

## Structural bioinformatics

# pyCapsid: identifying dominant dynamics and quasi-rigid mechanical units in protein shells

Colin Brown<sup>1,2</sup>, Anuradha Agarwal<sup>1,3</sup>, Antoni Luque <sup>1,3,4,5,\*</sup>

<sup>1</sup>Viral Information Institute, San Diego State University, San Diego, CA 92116, United States

<sup>2</sup>Department of Physics, San Diego State University, San Diego, CA 92116, United States

<sup>3</sup>Computational Science Research Center, San Diego State University, San Diego, CA 92116, United States

<sup>4</sup>Department of Mathematics and Statistics, San Diego State University, San Diego, CA 92116, United States

<sup>5</sup>Department of Biology, University of Miami, Coral Gables, FL 33146, United States

\*Corresponding author. Department of Biology, University of Miami, 1301 Memorial Drive, Coral Gables, FL 33146, United States.

E-mail: antoni.luque@miami.edu (A.L.)

Associate Editor: Arne Elofsson

## Abstract

**Summary:** pyCapsid is a Python package developed to facilitate the characterization of the dynamics and quasi-rigid mechanical units of protein shells and other protein complexes. The package was developed in response to the rapid increase of high-resolution structures, particularly capsids of viruses, requiring multiscale biophysical analyses. Given a protein shell, pyCapsid generates the collective vibrations of its amino-acid residues, identifies quasi-rigid mechanical regions associated with the disassembly of the structure, and maps the results back to the input proteins for interpretation. pyCapsid summarizes the main results in a report that includes publication-quality figures.

**Availability and implementation:** pyCapsid's source code is available under MIT License on [GitHub](#). It is compatible with Python 3.8–3.10 and has been deployed in two leading Python package-management systems, [PIP](#) and [Conda](#). Installation instructions and tutorials are available in the [online documentation](#) and in the [pyCapsid's YouTube playlist](#). In addition, a cloud-based implementation of pyCapsid is available as a [Google Colab notebook](#). pyCapsid Colab does not require installation and generates the same report and outputs as the installable version. Users can post issues regarding pyCapsid in the repository's [issues section](#).

## 1 Introduction

Viruses protect their infective genomes in protein shells called capsids (Twarock and Luque 2019). The number of capsid structures solved at high resolution has increased exponentially in the last two decades, partly thanks to cryo-electron microscopy advances (Callaway 2020, Johnson and Olson 2021, Montiel-Garcia *et al.* 2021). These 3D reconstructions combined with computational algorithms and complementary experimental techniques are leading to a mechanistic characterization of the assembly, dynamics, and stability of viral capsids, opening the doors to new antiviral strategies (Yeager *et al.* 1990, Li *et al.* 2008, Luque and Reguera 2013, Zhang *et al.* 2013, Grime *et al.* 2016, Hua *et al.* 2017, Organtini *et al.* 2017, Qazi *et al.* 2018, Twarock and Stockley 2019, Bayfield *et al.* 2020, Kizziah *et al.* 2020, Mata *et al.* 2023, Podgorski *et al.* 2020, Bruinsma *et al.* 2021, Johnson *et al.* 2021, Montiel-Garcia *et al.* 2021, Plavec *et al.* 2021, Wilson and Roof 2021, de Pablo and San Martín 2022, Lee *et al.* 2022, Mohajerani *et al.* 2022). Among the computational methods, molecular dynamics algorithms have improved dramatically in the last decades and can infer the dynamics of large protein complexes. However, they resolve relatively short timescales ( $\leq 1$   $\mu$ s) and require specialized computational resources (Perilla and Schulten 2017, Hadden *et al.* 2018, Jana and May 2021, Bryer *et al.* 2022). The fact that

capsids are assembled from 60 to more than 60 000 proteins further limits the application of molecular dynamics (Twarock and Luque 2019, Luque *et al.* 2020, Berg and Roux 2021). Alternatively, the combination of normal mode analysis (NMA), molecular coarse-graining, and elastic network models (ENM) offers a more scalable solution (Bahar *et al.* 2010, Romo and Grossfield 2011). This approach has successfully estimated the collective motion of proteins in complexes (Bahar *et al.* 2010) or identified structural conformational changes in capsids (Tama and Brooks, 2005). Nonetheless, no easy-to-use computational packages for non-computational-based specialists are currently available to characterize the dynamics and mechanical properties of protein shells. The bioinformatics software presented here, pyCapsid, aims to address this issue.

pyCapsid is inspired by prior publications that applied ENM, NMA, and clustering methods to extract the quasi-rigid regions of protein shells (Polles *et al.* 2013, Ponzoni *et al.* 2015). These methods can identify the mechanical units involved in the assembly or disassembly of protein shells (Polles *et al.* 2013). However, obtaining these results using packages such as PRINQ++ (Polles *et al.* 2013), SPECTRUS (Ponzoni *et al.* 2015), NRGTEEN (Mailhot and Najmanovich 2021), ClustENMD (Kaynak *et al.* 2021), WebPSN (Seeber *et al.* 2015), or the popular ProDy (Zhang *et al.* 2021), is not

Received: 27 February 2023; Revised: 1 November 2023; Editorial Decision: 28 November 2023; Accepted: 15 December 2023

© The Author(s) 2023. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

trivial and limited to relatively small complexes. To address this issue, we introduce pyCapsid as an accessible and efficient Python package that identifies the dominant dynamics and quasi-rigid regions of protein shells. The underlying methods used in pyCapsid are generic and can be applied to other protein complexes, yet, in this first release of pyCapsid, we have focused on the characterization of protein shells, such as viral capsids, cellular protein compartments like encapsulins, and gene-transfer agents (Giessen *et al.* 2019, Bárdy *et al.* 2020, Johnson and Olson 2021, Montiel-Garcia *et al.* 2021).

## 2 Methods and features

pyCapsid's Python package is divided into five independent modules (Fig. 1): the PDB (protein data bank) module, the CG (coarse-graining) module, the NMA (normal mode analysis) module, the QRC (quasi-rigid clustering) module, and the VIS (visualization) module. The role and technical aspects of each module are briefly described below.

**PDB module.** This module retrieves and loads structural data from the Protein Data Bank (using the PDB ID) or a local file in PDB or PDBx/mmCIF formats (Berman *et al.* 2000, Westbrook *et al.* 2022). The PDB module builds on functions from the Python package Biotite (Kunzmann and Hamacher 2018).

**CG module.** This module coarse-grains the proteins at the amino-acid level and establishes an elastic force field between amino acids. It offers four different elastic models: the anisotropic network model (ANM), the Gaussian network model

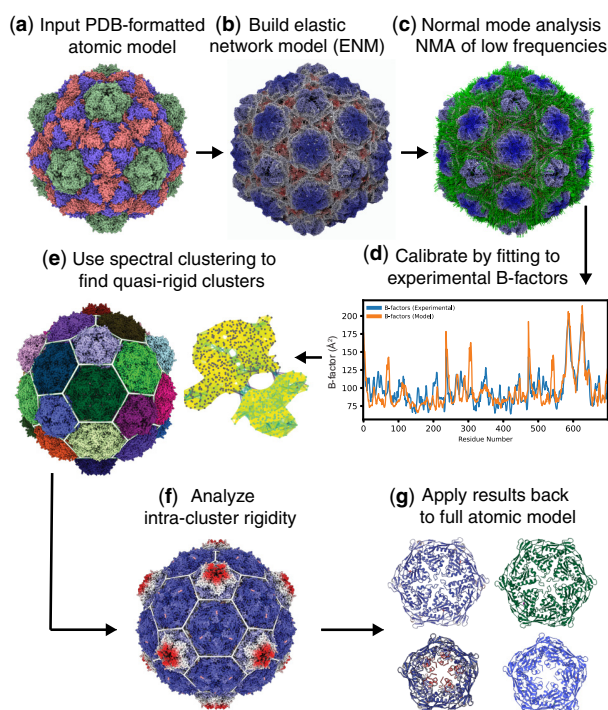
(GNM), the generalized anisotropic, or unified elastic network model (G-ANM or U-ENM), which is the default, and the backbone-enhanced elastic network model (bb-ENM). Each amino acid is coarse-grained as a point mass on the alpha-carbon. The links in the network connect amino acids that are within a threshold. The default value is 15 Å for ANM and 7.5 Å for the other models. These values are based on prior studies of elastic models reproducing empirical molecular thermal fluctuations (B-factors) (Micheletti *et al.* 2004, Eyal *et al.* 2006, Zheng 2008, Romo and Grossfield 2011). The small threshold distance leads to a sparse network. This network, combined with the elastic strength values of the elastic model, defines the Hessian matrix. The calculations to build the matrix are accelerated using Numba (Lam *et al.* 2015).

**NMA module.** This module obtains the motions of the macromolecular complex by decomposing the dynamics into independent sinusoidal motions called normal modes (Goldstein *et al.* 2002). The normal modes and associated frequencies are obtained from the Hessian matrix derived in the CG module. It is well established that only low-frequency modes are relevant to the global dynamics of macromolecules (Bahar *et al.* 2010). The default number of modes calculated in pyCapsid is 200. This number was selected by comparing the results with simulations using a larger number of modes (as many modes as 1/100 of the number of residues, i.e. 1000 modes for a structure containing 100 000 residues). pyCapsid also provides an optional dependency to accelerate the calculations in GPUs using CUDA via solvers in the cupy package (Okuta *et al.* 2017).

**QRC module.** This module estimates the amino acids that tend to fluctuate as a single mechanical unit (quasi-rigid cluster) using the SPECTRUS algorithm (Ponzoni *et al.* 2015). A cluster contains groups of residues that minimize the distance fluctuations between residues. The clustering method used in pyCapsid is the default discretize method from scikit-learn (Pedregosa *et al.* 2012), and k-means clustering is offered as an alternative (Yu and Shi 2003). pyCapsid explores a range of clusters from four to a maximum number (cluster\_max) set by the user. The quality score defined in (Ponzoni *et al.* 2015) is calculated for each number of clusters, and the set of clusters with the maximum score is selected. If multiple local maxima are observed, each can be selected via the application programming interface (API) for further analysis. These alternative maxima can also be obtained by restricting the number of clusters when running pyCapsid.

**VIS module.** The results obtained from pyCapsid are stored and organized in data files and figures in the same running folder. pyCapsid's [online tutorial](#) and [YouTube channel](#) provide instructions and scripts to visualize the results using the molecular visualization tools NGLview and ChimeraX (Nguyen *et al.* 2018, Pettersen *et al.* 2021). The results from the cloud-based Colab notebook include a script that can be run directly in Chimera X to generate the 3D visualizations. To obtain high-quality visualizations and animations when running pyCapsid locally, users must have installed ChimeraX (version 1.5 or above) and indicate the path in the pyCapsid's configuration file.

**Execution.** Users running pyCapsid locally can specify the necessary parameters in a configuration file (TOML format), run it, and obtain the results. Users running pyCapsid on the cloud as a Colab notebook can modify parameters as indicated in the notebook's quick-start guide. Providing the



**Figure 1.** pyCapsid's key steps. (a) The protein shell (atom coordinates, atom types, and B-factors) is inputted using a PDB file (or a PDB ID). (b) The elastic network model (ENM) is calibrated (parametrized). (c) Normal mode analysis (NMA) determines the distance fluctuations between residues. (d) The correlation coefficient of predicted and experimental B-factors is generated for quality control. (e) Spectral clustering is applied to obtain the quasi-rigid molecular domains in the capsid. (f) The fluctuations between residues within the rigid clusters are calculated. (g) The results are mapped back to the capsid protein atomic model for structural interpretation.

pyCapsid Colab option follows the accessibility trend of cloud-based bioinformatics tools in this environment, like ColabFold (Mirdita *et al.* 2022) and other tools predicting protein complexes (Bryant *et al.* 2022). For advanced users, pyCapsid provides an API to access the objects in the five separate modules described above.

## 3 Applications

### 3.1 Performance and accuracy

pyCapsid's performance and accuracy were obtained by studying 25 protein shells (see Data availability), which contained from 16 000 to 400 000 amino acid residues, displayed icosahedral symmetry with T-numbers spanning from  $T=1$  to  $T=16$ , and had resolutions ranging from 2 to 5.2 Å. The baseline performance was obtained using an HPC cluster core with Intel Xeon CPU E5-2650 v4 (2.20 GHz) and 128 GB of RAM. The peak memory usage ranged from 800 MB to 90 GB and increased with the number of residues following a power law (exponent =  $1.46 \pm 0.06$  and  $R^2 = 0.97$ ). The runtime ranged from 2 min to 36 h and increased with the number of residues following a power law (exponent =  $2.20 \pm 0.10$  and  $R^2 = 0.95$ ). The correlation coefficient between the simulated and empirical thermal motions (B-factors) of the amino acids was used as a proxy to evaluate the accuracy of the selected elastic network model (ENM) and normal mode analysis (NMA). The correlation coefficients using the default number of modes ranged from 0.10 to 0.88 out of 1.00. The distribution of correlation coefficients was consistent with the correlations observed for B-factors predicted using ENM and NMA in smaller protein complexes (Eyal *et al.* 2006). The accuracy decreased linearly for structures with lower experimental resolution (slope =  $-0.20 \pm 0.05$  1/Å and  $R^2 = 0.40$ ), with a regression projecting perfect accuracy for structures with an ideal experimental resolution of 0 Å (intercept =  $1.23 \pm 0.18$ ). The accuracy was independent of the number of residues (Spearman's coefficient =  $-0.09$  and  $P$ -value = 0.66). pyCapsid's performance was also assessed by analyzing five small-to-medium capsids in a free cloud-based Colab account and three personal computers with different configurations. The execution times were on the same order of magnitude as in the HPC, with average relative runtime factors ranging from 0.98 (faster) to 2.66 (slower). The main bottleneck was memory, which placed an upper limit to the largest capsid size that could be analyzed. See Data availability for further details.

### 3.2 Benchmarking

The five smallest protein shells were used to benchmark the speed and accuracy of pyCapsid with respect to ProDy, a well-established Python package for protein structural dynamics analysis (Zhang *et al.* 2021). Since ProDy does not generate the quasi-rigid domain decomposition, the comparison focused on the modules responsible for loading the PDB and generating the normal modes analysis (NMA). The anisotropic network model (ANM) was available in both ProDy and pyCapsid and yielded the same B-factors. The unified elastic network model (G-ANM or U-ENM), which is available in pyCapsid but not ProDy, improved by five times the average correlation coefficient (from 0.11 for ANM to 0.56 for U-ENM) of the B-factors using the default number of modes. pyCapsid displayed an average speed increase of  $3.0 \pm 1.5$  with respect to ProDy. This increase was independent of capsid size (Spearman's coefficient = 0.11 and

$P$ -value = 0.76). The increase in speed was due to the use of Numba and the invert shift mode in SciPy. This, however, caused a similar increase in memory usage. In any case, ANM or U-ENM did not impact pyCapsid's speed performance when using the same number of modes. Thus, U-ENM was selected as the default and recommended model in the pyCapsid package. pyCapsid was not benchmarked quantitatively with tools other than ProDy because even the smallest protein shells exceeded their capacity, or we encountered technical barriers when deploying them locally to analyze such protein shells. Nonetheless, four additional small capsids (PDB IDs 2ms2, 1za7, 1a34, and 3nap) were investigated by pyCapsid recovering the quasi-rigid domains published previously with PISQRD++ (Polles *et al.* 2013) and SPECTRUS (Ponzoni *et al.* 2015). See Data availability for further details.

### 3.3 Capsid disassembly

An application of pyCapsid is the ability to identify the capsid regions and protein clusters that are most likely to be involved in the initial steps of capsid disassembly, which is crucial for the development of antiviral strategies. A current experimental issue is the fact that capsids are very stable, so generic stressors and denaturants are used to accelerate the disassembly to make studies *in vitro* possible (Zhao *et al.* 2012, Zhou *et al.* 2022). The mechanical units predicted by pyCapsid could be tested experimentally, e.g. using atomic force microscopy (Ortega-Esteban *et al.* 2013, Martín-González *et al.* 2023), mass spectrometry (Utrecht *et al.* 2011, Bond *et al.* 2020), and light scattering (Garmann *et al.* 2019, Timmermans *et al.* 2022). Here, this potential application of pyCapsid was tested by analyzing seven diverse capsids (SSV, TrV, HBV, CCMV, MS2, P22, and HK97) that have been observed to disassemble experimentally (Anobom *et al.* 2003, Lavelle *et al.* 2007, Duda *et al.* 2009, Agirre *et al.* 2013, Kant *et al.* 2018, Strauss *et al.* 2018, Starr *et al.* 2022). The quasi-rigid domains identified in these capsids were consistent with the disassembly units observed empirically (see the [online/results/disassembly](#) folder in Data availability). pyCapsid's predictions identify molecular targets that will guide more mechanistic and capsid-specific disassembly experiments, helping bridge *in vivo* and *in vitro* conditions.

## 4 Concluding remarks

pyCapsid can generate the collective motion and extract the quasi-rigid functional regions involved in the stability and disassembly of protein shells and other protein complexes. The underlying algorithm of pyCapsid generates the dynamical modes faster than established protein dynamics packages and can handle the quasi-rigid domain decomposition of large protein complexes in the range of minutes to over a day, even in regular computers. The computational efficiency of pyCapsid, combined with its accessibility via Python distribution packages, Google Colab, and online tutorials, will benefit researchers in physical virology, structural bioinformatics, and related fields and will facilitate the prediction of disassembly units in protein shells, fostering new antiviral and drug delivery strategies.

## Acknowledgements

The authors thank the insights from the SDSU Biomath Group, in particular, professors Arlette Baljon and Parag Katira. The authors also thank the SDSU Computational



Research Science Center for the HPC used to test and benchmark pyCapsid.

## Conflict of interest

The authors declare no conflicts of interest.

## Funding

This work was supported by the National Science Foundation [1951678]; and the Gordon and Betty Moore Foundation [GBMF9871] (<https://doi.org/10.37807/GBMF9871>). The HPC facilities were supported by the NSF Office of Advanced Cyberinfrastructure [1659169].

## Data availability

All data are incorporated into the article and its online supplementary material. The data files regarding the performance results reported in this article are available in the GitHub repository's [results folder](#). A gallery displaying the application of pyCapsid to various protein shells is available in the online documentation [gallery section](#).

## References

- Agirre J, Goret G, LeGoff M *et al.* Cryo-electron microscopy reconstructions of triatoma virus particles: a clue to unravel genome delivery and capsid disassembly. *J Gen Virol* 2013;**94**:1058–68.
- Anobom CD, Albuquerque SC, Albernaz FP *et al.* Structural studies of MS2 bacteriophage virus particle disassembly by nuclear magnetic resonance relaxation measurements. *Biophys J* 2003;**84**:3894–903.
- Bahar I, Lezon TR, Yang L-W *et al.* Global dynamics of proteins: bridging between structure and function. *Annu Rev Biophys* 2010;**39**:23–42.
- Bárdy P, Füzik T, Hrebík D *et al.* Structure and mechanism of DNA delivery of a gene transfer agent. *Nat Commun* 2020;**11**:3034.
- Bayfield OW, Steven AC, Antson AA. Cryo-EM structure in situ reveals a molecular switch that safeguards virus against genome loss. *eLife* 2020;**9**:e55517. <https://doi.org/10.7554/eLife.55517>.
- Berg M, Roux S. Extreme dimensions—how big (or small) can tailed phages be? *Nat Rev Microbiol* 2021;**19**:407.
- Berman HM, Westbrook J, Feng Z *et al.* The protein data bank. *Nucleic Acids Res* 2000;**28**:235–42.
- Bond KM, Lyktey NA, Tsvetkova IB *et al.* Disassembly intermediates of the brome mosaic virus identified by charge detection mass spectrometry. *J Phys Chem B* 2020;**124**:2124–31.
- Bruinsma RF, Wuite GJL, Roos WH. Physics of viral dynamics. *Nat Rev Phys* 2021;**3**:76–91.
- Bryant P, Pozzati G, Zhu W *et al.* Predicting the structure of large protein complexes using AlphaFold and Monte Carlo tree search. *Nat Commun* 2022;**13**:6028.
- Bryer AJ, Reddy T, Lyman E *et al.* Full scale structural, mechanical and dynamical properties of HIV-1 liposomes. *PLoS Comput Biol* 2022;**18**:e1009781.
- Callaway E. Revolutionary cryo-EM is taking over structural biology. *Nature* 2020;**578**:201.
- de Pablo PJ, San Martín C. Seeing and touching adenovirus: complementary approaches for understanding assembly and disassembly of a complex virion. *Curr Opin Virol* 2022;**52**:112–22.
- Duda RL, Ross PD, Cheng N *et al.* Structure and energetics of encapsidated DNA in bacteriophage HK97 studied by scanning calorimetry and cryo-electron microscopy. *J Mol Biol* 2009;**391**:471–83.
- Eyal E, Yang L-W, Bahar I. Anisotropic network model: systematic evaluation and a new web interface. *Bioinformatics* 2006;**22**:2619–27.
- Garman RF, Goldfain AM, Manoharan VN. Measurements of the self-assembly kinetics of individual viral capsids around their RNA genome. *Proc Natl Acad Sci USA* 2019;**116**:22485–90.
- Giessen TW, Orlando BJ, Verdegaal AA *et al.* Large protein organelles form a new iron sequestration system with high storage capacity. *eLife* 2019;**8**:e46070. <https://doi.org/10.7554/eLife.46070>.
- Goldstein H, Poole C, Safko J. *Classical Mechanics*. 3rd Ed, London, England: Pearson 2001.
- Grime JMA, Dama JF, Ganser-Pornillos BK *et al.* Coarse-grained simulation reveals key features of HIV-1 capsid self-assembly. *Nat Commun* 2016;**7**:11568.
- Hadden JA, Perilla JR, Schlitsch CJ *et al.* All-atom molecular dynamics of the HBV capsid reveals insights into biological function and cryo-EM resolution limits. *eLife* 2018;**7**:e32478. <https://doi.org/10.7554/eLife.32478>.
- Hua J, Huet A, Lopez CA *et al.* Capsids and genomes of jumbo-sized bacteriophages reveal the evolutionary reach of the HK97 fold. *mBio* 2017;**8**:e01579–17. <https://doi.org/10.1128/mBio.01579-17>.
- Jana AK, May ER. Atomistic dynamics of a viral infection process: release of membrane lytic peptides from a non-enveloped virus. *Sci Adv* 2021;**7**:eabe1761. <https://doi.org/10.1126/sciadv.abe1761>.
- Johnson JE, Domitrovic T, Matsui T *et al.* Dynamics and stability in the maturation of a eukaryotic virus: a paradigm for chemically programmed large-scale macromolecular reorganization. *Arch Virol* 2021;**166**:1547–63.
- Johnson JE, Olson AJ. Icosahedral virus structures and the protein data bank. *J Biol Chem* 2021;**296**:100554.
- Kant R, Llauro A, Rayaprolu V *et al.* Changes in the stability and biomechanics of P22 bacteriophage capsid during maturation. *Biochim Biophys Acta Gen Subj* 2018;**1862**:1492–504.
- Kaynak BT, Zhang S, Bahar I *et al.* ClustENMD: efficient sampling of biomolecular conformational space at atomic resolution. *Bioinformatics* 2021;**37**:3956–8.
- Kizziah JL, Rodenburg CM, Dokland T. Structure of the capsid Size-Determining scaffold of "satellite" bacteriophage P4. *Viruses* 2020;**12**:953.
- Kunzmann P, Hamacher K. Biotite: a unifying open source computational biology framework in python. *BMC Bioinformatics* 2018;**19**:346.
- Lam SK, Pitrou A, Seibert S. Numba: a LLVM-Based python JIT compiler. In: *Proceedings of the Second Workshop on the LLVM Compiler Infrastructure in HPC*. New York, NY: Association for Computing Machinery, 2015.
- Lavelle L, Michel J-P, Gingery M. The disassembly, reassembly and stability of CCMV protein capsids. *J Virol Methods* 2007;**146**:311–6.
- Lee DY, Bartels C, McNair K *et al.* Predicting the capsid architecture of phages from metagenomic data. *Comput Struct Biotechnol J* 2022;**20**:721–32.
- Li L, Lok S-M, Yu I-M *et al.* The flavivirus precursor membrane-envelope protein complex: structure and maturation. *Science* 2008;**319**:1830–4.
- Luque A, Benler S, Lee DY *et al.* The missing tailed phages: prediction of small capsid candidates. *Microorganisms* 2020;**8**:1944.
- Luque A, Reguera D. Theoretical studies on assembly, physical stability and dynamics of viruses. In: Mateu M (Ed.), *Structure and Physics of Viruses*. Vol 68. Dordrecht: Springer, 2013, 553–95.
- Mailhot O, Najmanovich R. The NRGTE Python package: an extensible toolkit for coarse-grained normal mode analysis of proteins, nucleic acids, small molecules and their complexes. *Bioinformatics* 2021;**37**:3369–71.
- Martín-González N, Gómez-González A, Hernando-Pérez M *et al.* Adenovirus core protein V reinforces the capsid and enhances genome release from disrupted particles. *Sci Adv* 2023;**9**:eade9910. <https://doi.org/10.1126/sciadv.ade9910>.
- Mata CP, Rodríguez JM, Suzuki N *et al.* Structure and assembly of double-stranded RNA mycoviruses. In: *Advances in Virus Research*, Vol. 108, Academic Press pp. 213–47 2023.
- Micheletti C, Carloni P, Maritan A. Accurate and efficient description of protein vibrational dynamics: comparing molecular dynamics and Gaussian models. *Proteins Struct Funct Bioinf* 2004;**55**:635–45.

- Mirdita M, Schütze K, Moriwaki Y *et al.* ColabFold: making protein folding accessible to all. *Nat Methods* 2022;19:679–82.
- Mohajerani F, Tyukodi B, Schlicksup CJ *et al.* Multiscale modeling of hepatitis B virus capsid assembly and its dimorphism. *ACS Nano* 2022;16:13845–59.
- Montiel-García D, Santoyo-Rivera N, Ho P *et al.* VIPERdb v3.0: a structure-based data analytics platform for viral capsids. *Nucleic Acids Res* 2021;49:D809–16.
- Nguyen H, Case DA, Rose AS. NGLview—interactive molecular graphics for Jupyter notebooks. *Bioinformatics* 2018;34:1241–2.
- Okuta R, Unno Y, Nishino D *et al.* CuPy: A NumPy-compatible library for NVIDIA GPU calculations. In: *Workshop on Machine Learning Systems (LearningSys) at Neural Information Processing Systems (NIPS)*, Long Beach, CA, USA, 2017.
- Organtini LJ, Shingler KL, Ashley RE *et al.* Honey bee deformed wing virus structures reveal that conformational changes accompany genome release. *J Virol* 2017;91:e01795–16. <https://doi.org/10.1128/JVI.01795-16>.
- Ortega-Esteban A, Pérez-Berná AJ, Menéndez-Conejero R *et al.* Monitoring dynamics of human adenovirus disassembly induced by mechanical fatigue. *Sci Rep* 2013;3:1434.
- Pedregosa F, Varoquaux G, Gramfort A *et al.* *Scikit-learn: Machine Learning in Python*. arXiv:1201.0490. 2012.
- Perilla JR, Schulten K. Physical properties of the HIV-1 capsid from all-atom molecular dynamics simulations. *Nat Commun* 2017;8:15959.
- Pettersen EF, Goddard TD, Huang CC *et al.* UCSF ChimeraX: structure visualization for researchers, educators, and developers. *Protein Sci* 2021;30:70–82.
- Plavec Z, Pöhner I, Poso A *et al.* Virus structure and structure-based antivirals. *Curr Opin Virol* 2021;51:16–24.
- Podgorski J, Calabrese J, Alexandrescu L *et al.* Structures of three actinobacteriophage capsids: roles of symmetry and accessory proteins. *Viruses* 2020;12:294.
- Polles G, Indelicato G, Potestio R *et al.* Mechanical and assembly units of viral capsids identified via quasi-rigid domain decomposition. *PLoS Comput Biol* 2013;9:e1003331. <https://doi.org/10.1371/journal.pcbi.1003331>.
- Ponzoni L, Polles G, Carnevale V *et al.* SPECTRUS: a dimensionality reduction approach for identifying dynamical domains in protein complexes from limited structural datasets. *Structure* 2015;23:1516–25.
- Qazi S, Schlicksup CJ, Rittichier J *et al.* An assembly-activating site in the hepatitis B virus capsid protein can also trigger disassembly. *ACS Chem Biol* 2018;13:2114–20.
- Romo TD, Grossfield A. Validating and improving elastic network models with molecular dynamics simulations. *Proteins Struct Funct Bioinf* 2011;79:23–34.
- Seeber M, Felling A, Raimondi F *et al.* WebPSN: a web server for high-throughput investigation of structural communication in biomacromolecules. *Bioinformatics* 2015;31:779–81.
- Starr CA, Barnes LF, Jarrold MF *et al.* Hysteresis in hepatitis B virus (HBV) requires assembly of near-perfect capsids. *Biochemistry* 2022;61:505–13.
- Strauss M, Jayawardena N, Sun E *et al.* Cryo-electron microscopy structure of seneca valley virus procapsid. *J Virol* 2018;92:e01927–17. <https://doi.org/10.1128/JVI.01927-17>.
- Tama F, Brooks CL. Diversity and identity of mechanical properties of icosahedral viral capsids studied with elastic network normal mode analysis. *J Mol Biol* 2005;345:299–314.
- Timmermans SBPE, Ramezani A, Montalvo T *et al.* The dynamics of viruslike capsid assembly and disassembly. *J Am Chem Soc* 2022;144:12608–12.
- Twarock R, Luque A. Structural puzzles in virology solved with an overarching icosahedral design principle. *Nat Commun* 2019;10:4414.
- Twarock R, Stockley PG. RNA-Mediated virus assembly: mechanisms and consequences for viral evolution and therapy. *Annu Rev Biophys* 2019;48:495–514.
- Utrecht C, Barbu IM, Shoemaker GK *et al.* Interrogating viral capsid assembly with ion mobility–mass spectrometry. *Nat Chem* 2011;3:126–32.
- Westbrook JD, Young JY, Shao C *et al.* PDBx/mmCIF ecosystem: foundational semantic tools for structural biology. *J Mol Biol* 2022;434:167599.
- Wilson DP, Roof DA. Viral phrenology. *Viruses* 2021;13:2191.
- Yeager M, Dryden KA, Olson NH *et al.* Three-dimensional structure of rhesus rotavirus by cryoelectron microscopy and image reconstruction. *J Cell Biol* 1990;110:2133–44.
- Yu SX, Shi J. Multiclass spectral clustering. In: *Proceedings Ninth IEEE International Conference on Computer Vision, Nice, France*, Vol. 1, 2003, 313–9.
- Zhang S, Krieger JM, Zhang Y *et al.* ProDy 2.0: increased scale and scope after 10 years of protein dynamics modelling with Python. *Bioinformatics* 2021;37:3657–9.
- Zhang X, Guo H, Jin L *et al.* A new topology of the HK97-like fold revealed in Bordetella bacteriophage by cryoEM at 3.5 Å resolution. *eLife* 2013;2:e01299. <https://doi.org/10.7554/eLife.01299>.
- Zhao Q, Modis Y, High K *et al.* Disassembly and reassembly of human papillomavirus virus-like particles produces more virion-like antibody reactivity. *Virol J* 2012;9:52.
- Zheng W. A unification of the elastic network model and the gaussian network model for optimal description of protein conformational motions and fluctuations. *Biophys J* 2008;94:3853–7.
- Zhou J, Zlotnick A, Jacobson SC. Disassembly of single virus capsids monitored in real time with multicycle resistive-pulse sensing. *Anal Chem* 2022;94:985–92.