**Electronic supplementary materials** *Mendeleev Commun.*, 2023, **33**, 225–227

# Identification of the precursor cluster in thermolysin crystallization solution by molecular dynamics methods

**Yuliya V. Kordonskaya, Vladimir I. Timofeev, Yulia A. Dyakova, Margarita A. Marchenkova, Yury V. Pisarevsky, Svetlana Yu. Silvestrova and Mikhail V. Kovalchuk**

## MATERIALS AND METHODS

*Preparation of hexamer models.* Molecular models of possible growth units of thermolysin crystals were isolated from the hexagonal crystal structures (PDB ID: 3DNZ). The thermolysin crystal belongs to space group $P6_122$ with unit cell parameters a = b = 92.535 Å, c = 128.628 Å, α = β = 90.00°, γ = 120.00°. Using PyMOL program [S1], the appropriate symmetry operators were applied to reconstruct a fragment of the protein crystal lattice and the coordinates of 4 hexamers (A–D) were obtained. The calcium and zinc ions associated with thermolysin (4 calcium ions and 1 zinc ion per protein molecule), were retained in the structure, and crystalline water was removed.

Calculation of the amino acid residues ionization states at pH 6.0 for thermolysin, (in accordance with the pH values of corresponding crystallization solutions) was performed using PROPKA server (Version 2.0.0 [S2]). The residues THR 49, SER 53, ASN 89, MET 120, TYR 157, SER 161 in 3DNZ PDB file exist in the crystal in two conformations with the same occupancy factors (0.5) but only one conformation was retained.

*Molecular dynamics protocol*. The calculations were performed via the GROMACS software package version 2021 [S3]. Molecular dynamics was modeled in the Amber ff99SB-ILDN field [S4], since it contains accurate torsion potentials.

Each hexamer was placed in the center of a cubic simulation box and three-dimensional periodic boundary conditions were applied. The dimensions of the boxes were set in such a way that the minimum distance between their facet and any protein atom was 1 nm. To fill the boxes with an explicit solvent, a 4-site water model designed to use Ewald summation methods (TIP4P-Ew [S5]) was chosen. In order to reproduce the crystallization conditions (solution with

the precipitant), water molecules were replaced by ammonium and sulfate ions so that the salt concentration in the box was 0.75M. Since there were no parameters for $NH_4^+$ and $SO_4^{2-}$ ions in the ff99SB-ILDN force field, the 3D structures of ammonium were taken from PDBeChem (code: NH4) while the sulfate 3D structure was obtained from the PLMD (Peptide Ligand Molecular Dynamics) python module and converted from .mol2 to .pdb format using the Antechamber algorithm [S6]. The total charge of each box was neutralized by adding a negligible number of chloride ions as it is required to use the PME algorithm for calculating long-range electrostatic interactions.

Before each start of the productive MD calculations, the energy of the systems was minimized by the steepest descent method (50,000 steps) until the force acting on any atom became less than 1000 kJ/(M·nm$^{-2}$) so that atomic overlaps existed in the crystal structures and produced by the addition of water were eliminated. Then, the boxed were thermostated for 100 ps by the modified Berendsen (V-rescale) method [S7] in NVT-ensemble and barostated for 100 ps by the Parrinello-Raman algorithm [S8] in NPT-ensemble.

The productive MD simulation was carried out in an isothermal-isobaric ensemble using the V-rescale thermostat ($T$ = 293 K) and a Parinello-Raman barostat ($P$ = 1 atm) again. Integration was performed using a standard leap-frog algorithm [S9] with the integration time step set at 2 fs. Noncovalent interactions were taken into account only for atoms located within a radius of 1 nm. The long-range electrostatic interactions were processed by the smooth particle mesh Ewald (PME) summation method [S10] with cubic interpolation and grid spacing in Fourier space of 0.16 nm. The bonds lengths of the hexamers were constrained using the LINCS algorithm [S11]. Each of the calculated trajectories lasts 100 ns. Other parameters of the productive MD simulations are listed in Table S1.

Before the trajectory analysis, the *gmx trjconv* command with the *−pbc nojump* flag was run to check whether atoms jump across the box and put them back so that all molecules remain whole. The reference protein structure for the structural alignment of the trajectories was taken from the file with the atomic positions being after equilibration. Molecules' trajectories were fitted to the reference ones by performing the command *gmx trjconv* with the *−fit rot+trans*. RMSF, RMSD and $R_g$ were computed by using the commands *gmx rmsf*, *gmx rms* and *gmx gyrate*, respectively.

**Table S1. General run parameters of the productive MD simulations.**

| Parameter | Value |
|---|---|
| integrator | md |
| nsteps | 50000000 |
| dt | 0.002 |
| constraint_algorithm | lincs |
| constraints | h-bonds |
| lincs_iter | 1 |
| lincs_order | 4 |
| cutoff-scheme | Verlet |
| ns_type | grid |
| nstlist | 10 |
| rcoulomb | 1.0 |
| rvdw | 1.0 |
| coulombtype | PME |
| pme_order | 4 |
| fourierspacing | 0.16 |
| tcoupl | V-rescale |
| tc-grps | Protein Non-Protein |
| tau_t | 0.1 |
| pcoupl | Parrinello-Rahman |
| pcoupltype | isotropic |
| tau_p | 2.0 |
| ref_p | 1.0 |
| pbc | xyz |
| DispCorr | EnerPres |

## REFERENCES


S1. Schrödinger, L. The PyMol Molecular Graphics System, Versión 1.8. *Thomas Holder* 2015.

S2. Dolinsky, T.J.; Nielsen, J.E.; McCammon, J.A.; Baker, N.A. PDB2PQR: An Automated Pipeline for the Setup of Poisson-Boltzmann Electrostatics Calculations. *Nucleic Acids Res* **2004**, doi:10.1093/nar/gkh381.

S3. van der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A.E.; Berendsen, H.J.C. GROMACS: Fast, Flexible, and Free. *J Comput Chem* 2005, *26*, 1701–1718.

S4. Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J.L.; Dror, R.O.; Shaw, D.E. Improved Side-Chain Torsion Potentials for the Amber Ff99SB Protein Force Field. *Proteins: Structure, Function and Bioinformatics* **2010**, *78*, 1950–1958, doi:10.1002/prot.22711.

S5. Horn, H.W.; Swope, W.C.; Pitera, J.W.; Madura, J.D.; Dick, T.J.; Hura, G.L.; Head-Gordon, T. Development of an Improved Four-Site Water Model for Biomolecular Simulations: TIP4P-Ew. *Journal of Chemical Physics* **2004**, *120*, 9665–9678, doi:10.1063/1.1683075.

S6. Sousa Da Silva, A.W.; Vranken, W.F. ACPYPE - AnteChamber PYthon Parser InterfacE. *BMC Res Notes* **2012**, *5*, doi:10.1186/1756-0500-5-367.

S7. Berendsen, H.J.C.; Postma, J.P.M.; van Gunsteren, W.F.; Dinola, A.; Haak, J.R. Molecular Dynamics with Coupling to an External Bath. *J Chem Phys* **1984**, *81*, 3684–3690, doi:10.1063/1.448118.

S8. Parrinello, M.; Rahman, A. Strain Fluctuations and Elastic Constants. *J Chem Phys* **1982**, *76*, 2662–2666, doi:10.1063/1.443248.

S9. van Gunsteren, W.F.; Berendsen, H.J.C. A Leap-Frog Algorithm for Stochastic Dynamics. *Mol Simul* **1988**, *1*, 173–185, doi:10.1080/08927028808080941.

S10. Essmann, U.; Perera, L.; Berkowitz, M.L.; Darden, T.; Lee, H.; Pedersen, L.G. A Smooth Particle Mesh Ewald Method. *J Chem Phys* **1995**, *103*, 8577–8592, doi:10.1063/1.470117.

S11. Hess, B.; Bekker, H.; Berendsen, H.J.C.; Fraaije, J.G.E.M. LINCS: A Linear Constraint Solver for Molecular Simulations. *J Comput Chem* **1997**, *18*, 1463–1472, doi:10.1002/(SICI)1096-987X(199709)18:12<1463::AID-JCC4>3.0.CO;2-H.