

Treinamento em Biologia Computacional - Módulo 2

Equipe de Dados Biológicos (EDB)

Grupo de Design de Proteínas (GDP)

Laboratório de Biologia Computacional (LBC)

Dia 2 (07/11/2025)



Módulo 2

Detecção e caracterização de sítios de ligação em biomoléculas

Equipe de Dados Biológicos (EDB)

João V. S. Guerra

Grupo de Design de Proteínas (GDP)

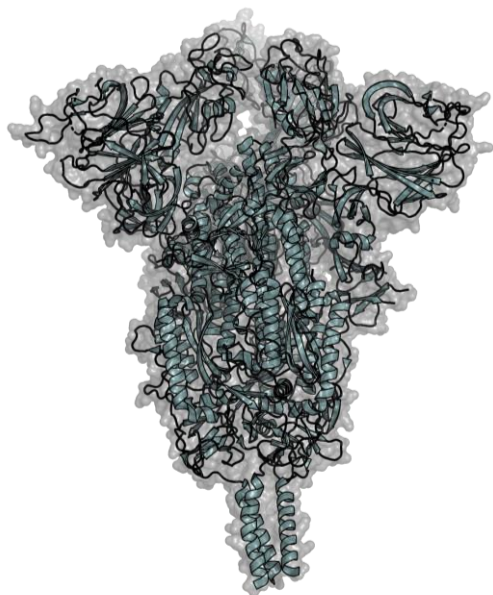
Helder V. Ribeiro-Filho

Dia 2 (07/11/2025)



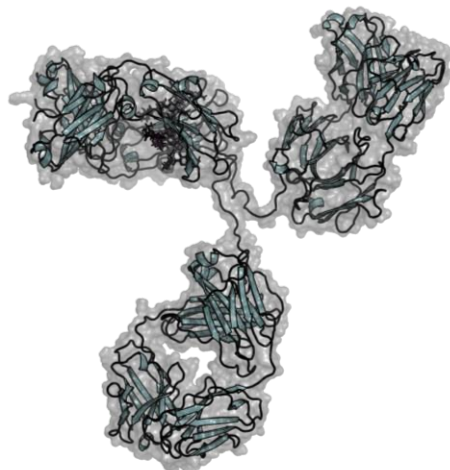
- » Usualmente, os processos biológicos são modulados pela interação entre biomoléculas, que variam de íons, como ferro e fosfato, a macromoléculas, como proteínas e ácidos nucleicos.

Invasão viral



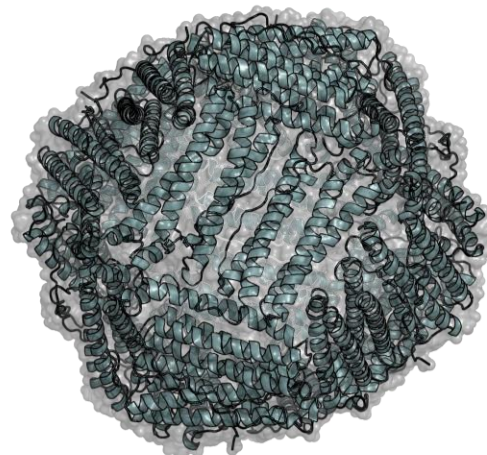
Espícula do SARS-CoV-2
(PDB ID: 7VX1)

Defesa



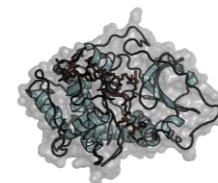
Imunoglobulina
(PDB ID: 1IGT)

Armazenamento



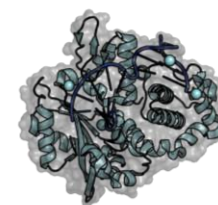
Ferritina
(PDB ID: 1FHA)

Enzima



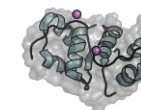
Quinase A
(PDB ID: 1FMO)

Replicação DNA



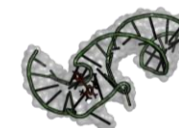
DNA Polimerase
(PDB ID: 5TYX)

Comunicação

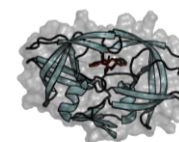


Insulina
(PDB ID: 4INS)

Replicação viral

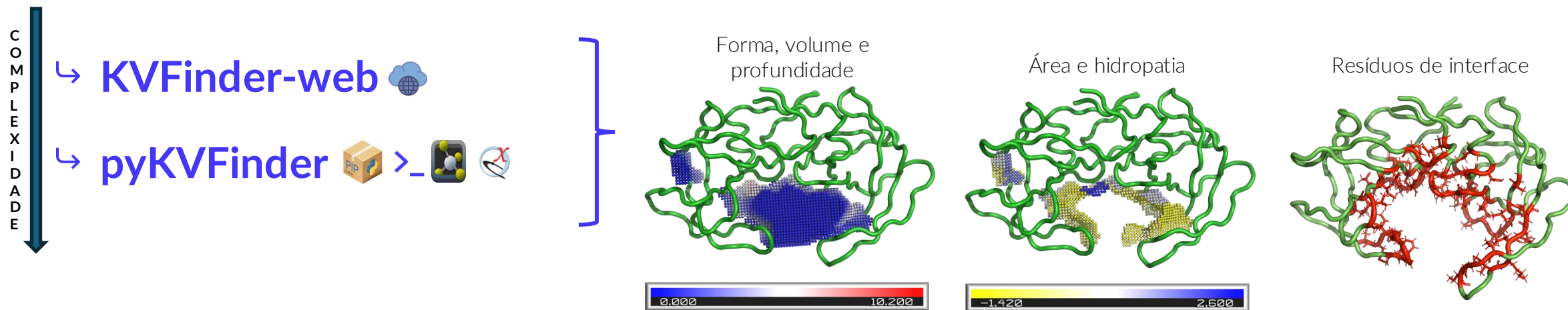


Enterovírus A71 RNA
(PDB ID: 6XB7)



HIV-1 Protease
(PDB ID: 1HVR)

- » A plataforma **KVFinder suite** conta 2 ferramentas computacionais para detecção e caracterização de sítios de ligação em biomoléculas:



» O **KVFinder-web** (<https://kvfinder-web.cnpem.br>) é uma aplicação web para detecção e caracterização de cavidades em qualquer tipo de estrutura biomolecular, que consiste em dois componentes independentes:

JOURNAL ARTICLE

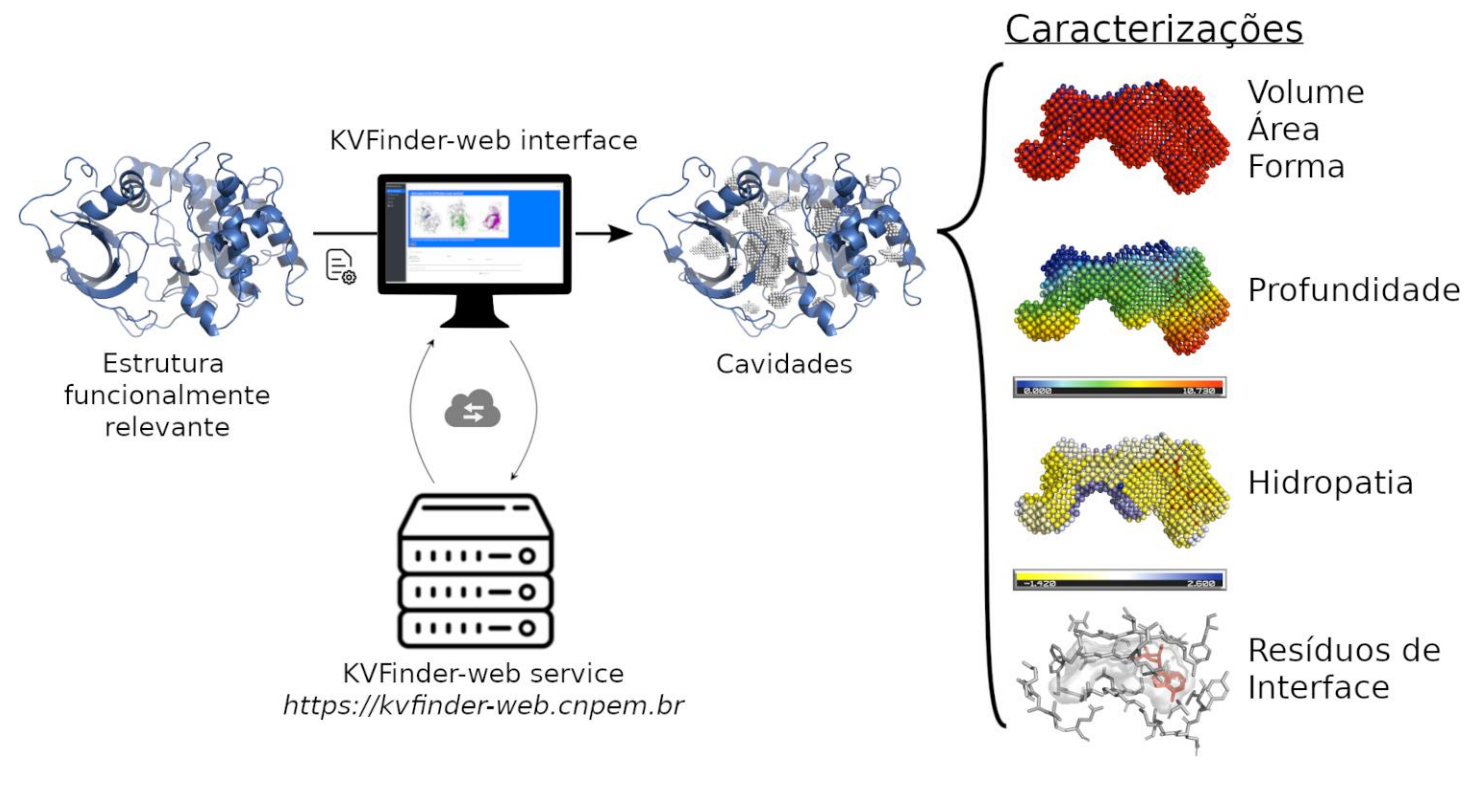
KVFinder-web: a web-based application for detecting and characterizing biomolecular cavities

João V S Guerra, Helder V Ribeiro-Filho, José G C Pereira ✉, Paulo S Lopes-de-Oliveira ✉ [Author Notes](#)

Nucleic Acids Research, Volume 51, Issue W1, 5 July 2023, Pages W289–W297,

<https://doi.org/10.1093/nar/gkad324>

Published: 04 May 2023 [Article history](#)



» O Python-C parallel KVFinder (pyKVFinder) é um pacote Python para detectar e caracterizar cavidades biomoleculares em protocolos automatizados e aplicações de ciências de dados.

Software | [Open Access](#) | [Published: 20 December 2021](#)

pyKVFinder: an efficient and integrable Python package for biomolecular cavity detection and characterization in data science

[João Victor da Silva Guerra](#) ✉, [Helder Veras Ribeiro-Filho](#), [Gabriel Ernesto Jara](#),
[Leandro Oliveira Bortot](#), [José Geraldo de Carvalho Pereira](#) & [Paulo Sérgio Lopes-de-Oliveira](#) ✉

BMC Bioinformatics **22**, Article number: 607 (2021) | [Cite this article](#)

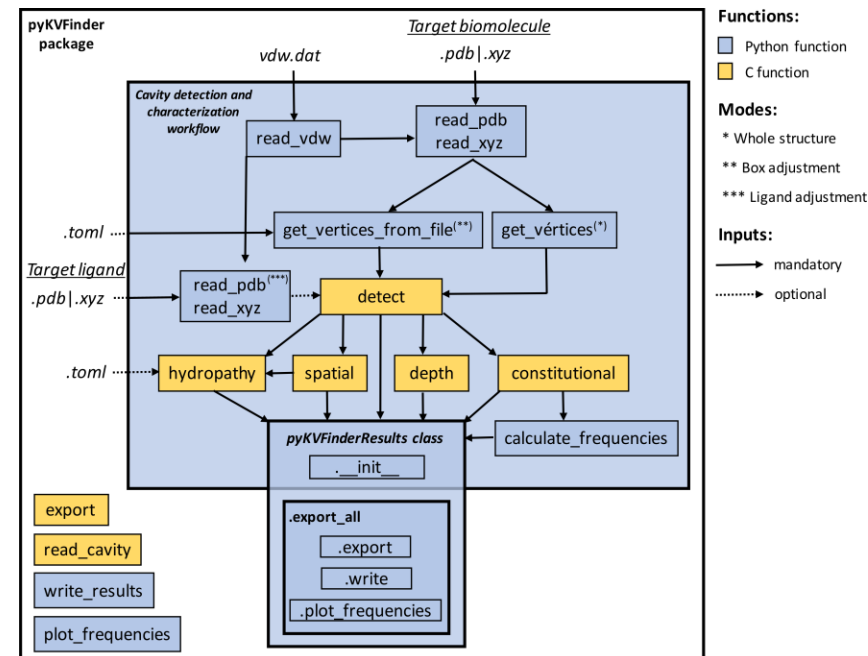
Código-fonte: <https://github.com/LBC-LNBio/pyKVFinder>

Documentação: <https://lbc-lnbio.github.io/pyKVFinder/>

Tutorial: <https://lbc-lnbio.github.io/pyKVFinder/package/tutorial/index.html>

Plugins gráficos:

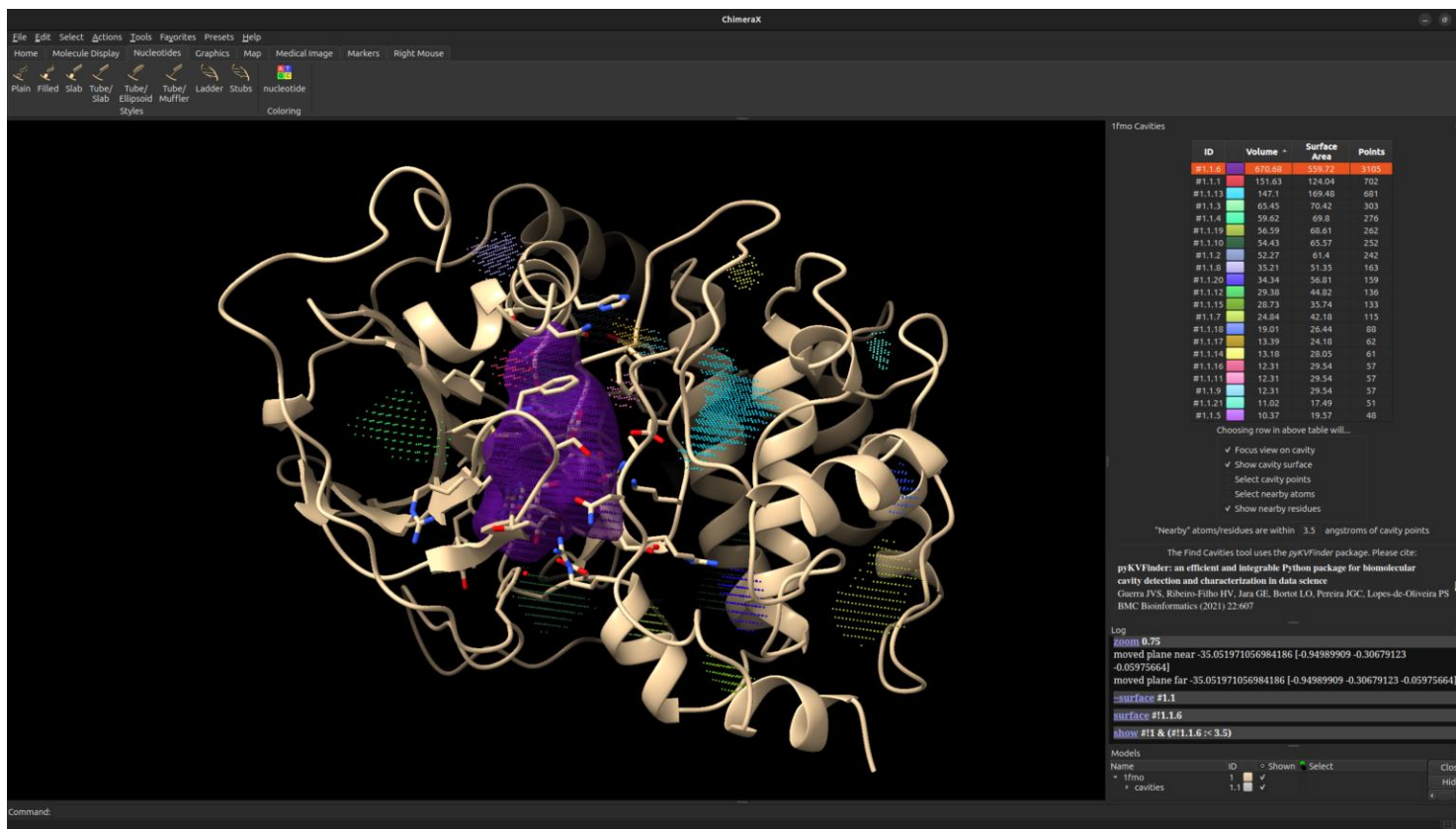
- PyMOL pyKVFinder Tools: [link](#)
- ChimeraX: [link](#)



Atualizações e melhorias:

- v0.4.2: Aberturas
- v0.5.0: Modelagem de moléculas
- v0.7.0: plugin gráfico para PyMOL e ChimeraX

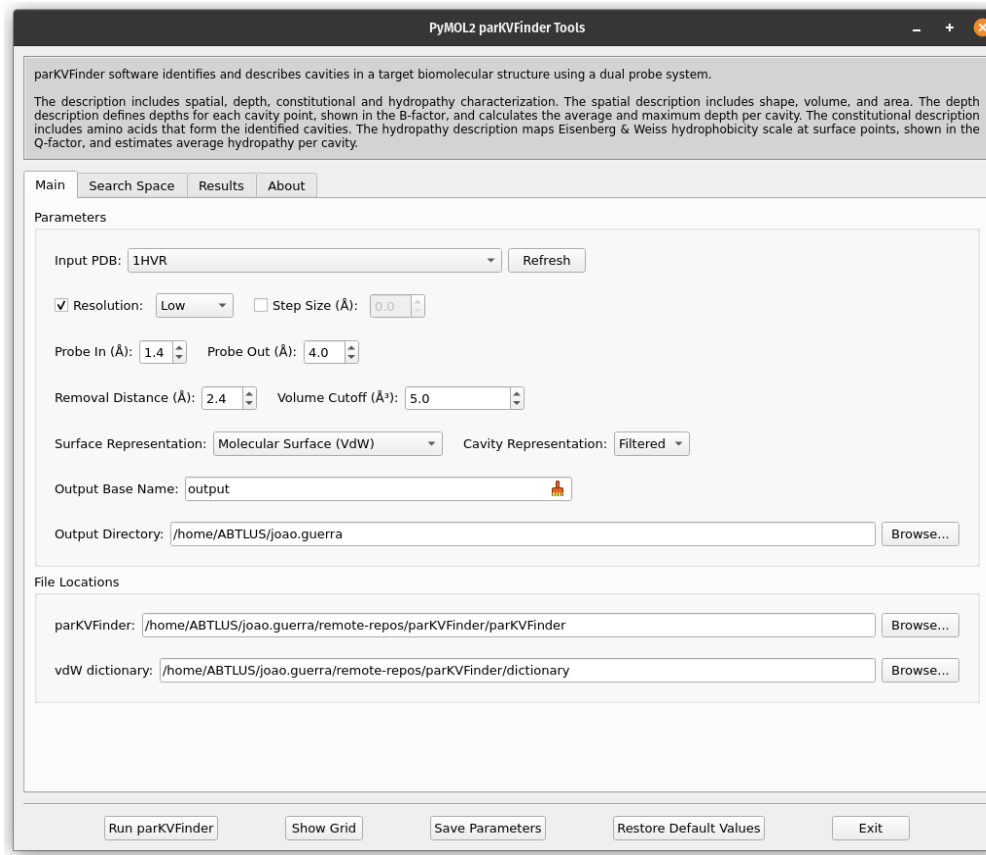
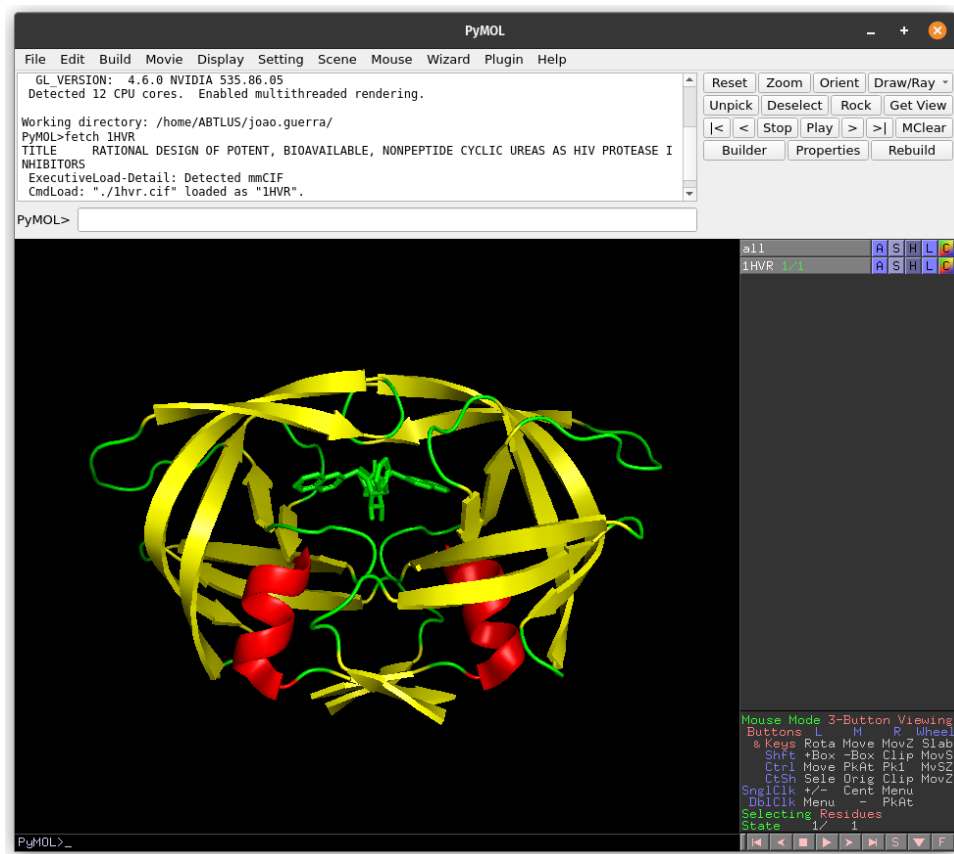
Plugin gráfico - ChimeraX



Find Cavities tool: <https://www.cgl.ucsf.edu/chimerax/docs/user/tools/findcavities.html>

KVFinder command: <https://www.cgl.ucsf.edu/chimerax/docs/user/commands/kvfinder.html>

Plugin gráfico - PyMOL



PyMOL2 parkFinder Tools interface showing parameters for cavity detection.

parKFinder software identifies and describes cavities in a target biomolecular structure using a dual probe system.

The description includes spatial, depth, constitutional and hydrophobicity characterization. The spatial description includes shape, volume, and area. The depth description defines depths for each cavity point, shown in the B-factor, and calculates the average and maximum depth per cavity. The constitutional description includes amino acids that form the identified cavities. The hydrophobicity description maps Eisenberg & Weiss hydrophobicity scale at surface points, shown in the Q-factor, and estimates average hydrophobicity per cavity.

Main Search Space Results About

Parameters

Input PDB: 1HVR Refresh

☒ Resolution: Low ☐ Step Size (Å): 0.0

Probe In (Å): 1.4 Probe Out (Å): 4.0

Removal Distance (Å): 2.4 Volume Cutoff (Å³): 5.0

Surface Representation: Molecular Surface (VdW) Cavity Representation: Filtered

Output Base Name: output

Output Directory: /home/ABTLUS/joao.guerra Browse...

File Locations

parKFinder: /home/ABTLUS/joao.guerra/remote-repos/parKFinder/parKFinder Browse...

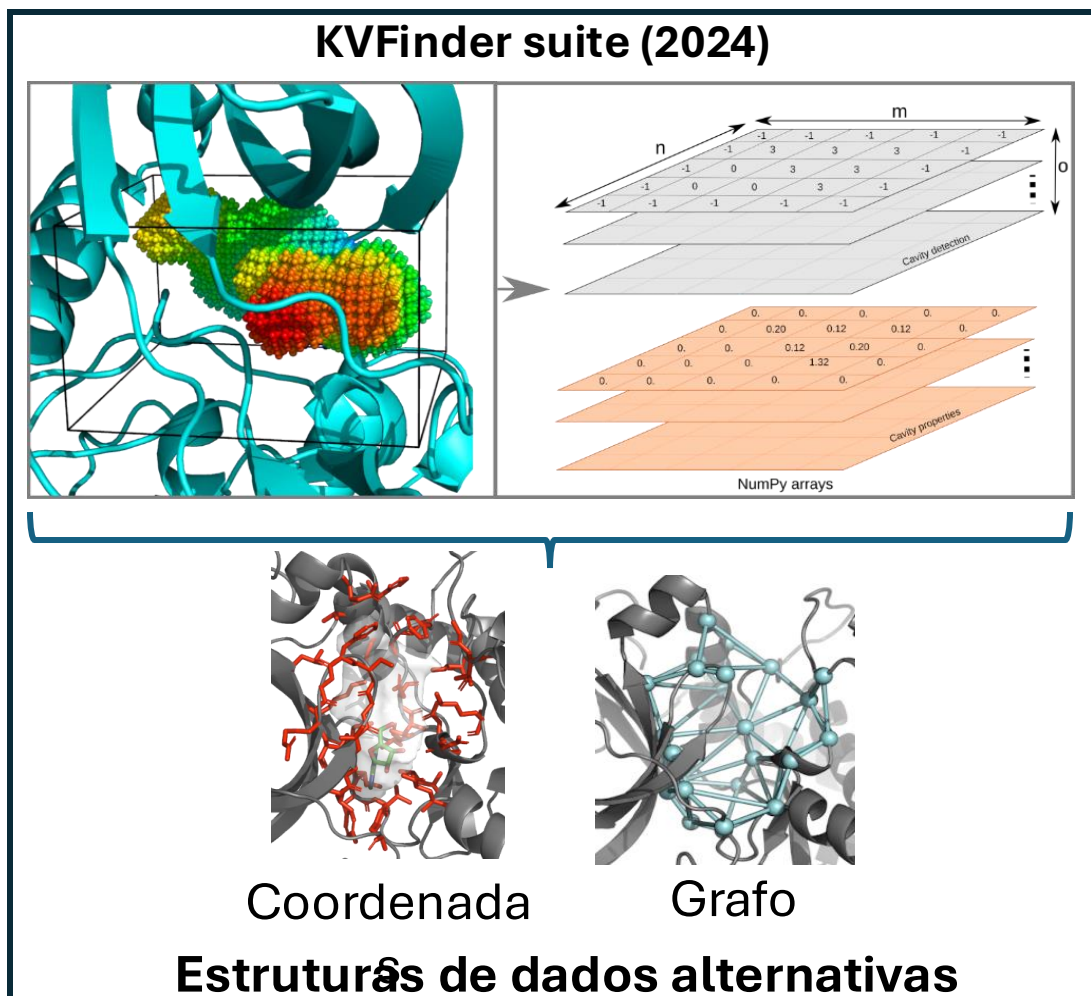
vdW dictionary: /home/ABTLUS/joao.guerra/remote-repos/parKFinder/dictionary Browse...

Run parKFinder Show Grid Save Parameters Restore Default Values Exit

Instalação: <https://lbc-lnbio.github.io/pyKVFinder/plugins/pymol/index.html#installation>

Tutorial: <https://lbc-lnbio.github.io/pyKVFinder/plugins/pymol/tutorial/biomolecular-cavity-detection.html>

- » Uma plataforma computacional robusta para biologia estrutural requer um conjunto de ferramentas integradas com estruturas de dados de fácil manipulação para aplicações em ciência de dados, criando um ambiente propício para a adoção de IA.



Exemplos de aplicações avançadas:

- Dinâmica molecular: [link](#)
- Conservação estrutural: [link](#)

Citações de aplicações avançadas:

- ActSeek (busca por sítios ativos): [paper](#)
- AR-V7 (busca por sítios de ligação em proteínas parcialmente desordenadas): [paper](#)
- Desenvolvimento de caracterização de cavidades (área da seção transversal da cavidade): [paper](#)

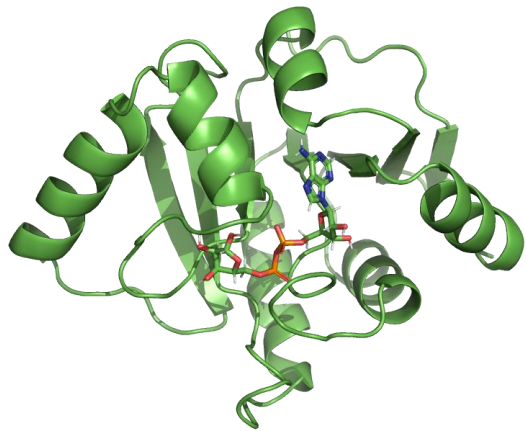
Aplicações avançadas do pyKVFinder



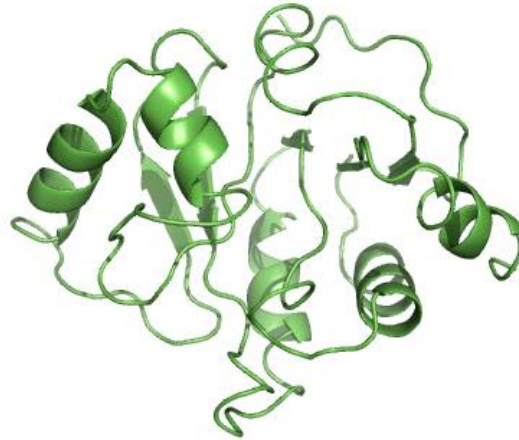
CNPq

- » Dinâmica molecular do domínio ADRP do SARS-CoV-2, importante na virulência e na regulação da imunidade inata à infecção.

↳ <https://github.com/LBC-LNBio/pyKVFinder/blob/master/examples/md-analysis/md-analysis.ipynb>



ADRP SARS-CoV-2
(PDB ID: 6W02)



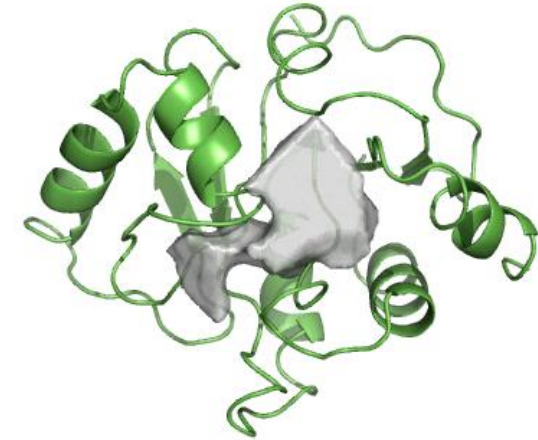
AMBER 18



600 ns



600 quadros



pyKVFinder v0.8.2

Aplicações avançadas do pyKVFinder

» ADRP do SARS-CoV-2 e proteínas homólogas, importante na virulência e na regulação da imunidade inata à infecção.

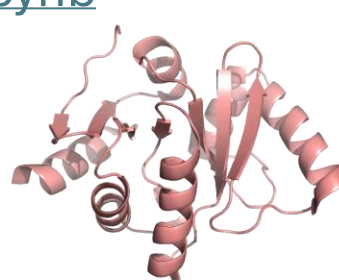
↳ <https://github.com/LBC-LNBio/pyKVFinder/blob/master/examples/conservation-analysis/conservation-analysis.ipynb>



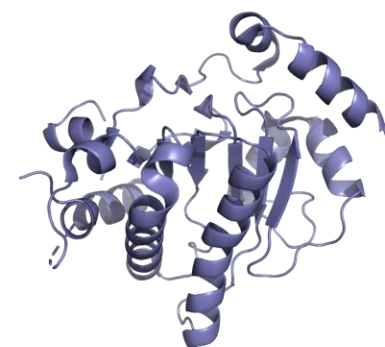
SARS-CoV-2
(PDB ID: 6WEN, cadeia A)



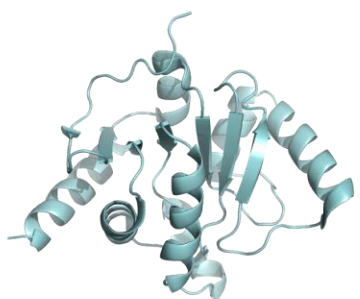
MERS-CoV
(PDB ID: 5HIH, cadeia A)



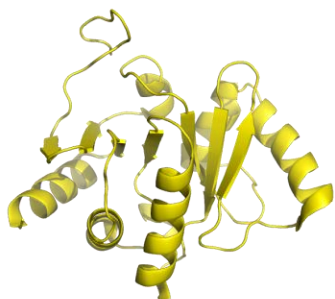
HCoV-229E
(PDB ID: 3EJG, cadeia A)



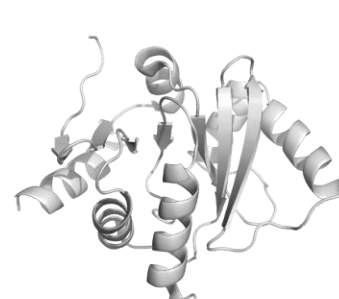
macroD1
(PDB ID: 6WEN, cadeia A)



SARS-CoV-1
(PDB ID: 2ACF, cadeia A)



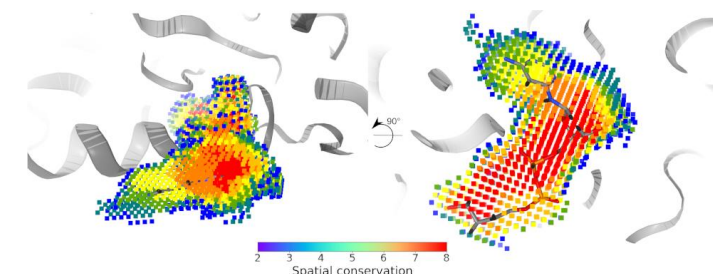
HCoV-NL63
(PDB ID: 2VRI, cadeia A)



FCoV
(PDB ID: 3ETI, cadeia B)



macroD2
(PDB ID: 6Y73, cadeia D)

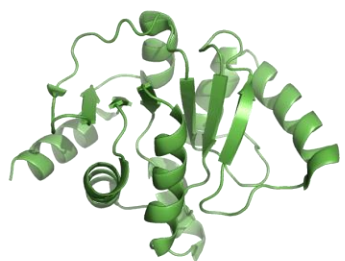


Cavidade do domínio ADRP contém uma parte central altamente conservada, ocupada pelo difosfato e ribose do ADP, e outra região menos conservada que pode mudar para acomodar o substrato ADP-ribose em diferentes espécies.

Aplicações avançadas do pyKVFinder

» ADRP do SARS-CoV-2 e proteínas homólogas, importante na virulência e na regulação da imunidade inata à infecção.

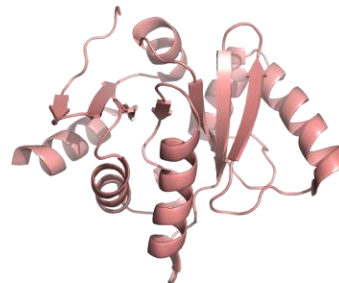
↳ <https://github.com/LBC-LNBio/pyKVFinder/blob/master/examples/conservation-analysis/conservation-analysis.ipynb>



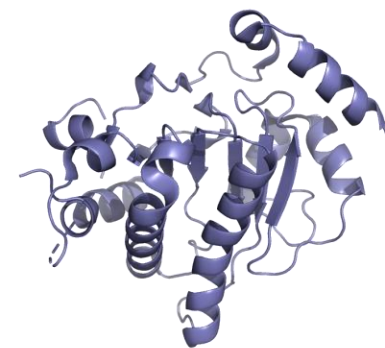
SARS-CoV-2
(PDB ID: 6WEN, cadeia A)



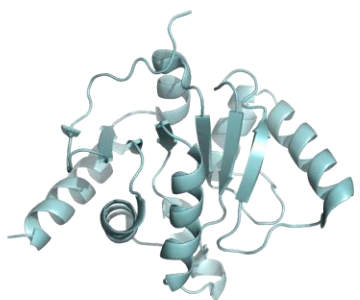
MERS-CoV
(PDB ID: 5HIH, cadeia A)



HCoV-229E
(PDB ID: 3EJG, cadeia A)



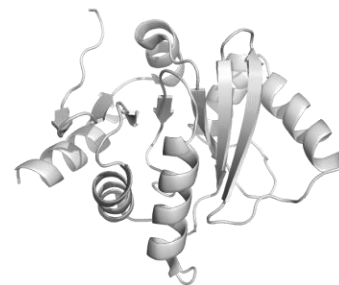
macroD1
(PDB ID: 6WEN, cadeia A)



SARS-CoV-1
(PDB ID: 2ACF, cadeia A)



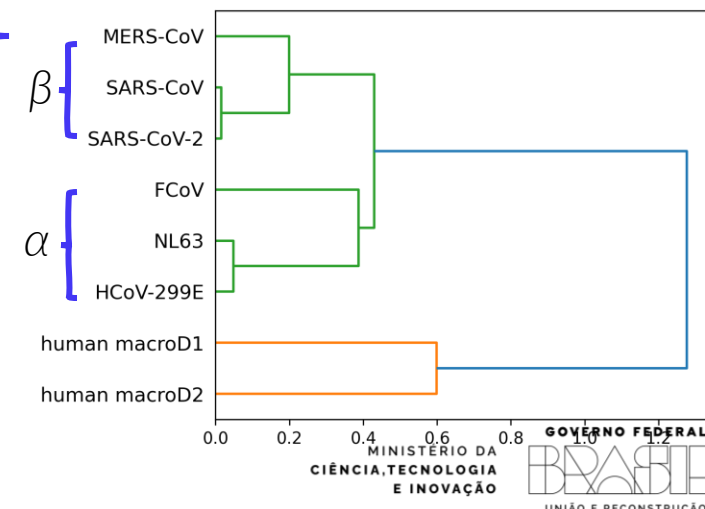
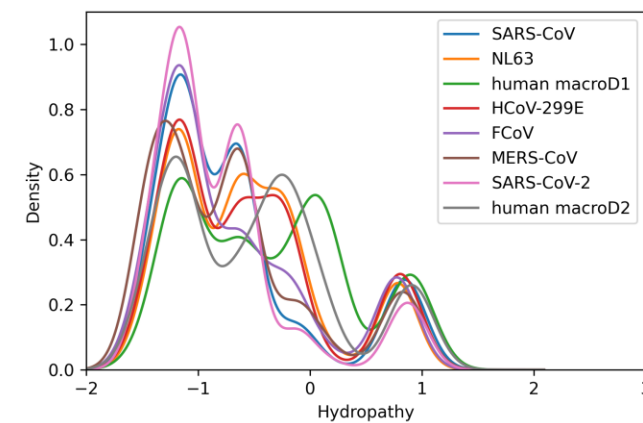
HCoV-NL63
(PDB ID: 2VRI, cadeia A)



FCoV
(PDB ID: 3ETI, cadeia B)



macroD2
(PDB ID: 6Y73, cadeia D)



Obrigado!

Equipe de Dados Biológicos (EDB)

edb@lnbio.cnpem.br

+55 (19) 3512-1113

Grupo de Design de Proteínas (GDP)

+55 (19) 3512-2389

Laboratório de Biologia Computacional (LBC)

+55 (19) 3512-1255



Sign up to receive
newsletters
about **CNPEN**
and its units

cnpem.br



CNPEN
Brazilian Center for Research
in Energy and Materials

MINISTRY OF
SCIENCE TECHNOLOGY
AND INNOVATION



Material suplementar: Detecção e caracterização de sítios de ligação em biomoléculas

Equipe de Dados Biológicos (EDB)

João V. S. Guerra

Grupo de Design de Proteínas (GDP)

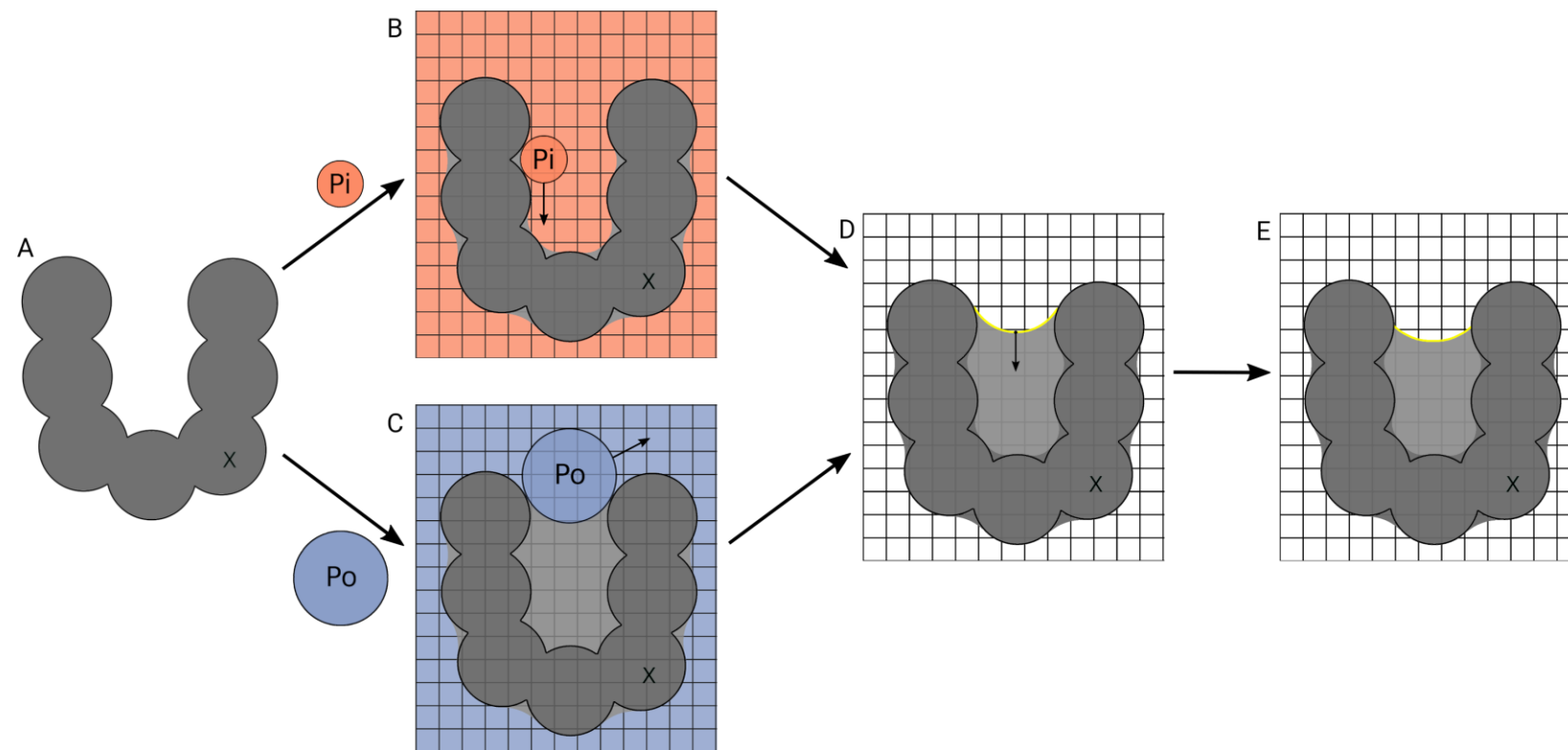
Helder V. Ribeiro-Filho

Dia 2 (03/11/2025)



KVFinder suite: Algoritmo de detecção

- » KVFinder suite aplica um algoritmo geométrico baseado em grade 3D e esferas baseado em um sistema de dupla sonda para detectar cavities



Representação esquemática do algoritmo de detecção de cavidades no parkKVFinder.

(A) Uma estrutura biomolecular X, composta por átomos modelados como esferas rígidas com raios de van der Waals, é inserida em uma grade 3D.

(B) A sonda Probe In (Pi) percorre a superfície da estrutura, movendo-se pelos pontos da grade (laranja).

(C) Em seguida, a sonda Probe Out (Po) percorre os pontos acessíveis em azul.

(D) Os pontos de cavidade (cinza claro) são definidos como a diferença entre os pontos acessíveis das sondas. Os pontos não alcançados por Pi (cinza escuro) denotam a SES (padrão) ou a SAS, dependendo da representação de superfície escolhida pelo usuário.

(E) Por fim, é aplicado um procedimento de remoção por distância para eliminar os pontos de cavidade que estão próximos à fronteira da cavidade-solvente (linha amarela).

KVFinder suite: pKA

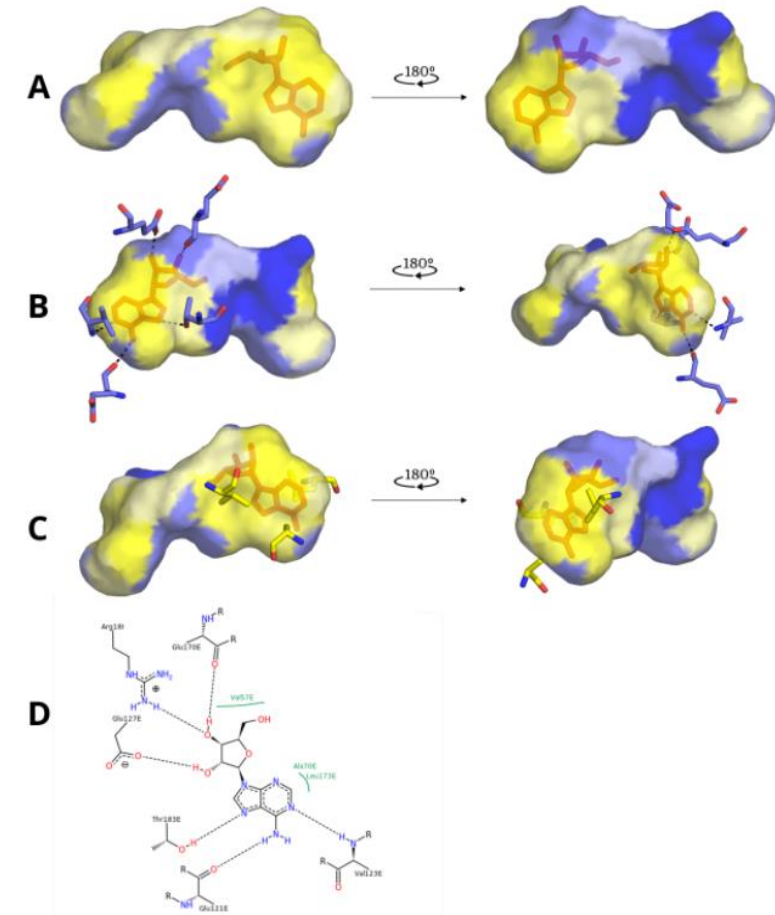
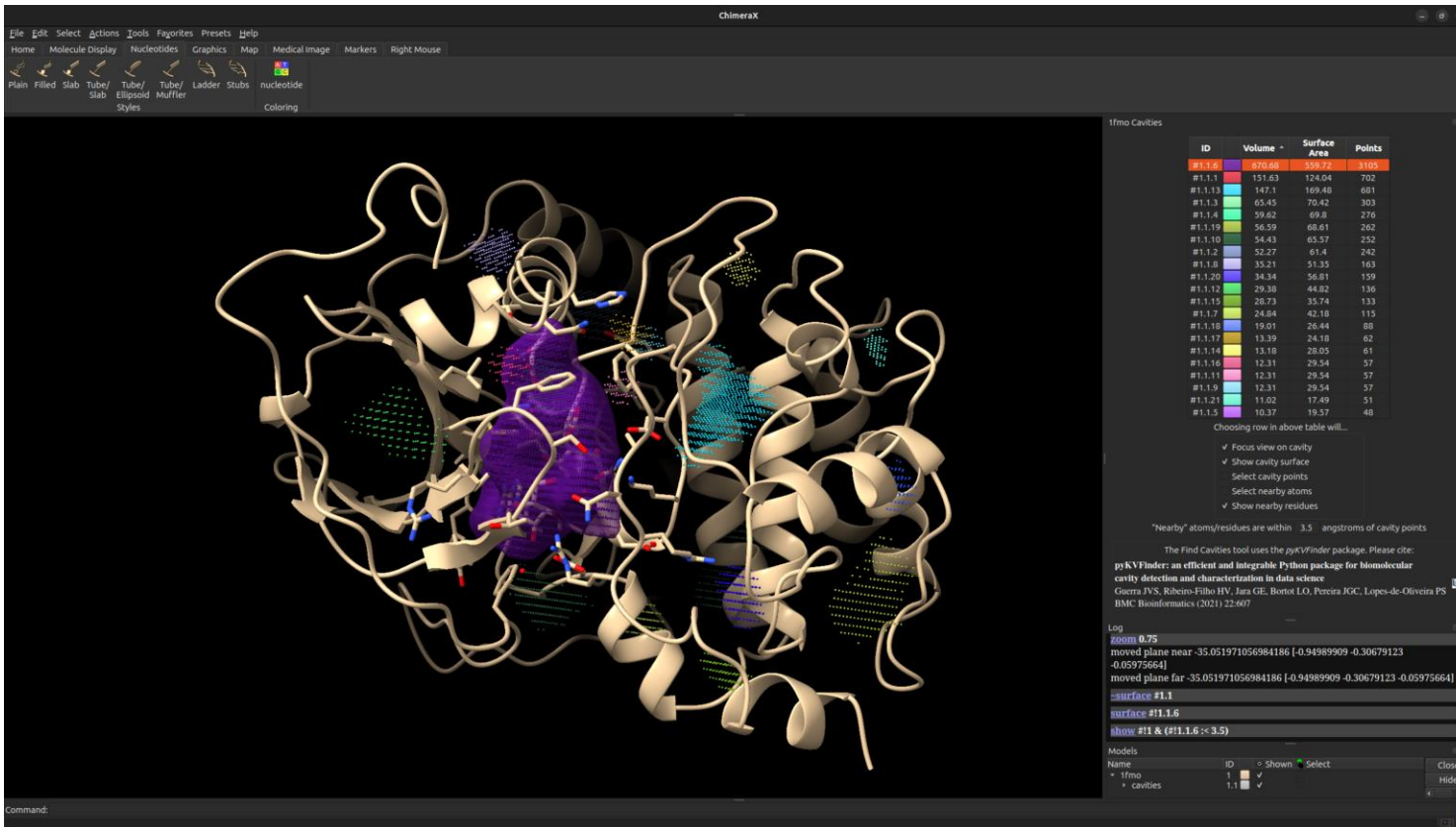


Figura 38: Determinação da hidropatia do sítio de ligação da adenosina da quinase A. Resultados da prospecção do sítio da adenosina na estrutura da quinase A (PDB 1FMO) com os pontos de superfície coloridos para a escala de hidrofobicidade de Hessa & Heijne na paleta amarelo-branco-azul. **(A)** Pontos de superfície da cavidade do sítio de ligação da adenosina, na forma de *surface* com a adenosina na forma de *sticks* (vermelho). **(B)** Interações hidrofílicas entre os resíduos (roxo) e a adenosina (vermelho). **(C)** Interações hidrofóbicas entre os resíduos (amarelo) e a adenosina (vermelho). **(D)** Interações entre os resíduos do sítio de ligação e a adenosina retirado do RCSB PDB. Linhas tracejadas pretas indicam ligações de hidrogênio, pontes salinas e interações metálicas. Linhas verdes mostram interações hidrofóbicas.

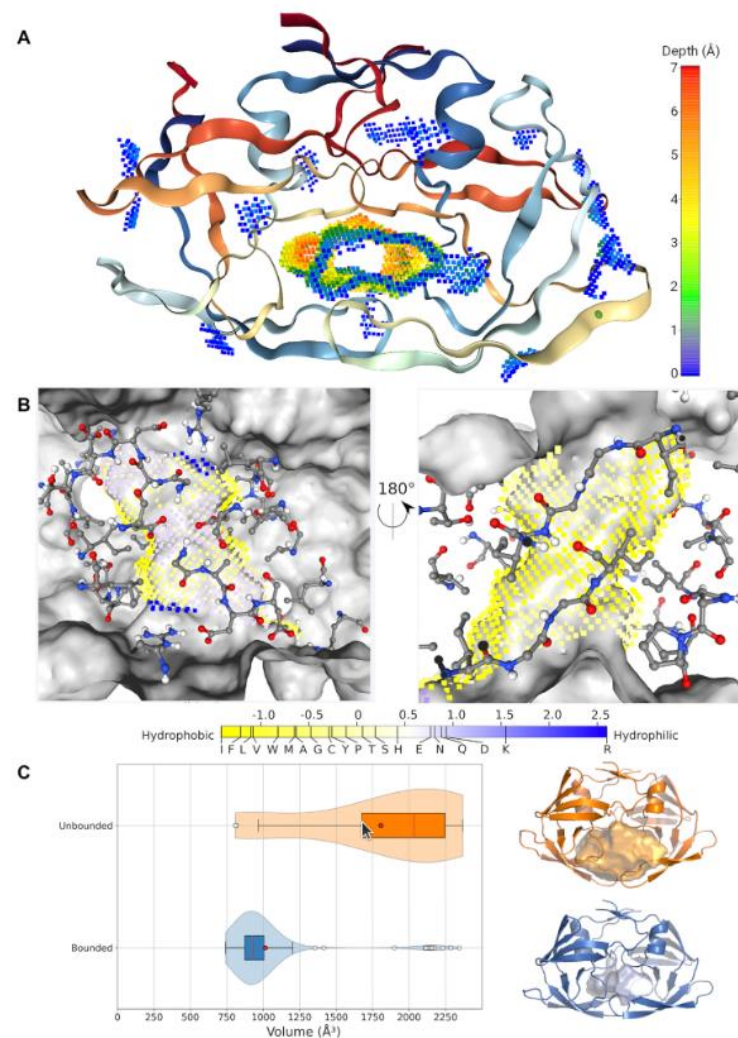


Figure 3. Illustrative example of HIV-1 protease cavity detection. Molecular structure of HIV-1 protease (PDB ID: 1HVR) with detected cavities. (A) Cavity detection throughout the protein structure (cartoon). Cavity points are colored by depth (points colored in rainbow scale). (B) Hydropathy mapped to surface cavity points in the regions around the catalytic aspartic acids (left panel) and around the β -hairpins (right panel). Eisenberg & Weiss hydrophobicity scale ranges from -1.42 (highly hydrophobic) to 2.6 (highly hydrophilic). Protein is shown as a gray surface and interface residues are shown as atom-colored sticks. (C) Violin plot of active site volume of HIV-1 protease structures from the RCSB PDB for the structures with ligands bound in the active site (blue) and the structures without ligands (orange). The structures with a median volume and the corresponding cavity are shown as cartoon and surface, respectively.