDB] Protein-Ligand [EDB]
○ F1	12.4	t4	View	Ligand [MOL3] [EDB] Protein-Ligand [EDB]
○ F1	46.0	t1	View	Ligand [MOL3] [EDB] Protein-Ligand [EDB]

d

Templates' information

Template ID	Template Protein	Template Ligand	FP2	Pocket Identity	Pocket RMSD
t1	2bxg	IBP	1.0	1.0	0.46
t2	6oci	IBP	1.0	0.9	0.44
t3	2vdb	NPS	0.64	1.0	0.61
t4	4er0	NPS	0.64	0.92	0.88
t5	5dby	NPS	0.64	0.9	0.48

Figura 56 – Subseções da Seção Docagem cega baseada em modelo (*Template-based blind docking*).

4.1. Subseção 1: Informações gerais

Esta primeira subseção (Figura 57) apresenta as informações apresentadas a seguir:

- **Proteína submetida (*Submitted protein*) (a)**: apresenta o código PDB do arquivo da proteína submetida;
- **Ligante submetido (*Submitted ligand*) (b)**: apresenta o arquivo do ligante submetido;
- **Cavidades detectadas (*Detected FitPockets*) (c)**: apresenta o número de cavidades detectadas, definida pelo programa FitPocket®;
- **Botão ocultar (*Hide*) (d)**: oculta os resultados de Docagem cega baseada em modelo (*Template-based blind docking*).

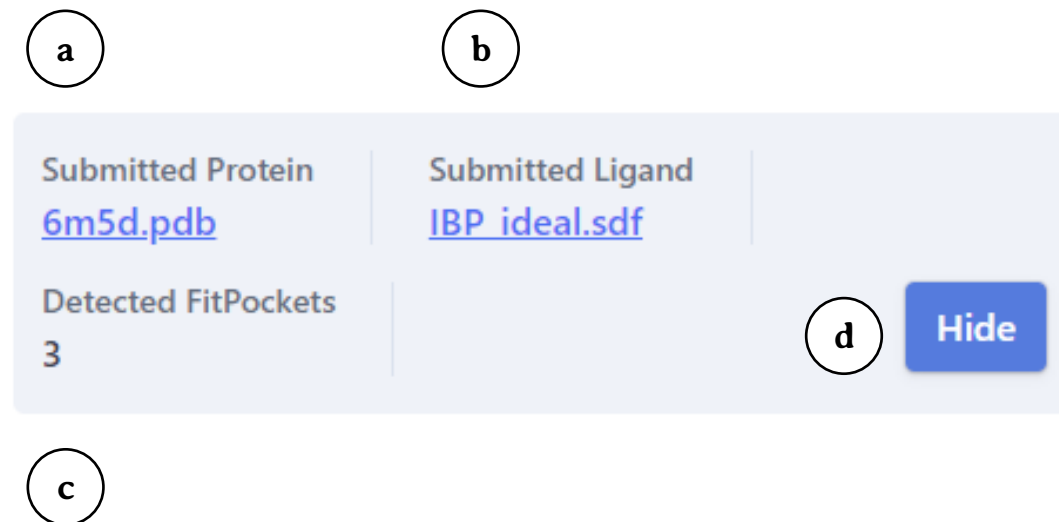


Figura 57 – Subseção 1: Informações gerais.

4.2. Subseção 2: Visualizador NGL® (NGL®-viewer)

O visualizador NGL® (Figura 58) apresenta os botões:

- **Centro (Center) e Tela cheia (Fullscreen);**
- **Exibir ligante modelo (Show template(Lig));**
- **Estilo do ligante (Ligand style):** Bola e varetas (*Ball&Stick*), Preenchimento de espaço (*Spacefill*), Varetas (*Licorice*) e Superfície (*Surface*);
- **Cor do ligante (Color ligand):** por elemento químico (*by element*), ligante (*by ligand*) ou monocromático (*uniform*);
- **Estilo do receptor (Receptor style):** Desenho (*Cartoon*), Cadeia principal (*Backbone*), Preenchimento de espaço (*Spacefill*), Varetas (*Licorice*), Superfície (*Surface*) e Esconder o receptor (*Hide receptor*).
- **Cor do receptor (Receptor color):** por cadeia (*By chain*), estrutura secundária (*By secondary structure*), hidrofobicidade (*By hydrophobicity*), elemento químico (*By element*), fator-B (*By B-factor*) ou monocromático (*Uniform*) (ver Anexo 3);
- **Botões com ícones de Configurações (Settings), Salvar figura (Save picture) e Menu de ajuda.**

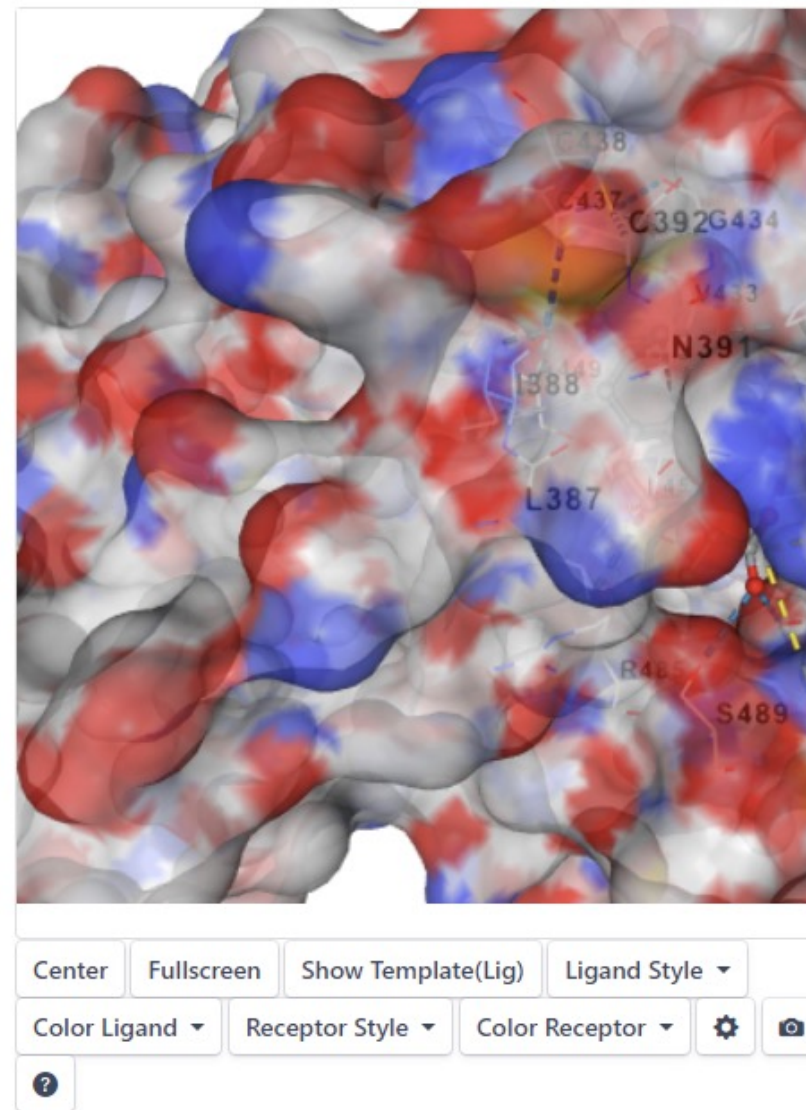



Figura 58 – Subseção 2: Visualizador NGL.

4.3. Subseção 3: Quadro de resultados

O Quadro de Resultados (Figura 59) apresenta as informações:

- **Código de identificação da cavidade (*FitPocket ID*) (a):** indica a numeração da cavidade;
- **Pontuação *FitDock* (*FitDock score*) (b):** corresponde ao resultado de Docagem cega baseada em modelo (*Template-based blind docking*) obtido pelo *FitDock*[®];
- **Código de identificação do modelo (*Template ID*) (c):** indica a numeração do modelo usado para comparação;
- **Resíduos de contato (*Contact residues*) (d):** ao clicar no *link View*, exibe todos os resíduos da cavidade que interagem com o ligante;
- **Baixar (*Download*) (e):** permite baixar um arquivo no formato MOL2 ou PDB do ligante (*Ligand [MOL2], [PDB]*) ou do complexo proteína-ligante no formato PDB (*Protein-Ligand [PDB]*).




FitPocket ID 	FitDock ^{TF} score	Template ^{TF} ID	Contact residues	Download
F2	-5.7	t2	View	Ligand [MOL2] , [PDB] Protein-Ligand [PDB]
F2	-5.1	t5	View	Ligand [MOL2] , [PDB] Protein-Ligand [PDB]
F3	-4.8	t3	View	Ligand [MOL2] , [PDB] Protein-Ligand [PDB]
F1	12.4	t4	View	Ligand [MOL2] , [PDB] Protein-Ligand [PDB]
F1	46.0	t1	View	Ligand [MOL2] , [PDB] Protein-Ligand [PDB]

Figura 59 – Subseção 3: Quadro de resultados.

4.4. Subseção 4: Informações dos modelos

Esta subseção (Figura 60) apresenta as seguintes informações:

- **Código de identificação do modelo (*Template ID*) (a):** indica o código do modelo;
- **Proteína modelo (*Template protein*) (b):** corresponde ao arquivo PDB da proteína usada como modelo;
- **Ligante modelo (*Template ligand*) (c):** corresponde ao código do ligante usado como modelo no *site* RCSB PDB;
- **Similaridade de topologia (*FP2*) (d):** indica a similaridade entre o ligante modelo e o submetido, sendo que o valor 1,0 corresponde a 100% de similaridade e assim por diante;
- **Identidade da cavidade (*Pocket identity*) (e):** indica a similaridade na sequência de resíduos na região da cavidade da proteína modelo e a consultada, sendo que o valor 1,0 corresponde a 100% de similaridade e assim por diante;
- **RMSD da cavidade (*Pocket RMSD*) (f):** indica o valor da raiz quadrada do desvio quadrático médio das posições atômicas (RMSD).

<i>Templates' information</i>					
Template ID	Template Protein ?	Template Ligand ?	FP2 ?	Pocket Identity ?	Pocket RMSD ?
t1	2bxg	IBP	1.0	1.0	0.46
t2	6oci	IBP	1.0	0.9	0.44
t3	2vdb	NPS	0.64	1.0	0.61
t4	4or0	NPS	0.64	0.92	0.88
t5	5dby	NPS	0.64	0.9	0.48

(a)
(b)
(c)
(d)
(e)
(f)

Figura 60 – Subseção 4: Informações dos modelos.

4.5. Interpretando os resultados de docagem cega baseada em modelo (*Template-based blind docking*)

A **Docagem cega baseada em modelo (*Template-based blind docking*)** corresponde a uma investigação inicial da interação entre um ligante e um substrato, onde o **FitPocket**[®] determina quais são as cavidades das proteínas que possuem maior chance de servirem como sítio de ligação e o **FitDock**[®] apresenta um valor denominado **FitDock score** que irá indicar qual cavidade da proteína apresenta a maior chance de interagir com o ligante.

No exemplo apresentado (Figura 62), o resultado obtido com este método de docagem ou seja, a cavidade que apresenta a maior chance de interação com o ligante corresponde à F2, pois apresenta um menor valor para o **FitDock score** (a).

Entretanto, vale ressaltar que os resultados de **docagem cega baseada em modelo** é apenas uma investigação inicial e que a proteína apresenta outros sítios de interação que devem ser buscados por meio de outros recursos.

a


FitPocket ID 	FitDock ^{TF} score	Template ^{TF} ID	Contact residues	Download
F2	-5.7	t2	View	Ligand [MOL2] , [PDB] Protein-Ligand [PDB]
F2	-5.1	t5	View	Ligand [MOL2] , [PDB] Protein-Ligand [PDB]
F3	-4.8	t3	View	Ligand [MOL2] , [PDB] Protein-Ligand [PDB]
F1	12.4	t4	View	Ligand [MOL2] , [PDB] Protein-Ligand [PDB]
F1	46.0	t1	View	Ligand [MOL2] , [PDB] Protein-Ligand [PDB]

Figura 62 – Na análise de resultados da docagem cega baseada em modelo, o usuário deve considerar os valores apresentados no **FitDock Score**.

4.6. Comparando os resultados de docagem cega baseada em estrutura (*Structure-based blind docking*) e docagem cega baseada em modelo (*Template-based blind docking*)

Para comparar os resultados de **docagem cega baseada em estrutura** (*Structure-based blind docking*) e de **docagem cega baseada em modelo** (*Template-based blind docking*), o usuário pode usar os visualizadores moleculares para verificar a posição dos sítios de ligação ou clicar no link **View** (Visualizar) (Figura 63) para visualizar os resíduos de aminoácidos de cada cavidade.

No exemplo apresentado, observa-se que a cavidade 5 indicada pela **docagem cega baseada em estrutura** (*Structure-based blind docking*) e a cavidade F2 indicada pela **docagem cega baseada em modelo** (*Template-based blind docking*), compartilham 18 aminoácidos idênticos, indicando que ambas docagens sugerem que a mesma cavidade é a mais provável para estabelecer ligação entre o IBP e a HSA.

a

Pocket: C5 & Score: -9.0

Chain A: LEU387 ILE388 LYS389 GLN390 ASN391
CYS392 PHE395 PHE403 LEU407 ARG410 TYR411
LYS414 LEU430 VAL433 GLY434 SER435 CYS437
CYS438 ALA449 LEU453 LEU457 ARG485 SER489

b

Pocket: F2 & Template: t2 & Score: -5.7

Chain A: LEU387 ILE388 ASN391 CYS392 PHE403
LEU407 ARG410 TYR411 LYS414 LEU430
VAL433 GLY434 CYS437 CYS438 ALA449 LEU453
ARG485 SER489

Figura 62 – Na análise de resultados da docagem cega baseada em modelo, o usuário deve considerar os valores apresentados no *FitDock Score*.



CRÉDITOS DAS FIGURAS

Créditos das figuras

Capa, páginas capitulares e contracapa.

Fundo. Fonte: Modificado de: Negative space. Pink white black purple blue textile web scripts [Internet]. 2017 Feb 09 [acesso 08 jan 2023]. Disponível em: <https://www.pexels.com/photo/pink-white-black-purple-blue-textile-web-scripts-97077/>. Figura registrada como: *Free to use. No attribution required. Creative Commons Zero (CC0) license.*

Molécula. Fonte: Modificado de: CB-Dock2: cavity detection guided blind docking [Internet]. 2023 [acesso 08 jan 2023]. Disponível em: <https://cadd.labshare.cn/cb-dock2/php/index.php>. Figura utilizada com a permissão do autor.

Página 4: Agradecimentos.

Fonte: Pixabay. A pirate ship sailing on sea during golden hour [Internet]. 2017 Oct 01 [acesso 09 jan 2023]. Disponível em: <https://www.pexels.com/photo/a-pirate-ship-sailing-on-sea-during-golden-hour-37730/>. Figura registrada como: *Free to use. No attribution required. Creative Commons Zero (CC0) license.*

Página 5: Epígrafe.

Fonte: Woroniecki J. Clear glass bottle on brown sand [Internet]. 2020 Jun 30 [acesso 09 jan 2023]. Disponível em: <https://www.pexels.com/photo/clear-glass-bottle-on-brown-sand-4724049/>. Figura registrada como: *Free to use. No attribution required. Creative Commons Zero (CC0) license.*

Página 6: Apresentação.

Fonte: Mohamad Z. Beach during golden hour [Internet]. 2015 Dec 16 [acesso 09 jan 2023]. Disponível em: <https://www.pexels.com/photo/beach-during-golden-hour-21787/>. Figura registrada como: *Free to use. No attribution required. Creative Commons Zero (CC0) license.*

Página 11: Figura 1 – O termo ancoragem é utilizado em referência ao lançamento da âncora, uma manobra essencial para o aporte de uma embarcação.

Fonte: Damboldt M. Photography of ship [Internet]. 2018 Jan 22 [acesso 09 jan 2023]. Disponível em: <https://www.pexels.com/photo/photography-of-ship-799091/>. Figura registrada como: *Free to use. No attribution required. Creative Commons Zero (CC0) license.*

Página 12: Figura 2 – Programas de docagem cega, como o CB-Dock2®, podem ser utilizados no desenvolvimento de novos fármacos.

Fonte: Ready made. Medicines on the table [Internet]. 2020 Mar 07 [acesso 09 jan 2023]. Disponível em: <https://www.pexels.com/photo/medicines-on-the-table-3850711/>. Figura registrada como: *Free to use. No attribution required. Creative Commons Zero (CC0) license.*

Página 13: Figura 3 – Na DC, a interação de uma proteína com diferentes tipos de candidatos a medicamentos pode ser testada.

Fonte: Pixabay. Photo of medicinal drugs [Internet]. 2016 Jul 25 [acesso 28 fev 2024]. Disponível em: <https://www.pexels.com/photo/close-up-photo-of-medicinal-drugs-159211/>. Figura registrada como: *Free to use. No attribution required. Creative Commons Zero (CC0) license.*

Página 15: Figura 5 – O CB-Dock2® pode ser utilizado em várias pesquisas na investigação de mecanismos de ação de fitoterápicos.

Fonte: Khongchum C. Person holding green-leafed plant [Internet]. 2019 Jun 13 [acesso 10 jan 2023]. Disponível em: <https://www.pexels.com/photo/person-holding-green-leafed-plant-2280551/>. Figura registrada como: *Free to use. No attribution required. Creative Commons Zero (CC0) license.*

Página 16: Figura 6 – O CB-Dock2® pode ser utilizado em pesquisa para investigar a ação de diferentes fármacos.

Fonte: Spinetta D. Close-up photo of assorted pills [Internet]. 2022 Nov 30 [acesso 12 jan 2023]. Disponível em: <https://www.pexels.com/photo/close-up-photo-of-assorted-pills-14586162/>. Figura registrada como: *Free to use. No attribution required. Creative Commons Zero (CC0) license.*

Página 18: Figura 7 – Para usar um programa de docagem cega, o usuário deve escolher primeiramente a proteína e ligante a serem estudados.

Fonte: Fauxels. Person typing [Internet]. 2019 Nov 05 [acesso 28 jul 2023]. Disponível em: <https://www.pexels.com/photo/person-typing-3183199/>. Figura registrada como: *Free to use. No attribution required. Creative Commons Zero (CC0) license.*

Página 19: Figura 8 – O termo albumina refere-se à cor branca da clara do ovo após cozimento.

Fonte: Nielsen K. Person holding pan with fried egg [Internet]. 2020 Dec 26 [acesso 15 jan 2023]. Disponível em: <https://www.pexels.com/photo/person-holding-pan-with-fried-egg-6294467/>. Figura registrada como: *Free to use. No attribution required. Creative Commons Zero (CC0) license.*

Página 20: Figura 9 – Estrutura secundária da albumina, com indicação dos subdomínios e sítios de ligação a ácidos graxos.

Fonte: Modificado de: Vicente MS, Hassunuma RM. Ibuprofeno: interações com a albumina sérica e ciclo-oxigenase-2. 1ª ed. Bauru: Canal 6 Editora; 2017. [acesso 15 jan 2023] Capítulo 2, Albumina; p. 14-7. Disponível em: <https://canal6.com.br/livreacesso/livro/ibuprofeno-interacoes-com-albumina-serica-e-ciclo-oxigenase-2/>. Figura utilizada com a permissão dos autores.

Página 21: Figura 10 – O ibuprofeno pode ser comprado no formato de comprimidos, sem necessidade de prescrição médica.

Fonte: Coetzee D. Generic Safeway tablets, 200 mg. File:Ibuprofen tablets in palm of hand.jpg [Internet]. 2013 Feb 18 [acesso 12 jan 2023]. Disponível em: https://commons.wikimedia.org/wiki/File:Ibuprofen_tablets_in_palm_of_hand.jpg. Figura registrada como: *Creative Commons CC0 1.0 Universal Public Domain Dedication.*

Página 22: Figura 11 – Estrutura bioquímica do ibuprofeno representada no esquema de fórmula estrutural plana (a) e no modelo de bolas e varetas (b).

Figura 11a: Fonte: Jü. Dexibuprofen_Structural_Formulae. File:Dexibuprofen Structural Formulae V.1.svg [Internet]. 2011 Feb 13 [acesso 12 jan 2023]. Disponível em: https://commons.wikimedia.org/wiki/File:Dexibuprofen_Structural_Formulae_V.1.svg. Figura registrada como: *Creative Commons CC0 1.0 Universal Public Domain Dedication.*

Figura 11b: Fonte: Mills B. Ball-and-stick model of the (R)-ibuprofen molecule, based on neutron diffraction data from N. Shankland, C. C. Wilson, A. J. Florence and P. J. Cox (1997). "Refinement of Ibuprofen at 100K by Single-Crystal Pulsed Neutron Diffraction". *Acta Crystallographica Section C* 53: 951-954. DOI:10.1107/S0108270197003193. File:Ibuprofen-3D-balls.png [Internet]. 2008 Jan 28 [acesso 12 jan 2023]. Disponível em: <https://commons.wikimedia.org/wiki/File:Ibuprofen-3D-balls.png>. Figura registrada como: *Creative Commons CC0 1.0 Universal Public Domain Dedication*.

Página 24: Figura 12 – Os arquivos da proteína e ligante podem ser obtidos gratuitamente na internet.

Fonte: Grabowska K. Crop person with smartphone and notebook using touchpad on laptop [Internet]. 2020 Apr 19 [acesso 16 jul 2023]. Disponível em: <https://www.pexels.com/photo/crop-person-with-smartphone-and-notebook-using-touchpad-on-laptop-4195409/>. Figura registrada como: *Free to use. No attribution required. Creative Commons Zero (CC0) license*.

Página 25: Figura 13 – Obtendo do arquivo PDB de uma proteína.

Fonte: Ito S, Senoo A, Nagatoishi S, Yamamoto M, Tsumoto K, Wakui N. Crystal structure of human serum albumin [Internet]. 2020 Dec 09 [acesso 11 jul 2023]. doi: 10.2210/pdb6m5d/pdb. Páginas e imagens do *site PDB* registradas como: *CC0 1.0 Universal (CC0 1.0) Public Domain*.

Página 26: Figura 14 – Baixando o arquivo referente ao IBP no *site Protein Data Bank*.

Fonte: Ghuman J, Zunszain PA, Petitpas I, Bhattacharya AA, Curry S. Human serum albumin complexed with ibuprofen [Internet]. 2011 Jul 13 [acesso 28 fev 2023]. doi: 10.2210/pdb2bxg/pdb. Páginas e imagens do *site PDB* registradas como: *CC0 1.0 Universal (CC0 1.0) Public Domain. Dedication*.

Página 27: Figura 15 – Baixando o arquivo referente ao IBP na Seção *Ligand*.

Fonte: IBP: Ibuprofen [Internet]. 2023 [acesso 28 fev 2023]. Disponível em: <https://www.rcsb.org/ligand/IBP>. Páginas e imagens do *site PDB* registradas como: *CC0 1.0 Universal (CC0 1.0) Public Domain*.

Página 28: Figura 16 – Os arquivos para ligantes podem ser baixados em vários sites disponíveis na internet.

Fonte: Shvets A. Serious man working on laptop in office [Internet]. 2020 Oct 14 [acesso 28 fev 2023]. Disponível em: <https://www.pexels.com/photo/serious-man-working-on-laptop-in-office-5324926/>. Figura registrada como: *Free to use. No attribution required. Creative Commons Zero (CC0) license.*

Página 39: Figura 26 – Caso o serviço solicitado no CB-Dock2[®] demore mais do que o esperado, é sugerido que o usuário atualize a página.

Fonte: Subiyanto K. Crop faceless female entrepreneur doing multitasking work on different devices [Internet]. 2020 May 25 [acesso 16 jul 2023]. Disponível em: <https://www.pexels.com/photo/crop-faceless-female-entrepreneur-doing-multitasking-work-on-different-devices-4474033/>. Figura registrada como: *Free to use. No attribution required. Creative Commons Zero (CC0) license.*

Página 65: Figura 50 – Na seção Docagem cega baseada em estrutura, o usuário encontra resultados fornecidos pelo CurPocket[®] e AutoDock Vina[®].

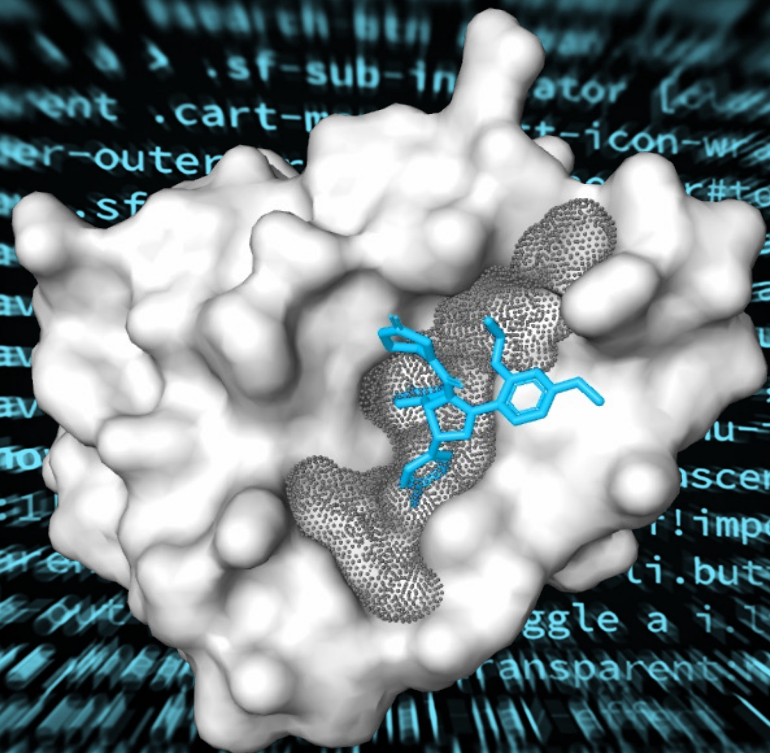
Fonte: Piacquadio A. Man in black jacket sitting on brown wooden folding chair using macbook [Internet]. 2020 Mar 01 [acesso 22 jul 2023]. Disponível em: <https://www.pexels.com/photo/man-in-black-jacket-sitting-on-brown-wooden-folding-chair-using-macbook-3837264/>. Figura registrada como: *Free to use. No attribution required. Creative Commons Zero (CC0) license.*

Página 70: Figura 55 – Na seção Docagem cega baseada em modelo, o usuário encontra resultados fornecidos pelo algoritmo FitDock[®].

Fonte: Piacquadio A. Woman in blue floral top sitting while using laptop [Internet]. 2018 Jan 17 [acesso 25 jul 2023]. Disponível em: <https://www.pexels.com/photo/woman-in-blue-floral-top-sitting-while-using-laptop-806835/>. Figura registrada como: *Free to use. No attribution required. Creative Commons Zero (CC0) license.*

Demais figuras não listadas anteriormente:

Fonte: CB-Dock2: cavity detection guided blind docking [Internet]. 2023 [acesso 08 jan 2023]. Disponível em: <https://cadd.labshare.cn/cb-dock2/php/index.php>. Todas figuras obtidas no *site* CB-Dock2[®] (incluindo o logo do servidor) foram utilizadas com a permissão do Dr. Yang Cao e equipe YANG CAO LAB[©].



REFERÊNCIAS

Referências

- AHFS patient medication information. Bethesda (MD): American Society of Health-System Pharmacists, Inc. Ibuprofen [Internet]. 2022 Jan 15 [acesso 12 jan 2023]. Disponível em: <https://medlineplus.gov/druginfo/meds/a682159.html>.
- Ascenzi P, di Masi A, Fanali G, Fasano M. Heme-based catalytic properties of human serum albumin. *Cell Death Discovery* [Internet]. 2015 Sep 07 [acesso 15 jan 2023];1(15025). doi: 10.1038/cddiscovery.2015.25.
- Bai Y, Li J, Wang G, Liang G, He Y, Hu L, Chen C. Mass spectrometric behaviors and molecular mechanisms of Amadori compounds between acute lung injury and diabetes based on network pharmacology. *Ann Palliat Med* [Internet]. 2021 Dec [acesso 19 dez 2022];10(12):12441-55. doi: 10.21037/apm-21-3298.
- Bai Y, Wang G, Lan J, Wu P, Liang G, Huang J, Wu Z, Wang Y, Chen C. Mass spectrometric behavior and molecular mechanisms of fermented deoxyanthocyanidins to alleviate ulcerative colitis based on network pharmacology. *Int J Anal Chem* [Internet]. 2022 [acesso 19 dez 2022];9293208. doi:10.1155/2022/9293208.
- Blanes-Mira C, Fernández-Aguado P, de Andrés-López J, Fernández-Carvajal A, Ferrer-Montiel A, Fernández-Ballester G. Comprehensive survey of consensus docking for high-throughput virtual screening. *Molecules* [Internet]. 2022 Dec 25 [acesso 14 jul 2023];28(1):175. doi: 10.3390/molecules28010175.
- CB-Dock2: Cavity-detection guided blind docking [Internet]. 2019 [acesso 18 dez 2022]. Disponível em: <https://cadd.labshare.cn/cb-dock2/php/index.php>
- CB-Dock: Cavity-detection guided blind docking [Internet]. 2022 [acesso 18 dez 2022]. Disponível em: <http://clab.labshare.cn/cb-dock/php/index.php>.

- Collins English Dictionary: docking [Internet]. 2023 [acesso 09 jan 2023]. Disponível em: <https://www.collinsdictionary.com/dictionary/english/docking>.
- Czub MP, Stewart AJ, Shabalín IG, Minor W. Organism-specific differences in the binding of ketoprofen to serum albumin. *IUCrJ* [Internet]. 2022 Jul 16 [acesso 21 abr 2024];9(Pt 5):551-561. doi: 10.1107/S2052252522006820.
- de Sá Ribeiro F, Lima LMTR. Linking B-factor and temperature-induced conformational transition. *Biophys Chem* [Internet]. 2023 Jul [acesso 19 jul 2023];298:107027. doi: 10.1016/j.bpc.2023.107027.
- Fanali G, di Masi A, Trezza V, Marino M, Fasano M, Ascenzi P. Human serum albumin: from bench to bedside. *Mol Aspects Med* [Internet]. 2012 Jun [acesso 15 jan 2023];33(3):209-90. doi: 10.1016/j.mam.2011.12.002.
- Ferreira LG, Dos Santos RN, Oliva G, Andricopulo AD. Molecular docking and structure-based drug design strategies. *Molecules* [Internet]. 2015 Jul 22 [acesso 18 dez 2022];20(7):13384-421. doi: 10.3390/molecules200713384.
- Ganai SA. Characterizing binding intensity and energetic features of histone deacetylase inhibitor pracinostat towards class I HDAC isozymes through futuristic drug designing strategy. *In Silico Pharmacol* [Internet]. 2021 [acesso 19 dez 2022];9(1):18. doi:10.1007/s40203-021-00077-y.
- Ghuman J, Zunszain PA, Petitpas I, Bhattacharya AA, Curry S. Human serum albumin complexed with ibuprofen [Internet]. 2011 Jul 13 [acesso 28 fev 2023]. doi: 10.2210/pdb2bxg/pdb.
- Halford GM, Lordkipanidzé M, Watson SP. 50th anniversary of the discovery of ibuprofen: an interview with Dr Stewart Adams. *Platelets* [Internet]. 2012 [acesso 21 dez 2022];23(6):415-22. doi: 10.3109/09537104.2011.632032.

- Hassan NM, Alhossary AA, Mu Y, Kwoh CK. Protein-ligand blind docking using Quickvina-w with inter-process spatio-temporal integration. *Sci Rep* [Internet]. 2017 Nov 13 [acesso 02 fev 2024];7(1):15451. doi: 10.1038/s41598-017-15571-7.
- Hassunuma RM, Souza AR. Desenvolvimento de scripts em software de simulação computacional para visualização de biomoléculas [Internet]. 1ª ed. São Paulo: Cultura Acadêmica; 2016 [acesso 13 jul 2023]. Capítulo 5, Ligação peptídica e resíduos de aminoácidos; p. 73-92. Disponível em: <https://www.culturaacademica.com.br/catalogo/desenvolvimento-de-scripts-em-software-de-simulacao-computacional-para-visualizacao-de-biomoleculas/>.
- Hetényi C, van der Spoel D. Blind docking of drug-sized compounds to proteins with up to a thousand residues. *FEBS Lett* [Internet]. 2006 Feb 20 [acesso 28 fev 2024];580(5):1447-50. doi: 10.1016/j.febslet.2006.01.074.
- Ito S, Senoo A, Nagatoishi S, Yamamoto M, Tsumoto K, Wakui N. Crystal structure of human serum albumin [Internet]. 2020 Dec 09 [acesso 11 jul 2023]. doi: 10.2210/pdb6m5d/pdb.
- Jewett AI, Huang CC, Ferrin TE. MINRMS: an efficient algorithm for determining protein structure similarity using root-mean-squared-distance. *Bioinformatics* [Internet]. 2003 Mar 22 [acesso 14 jul 2023];19(5):625-34. doi: 10.1093/bioinformatics/btg035.
- Liu Y, Grimm M, Dai WT, Hou MC, Xiao ZX, Cao Y. CB-Dock: a web server for cavity detection-guided protein-ligand blind docking. *Acta Pharmacol Sin* [Internet]. 2020 [acesso 18 dez 2022];41(1):138-44. doi:10.1038/s41401-019-0228-6.
- Liu Y, Yang X, Gan J, Chen S, Xiao ZX, Cao Y. CB-Dock2: improved protein-ligand blind docking by integrating cavity detection, docking and homologous template fitting. *Nucleic Acids Res* [Internet]. 2022 [acesso 18 dez 2022];50(W1):W159-64. doi:10.1093/nar/gkac394.
- Lu ST, Tang LL, Zhou LH, Lai YT, Liu LX, Duan Y. Study on the multitarget mechanism and active compounds of essential oil from *Artemisia argyi* treating pressure injuries based on network pharmacology. *Evid Based Complement Alternat Med* [Internet]. 2022 Jan 19 [acesso 09 jan 2023];2022:1019289. doi: 10.1155/2022/1019289.

- Lv H, Wang J, Zhu Y, Jiang T. Study on the mechanism of compound kidney-invigorating granule for osteoporosis based on network pharmacology and experimental verification. *Evid Based Complement Alternat Med* [Internet]. 2022 Jan 4 [acesso 19 dez 2022];2022:6453501. doi: 10.1155/2022/6453501.
- Mishra V, Heath RJ. Structural and biochemical features of human serum albumin essential for eukaryotic cell culture. *Int J Mol Sci* [Internet]. 2021 Aug 5 [acesso 15 jan 2023];22(16):8411. doi: 10.3390/ijms22168411.
- Moman RN, Gupta N, Varacallo M. Physiology, Albumin. 2022 Jan 4 [acesso 15 jan 2023]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing. Disponível em: <https://www.ncbi.nlm.nih.gov/books/NBK459198/>.
- Morris GM, Lim-Wilby M. Molecular docking. *Methods Mol Biol* [Internet]. 2008 [acesso 18 dez 2022];443:365-82. doi: 10.1007/978-1-59745-177-2_19.
- Meng XY, Zhang HX, Mezei M, Cui M. Molecular docking: a powerful approach for structure-based drug discovery. *Curr Comput Aided Drug Des* [Internet]. 2011 [acesso 18 dez 2022];7(2):146-157. doi:10.2174/157340911795677602.
- NLG. 2022 Aug. 21 [acesso 19 jul 2023]. Disponível em: <https://www.rcsb.org/docs/3d-viewers/nlg>.
- Okulik N, Jubert AH. Theoretical study on the structure and reactive sites of three non-steroidal anti-inflammatory drugs: ibuprofen, naproxen and tolmetin acids. *THEOCHEM* [Internet]. 2006 Sep 14 [acesso 13 jan 2023];769(1-3):135-41. doi: 10.1016/j.theochem.2005.10.061.
- Orlando BJ, Lucido MJ, Malkowski MG. The structure of ibuprofen bound to cyclooxygenase-2. *J Struct Biol* [Internet]. 2015 Jan [acesso 13 jan 2023];189(1):62-6. doi: 10.1016/j.jsb.2014.11.005.
- Quiroga R, Villarreal MA. Vinardo: A scoring function based on autodock vina improves scoring, docking, and virtual screening. *PLoS One* [Internet]. 2016 May 12 [acesso 25 jul 2023];11(5):e0155183. doi: 10.1371/journal.pone.0155183.

- Raoufinia R, Mota A, Keyhanvar N, Safari F, Shamekhi S, Abdolalizadeh J. Overview of albumin and its purification methods. *Adv Pharm Bull* [Internet]. 2016 Dec [acesso 15 jan 2023];6(4):495-507. doi: 10.15171/apb.2016.063.
- RCSB Protein Data Bank (RCSB PDB) [Internet]. 2023 [acesso 20 jan 2023]. Disponível em: <https://www.rcsb.org/>.
- Rose AS, Bradley AR, Valasatava Y, Duarte JM, Prlic A, Rose PW. NGL viewer: web-based molecular graphics for large complexes. *Bioinformatics* [Internet]. 2018 Nov 1 [acesso 15 jul 2023];34(21):3755-3758. doi: 10.1093/bioinformatics/bty419.
- Sekula B, Zielinski K, Bujacz A, Bujacz G. Crystal structure of equine serum albumin in complex with ibuprofen [Internet]. 2020 Feb 26 [acesso 14 jul 2023]. doi: 10.2210/pdb6oci/pdb.
- Sugio S, Mochizuki S, Noda M, Kashima A. Crystal structure of human serum albumin [Internet]. 2011 Jul 13 [acesso 20 jan 2023]. doi: 10.2210/pdb1ao6/pdb.
- Tao YG, Huang XF, Wang JY, Kang MR, Wang LJ, Xian SX. Exploring molecular mechanism of huangqi in treating heart failure using network pharmacology. *Evid Based Complement Alternat Med* [Internet]. 2020 Apr 23 [acesso 09 jan 2023];2020:6473745. doi: 10.1155/2020/6473745.
- Wang F, An SJ, Yin Y, Li JJ, Sun CH, Lan J, Zhao WJ, Li CQ. Triptolide is a promising therapeutic approach in treating thyroid cancer based on *in silico* and *in vitro* experiment. *Drug Des Devel Ther* [Internet]. 2021 [acesso 19 dez 2022];15:4275-87. doi: 10.2147/DDDT.S322502.
- Wang X, Liu Q, Wu S, Xu N, Li H, Feng A. Identifying the effect of celastrol against ovarian cancer with network pharmacology and *in vitro* experiments. *Front Pharmacol* [Internet]. 2022 Mar 18 [acesso 18 dez 2022];13:739478. doi: 10.3389/fphar.2022.739478.

- Xia LJ, Zhang LM, Yang K, Chen T, Ye XW, Yan ZJ. The pharmacological mechanism of xiyanping injection for the treatment of novel coronavirus pneumonia (Covid-19): based on network pharmacology strategy. *Evid Based Complement Alternat Med* [Internet]. 2022 Jul 8 [acesso 19 dez 2022];2022:9152201. doi: 10.1155/2022/9152201.
- Xue Q, Liu X, Russell P, Li J, Pan W, Fu J, Zhang A. Evaluation of the binding performance of flavonoids to estrogen receptor alpha by Autodock, Autodock Vina and Surflex-Dock. *Ecotoxicol Environ Saf* [Internet]. 2022 Mar 15 [acesso 25 jul 2023];233:113323. doi: 10.1016/j.ecoenv.2022.113323.
- Yang Cao Lab: Structural Bioinformatics Research Group [Internet]. 2021 [acesso 18 dez 2022]. Disponível em: <http://cao.labshare.cn/clab/index.html>.
- Yang X, Liu Y, Gan J, Xiao ZX, Cao Y. FitDock: protein-ligand docking by template fitting. *Brief Bioinform* [Internet]. 2022 May 13 [acesso 26 jul 2023];23(3):bbac087. doi: 10.1093/bib/bbac087.
- Ye XW, Deng YL, Zhang X, Liu MM, Liu Y, Xie YT, Wan Q, Huang M, Zhang T, Xi JH, Zhang JL. Study on the mechanism of treating COVID-19 with Shenqi Wan based on network pharmacology. *Drug Dev Ind Pharm* [Internet]. 2021 Aug [acesso 18 dez 2022];47(8):1279-89. doi: 10.1080/03639045.2021.1989453.
- Zhang M, Hu G, Shao N, Qin Y, Chen Q, Wang Y, Zhou P, Cai B. Thioredoxin-interacting protein (TXNIP) as a target for Alzheimer's disease: flavonoids and phenols. *Inflammopharmacology* [Internet]. 2021 Oct [acesso 18 dez 2022];29(5):1317-29. doi: 10.1007/s10787-021-00861-4.



ANEXOS

Anexo 1 – Abreviaturas de uma e três letras dos aminoácidos

O Quadro 1 apresenta as abreviaturas de uma e três letras dos aminoácidos.

Abreviatura com 1 letra	Abreviatura com 3 letras	Nome
A	Ala	Alanina
C	Cys, Cis	Cisteína
D	Asp	Ácido aspártico
E	Glu	Ácido glutâmico
F	Phe ou Fen	Fenilalanina
G	Gly, Gli	Glicina
H	His	Histidina
I	Ile	Isoleucina
K	Lys, Lis	Lisina
L	Leu	Leucina
M	Met	Metionina
N	Asn	Asparagina
P	Pro	Prolina
Q	Gln	Glutamina
R	Arg	Arginina
S	Ser	Serina
T	Thr, The	Treonina
V	Val	Valina
W	Trp, Tri	Triptofano
Y	Tyr, Tir	Tirosina

Fonte: Modificado de: Hassunuma, Souza (2016).

Anexo 2 – Representação das interações químicas no visualizador NGL®

O Quadro 2 apresenta as formas de exibição das interações químicas no visualizador NGL® (NGL-viewer®).




Tipo	Cor	Estilo	Descrição
Ligação de hidrogênio	Azul	---	Ligação de hidrogênio entre átomos doadores e receptores fortes.
Ligação de hidrogênio fraca	Azul claro	---	Ligação de hidrogênio entre um átomo doador de carbono e um aceptor, ou um grupo fosfato inorgânico (Pi) e um átomo doador.
Interação hidrofóbica	Cinza	---	Interações entre grupos alquil, ou um grupo alquil e um grupo Pi.
ligação de halogênio	Ciano	---	Interações com átomos de flúor, cloro, bromo ou iodo.
Interação Iônica	Amarelo	---	Interações entre pares de grupos com cargas opostas.
Interação cátion-Pi	Laranja	---	Interações entre um átomo carregado positivamente e os elétrons de um sistema Pi deslocalizado.
Empilhamento Pi-Pi	Verde	---	Interações entre sistemas Pi deslocalizados.

Fonte: Modificado de: Rose, Bradley, Valasatava, Duarte, Prlic, Rose, 2018.

Anexo 3 – Esquemas de cores do visualizador NGL®

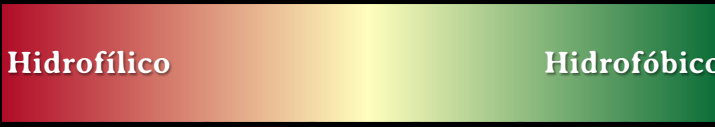
O Quadro 3 apresenta os seguintes esquemas de cores:

- **Arco-íris (*Rainbow*):** o terminal amina da proteína é colorida em azul e o terminal carboxila em vermelho, sendo que o restante da molécula é colorida em degradê;
- **Cadeia (*Chain*):** a primeira cadeia da proteína é colorida em vermelho e a última em azul, e o restante em degradê (Liu, Yang, Gan, Chen, Xiao, Cao, 2022);
- **Fator B (*B-factor*):** em cristalografias, existe um grau de incerteza em relação à posição correta dos átomos. Este grau pode ser medido pelo fator B ou fator de temperatura. Este fator corresponde à dispersão de raios X para cada átomo ou para grupos de átomos que descreve o grau em que a densidade eletrônica é espalhada (de Sá Ribeiro, Lima, 2023). No CB-Dock2®, os átomos com fator-B alto estão representados na cor vermelha, indicando que suas posições são mais incertas. Os átomos com fator-B mais baixo estão representados em bronze, indicando que suas posições são mais certas. Valores intermediários são representados em degradê (Liu, Yang, Gan, Chen, Xiao, Cao, 2022);

Esquema de cores	Cor
Arco-íris (<i>Rainbow</i>)	Terminal amina  Terminal carboxila
Cadeia (<i>Chain</i>)	Primeira cadeia  Última cadeia
Fator-B (<i>B-factor</i>)	Alto  Baixo

Fonte: Modificado de: NLG, 2022.

- **Hidrofobicidade (*Hydrophobicity*):** os resíduos mais hidrofílicos são coloridos em vermelho e os mais hidrofóbicos em verde. Valores intermediários são representados em degradê;
- **Elemento (CPK):** os átomos são representados em cores diferentes, sendo indicados no Quadro 4, os átomos mais frequentes em proteínas e ácidos nucleicos: hidrogênio (H), carbono (C), nitrogênio (N), oxigênio (O), flúor (F), fósforo (P) e enxofre (S);
- **Resíduo de aminoácido (*Residue – Amino acids*):** os átomos são representados em cores diferentes. Em caso de dúvidas na abreviatura dos aminoácidos, consulte o Anexo 1;

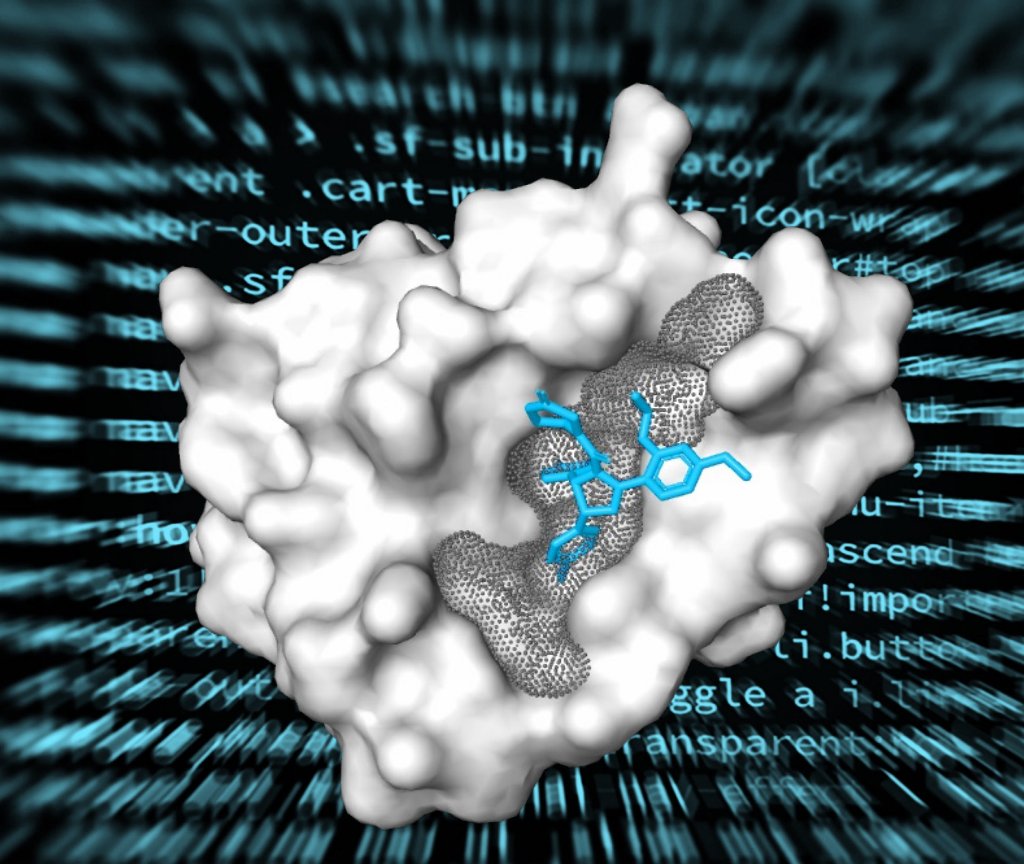
Esquema de cores	Cor						
Hidrofobicidade	Hidrofílico  Hidrofóbico						
Elemento (CPK)	H	C	N	O	F	P	S
Resíduo de aminoácido (<i>Residue - Amino acids</i>)	Ala	Arg	Asn	Asp	Cys		
	Gln	Glu	Gly	His	Ile		
	Leu	Lys	Met	Phe	Pro		
	Ser	Thr	Trp	Tyr	Val		

Fonte: Modificado de: NLG, 2022.

- **Nucleotídeos (*Nucleotides*):** os nucleotídeos são representados em cores diferentes. No Quadro 5 estão indicadas as bases nitrogenadas dos nucleotídeos mais frequentes: (desoxi)adenosina (A), (desoxi)guanosina (G), (desoxi)inosina, (desoxi)xantosina, (desoxi)citidina, (desoxi)timidina e (desoxi)uridina;
- **Estrutura secundária e DNA, RNA e carboidrato (*Secondary structure plus DNA, RNA and carbohydrate*):** representa as regiões da proteína ou ácido nucleico de acordo com sua estrutura secundária ou composição (NLG, 2022).

Esquema de cores	Cor						
Nucleotídeo (<i>Nucleotides</i>)	A	G	I	X	C	T	U
Estrutura secundária e DNA, RNA e carboidrato (<i>Secondary structure plus DNA, RNA and carbohydrate</i>)	Alfa-hélice (<i>alpha-helix</i>)		Hélice 3/10 (<i>3/10 helix</i>)		Hélice Pi (<i>Pi helix</i>)		
	Fita beta (<i>Beta strand</i>)		Volta beta (<i>Beta turn</i>)		Espiral beta (<i>coil</i>)		
	DNA		RNA		Carboidrato		

Fonte: Modificado de: NLG, 2022.



A docagem cega corresponde a um termo utilizado em Bioinformática para a localização de possíveis sítios de ligação em cavidades não conhecidas.

Os métodos tradicionais de docagem cega geralmente utilizam programas computacionais direcionados para pesquisadores, com uma certa complexidade para usuários iniciantes.

Assim, esta obra apresenta o **servidor web CB-Dock2[®]**, uma ferramenta gratuita e simples para docagem cega guiada por detecção de cavidade. O livro apresenta uma sequência de redocagem cega entre o anti-inflamatório Ibuprofeno e a proteína plasmática Albumina.

Desta forma, você, leitor, professores e alunos do Ensino Médio ou Superior podem utilizar este livro como um guia rápido para aprender a realizar a docagem cega em sua residência ou em qualquer laboratório de informática, tornando o recurso democrático e acessível a toda comunidade acadêmica e para a população em geral.

Boa leitura!