Co-translocational Unfolding of HP35 in MIL-101(Cr) MOF

Oishika Jash,[†] Anand Srivastava,[‡] and Sundaram Balasubramanian[*,†]

†*Chemistry and Physics of Materials Unit, Jawaharlal Nehru Centre for Advanced Scientific Research, Bangalore 560 064, India*

‡*Molecular Biophysics Unit, Indian Institute of Science, Bangalore, 560 012, India*

E-mail: bala@jncasr.ac.in

Abstract

The immobilization of enzymes in the cages of Metal-Organic Frameworks (MOF) is important in biotechnology. In this context, the mechanism of translocation of proteins through the cavities of the MOF and the roles played by confinement and MOF chemistry in giving rise to stable protein intermediates that are otherwise transiently populated in physiological environment are important questions to be addressed. Herein, these unexplored aspects are examined with Villin Headpiece (HP35) as a model protein confined within a mesopore of MIL-101(Cr) using molecular dynamics simulations. At equilibrium, the protein is located farther from the center of the cavity and closer to the MOF surface. Molecular interactions with the MOF leads to a partial unfolding of helix-1 at its N-terminus. Umbrella sampling simulations inform the range of conformations that HP35 undertakes during the translocation from one cavity to another and also provide the free energy differences at various stages of translocation. Relative to its equilibrium state within the cavity, the free energy barrier for the unfolded protein at the cage window is estimated to be 16 kcal/mol. This study of MOF-based protein conformation can also serve as a general approach to observing intermediates in folding-unfolding pathways.

Keywords Metal-Organic Frameworks (MOF), Protein immobilization, Co-translocational unfolding, Molecular Dynamics Simulations

1 Introduction

Misfolded and unfolded proteins inside a cell are refolded and sometimes degraded by enzymes and chaperones to maintain protein homeostasis.[1,2] In certain aberrant cases, these non-native protein intermediates are known to form pathologically toxic aggregates as in Alzheimer's and Parkinson's diseases.[3] Understanding the mechanism of these aggregate formations is often challenging experimentally because they are driven by conformations that are short-lived.[4] Knowledge about these short-lived intermediates could shed tremendous insights into the mechanisms leading to these pathological aggregates and other protein activities driven by non-native conformations. Molecular dynamics simulations on these systems can assist in addressing these problems by providing information on the structures and dynamics of these proteins.[5] Much like chaperones in a cellular context, non-protein matrices such as metal-organic framework (MOF) can be used as laboratory-reconstituted assays to trap protein intermediates in their folding-unfolding pathways. This reconstituted pipeline provides a tremendous opportunity to explore the area related to the immobilization of enzymes (proteins) and the functional implications of the process. This makes the immobilization of enzymes (proteins) a topic of vital theoretical interest.

Immobilized enzymes are important for biotechnological industries. In an industrial environment, exposure to various stresses such as high temperature, pressure, organic solvents, proteolytic, and chaotropic agents makes enzymes unstable apart from issues related to product separation and purification.[6] Approaches to mitigate such problems include the entrapment of enzymes in different materials starting from bulk materials,[7] nanoparticles,[8–10] hydrogels,[11] carbon nanotube,[12] DNA origami,[13] covalent organic frameworks,[14] hydrogen-bonded organic frameworks,[15] silica[16] and metal-organic frameworks (MOF) utilizing covalent and non-covalent interfacial interactions. Versatility in combining inorganic (metal nodes) and organic (linkers) chemistry make MOFs one of the most used host platforms.[6,15,17–22] Such confinement also allows one to study transient intermediates under confinement (as in MOFs

with porous cages). Due to the presence of multiple cavities, the MOF platform also allows one to study the translocation of proteins from one cavity to the neighboring one. Translocation phenomenon such as crossing the blood-brain barrier[23,24] has its importance in the spread of diseases along with the unfolding of proteins through biological nanopores. One of the possible ways to understand these phenomena in a reconstituted system is to study with a model protein.

One of the model systems extensively studied in protein folding-unfolding processes is the chicken villin headpiece (HP35). HP35, which is an actin-bundling protein, contains two domains: "core" and "headpiece". The headpiece is the F actin-binding domain lying in the C-TER of super villin.[25,26] HP35 is the 35-residue subdomain within the headpiece domain. It is one of the smallest monomeric polypeptides, made of naturally occurring amino acids that fold independently and autonomously into a unique and thermostable structure with a melting temperature of 342 K.[27–29] Subdomains that fold independently are important for studying protein folding. As shown by McKnight and co-workers, HP35 undergoes a cooperative thermal unfolding transition (unlike molten globules).[27] Like other small proteins that fold at sub-microsecond timescale, HP35 also shows low folding cooperativity and a low energy barrier.[30] It displays the properties of a fully folded protein with a unique structure (i.e. a clear secondary structure and a well-packed core). It has a three-helix (alpha helices) topology with a closely packed hydrophobic core involving three phenylalanine residues. This structural architecture was revealed by both NMR and XRD experiments.[29,31]

Although there are many experimental[32–43] and a few simulation studies[43–49] on protein-MOF composites, these have largely focused on studying interfacial interactions[42,45–48,50] and structural influences at the active site of the enzymes.[44,49] The current work, however, has a distinctive focus, i.e., to model protein translocation in such a system. In this work, where we model the co-translocational unfolding of HP35 in MIL-101(Cr), we explore (i)

how MOF confinement affects the structure of HP35 while it moves from one cavity to the neighboring one, (ii) how confined waters access the protein surface, (iii) and the free energy barrier for protein translocation from the center of the cavity towards the window that separates two neighboring cavities of the MOF.

The outline of the paper is as follows. Following this introduction, we describe the methodology in detail, which includes equilibrium MD simulations as well as Umbrella Sampling runs to obtain the free energy profile. In the Results and Discussion section, we provide detailed analyses of protein conformations and interactions with the MOF framework and also report on the free energy profile for protein translocation. We summarize our work in the Conclusions section. Notations that have been used in the following sections are as follows - center of mass of the protein = P_{COM}, center of mass of the cavity = C_{COM}, and reference NMR structure = S_{NMR} and the corresponding definitions are explained in the Supporting Information.

2 Materials and Methods

2.1 A hierarchically porous MOF: MIL-101(Cr)

MIL-101(Cr)[51] possesses a MTN zeotype architecture formed by corner-shared super tetrahedra (ST) comprising 1,4-benzenedicarboxylate (BDC) linkers and Cr atoms; each Cr atom has an octahedral environment consisting of four oxygen atoms from BDC, one μ_3-O atom, and one water molecule (Figure S1). The STs are microporous with an 8.6 Å window aperture. Through corner-sharing of STs, two types of mesoporous cages are formed with internal diameters of 29 Å and 34 Å, respectively. The former has pentagonal windows with a free aperture of 12 Å diameter while the latter cages are accessible through either pentagonal or hexagonal windows (14.5 x 16 Å free aperture). MIL-101 has been used for the infiltration of microperoxidase-8 (MP-8)[43] and *Aspergillus saitoi* proteinase[52] enzymes.

2.2 HP35 as a model system for co-translocational unfolding

HP-35 has radii of gyration of 6.2 Å x 7.4 Å x 8.3 Å along its principal axes and cannot translocate across the cavities of MIL-101(Cr) without unfolding due to the smaller dimensions of the intervening window (Figure S2). Hence it is a suitable model system to study the process of co-translocational unfolding.

2.3 Preparation of protein containing MOF supercell

The initial structure of the protein for the simulations was the one with the lowest energy among those solved via NMR experiments (S_{NMR}) from RCSB (PDB ID: 1UNC).[53] The primitive unit cell of the MOF crystal structure (OCUNAC_manual; CCDC 605510) was taken from a GitHub repository (`https://github.com/scidatasoft/mof,03.01.2023`), where partial charges of cleaned MOF structures from the CoRE MOF database[54,55] are provided. As the sixth coordination of the metal (Cr) was missing in the primitive unit cell, we added the water molecules manually to the secondary building unit (SBU) of the MOF using Gaussview (version: 5.0.9), and subsequently performed geometry optimization using periodic density functional theory (DFT) implemented in the QUICKSTEP module[56] of CP2K package (version: 7.1).[57] In this method, a linear combination of atom-centered Gaussian-type orbitals are used to describe the Kohn-Sham orbitals and the electron density is described in an auxiliary plane-wave basis set in conjunction with Goedecker-Teter-Hutter (GTH) pseudopotentials.[58] An accuracy of 10^{-7} was used for the self-consistent field (SCF) convergence of both inner and outer loops. PBE exchange-correlation functional[59] was used and dispersion corrections were incorporated using the DFT-D3 approach.[60] The geometry-optimized distance between the metal and the water was 2.30 Å (Figure S1). We added the water molecules at this distance to the primitive unit cell and replicated it to a supercell of

size 2x2x2. The OBGMX code[61] was used to generate the topology of the MOF supercell. Subsequently, solvent water molecules were added using the GROMACS solvation module to fill the supercell using a scaling factor of 0.467. The number of water molecules required to fill the supercell (hence the scaling factor) was approximated from the void volume of the supercell calculated using Biovia Materials Studio 2020.[62] The protein was packed at the center of a central big cavity in the equilibrated MOF supercell structure using PACKMOL (version: 20.2.2)[63] with a distance tolerance of 1 Å. This was followed by the removal of water molecules within 2.5 Å of the protein (Figure 1).

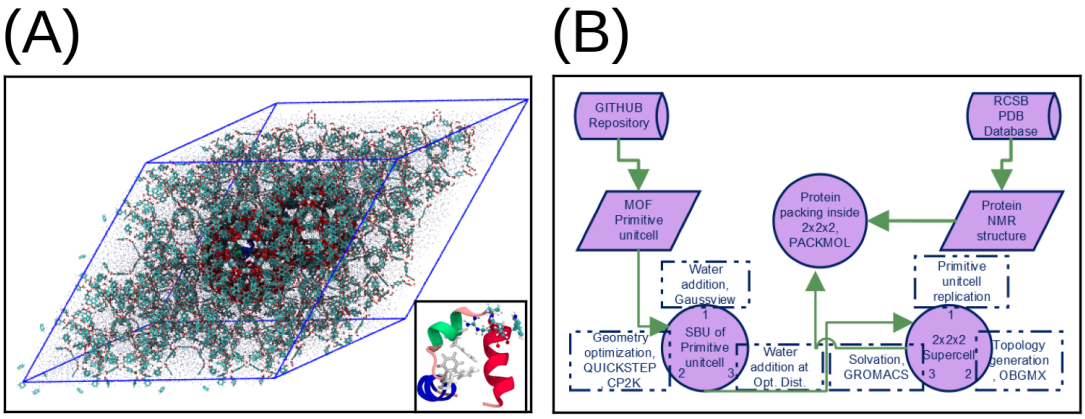


Figure 1: (A) Simulation box (supercell) of MIL-101(Cr) with HP35 present in one of the large cavities (highlighted in vdW representation along with the neighbouring cavity) having dimensions 12.568 nm x 12.568 nm x 12.568 nm and 60°cell angles. Water molecules are shown in CPK representation with reduced scale in Iceblue color. Counterion is not shown. Inset shows geometric motifs of HP35. The hydrophobic core (residues 6,10,17) are highlighted in White Licorice representation. The PXWK motif (residues 21-24) is highlighted in CPK representation. Helix-1 (residue 3-10), 2 (residue 15-19), and 3 (residue 22-33) are in Blue, Green, and Red, respectively. (B) Flowchart for system preparation.

2.4 General protocol for the simulations

All molecular dynamics simulations were performed using GROMACS 2022.3.[64–68] We employed the all-atom AMBER99sb-star-ildn[69–71] bonded and non-bonded parameters for the protein, UFF[72–74] bonded and non-bonded parameters (except partial charges) for the MOF and the rigid TIP3P model[75] for water molecules. Periodic boundary conditions (PBC) were applied in all the three directions. PBC for the MOF involved the inclusion of intramolecular potential terms (bond, angle, torsion) across minimum images. Partial charges for MOF atoms (except metal-ligated water) were taken from a repository (`https://github.com/scidatasoft/mof,03.01.2023`), wherein they were assigned using a machine learning model[76] which had combined the high accuracy of density-derived electrostatic and chemical charge (DDEC) method and the scalability of the charge equilibration (Qeq) method. TIP3P charges were assigned to the metal-ligated water. For non-bonded interactions, real space cutoffs for the Lennard-Jones (LJ) and Coulomb interactions were set to 10 Å. Lorentz-Berthelot mixing rules were used to obtain the LJ parameters between two different atom types. The particle Mesh Ewald (PME)[77] method with an interpolation order of 4 and a relative tolerance of 10^{-5} was used to calculate the electrostatic interactions for distances above 10 Å. Scaling factors for 1-4 non-bonded interactions were set following the AMBER force field. Long range dispersion corrections were applied for energy and pressure.

Energy minimization was done using the steepest-descent algorithm with an initial step size of 0.01 nm and force tolerance of 10 $kJmol^{-1}nm^{-1}$. Simulations were performed in the isochoric-isothermal ensemble (NVT). For temperature coupling, Bussi-Donadio-Parrinello velocity-rescaling thermostat[78] with a coupling constant of 0.5 ps was used. The temperature was set to 298 K. Covalent bonds to hydrogen atoms were constrained using the LINCS algorithm[79,80] with order 4 and warn angle 30°. During equilibration, heavy atoms were position-restrained with a force constant of 10^3 $kJmol^{-1}nm^{-2}$. Equations of motion were integrated using the leap-frog algorithm with a time step of 1.0 fs.

For equilibrium MD simulations, the following procedures were followed. Prior to the insertion of the protein into the MOF cavity, the MOF supercell (cavities filled with water molecules) was equilibrated. The steps followed were (i) energy minimization for solvent water, (ii) NVT annealing from 0 K to 298 K over 2 ns followed by equilibration at 298K for 1 ns. (iii) A short production run for 5 ns under constant NVT conditions followed. For the equilibration of the protein@MOF system (the complete system under study), an intermediate step of energy minimization of the protein alone was carried out. During energy minimization of the protein, water was position restrained and vice versa. Non-hydrogen atoms of the MOF were position restrained in all equilibration steps. So, the general protocol was: (i) energy minimization of water, (ii) energy minimization of protein (iii) and NVT equilibration.

2.5 Prescription for performing translocation experiment in MOF

Initial configurations of the protein for translocation through the hexagonal window of the MOF were generated using Steered Molecular Dynamics (SMD) simulations. As a starting point, HP35 was placed at the center of one of the big cavities of the supercell. During equilibration, a harmonic potential was applied between the P_{COM} and C_{COM} along with a restraining potential for 'MOF and protein', 'MOF and water' and MOF alone, respectively in the first three stages of the general protocol (described earlier). This was followed by an additional step of NVT run at 298 K removing position restraints on the MOF. During the SMD runs, the non-hydrogen atoms of MOF were position-restrained. In these runs, a harmonic spring with a force constant of 10^5 kJmol^{-1}nm^{-2} attached to the P_{COM}, was pulled with a speed of 0.02 nm/ns along a defined (direction) unit vector towards the hexagonal window. This direction vector was obtained as the cross product of two edges (vectors) of the hexagonal window. 15 SMD runs, each initiated with a different seed for generating

Table 1: Details of Simulation of HP35 in MIL-101(Cr) water system

System name	Run type	# of MOF atoms	# Protein atoms	# of Water molecules	# of Ions	Total atoms	Production run
MOF	NVT	33184	0	36935	0	143989	5 ns
Protein-MOF	Equilibrium (NVT)	33184	574	36629	1	143646	680 ns
Protein-MOF	Umbrella Sampling	33184	574	36629	1	143646	Table S1

initial random velocities of the atoms, were generated. The work profiles from each were examined and the one displaying the lowest work value was chosen as the putative 'path' for the Umbrella Sampling runs.

Taking the initial positions and coordinates of the protein from the chosen SMD run, 54 umbrella windows that covered the distance from the center of the cavity to the hexagonal window (that connects two adjacent cavities of the MOF) were identified. A harmonic potential was applied to the distance between the C_{COM} and P_{COM}, with a force constant of 10^5 kJmol^{-1}nm^{-2} at each of these positions. The non-hydrogen atoms of the MOF were position restrained for all the windows during umbrella sampling simulations with a force constant of 10^3 kJmol^{-1}nm^{-2}. The run lengths in different umbrella windows are presented in Table S1. The cumulative MD run length over all the Umbrella Sampling windows amounts to 6.54 μs.

At each window, the run length was chosen such that the amplitude of fluctuation of the backbone RMSD of the protein lied between 0.5 to 1.0 Å around a mean value for a duration of at least 5 ns. Thus, the last 5 ns of the molecular dynamics trajectory for each umbrella window was used for further analysis. This procedure ensured that the PMF value at each window was obtained from a converged orientation and conformation of the protein. The free energy surface was reconstructed by reweighting configurations from umbrella windows using the weighted histogram analysis method (WHAM)[81] implemented within gmx_mpi wham code. The number of bins for the histogram was 200 and for estimating error in the potential of mean force (PMF) or the free energy profile, Bayesian bootstrapping was carried out with 200 bootstraps. Table 1 presents a summary of the simulations performed.

3 Results and Discussion

3.1 HP35 is located near the surface of the MOF cavity and not at its center

Although this remains to be verified, an enzyme within the MOF is invariably assumed to be located at the centre of the MOF cavity.[82,83] To understand the behavior of HP35@MIL-101(Cr), we first performed equilibrium MD simulations at 298 K. The initial location of the protein was with P_{COM} at the centre of one of the larger cavities (Figure 2A). Even during the equilibration stage (Figure 2B inset), the protein moved around 4 Å from this location. Subsequently, this distance, $d_{MOF-Protein}$, reached a value of 7.5 Å in about 400 ns (Figure 2B), and maintained the same for a further duration of 280 ns. This observation demonstrates that the equilibrium position of the protein is not the center of the cavity but closer to the cavity surface of the MOF as seen from Figure 2C.

Our analyses show that HP35 broadly maintains its native structure upon confinement. The root mean squared deviation of the positions of backbone atoms (N, CA, C) of the protein relative to S_{NMR} remains largely (77% of the complete trajectory) within 3 Å over 300 ns (see the 350-650 ns segment of Figure 2D). Marginal differences in the conformation of helix-1 of HP35@MOF with that of the reference structure (S_{NMR}) are seen (2F). Inter-molecular interactions between the MOF sites and HP35 were examined to identify the origin of the differences.

One of the ways to quantify the extent of van der Waals interactions between the protein and the MOF is to count the number of heavy atom-heavy atom contacts (i.e., non-hydrogen) that lie within a distance of 4 Å. This quantity too increased over the length of the equi-

librium trajectory and saturated to a value of around 40 (Figure 2E inset). Furthermore, as seen from Figure 2E, residues of helix-1 and helix-2 interacted more often with the MOF than those of helix-3, as also those in the non-helical region (N-TER, residues joining helix1 and helix2, and residues joining helix-2 and helix-3). Within helix-1 and helix-2, no difference in interaction with the MOF between polar and non-polar residues was observed. An overlay of the backbone atoms (N, CA, and C) of the protein in the last time frame of the production run with S_{NMR} is presented in Figure 2F. Along with marginal reduction in the alignment of helix-1, a different orientation for the three F residues of the hydrophobic core is seen; however, the overall secondary structures and, in particular, helix-2 and helix-3 align well.

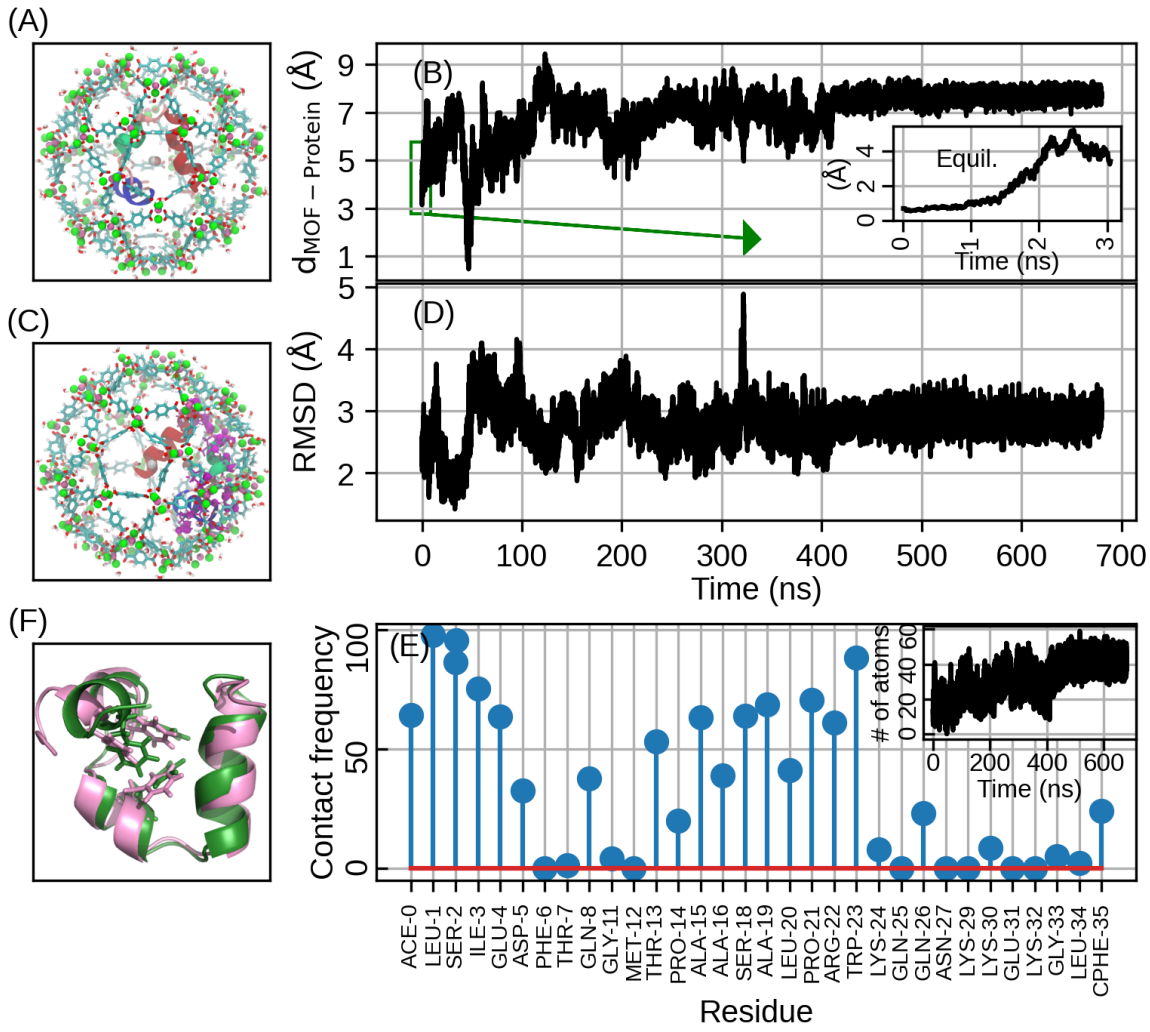


Figure 2: Results from the equilibrium MD simulation. (A) HP35 in its initial configuration inside one of the MOF cavities. HP35 is shown in the New Cartoon representation with helix-1, helix-2, and helix-3 in Blue, Green, and Red, respectively. MOF: Cr atoms and μ3-Oxygens in Green and Mauve with vdW representation with reduced scale, metal-ligated waters, and organic ligands in Licorice representation. Water molecules that fill the cavity in the simulations are not shown for clarity. (B) Distance between the protein center of mass (COM) and the cavity center of the MOF. The same but during equilibration stage is in inset. (C) Same as in (A) but for the last time frame of the equilibrium MD run. (D) Fraction of time spent by a residue (non-hydrogen atoms) within 4 Å of any MOF non-hydrogen atom. Inset: Total number of protein-MOF atom contacts vs simulation time. (E) Backbone RMSD with respect to solution NMR structure. (F) Overlay of NMR structure of the protein (Green) and that of the last time frame of the protein@MOF run. (Backbone atoms N, CA, C, and O have been used for alignment). The hydrophobic core is highlighted in Licorice representation.

3.2 Existence of a "constriction region" in the protein translocation pathway

In order to understand the conformational heterogeneity of the protein inside the MOF cavity during translocation, we examined various geometrical parameters of the protein along its path of migration from one cavity to the other. The progress of the protein on this path is captured herein from MD trajectories through the sequence of Umbrella Sampling windows. Figure S5 provides milestones on this path, which are used in the present discussion. The window index increases as the protein moves from the center of the cavity towards the hexagonal aperture that connects two neighboring MOF cavities. Here, we describe the conformations adopted by HP35 during its translocation.

For a few central umbrella windows (windows 18-29), the root mean square deviation of protein backbone from S_{NMR} (Figure 3) lies between 2-3.5 Å. Here, the protein explores its conformational space around its native state. The same is reflected in the plateau region of the helicity plot (Figure 3) wherein the alpha-helical content of the protein was the highest and has nearly the same value as in its native state, S_{NMR} (Blue horizontal dotted line). The medians of solvent accessible surface area of the protein (Figure 4) remain close to the two specific SASA values (3056 and 3109 $Å^2$) corresponding to two folded conformations of HP35 in neat water reported in an earlier study.[84] The latter was described as a "dry molten globule".[85] These two values are labelled in the SASA plot (Figure 4) with dashdotted and dotted line in Blue, respectively. This central zone is termed as the "constriction region" in our study.

Before the constriction region, i.e., closer to the center of the cavity (windows 1-17), the helicity of the protein is reduced (Figure 3). The median values of the protein SASA displays an undulated profile in this zone (Figure 4). Further, the deviation of the backbone atoms (RMSD) from S_{NMR} is within 2-5 Å (Figure 3). On the other hand, the number of non-

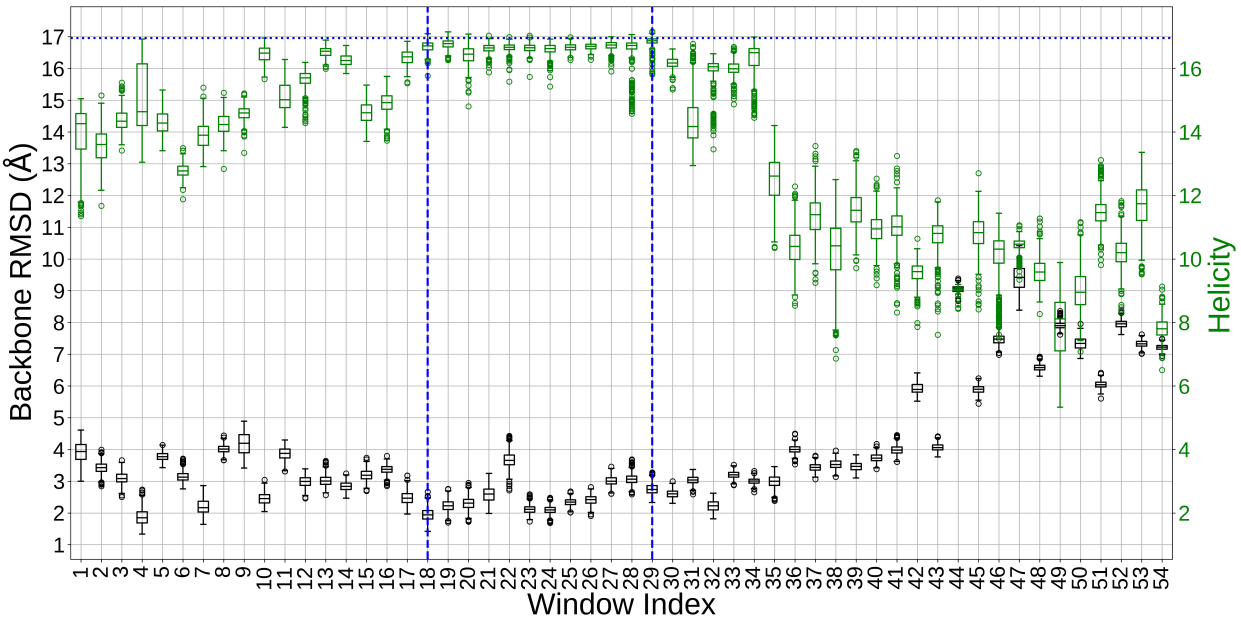


Figure 3: Black: RMSD of protein backbone with respect to S_{NMR} across Umbrella Sampling windows. Green: Alpha Helical content across windows. The horizontal dotted line in Blue is the value of helicity of S_{NMR}. The vertical dotted lines in Blue enclose the constriction region. The five number summary statistic, namely Box Plot[86–89] has been used for representation of the distribution of different quantities across umbrella sampling windows.

hydrogen contacts between the protein and the MOF (Figure 4) displays a steady increase. As the protein is farther from the MOF surface in this zone, the former is more accessible to water molecules. Hence, both the facts – (i) the spontaneous tendency of the protein to have van der Waals contact with the cavity surface of the MOF and (ii) interaction of the confined water molecules with the surface of the protein results in a zig-zag traversal path of the protein (Figure 5(A)). However, the protein did remain intact in its native state, as the medians of SASA of the hydrophobic core (Figure S10) were close to that of the S_{NMR}.

Beyond the constriction region, i.e., towards the hexagonal window (windows 30-54), the helicity of the protein decreased drastically, while the RMSD of the backbone atoms as well as the SASA of the protein increased. The non-hydrogen contacts between the protein and the MOF atoms reached the highest value at the hexagonal window. In this region, there existed a sub-zone (approximately, windows 30-41) where the RMSD of the backbone atoms

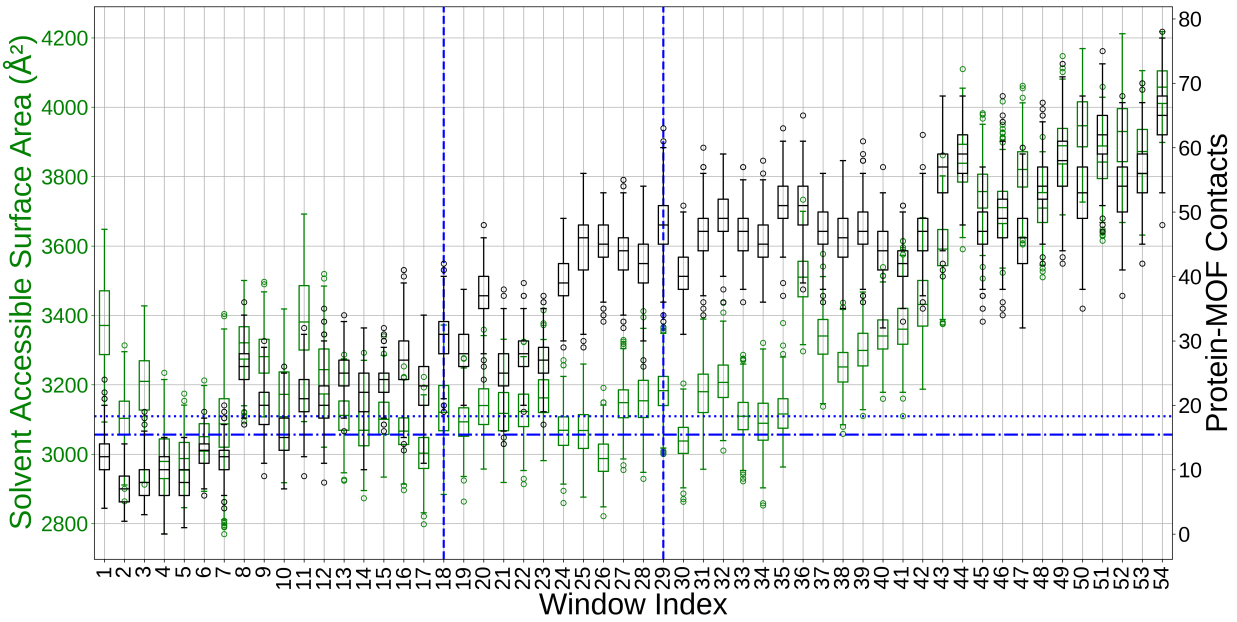


Figure 4: Black: Number of Protein-MOF non-hydrogen atom contacts (with a cutoff of 4 Å) across Umbrella Sampling windows. Solvent accessible surface area of the protein across windows is shown in Green (for the horizontal dashdotted and dotted lines in Blue, see text). The vertical dotted lines in Blue enclose the constriction region. The five number summary statistic, namely Box Plot [86–89] has been used for representation of the distribution of different quantities across umbrella sampling windows.

lay within 4 Å; the number of protein-MOF contacts was steady, while the helicity reduced and SASA increased. This region marks the initiation of protein unfolding due to its translocation. The unfolding exposes the hydrophobic core which interacts with the non-polar linker groups of the MOF. Beyond this sub-zone, the protein accesses more extended structures, driven by the favorable van der Waals contacts with the MOF surface (SASA of the protein accessing values greater than 3700 Å^2).

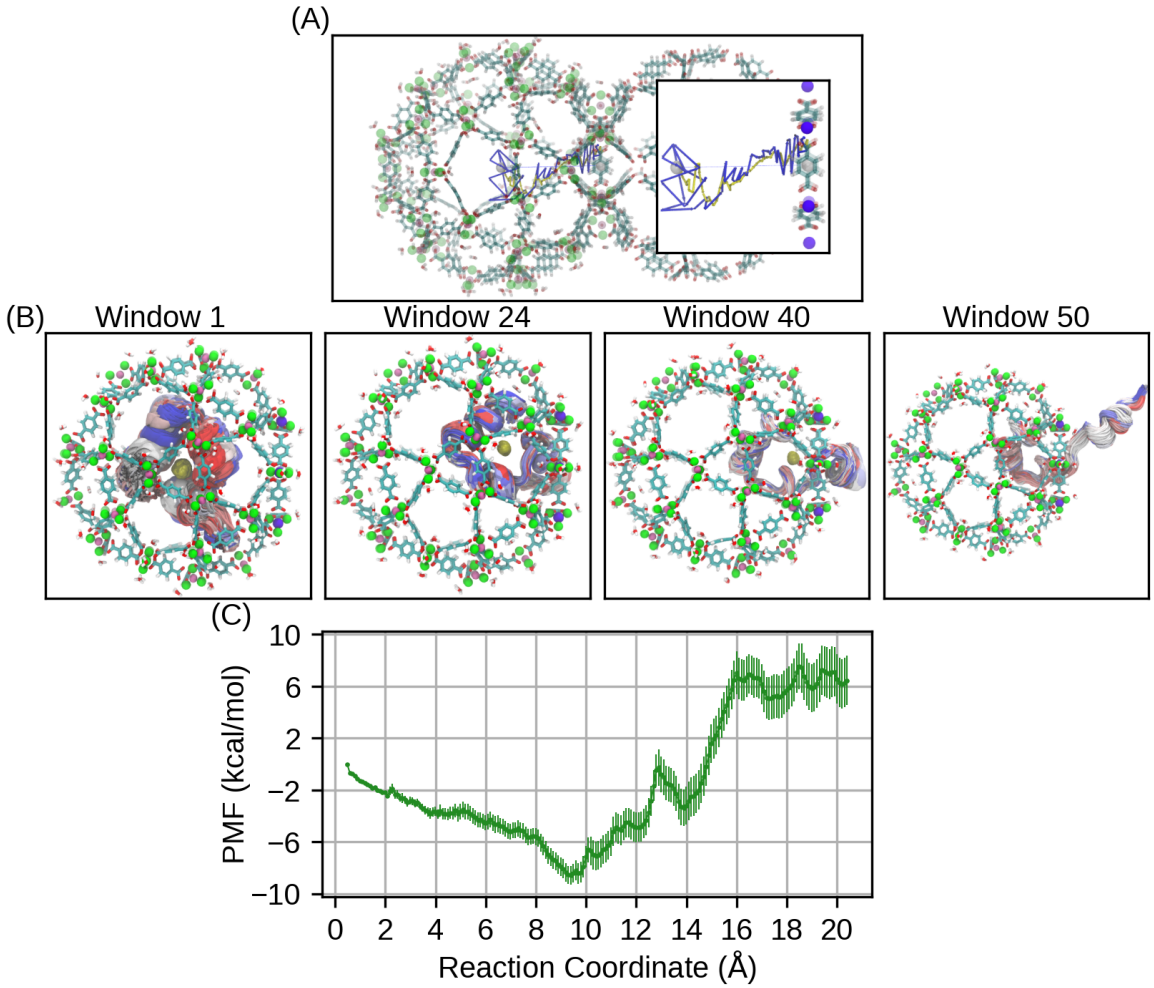


Figure 5: (A) Zig-zag excursions of the protein shown as Blue arrows obtained by concatenating the last time frame of all Umbrella Sampling windows. It provides a glimpse of the reaction coordinate. The path traversed during the SMD run (one-way trip) is shown with Yellow arrows. Inset: Zoom-in image of the path. Water molecules and ion are not shown for clarity.[90] (B) New Cartoon representation of the secondary structure of HP35 in a few Umbrella Sampling windows during its translocation. Red to Blue show structures with increasing time. (C) Free Energy profile for the translocation of HP35 from the center of the cavity of MIL-101(Cr) to the hexagonal aperture connecting the neighboring cavity.

3.3 Potential of mean force reveals that unfolding of HP35 during translocation is regulated by both cage geometry and confined waters

The free energy profile of the system during the translocation process obtained through umbrella sampling simulations shows an interesting behavior. Here, the reaction coordinate is the distance between P_{COM} and C_{COM}. The potential of mean force (PMF) displays a minimum at a distance of around 9.5 Å, which lies within the constriction region (approximately 24^{th} window in umbrella sampling). In this region, the conformations of the protein are influenced both by solvent water molecules as well as by the atoms of the MOF framework. An optimal arrangement leads to the minimum in the PMF, where the free energy value with respect to the protein at the cavity center is around -9 kcal/mol. Notably, the distance from the cavity center where the PMF minimum is observed matches that obtained from the equilibrium MD simulations and thus serves as an internal consistency check between these two categories of MD runs. The PMF attains its highest value at the hexagonal window where HP35 is in the most extended conformation. With respect to its equilibrium position, the free energy barrier for the co-translocational unfolding of the protein is estimated to be 16 kcal/mol (Figure 5).

4 Conclusions

While the translocations of ions, molecules, and macromolecules across lipid membranes have been studied using MD simulations,[91] that of a protein (or an enzyme) through a porous inorganic host such as a metal-organic framework has not been examined so far through computational methods, despite the vast amount of experimental reports on enzyme@MOF as functional biocatalytic platforms. The current work addresses this problem and provides considerable insights into the changes in the secondary structure of the protein as well as

that in the free energy along the transport path. This ensemble of structures of the HP35 protein and their interactions with the surface of the MIL-101(Cr) MOF cavity allows us to draw a possible general mechanism for co-translocational unfolding. As shown in the schematic (Figure 6), the protein (enzyme) closely resembles its native state when it is present somewhere between the cavity center and the hexagonal window of the MOF (label "2" in Figure 6). This key result observed in our equilibrium MD simulations is further confirmed by the presence of a free energy minimum away from the cavity center as seen in the umbrella sampling MD runs. At least concerning HP35@MIL-101(Cr), the center of the MOF cavity is *not* the equilibrium position for the protein. While the general applicability of this observation needs to be verified, it is likely to be followed by proteins that are smaller than the size of the cavity; after all, the macromolecule would prefer to interact with the atoms of the MOF, if possible without unfolding, via dispersion interactions.

When restrained to stay at the center of the cavity, HP35 undergoes marginal changes in its conformation towards partially unfolded structures in such a manner as to interact with the MOF surface. While translocating through the aperture connecting two cavities, the protein is present in an extended form stabilized by van der Waals contacts with the surfaces of both the MOF cavities. Since the initial configurations for the umbrella sampling simulations performed here were chosen from one of the SMD runs in which helix-1 of the protein translocated through the hexagonal window first, the same is seen in the US runs as well. However, in other SMD trajectories (not reported here), we observed no preference for any specific helix to cross through the hexagonal window of the MOF first.

The confinement of HP35 in the MOF, enabled conformations of the protein that were different from those observed in neat liquid water. This was reflected in the SASA (Figure 4). Thus, the solvent water confined inside the MOF could access the protein surface to a greater extent which was in turn made possible through increased protein-MOF direct

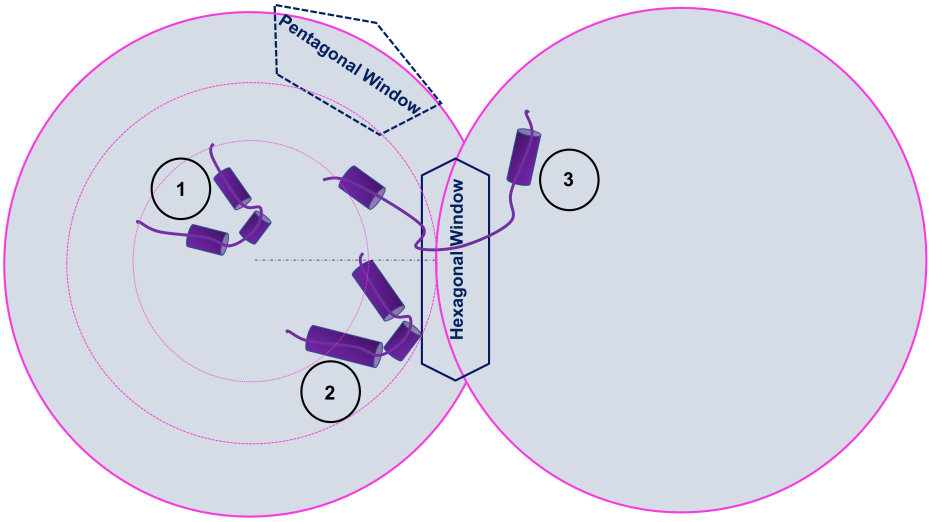


Figure 6: A schematic for co-translocational unfolding of protein inside MOF pores. In the context of HP35@MIL-101(Cr), when the protein is near the center of the cavity ("1" in figure), its native structure is marginally affected due to the competing interactions with the internal surface of the cavity. Near the aperture joining neighbouring cavities, the protein accesses extended structures being partitioned between the two cavities as represented by label "3". In between these two regions (labeled "2" in figure), the protein maintains its native ensemble of structures fairly well in what we have defined as the "constriction region".

contact. MIL-101(Cr) lacks groups that can hydrogen bond with the polar side chains of the HP35; thus, van der Waals is the dominant interaction type between the MOF and the protein. This interaction energy displayed a monotonic increase in magnitude (Figure S12) as the protein approached the hexagonal window during its translocation unlike the Coulombic interaction (Figure S11). Our observation that van der Waals is the major interaction in the HP35@MIL-101(Cr) system aligns with earlier studies of biomolecules confined in MOF channels.[46,47]

The free energy barrier for the co-translocational unfolding of HP35 across the hexagonal window of two neighboring cavities of MIL-101(Cr) MOF is estimated to be 16 kcal/mol at ambient conditions. The study enabled us to examine partially unfolded structures of the protein as well. Simulations of other proteins with different topologies in such porous hybrid materials could further our understanding of unfolded intermediates present under

confinement mimicking living cell milieu.

Supporting Information

The Supporting Information contains definitions and abbreviations being used in the paper, software used, and results from steered molecular dynamics (not shown in manuscript), and additional results from umbrella sampling along with with a Table of run lengths for umbrella sampling windows. The movie us.mp4 displays the translocation of HP35 through the hexagonal window of MIL-101(Cr).

Author Information

Corresponding Author

*E-mail: bala@jncasr.ac.in. Phone: +9180 22082808.

ORCID

Oishika Jash: 0009-0001-8465-7311

Anand Srivastava: 0000-0002-2757-1511

Sundaram Balasubramanian: 0000-0002-3355-6764

Notes

The authors declare no competing financial interest. The input files of our simulations are available in the GitHub repository: https://github.com/Oishika-1/CoTranslocationalUnfolding.

Acknowledgements

OJ thanks the Council of Scientific & Industrial Research (CSIR), India for financial support (Award Number: 09/733(0252)/2018-EMR-I). OJ thanks Giovanni Garberoglio for sending the static executable of OBGMX along with some basic instructions by email prior to them being made available through GitHub. OJ thanks Nimish Dwarkanath for help in setting up the MOF supercell structure. We are grateful for the support and resources provided by the "PARAM Yukti" facility under the National Supercomputing Mission, Government of India, at the Jawaharlal Nehru Centre For Advanced Scientific Research.

References

(1) Balch, W. E.; Morimoto, R. I.; Dillin, A.; Kelly, J. W. Adapting proteostasis for disease intervention. *Science* **2008**, *319*, 916–919.

(2) Cox, D.; Carver, J. A.; Ecroyd, H. Preventing α-synuclein aggregation: the role of the small heat-shock molecular chaperone proteins. *Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease* **2014**, *1842*, 1830–1843.

(3) Zlokovic, B. V. Neurovascular mechanisms of Alzheimer's neurodegeneration. *Trends in Neurosciences* **2005**, *28*, 202–208.

(4) Dobson, C. M.; Evans, P. A. Trapping folding intermediates. *Nature* **1988**, *335*, 666–667.

(5) Neudecker, P.; Robustelli, P.; Cavalli, A.; Walsh, P.; Lundström, P.; Zarrine-Afsar, A.; Sharpe, S.; Vendruscolo, M.; Kay, L. E. Structure of an intermediate state in protein folding and aggregation. *Science* **2012**, *336*, 362–366.

(6) Liang, W.; Wied, P.; Carraro, F.; Sumby, C. J.; Nidetzky, B.; Tsung, C.-K.; Falcaro, P.;

Doonan, C. J. Metal–organic framework-based enzyme biocomposites. *Chemical Reviews* **2021**, *121*, 1077–1129.

(7) Sheldon, R. A. Enzyme immobilization: the quest for optimum performance. *Advanced Synthesis & Catalysis* **2007**, *349*, 1289–1307.

(8) Kouassi, G. K.; Irudayaraj, J.; McCarty, G. Examination of cholesterol oxidase attachment to magnetic nanoparticles. *Journal of Nanobiotechnology* **2005**, *3*, 1–9.

(9) Zhang, F.; Cho, S. S.; Yang, S. H.; Seo, S. S.; Cha, G. S.; Nam, H. Gold Nanoparticle-Based Mediatorless Biosensor Prepared on Microporous Electrode. *Electroanalysis: An International Journal Devoted to Fundamental and Practical Aspects of Electroanalysis* **2006**, *18*, 217–222.

(10) Ansari, S. A.; Husain, Q. Potential applications of enzymes immobilized on/in nano materials: A review. *Biotechnology Advances* **2012**, *30*, 512–523.

(11) Lee, E. S.; Kwon, M. J.; Lee, H.; Kim, J. J. Stabilization of protein encapsulated in poly (lactide-co-glycolide) microspheres by novel viscous S/W/O/W method. *International Journal of Pharmaceutics* **2007**, *331*, 27–37.

(12) Feng, W.; Ji, P. Enzymes immobilized on carbon nanotubes. *Biotechnology Advances* **2011**, *29*, 889–895.

(13) Kong, G.; Xiong, M.; Liu, L.; Hu, L.; Meng, H.-M.; Ke, G.; Zhang, X.-B.; Tan, W. DNA origami-based protein networks: From basic construction to emerging applications. *Chemical Society Reviews* **2021**, *50*, 1846–1873.

(14) Li, M.; Qiao, S.; Zheng, Y.; Andaloussi, Y. H.; Li, X.; Zhang, Z.; Li, A.; Cheng, P.; Ma, S.; Chen, Y. Fabricating covalent organic framework capsules with commodious microenvironment for enzymes. *Journal of the American Chemical Society* **2020**, *142*, 6675–6681.

(15) Wied, P.; Carraro, F.; Bolivar, J. M.; Doonan, C. J.; Falcaro, P.; Nidetzky, B. Combining a Genetically Engineered Oxidase with Hydrogen-Bonded Organic Frameworks (HOFs) for Highly Efficient Biocomposites. *Angewandte Chemie International Edition* **2022**, *61*, e202117345.

(16) Magner, E. Immobilisation of enzymes on mesoporous silicate materials. *Chemical Society Reviews* **2013**, *42*, 6213–6222.

(17) Furukawa, H.; Cordova, K. E.; O'Keeffe, M.; Yaghi, O. M. The chemistry and applications of metal-organic frameworks. *Science* **2013**, *341*, 1230444.

(18) Zhang, X.; Chen, Z.; Liu, X.; Hanna, S. L.; Wang, X.; Taheri-Ledari, R.; Maleki, A.; Li, P.; Farha, O. K. A historical overview of the activation and porosity of metal–organic frameworks. *Chemical Society Reviews* **2020**, *49*, 7406–7427.

(19) Yaghi, O. M. The reticular chemist. *Nano Letters* **2020**, *20*, 8432–8434.

(20) Velásquez-Hernández, M. d. J.; Linares-Moreau, M.; Astria, E.; Carraro, F.; Alyami, M. Z.; Khashab, N. M.; Sumby, C. J.; Doonan, C. J.; Falcaro, P. Towards applications of bioentities@ MOFs in biomedicine. *Coordination Chemistry Reviews* **2021**, *429*, 213651.

(21) Chen, Z.; Kirlikovali, K. O.; Li, P.; Farha, O. K. Reticular chemistry for highly porous metal–organic frameworks: The chemistry and applications. *Accounts of Chemical Research* **2022**, *55*, 579–591.

(22) Wang, X.; He, L.; Sumner, J.; Qian, S.; Zhang, Q.; O'Neill, H.; Mao, Y.; Chen, C.; Al-Enizi, A. M.; Nafady, A.; et al. Spatially confined protein assembly in hierarchical mesoporous metal-organic framework. *Nature Communications* **2023**, *14*, 973.

(23) Wu, D.; Chen, Q.; Chen, X.; Han, F.; Chen, Z.; Wang, Y. The blood–brain barrier:

structure, regulation, and drug delivery. *Signal Transduction and Targeted Therapy* **2023**, *8*, 217.

(24) Theodorakis, P. E.; Müller, E. A.; Craster, R. V.; Matar, O. K. Physical insights into the blood–brain barrier translocation mechanisms. *Physical Biology* **2017**, *14*, 041001.

(25) Tang, Y.; Grey, M. J.; McKnight, J.; Palmer III, A. G.; Raleigh, D. P. Multistate folding of the villin headpiece domain. *Journal of Molecular Biology* **2006**, *355*, 1066–1077.

(26) Vardar, D.; Chishti, A.; Frank, B.; Luna, E. J.; Noegel, A.; Oh, S. W.; Schleicher, M.; McKnight, C. Villin-type headpiece domains show a wide range of F-actin-binding affinities. *Cell motility and the cytoskeleton* **2002**, *52*, 9–21.

(27) McKnight, J. C.; Doering, D. S.; Matsudaira, P. T.; Kim, P. S. A thermostable 35-residue subdomain within villin headpiece. *Journal of Molecular Biology* **1996**, *260*, 126–134.

(28) Kubelka, J.; Eaton, W. A.; Hofrichter, J. Experimental Tests of Villin Subdomain Folding Simulations. *Journal of Molecular Biology* **2003**, *329*, 625–630.

(29) McKnight, C. J.; Matsudaira, P. T.; Kim, P. S. NMR structure of the 35-residue villin headpiece subdomain. *Nature Structural Biology* **1997**, *4*, 180–184.

(30) Muñoz, V. Conformational dynamics and ensembles in protein folding. *Annu. Rev. Biophys. Biomol. Struct.* **2007**, *36*, 395–412.

(31) Chiu, T. K.; Kubelka, J.; Herbst-Irmer, R.; Eaton, W. A.; Hofrichter, J.; Davies, D. R. High-resolution x-ray crystal structures of the villin headpiece subdomain, an ultrafast folding protein. *Proceedings of the National Academy of Sciences* **2005**, *102*, 7517–7522.

(32) Pisklak, T. J.; Macías, M.; Coutinho, D. H.; Huang, R. S.; Balkus, K. J. Hybrid materials for immobilization of MP-11 catalyst. *Topics in Catalysis* **2006**, *38*, 269–278.

(33) Liang, K.; Ricco, R.; Doherty, C. M.; Styles, M. J.; Bell, S.; Kirby, N.; Mudie, S.; Haylock, D.; Hill, A. J.; Doonan, C. J.; others Biomimetic mineralization of metal-organic frameworks as protective coatings for biomacromolecules. *Nature Communications* **2015**, *6*, 7240.

(34) Zhu, W.; Guo, J.; Amini, S.; Ju, Y.; Agola, J. O.; Zimpel, A.; Shang, J.; Noureddine, A.; Caruso, F.; Wuttke, S.; others SupraCells: living mammalian cells protected within functional modular nanoparticle-based exoskeletons. *Advanced Materials* **2019**, *31*, 1900545.

(35) He, C.; Lu, K.; Liu, D.; Lin, W. Nanoscale metal–organic frameworks for the co-delivery of cisplatin and pooled siRNAs to enhance therapeutic efficacy in drug-resistant ovarian cancer cells. *Journal of the American Chemical Society* **2014**, *136*, 5181–5184.

(36) Chen, Y.; Han, S.; Li, X.; Zhang, Z.; Ma, S. Why does enzyme not leach from metal–organic frameworks (MOFs)? Unveiling the interactions between an enzyme molecule and a MOF. *Inorganic Chemistry* **2014**, *53*, 10006–10008.

(37) Chen, Y.; Lykourinou, V.; Hoang, T.; Ming, L.-J.; Ma, S. Size-selective biocatalysis of myoglobin immobilized into a mesoporous metal–organic framework with hierarchical pore sizes. *Inorganic Chemistry* **2012**, *51*, 9156–9158.

(38) Deng, H.; Grunder, S.; Cordova, K. E.; Valente, C.; Furukawa, H.; Hmadeh, M.; Gándara, F.; Whalley, A. C.; Liu, Z.; Asahina, S.; others Large-pore apertures in a series of metal-organic frameworks. *Science* **2012**, *336*, 1018–1023.

(39) Chen, Y.; Lykourinou, V.; Vetromile, C.; Hoang, T.; Ming, L.-J.; Larsen, R. W.; Ma, S. How can proteins enter the interior of a MOF? Investigation of cytochrome c translocation into a MOF consisting of mesoporous cages with microporous windows. *Journal of the American Chemical Society* **2012**, *134*, 13188–13191.

(40) Lian, X.; Chen, Y.-P.; Liu, T.-F.; Zhou, H.-C. Coupling two enzymes into a tandem nanoreactor utilizing a hierarchically structured MOF. *Chemical Science* **2016**, *7*, 6969–6973.

(41) Li, P.; Moon, S.-Y.; Guelta, M. A.; Lin, L.; Gómez-Gualdrón, D. A.; Snurr, R. Q.; Harvey, S. P.; Hupp, J. T.; Farha, O. K. Nanosizing a metal–organic framework enzyme carrier for accelerating nerve agent hydrolysis. *ACS Nano* **2016**, *10*, 9174–9182.

(42) Tai, T.-Y.; Sha, F.; Wang, X.; Wang, X.; Ma, K.; Kirlikovali, K. O.; Su, S.; Islamoglu, T.; Kato, S.; Farha, O. K. Leveraging Isothermal Titration Calorimetry to Explore Structure–Property Relationships of Protein Immobilization in Metal–Organic Frameworks. *Angewandte Chemie Int. Ed.* **2022**, *134*, e202209110.

(43) Gkaniatsou, E.; Sicard, C.; Ricoux, R.; Benahmed, L.; Bourdreux, F.; Zhang, Q.; Serre, C.; Mahy, J.-P.; Steunou, N. Enzyme encapsulation in mesoporous metal–organic frameworks for selective biodegradation of harmful dye molecules. *Angewandte Chemie Int. Ed.* **2018**, *130*, 16373–16378.

(44) Liang, J.; Bin Zulkifli, M. Y.; Yong, J.; Du, Z.; Ao, Z.; Rawal, A.; Scott, J. A.; Harmer, J. R.; Wang, J.; Liang, K. Locking the ultrasound-induced active conformation of metalloenzymes in metal–organic frameworks. *Journal of the American Chemical Society* **2022**, *144*, 17865–17875.

(45) Chapman, J.; Zoica Dinu, C. Assessment of Enzyme Functionality at Metal–Organic Framework Interfaces Developed through Molecular Simulations. *Langmuir* **2023**, *39*, 1750–1763.

(46) Zhang, H.; Lv, Y.; Tan, T.; van der Spoel, D. Atomistic simulation of protein encapsulation in metal–organic frameworks. *The Journal of Physical Chemistry B* **2016**, *120*, 477–484.

(47) Tuan Kob, T.; Ismail, M.; Abdul Rahman, M.; Cordova, K. E.; Mohammad Latif, M. Unraveling the structural dynamics of an enzyme encapsulated within a metal–organic framework. *The Journal of Physical Chemistry B* **2020**, *124*, 3678–3685.

(48) Li, P.; Modica, J. A.; Howarth, A. J.; Vargas, E.; Moghadam, P. Z.; Snurr, R. Q.; Mrksich, M.; Hupp, J. T.; Farha, O. K. Toward design rules for enzyme immobilization in hierarchical mesoporous metal-organic frameworks. *Chem* **2016**, *1*, 154–169.

(49) Chen, Y.; Jimenez-Angeles, F.; Qiao, B.; Krzyaniak, M. D.; Sha, F.; Kato, S.; Gong, X.; Buru, C. T.; Chen, Z.; Zhang, X.; others Insights into the enhanced catalytic activity of cytochrome c when encapsulated in a metal–organic framework. *Journal of the American Chemical Society* **2020**, *142*, 18576–18582.

(50) Pan, Y.; Li, H.; Li, Q.; Lenertz, M.; Zhu, X.; Chen, B.; Yang, Z. Site-directed spin labeling-electron paramagnetic resonance spectroscopy in biocatalysis: Enzyme orientation and dynamics in nanoscale confinement. *Chem Catalysis* **2021**, *1*, 207–231.

(51) Férey, G.; Mellot-Draznieks, C.; Serre, C.; Millange, F.; Dutour, J.; Surblé, S.; Margiolaki, I. A chromium terephthalate-based solid with unusually large pore volumes and surface area. *Science* **2005**, *309*, 2040–2042.

(52) Navarro-Sánchez, J.; Almora-Barrios, N.; Lerma-Berlanga, B.; Ruiz-Pernía, J. J.; Lorenz-Fonfria, V. A.; Tuñón, I.; Marti-Gastaldo, C. Translocation of enzymes into a mesoporous MOF for enhanced catalytic activity under extreme conditions. *Chemical Science* **2019**, *10*, 4082–4088.

(53) Vermeulen, W.; Vanhaesebrouck, P.; Van Troys, M.; Verschueren, M.; Fant, F.; Goethals, M.; Ampe, C.; Martins, J. C.; Borremans, F. A. Solution structures of the C-terminal headpiece subdomains of human villin and advillin, evaluation of headpiece F-actin-binding requirements. *Protein Science* **2004**, *13*, 1276–1287.

(54) Chung, Y. G.; Camp, J.; Haranczyk, M.; Sikora, B. J.; Bury, W.; Krungleviciute, V.; Yildirim, T.; Farha, O. K.; Sholl, D. S.; Snurr, R. Q. Computation-ready, experimental metal–organic frameworks: A tool to enable high-throughput screening of nanoporous crystals. *Chemistry of Materials* **2014**, *26*, 6185–6192.

(55) Chung, Y. G.; Haldoupis, E.; Bucior, B. J.; Haranczyk, M.; Lee, S.; Zhang, H.; Vogiatzis, K. D.; Milisavljevic, M.; Ling, S.; Camp, J. S.; others Advances, updates, and analytics for the computation-ready, experimental metal–organic framework database: CoRE MOF 2019. *Journal of Chemical & Engineering Data* **2019**, *64*, 5985–5998.

(56) VandeVondele, J.; Krack, M.; Mohamed, F.; Parrinello, M.; Chassaing, T.; Hutter, J. Quickstep: Fast and accurate density functional calculations using a mixed Gaussian and plane waves approach. *Computer Physics Communications* **2005**, *167*, 103–128.

(57) Kühne, T. D.; Iannuzzi, M.; Del Ben, M.; Rybkin, V. V.; Seewald, P.; Stein, F.; Laino, T.; Khaliullin, R. Z.; Schütt, O.; Schiffmann, F.; others CP2K: An electronic structure and molecular dynamics software package-Quickstep: Efficient and accurate electronic structure calculations. *The Journal of Chemical Physics* **2020**, *152*.

(58) Goedecker, S.; Teter, M.; Hutter, J. Separable dual-space Gaussian pseudopotentials. *Physical Review B* **1996**, *54*, 1703.

(59) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized gradient approximation made simple. *Physical Review Letters* **1996**, *77*, 3865.

(60) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *The Journal of Chemical Physics* **2010**, *132*.

(61) Garberoglio, G. OBGMX: A web-based generator of GROMACS topologies for molecular and periodic systems using the universal force field. *Journal of Computational Chemistry* **2012**, *33*, 2204–2208.

(62) BIOVIA, Dassault Systèmes, Materials Studio, 2020, San Diego: Dassault Systèmes, 2020.

(63) Martínez, L.; Andrade, R.; Birgin, E. G.; Martínez, J. M. PACKMOL: A package for building initial configurations for molecular dynamics simulations. *Journal of Computational Chemistry* **2009**, *30*, 2157–2164.

(64) Berendsen, H. J.; van der Spoel, D.; van Drunen, R. GROMACS: A message-passing parallel molecular dynamics implementation. *Computer Physics Communications* **1995**, *91*, 43–56.

(65) Hess, B.; Kutzner, C.; Van Der Spoel, D.; Lindahl, E. GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. *Journal of Chemical Theory and Computation* **2008**, *4*, 435–447.

(66) Pronk, S.; Páll, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M. R.; Smith, J. C.; Kasson, P. M.; Van Der Spoel, D.; others GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* **2013**, *29*, 845–854.

(67) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015**, *1*, 19–25.

(68) Páll, S.; Abraham, M. J.; Kutzner, C.; Hess, B.; Lindahl, E. In *Solving Software Challenges for Exascale*; Markidis, S., Laure, E., Eds.; Springer International Publishing: Cham, 2015; pp 3–27.

(69) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins: Structure, Function, and Bioinformatics* **2006**, *65*, 712–725.

(70) Best, R. B.; Hummer, G. Optimized molecular dynamics force fields applied to the helix- coil transition of polypeptides. *The Journal of Physical Chemistry B* **2009**, *113*, 9004–9015.

(71) Lindorff-Larsen, K.; Trbovic, N.; Maragakis, P.; Piana, S.; Shaw, D. E. Structure and dynamics of an unfolded protein examined by molecular dynamics simulation. *Journal of the American Chemical Society* **2012**, *134*, 3787–3791.

(72) Rappé, A. K.; Casewit, C. J.; Colwell, K.; Goddard III, W. A.; Skiff, W. M. UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations. *Journal of the American Chemical Society* **1992**, *114*, 10024–10035.

(73) Addicoat, M. A.; Vankova, N.; Akter, I. F.; Heine, T. Extension of the universal force field to metal–organic frameworks. *Journal of Chemical Theory and Computation* **2014**, *10*, 880–891.

(74) Coupry, D. E.; Addicoat, M. A.; Heine, T. Extension of the universal force field for metal–organic frameworks. *Journal of Chemical Theory and Computation* **2016**, *12*, 5215–5225.

(75) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics* **1983**, *79*, 926–935.

(76) Korolev, V. V.; Mitrofanov, A.; Marchenko, E. I.; Eremin, N. N.; Tkachenko, V.; Kalmykov, S. N. Transferable and extensible machine learning-derived atomic charges for modeling hybrid nanoporous materials. *Chemistry of Materials* **2020**, *32*, 7822–7831.

(77) Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An N log (N) method for Ewald sums in large systems. *The Journal of Chemical Physics* **1993**, *98*, 10089–10092.

(78) Bussi, G.; Donadio, D.; Parrinello, M. Canonical sampling through velocity rescaling. *The Journal of Chemical Physics* **2007**, *126*.

(79) Hess, B.; Bekker, H.; Berendsen, H. J.; Fraaije, J. G. LINCS: A linear constraint solver for molecular simulations. *Journal of Computational Chemistry* **1997**, *18*, 1463–1472.

(80) Hess, B. P-LINCS: A parallel linear constraint solver for molecular simulation. *Journal of Chemical Theory and Computation* **2008**, *4*, 116–122.

(81) Hub, J. S.; De Groot, B. L.; van der Spoel, D. g_wham A Free Weighted Histogram Analysis Implementation Including Robust Error and Autocorrelation Estimates. *Journal of Chemical Theory and Computation* **2010**, *6*, 3713–3720.

(82) Wang, X.; Lan, P. C.; Ma, S. Metal–organic frameworks for enzyme immobilization: beyond host matrix materials. *ACS Central Science* **2020**, *6*, 1497–1506.

(83) Xia, H.; Li, N.; Zhong, X.; Jiang, Y. Metal-organic frameworks: a potential platform for enzyme immobilization and related applications. *Frontiers in Bioengineering and Biotechnology* **2020**, *8*, 695.

(84) Saladino, G.; Marenchino, M.; Gervasio, F. Bridging the gap between folding simulations and experiments: The case of the villin headpiece. *Journal of Chemical Theory and Computation* **2011**, *7*, 2675–2680.

(85) Reiner, A.; Henklein, P.; Kiefhaber, T. An unlocking/relocking barrier in conformational fluctuations of villin headpiece subdomain. *Proceedings of the National Academy of Sciences* **2010**, *107*, 4955–4960.

(86) Spear, M. *Charting Statistics*; McGraw-Hill Book Company, Inc, 1952.

(87) McGill, R.; Tukey, J. W.; Larsen, W. A. Variations of box plots. *The American Statistician* **1978**, *32*, 12–16.

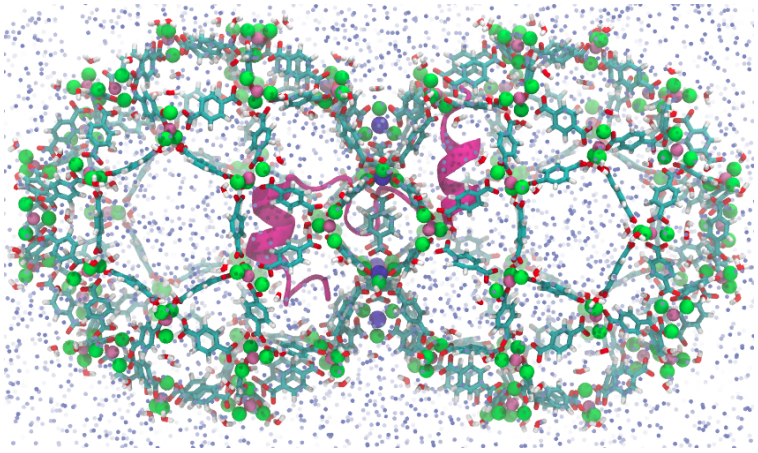
(88) Wickham, H.; Stryjewski, L. 40 years of boxplots. *Am. Statistician* **2011**, 2011.

(89) Tukey, J. W. *Exploratory data analysis: Limited preliminary Ed*; Addison-Wesley Publishing Company Ann Arbor, MI, USA, 1970.

(90) Zhang, Y.; Voth, G. A. Combined metadynamics and umbrella sampling method for the calculation of ion permeation free energy profiles. *Journal of Chemical Theory and Computation* **2011**, *7*, 2277–2283.

(91) Hu, Y.; Liu, X.; Sinha, S. K.; Patel, S. Translocation thermodynamics of linear and cyclic nonaarginine into model DPPC bilayer via coarse-grained molecular dynamics simulation: implications of pore formation and nonadditivity. *The Journal of Physical Chemistry B* **2014**, *118*, 2670–2682.

Table of Contents Graphics



For Table of Contents Only