

The ACPYPE web server for small-molecule MD topology generation

Kagami, Luciano; Wilter, Alan; Diaz, Adrian; Vranken, Wim

Published in:
Bioinformatics

DOI:
[10.1093/bioinformatics/btad350](https://doi.org/10.1093/bioinformatics/btad350)

Publication date:
2023

License:
CC BY

Document Version:
Final published version

[Link to publication](#)

Citation for published version (APA):

Kagami, L., Wilter, A., Diaz, A., & Vranken, W. (2023). The ACPYPE web server for small-molecule MD topology generation. *Bioinformatics*, 39(6), [btad350]. <https://doi.org/10.1093/bioinformatics/btad350>

Copyright

No part of this publication may be reproduced or transmitted in any form, without the prior written permission of the author(s) or other rights holders to whom publication rights have been transferred, unless permitted by a license attached to the publication (a Creative Commons license or other), or unless exceptions to copyright law apply.

Take down policy

If you believe that this document infringes your copyright or other rights, please contact openaccess@vub.be, with details of the nature of the infringement. We will investigate the claim and if justified, we will take the appropriate steps.

Structural bioinformatics

The ACPYPE web server for small-molecule MD topology generation

Luciano Kagami¹, Alan Wilter², Adrian Diaz^{1,3}, Wim Vranken ^{1,3,*}

¹Interuniversity Institute of Bioinformatics in Brussels, VUB/ULB, Brussels 1050, Belgium

²Phenopolis Ltd & Moorfields Eye Hospital, London, United Kingdom

³Structural biology Brussels, Vrije Universiteit Brussel, Brussels 1050, Belgium

*Corresponding author. Interuniversity Institute of Bioinformatics in Brussels, VUB/ULB, Brussels 1050, Belgium. E-mail: wim.vranken@vub.be (W.V.)

Associate Editor: Arne Elofsson

Abstract

Motivation: The generation of parameter files for molecular dynamics (MD) simulations of small molecules that are suitable for force fields commonly applied to proteins and nucleic acids is often challenging. The ACPYPE software and website aid the generation of such parameter files.

Results: ACPYPE uses OpenBabel and ANTECHAMBER to generate MD input files in Gromacs, AMBER, CHARMM, and CNS formats. It can now take a SMILES string as input, in addition to the original PDB or mol2 coordinate files, with GAFF2 support and GLYCAM force field conversion added. It can be installed locally via Anaconda, PyPI, and Docker distributions, while the web server at <https://bio2byte.be/acpype/> was updated with an API, and provides visualization of results for uploaded molecules as well as a pre-generated set of 3738 drug molecules.

Availability and implementation: The web application is freely available at <https://www.bio2byte.be/acpype/> and the open-source code can be found at <https://github.com/alanwilter/acpype>.

1 Introduction

Molecular dynamics (MD) simulations provide information about the time-dependent evolution of molecular systems at the atomic level. In biology, MD simulations are widely used to study the conformations, dynamics, and interactions of proteins and nucleic acids and provide important information about the variability in these systems (Kagami *et al.* 2017). While most of the commonly used MM force fields such as CHARMM (Vanommeslaeghe *et al.* 2009), OPLS-AA (Kaminski *et al.* 2001), GROMACS (Van Der Spoel *et al.* 2005), and AMBER (Wang *et al.* 2004) provide complete parameters for proteins and nucleic acids, the correct parameterization of many ligands remains a challenge with these force fields, as they are not optimized for small molecules. To simplify the inclusion of ligands in biological molecular systems, ACPYPE was developed in 2012. ACPYPE is a Python wrapper script around the ANTECHAMBER software that aids the generation of small-molecule topologies and parameters for a variety of MD programs by providing input files in AMBER, CHARMM, GROMACS, and CNS formats (Phillips *et al.* 2005). It is widely used in the community, with currently almost 1500 published papers referring to ACPYPE. During its extensive use by the community, ACPYPE has continually been updated in pace with GROMACS and AmberMD releases and we are now announcing the updated ACPYPE version and server (<https://bio2byte.be/acpype/>). Notably, in 2019, the ability to convert AMBER GLYCAM06 force field to GROMACS was added to ACPYPE by Bernardi *et al.* (2019). ACPYPE now also supports input in SMILES format, has GAFF2 support, and is included in the

GROMACS user guide (<https://manual.gromacs.org/current/user-guide/index.html>), as the tool to generate AMBER-based topologies. ACPYPE is now available for local install from GitHub, Anaconda, Pypi, and Docker Hub. It has in addition been embedded in MD simulation frameworks such as SimTK/mmttools (<https://simtk.org/home/mmttools>), MemGen (<http://memgen.uni-goettingen.de/>), LiGRO (Kagami *et al.* 2017), PolyParGen (<http://polypargen.com/>), Cosolvent Analysis Toolkit (Sabanés Zariquiey, de Souza and Bronowska 2019), BioExcel Building Blocks (<https://mmb.irbbarcelona.org/biobb/>), and Winmostar V10 (<https://winmostar.com>). The new server uses a token-based system, has API call functions, provides pre-generated files for common ligands, and supports direct access to all files as well as visualization of atom and coordinate information, including InChI information and links to online chemical resources (when available). In addition, the new functionality enables easier manual verification of the ACPYPE results, such as visualizing the coordinates to check, for example, for possible missing improper dihedrals in aromatic systems. We also refer to the GROMACS and AmberMD mailing lists when help in parametrization is required.

2 Materials and methods

2.1 Input

ACPYPE accepts 3 input atom coordinate formats, using the Openbabel package (<https://openbabel.org/>) to enable the upload of SYBYL MOL2 (<https://www.certara.com/>) and PDB (Protein Data Bank) formats (Shao *et al.* 2022), as well as text

Received: January 17, 2023. Revised: May 10, 2023. Editorial Decision: May 23, 2023. Accepted: May 29, 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.

Table 1. ACPYPE output files.

Amber	Charmm	CNS	GROMACS	Coordinates	Other
INPCRD	INP	INP	ITP ^a	MOL2	MDP ^b
LIB	PRM	PAR	TOP ^a	PDB ^c	LOG ^d
PMRTOP	RTF	TOP	GRO		

^a Amber and OPLS force field.

^b GROMACS parameters files to run minimization and dynamics.

^c Modified atom numbering.

^d ACPYPE run logs.

Table 2. ACPYPE REST API endpoints.

Method	Endpoint	Description
POST	/acpype/api/	Sends the data for processing.
GET	/acpype/api/queue/	Gets the data after processing. ^a

^a Using the alphanumeric hash ID provided during submission.

entry of SMILES (Simplified Molecular Input Line Entry System) (O'Boyle 2012). Entries of up to 200 atoms are accepted on the server.

2.2 Output

ACPYPE outputs 16 different file formats in a single run (Table 1), both from Amber and Charmm.

2.3 RESTful API

The ACPYPE API uses the Django REST Framework 3.11 and provides programmatic access to submission of single coordinate file in PDB and MOL2 in file format (recommended) or a single SMILES in string format, which is less reliable as 3D coordinates for the atoms must be generated through OpenBabel (O'Boyle 2012). Using two endpoints (Table 2), the API returns the results in JSON (JavaScript Object Notation) format. The result can be retrieved using the alphanumeric hash ID provided by the submission process, so protecting the users' data security and confidentiality. Batch processes are possible by submission of a single job at a time.

2.4 SMILES format recognition and ACPYPE performance

The ACPYPE functionality to directly use SMILES strings depends on converting this format to a three-dimensional model, which is performed by the embedded OpenBabel software. To verify how this new function handles the different SMILES, we ran ACPYPE for 4099 different drugs obtained from the DrugCentral server (Avram *et al.* 2021). For this, a TSV file was downloaded and used together with a Python script to convert SMILES to MOL2 format. The ability of ACPYPE to parameterize these MOL2 files was then tested, with all log files saved and analyzed.

2.5 Local installation availability

ACPYPE can also be installed locally from GitHub (<https://github.com/alanwilter/acpype>) as well as via Anaconda (<https://anaconda.org/conda-forge/acpype>), Pypi (<https://pypi.org/project/acpype/>), and Docker Hub (<https://hub.docker.com/t/acpype/acpype>). All ACPYPE functionality can be obtained using the local installation, including direct SMILES conversion. The source version also does not limit runtime or number of atoms.

Table 3. ACPYPE API success rate for SMILES format submission.

Status	Number of compounds	%
Success	3738	91.2
Antechamber failed	18	0.4
Intramolecular bonds	42	1.0
More than one molecule detected	120	2.9
Time exceeded	181	4.4
Total	4099	100

Figure 1. Visualization of an entry on the ACPYPE pages.

3 Results

The updates on the features, local installation options, and server extensions of ACPYPE are aimed at making the software easier to use and more accessible to users. Particularly useful might be the ACPYPE SMILES functionality, which was able to handle 91.2% of a set of small drug molecules (totalling 3738 compounds—Table 3). Failures are due to limitations of the ANTECHAMBER program (0.4%), or due to checks included in the ACPYPE program such as failures in the recognition of intramolecular bonds (1%), detection of multiple molecules (2.9%), or maximum processing time of 3 h exceeded (4.4%). The complete table with results is reported in Supplementary Tables S1 and S2.

The uploaded files and pre-calculated results can be browsed and visualized (Fig. 1), with various information shown as well as links to AlphaFill (Hekkelman *et al.* 2023), which can help the user find biologically relevant leads for their molecule(s).

4 Conclusion

ACPYPE is one of the most popular programs for building small-molecule topologies for MD, especially for use with the Amber force field. After a decade of use, ACPYPE is robust and has now been updated with new features. ACPYPE is free of charge and its results can be obtained through multiple approaches.

Supplementary data

Supplementary data is available at *Bioinformatics* online.

Conflict of interest: The authors have no conflicts of interest to declare. All authors have seen and agree with the contents of the manuscript and there is no financial interest to report. We certify that the submission is original work and is not under review at any other publication.

Funding

This work was supported by the Research Foundation Flanders [grant numbers G028821N and I000323N] to A.D. and W.V.

Data availability

The data underlying this article are available in the article and in its online supplementary material or is accessible via the ACPYPE web server at <https://bio2byte.be/acpype/>.

References

- Avram S, Bologna CG, Holmes J *et al.* DrugCentral 2021 supports drug discovery and repositioning. *Nucleic Acids Res* 2021;49: D1160–D1169. <https://doi.org/10.1093/nar/gkaa997>.
- Bernardi A, Faller R, Reith D *et al.* ACPYPE update for nonuniform 1–4 scale factors: conversion of the GLYCAM06 force field from AMBER to GROMACS. *SoftwareX* 2019;10:100241. <https://doi.org/10.1016/j.softx.2019.100241>.
- Hekkelman ML, de Vries I, Joosten RP *et al.* AlphaFill: enriching AlphaFold models with ligands and cofactors. *Nat Methods* 2023; 20:205–13. <https://doi.org/10.1038/s41592-022-01685-y>
- Kagami LP, das Neves GM, da Silva AWS *et al.* LiGRO: a graphical user interface for protein–ligand molecular dynamics. *J Mol Model* 2017; 23:304. <https://doi.org/10.1007/s00894-017-3475-9>.
- Kaminski GA, Friesner RA, Tirado-Rives J *et al.* Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides. *J Phys Chem B* 2001;105:6474–87. <https://doi.org/10.1021/jp003919d>.
- O’Boyle NM. Towards a universal SMILES representation—a standard method to generate canonical SMILES based on the InChI. *J Cheminform* 2012;4:22. <https://doi.org/10.1186/1758-2946-4-22>.
- Phillips JC, Braun R, Wang W *et al.* Scalable molecular dynamics with NAMD. *J Comput Chem* 2005;26:1781–802. <https://doi.org/10.1002/jcc.20289>.
- Sabanés Zariquiey F, de Souza JV, Bronowska AK. Cosolvent analysis toolkit (CAT): a robust hotspot identification platform for cosolvent simulations of proteins to expand the druggable proteome. *Sci Rep* 2019;9:19118. <https://doi.org/10.1038/s41598-019-55394-2>.
- Shao C, Westbrook JD, Lu C *et al.* Simplified quality assessment for small-molecule ligands in the protein data bank. *Structure* 2022;30: 252–62.e4. <https://doi.org/10.1016/j.str.2021.10.003>.
- Van Der Spoel D, Lindahl E, Hess B *et al.* GROMACS: fast, flexible, and free. *J Comput Chem* 2005;26:1701–18. <https://doi.org/10.1002/jcc.20291>.
- Vanommeslaeghe K, Hatcher E, Acharya C *et al.* CHARMM general force field: a force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J Comput Chem* 2009;31:671–90. <https://doi.org/10.1002/jcc.21367>.
- Wang J, Wolf RM, Caldwell JW *et al.* Development and testing of a general amber force field. *J Comput Chem* 2004;25:1157–74. <https://doi.org/10.1002/jcc.20035>.