I&EC research
Industrial & Engineering Chemistry Research

Article

Cite This: Ind. Eng. Chem. Res. XXXX, XXX, XXX−XXX

pubs.acs.org/IECR

Competition between Supramolecular Interaction and Protein−Protein Interaction in Protein Crystallization: Effects of Crystallization Method and Small Molecular Bridge

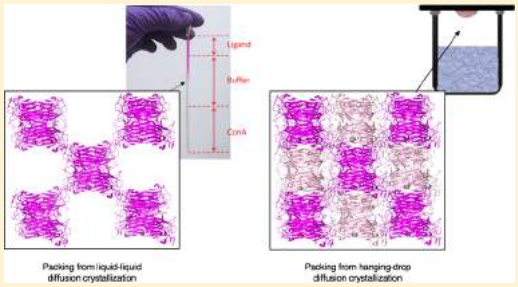
Rongting Hu,[†] Guang Yang,[†] Hong-ming Ding,*[,§] Jinbiao Ma,[‡] Yu-qiang Ma,[§] Jianhua Gan,*[,‡] and Guosong Chen*[,†]

[†]The State Key Laboratory of Molecular Engineering of Polymers and Department of Macromolecular Science and [‡]State Key Laboratory of Genetic Engineering, Collaborative Innovation Center of Genetics and Development, Department of Physiology and Biophysics, School of Life Sciences, Fudan University, Shanghai 200433, China

[§]Center for Soft Condensed Matter Physics and Interdisciplinary Research, Soochow University, Suzhou 215006, China

Supporting Information

ABSTRACT: We introduced a small molecular "inducing ligand" strategy to crystallize proteins via dual noncovalent interactions. Here we demonstrate that the variant protein-packing frameworks of concanavalin A (ConA) binding with ligands are controlled by different crystallization methods. Besides, the protein crystalline frameworks are also controlled by small-molecule inducing ligands **R*n*M** (n = 1−5), which varies in the number of ethylene oxide repeating units. To better understand the mechanism of different packing frameworks controlled by different inducing ligands, all-atom molecular dynamic (MD) simulations were performed. The MD simulation focused on the dynamic dimerization behavior of **ConA-R*n*M** system, which revealed clear relationships between the length of inducing ligands and ConA crystalline. In short, besides protein-packing framework control via different crystallization methods, inducing ligands **R*n*M** with suitable spacer lengths (n = 3, 4) also play an important role in desired and stable crystalline frameworks.



1. INTRODUCTION

Proteins are attractive building blocks for construction of functional materials because of their chemical and structural variety, and intrinsic functions. In laboratory, proteins self-assemble into different structures in one (1D), two (2D), and three dimensions (3D) with ordered-packing patterns.[1−6] For example, crystalline bacteria S-layers,[7] membranes containing porins or bacteriorhodopsin,[8] and 3D functional protein frameworks[9−12] have been obtained and inspired great interests in the research fields of biotechnology and nanotechnology.

These 2D/3D structures with protein highly ordered and tightly packed are protein crystals. Protein crystals, especially single crystals, are powerful tools to obtain molecular structure of proteins by using X-ray crystallography. From protein crystals, thousands of structures with important functions have been identified at the molecular level.[13] Crystallization is a complicated and tedious process, and it is one of the major bottlenecks of protein crystallography. Normally, it will take a few weeks or even months for the crystals to grow; the success of crystallization depends on many factors such as protein concentration, temperature, buffer, solvent pH and salt concentration, proportion of precipitant, and so on.[14]

From the view of material scientists, protein crystallization is a process of protein self-assembly in the solid state. Protein−protein, protein−solvent, and solvent−solvent interactions are the major interactions occurred during the crystallization process.[15] To promote protein crystallization, or to tune protein crystallization in "supramolecular way", other non-covalent interactions can be considered. In our previous study, we proposed and demonstrated that crystallization of model protein Concanavalin A (ConA) can be greatly accelerated under liquid−liquid diffusion condition by a small molecular ligand (Figure 1, **R*n*M**, n = 3, 4). This small molecule contains one mannopyranoside motif (Man; **M**, which binds to ConA) and one Rhodamine B motif (RhB; **R**, which dimerizes due to $\pi-\pi$ interaction);[10] ***n*** represents the repeating number of oligoethylene glycol units, the spacer which links Man and RhB together. In the crystal structure, it was found that ConA packed regularly into 3D frameworks, in which the molecular recognition between ConA and Man as well as the $\pi-\pi$ interaction between neighboring RhB play crucial roles in the framework structure. To be concise, this liquid−liquid diffusion crystallization method is named "supramolecular crystallization" in this paper, and the obtained crystals are called "supramolecular crystals" hereafter. Compared to the ConA crystal

Received: February 8, 2018
Revised: April 13, 2018
Accepted: April 18, 2018
Published: April 18, 2018

DOI: 10.1021/acs.iecr.8b00657
Ind. Eng. Chem. Res. XXXX, XXX, XXX−XXX

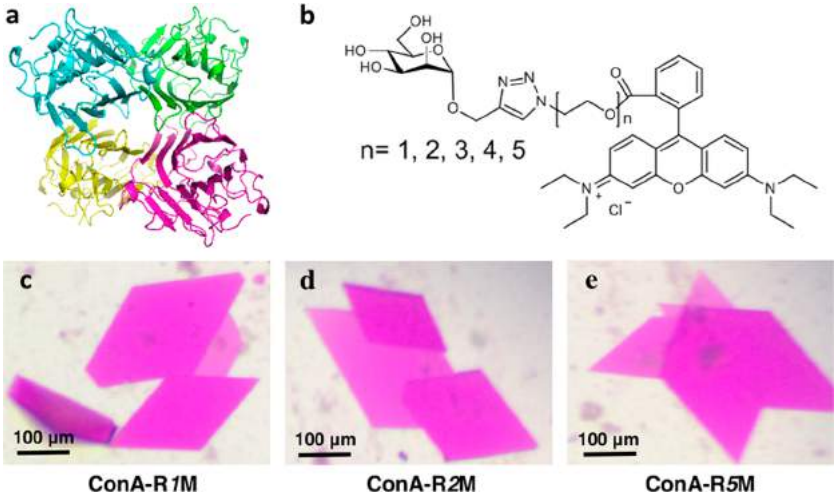


Figure 1. (a) Structure of tetrameric ConA from Protein Databank (PDB code: 1CVN). (b) Chemical structures of inducing ligands **R*n*M** (n = 1−5). (c−e) Cystals of **ConA-R*n*M** (n = 1, 2, 5) obtained through hanging-drop vapor diffusion method.

structures without **R*n*M** ligand,[16] much fewer protein−protein interactions were observed in the **ConA-R*n*M** complex structures. Interestingly, the packing modes of **ConA-R3M** and **ConA-R4M** are significantly different from each other in the crystal lattice, indicating the valuable contribution of the spacer length.

In this paper, we tried to answer two important questions following our previous work: whether this supramolecular framework can form under traditional hanging-drop vapor diffusion method and the detailed contribution of the spacer length in the supramolecular crystals. Toward this end, we recrystallized ConA in the presence of **R3M** or **R4M** using hanging-drop vapor diffusion technique; the resulting structures were then compared to our previous supramolecular ones. In addition to **R3M** and **R4M**, new small molecular ligands **R*n*M** were synthesized by incorporating different spacer lengths (n = 1, 2, 5) and were cocrystallized with ConA. To better decipher the crystallographic results, all-atom molecular simulation was performed in this study, which provided us a possible molecular view of the crystallization process.

2. EXPERIMENTAL SECTION

2.1. Materials and Synthesis of R*n*G (n = 1−5). Rhodamine B (≥95% purity) and pentamethyldiethylenetriamine (≥99% purity) were purchased from TCI (Shanghai). 2-[4-(2-Hydroxyethyl)-1-piperazinyl] ethanesulfonic acid (≥99% purity), (II) chloride tetrahydrate (≥99% purity), calcium chloride anhydrous (≥96% purity), sodium chloride (≥99.5% purity), and sodium hydroxide (≥96% purity) were purchased from Aladdin (Shanghai). D-Mannose (≥99% purity), thionyl chloride (≥99% purity), triethylene glycol (≥99% purity), tosyl chloride (≥99% purity), triethylamine (≥99% purity), sodium azido (≥99% purity), sodium sulfate pentahydrate (≥99% purity), hydrochloride acid (36−38% purity), dichloromethane (≥99.5% purity), tetrahydrofuran (≥99% purity), n-hexane (≥99% purity), and ethyl acetate (≥99.5% purity) were purchased from Sinopharm (China). Concanavalin A was purchased from Sigma-Aldrich.

R1M, **R2M**, and **R5M** were synthesized according to our previous synthesis protocols (Figure S1).[10,17] The synthesis scheme of **R*n*M** (n = 1−5) is shown in Figure S1. **R3M** and **R4M** were synthesized according to our previous literature.[10] The characterization details of ligands **R*n*M** (n = 1−5) were shown in Figure S2−S4.

2.2. Cocrystallization of ConA-R*n*G (n = 1−5). Concanavalin A (ConA), from *Canavalia ensiformis* (Jack bean), is a homotetrameric lectin of four identical 26 kDa subunits (Figure 1a). In our previous supramolecular crystallization, we injected the ligand (**R3M** or **R4M**) to the ConA solution in fine glass tubes (diameter about 3 mm and length about 130 mm), and the molar ratio between ligand and ConA (calculated as a monomer) was 1:1. Subsequently, we successfully obtained cocrystals, because of the spontaneous diffusion of ligands into the ConA solution (Figure S5).[10] Herein, five ligands **R*n*M** (n = 1−5) containing mannopyranoside and RhB are prepared and employed for protein crystallization (Figure 1b). Instead of liquid−liquid diffusion method, all the **ConA-R*n*M** crystals were cocrystallized using the hanging-drop vapor diffusion method. The crystallization conditions are composed of 20% polyethylene glycol 6000, 20 mM HEPES, 100 mM NaCl, 5 mM $CaCl_2$, and 5 mM $MnCl_2$ pH 7.2 for all the **ConA-R*n*M** complexes (n = 1−5). Crystallization drops (2 μL total = 1 μL protein−ligand mixture +1 μL precipitant) were set up with 24-well plates with sealant and siliconized glass cover slides (Hampton Research); the crystals grew at 18 °C over the course of 10 days. Interestingly, the crystals of **ConA-R3M** and **ConA-R4M** could grow using both liquid−liquid diffusion and hanging-drop vapor diffusion methods. Rhomboidal shaped crystals were grown to an average size of 0.05 mm × 0.3 mm× 0.4 mm (Figure 1c−e).

2.3. X-ray Diffraction and Data Collection. For data collection, crystals were flash frozen in the reservoir solution containing 4 M trimethylamineoxide[18] (TMAO) as a cryoprotectant. The diffraction data were collected from Shanghai Synchrotron Radiation Facility beamline BL17U at a wavelength of 0.97 Å and processed with program HKL2000 (http://www.hkl-xray.com/). All structures were solved using the molecular replacement program PHASER embedded in the CCP4 suite, the monomeric ConA crystal structure (PDB entry: 1JBC) was utilized as the search model. Model building were done using the COOT program and the structures were refined using the Refmac5 program of CCP4 suite. The bulk

DOI: 10.1021/acs.iecr.8b00657
Ind. Eng. Chem. Res. XXXX, XXX, XXX−XXX

water content was calculated with the MATTHEWS_COEF from CCP4 suite.

2.4. All-Atom Molecular Dynamics (MD) Simulation. As shown in Figure 2a, the all-atom molecular dynamics simulation system mainly consisted of ConA protein (PDB code: 1JBC) and the ligand **R*n*M** (n varies from 1−5). Here, for the sake of simplicity, we just considered one monomer of the ConA protein. Additionally, a number of water molecules were added to the system to separate the proteins from their mirrors. The ions (i.e., Na^+ and Cl^-) were also added to the simulation box to ensure the electric neutrality of the system. To model the process of crystallization, the following steps were carried out: (1) We took two adjacent proteins from the crystal structure, (2) separated them away, (3) bound the mannose moieties of **R*n*M** to the active sites of each protein since the binding of sugars to proteins took place before the dimerization,[10] (4) pulled them close to each other along the packing direction with very slow relative velocity (about 0.1 nm/ns), and (5) turned off the pulling (see Figure 2b).

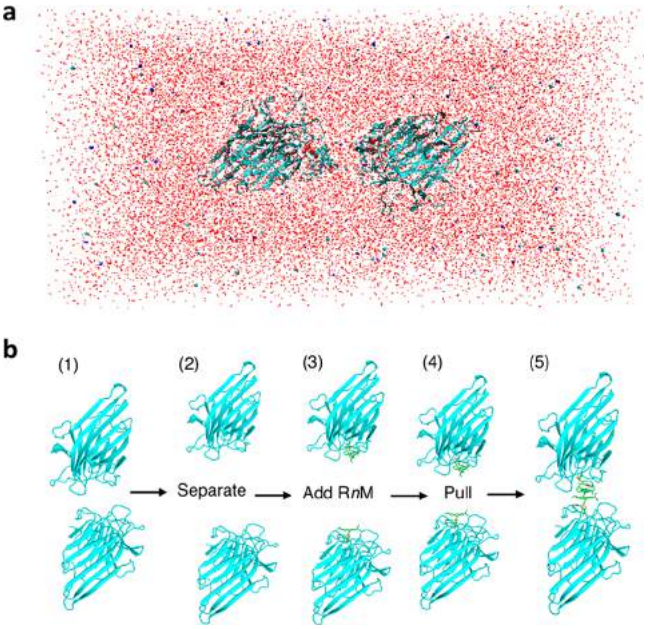


Figure 2. Setup of all-atom molecular dynamics (MD) simulation system. (a) Snapshot of the two **ConA-R*n*M** complexes immersed in aqueous solution. (b) Process of the interaction between two complexes in all-atom MD simulations.

All all-atom MD simulations were performed by using Gromacs 5.0.4 package[19] with the Amber force field[20] and the TIP3P water model[21] in the NPT ensemble. The force field parameter for **R*n*M** was built by using Antechamber tool.[22] During the simulation, the temperature was coupled at 291 K (18 °C) using Nosé−Hoover method,[23,24] and the pressure was fixed at 1 bar using Parrinello−Rahman method.[25] The Particle Mesh Ewald (PME) method was used to calculate the electrostatic interactions and the cutoff of Lennard-Jones (LJ) interaction was 1.2 nm.[26] The periodic boundary conditions were applied in all three dimensions. The time step was chosen as 2 fs, and each simulation was conducted for at least 50 ns (the first 40 ns for the pulling process, the last 10 ns for equilibrium MD simulation). Furthermore, the MD simulations were repeated three times for each system, starting from independent initial configurations. All of the figures of resulting structures were drawn using PYMOL (http://www.pymol.org).

3. RESULTS AND DISCUSSION

3.1. Crystal Structures of ConA-R*n*M (*n* = 1−5) via Hanging-Drop Method. All five **ConA-R*n*M** crystals diffracted to high resolution (2.0 Å for **ConA-R1M**, 2.5 Å for **ConA-R2M**, 2.15 Å for **ConA-R3M**, 1.89 Å for **ConA-R4M**, and 2.3 Å for **ConA-R5M**) and belonged to three different space groups with different unit cell dimensions. For **ConA-R*n*M** (*n* = 1, 2, 5) structures, each asymmetric unit contains four independent copies of the ConA protein, whereas it contains two ConA protein molecules for the **ConA-R*n*M** (*n* = 3, 4) crystals. The crystallographic data collection and refinement statistics of all **ConA-R*n*M** (*n* = 1−5) are summarized in Table 1.

3.2. Comparison of the Crystallographic Results of ConA-R3M and ConA-R4M via Different Crystallization Methods. Here the crystallographic results via two different methods (liquid−liquid diffusion and vapor diffusion) are compared. In our previous results, supramolecular crystals of **ConA-R*n*M** (*n* = 3, 4) were constructed through the molecular recognition between lectin and sugar (mannose) and the $\pi-\pi$ stacking of RhB between the ligands in neighboring proteins from liquid−liquid diffusion.[10] As shown in Figure 3a, it is quite clear that in our previous **ConA-R3M** crystal the ligand **R3M** induced a porous protein framework in one layer whose solvent fraction is up to 68.1%. The interactions between protein

Table 1. Parameters Determined by X-ray Diffraction Analysis of ConA-R*n*M (*n* = 1−5) via Hanging-Drop Vapor Diffusion

		ConA-R1M	ConA-R2M	ConA-R3M	ConA-R4M	ConA-R5M
			Data Collection			
space group		$P2_1\ 2_1\ 2_1$	$P2_1\ 2_1\ 2_1$	$P2\ 2_1\ 2_1$	$P2_1\ 2_1\ 2$	$P2_1$
unit cell	*a*, *b*, *c*	66.2, 118.8, 120.4	64.2, 119.9, 125.2	58.7, 61.9, 116.0	61.6, 116.1, 58.7	59.4, 64.1, 126.1
	α, *β*, *γ*	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90	90, 93, 90
resolution (Å)		2.0	2.5	2.15	1.89	2.30
reflections		60995	34047	84005	34408	42348
			Refinement			
R_{factor}		0.181	0.273	0.173	0.201	0.220
R_{free}		0.240	0.371	0.201	0.260	0.284
r.m.s. bound length (Å)		0.018	0.013	0.019	0.018	0.014
r.m.s. bound angles (deg)		1.93	1.71	2.11	1.87	1.78
solvent fraction		48.2%	49.2%	39.1%	39.1%	60.7%
no. of protein copies		4	4	2	2	4

DOI: 10.1021/acs.iecr.8b00657
Ind. Eng. Chem. Res. XXXX, XXX, XXX−XXX

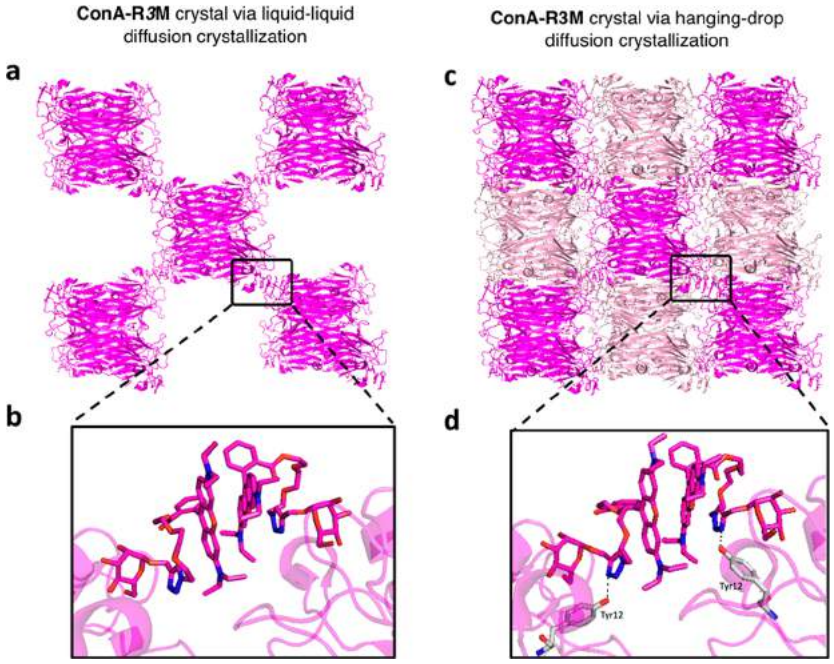


Figure 3. Packing structure of **ConA-R3M** crystal via (a) liquid−liquid diffusion method, with (b) close view of two **R3M**s dimerizing to each other (PDB code: 4P9W). (c) Packing structure of hanging-drop diffusion method and (d) zoom-in image of binding sites (PDB code: 5Z5P). Panels a and b are adapted with permission from ref 10. Copyright 2014 Nature Publishing Group.

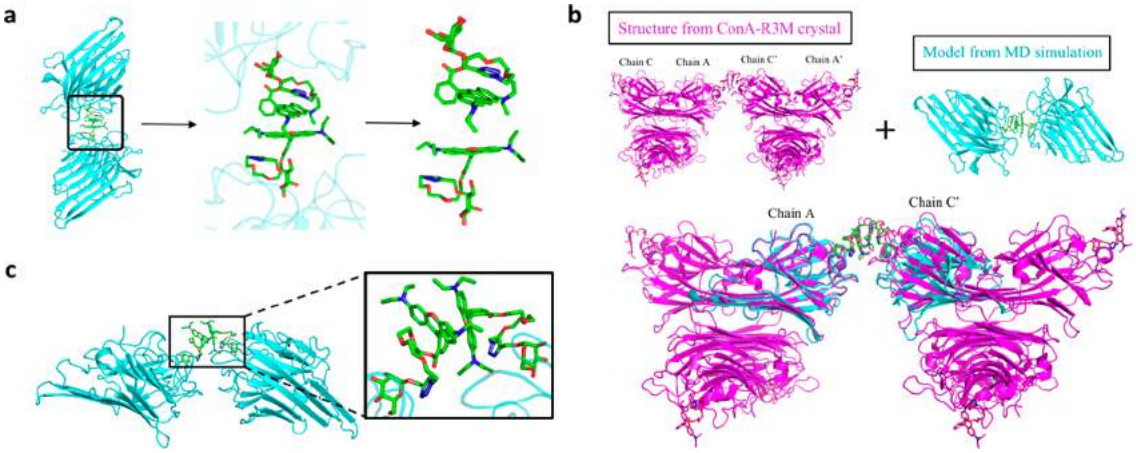


Figure 4. (a) Final snapshot of the packing of **ConA-R3M** in all-atom MD simulations, where the interaction energy between the benzene rings of the two ligands is about −37.1 kJ/mol. (b) Alignment of the chains in **ConA-R3M** in all-atom MD simulations (cyan) to that from experimental **ConA-R3M** crystal (magenta). (c) Final snapshot of the packing of **ConA-R4M** in all-atom MD simulations, where the interaction energy between the benzene rings of the two ligands is about −33.3 kJ/mol.

molecules were very limited, suggesting that the supramolecular interactions were the main driving force for crystallization, and the supramolecular interaction sites are shown in zoom-in image of Figure 3b (PDB code: 4P9W, redrawn from ref 9). In this work (Figure 3c,d), the results of ConA-α-mannopyranoside binding sites, detailing how saccharide is bound to ConA by hydrogen bonds and van der Waals interactions, are similar to the previous supramolecular one[10] and the examples reported in literature.[16] In the binding sites, the subsequent triazlole ring moiety interacts with Tyr12 of ConA via a hydrogen bond (2.7 Å, dashed line in zoom-in image Figure 3d), which indicates that the conformation of the chain from mannose to triazole ring is relatively rigid. The RhB group at the other end of the flexible triethylene glycol spacer dimerizes with the neighboring RhB moiety, which is very similar to supramolecular interactions of previous one (Figure 3b). Different from the local region, the packing of **ConA-R3M** complexes are quite different in the overall lattice of the crystals prepared by the two methods (Figure 3a, c). **ConA-R3M** crystals prepared from hanging-drop vapor diffusion induced two interpenetrating frameworks, depicted by two shades of magenta in Figure 3b, and the crystal belongs to $P22_12_1$ space group with unit cell parameters: a = 58.7 Å, b = 61.9 Å, c = 116.0 Å, $\alpha = \beta = \gamma = 90°$; whereas the previous supramolecular crystal belongs to $P2_1$ space group with unit cell parameters: a = 84.0 Å, b = 116.0 Å, c = 84.1 Å, $\alpha = 90°$, $\beta = 95°$, $\gamma = 90°$.

The packing frameworks of **ConA-R3M** and **ConA-R4M** via liquid−liquid diffusion method are shown in Figure 3a and Figure S6a, respectively. Interestingly, this interpenetrating framework packing of new **ConA-R3M** (Figure 3c) is very

DOI: 10.1021/acs.iecr.8b00657
Ind. Eng. Chem. Res. XXXX, XXX, XXX−XXX

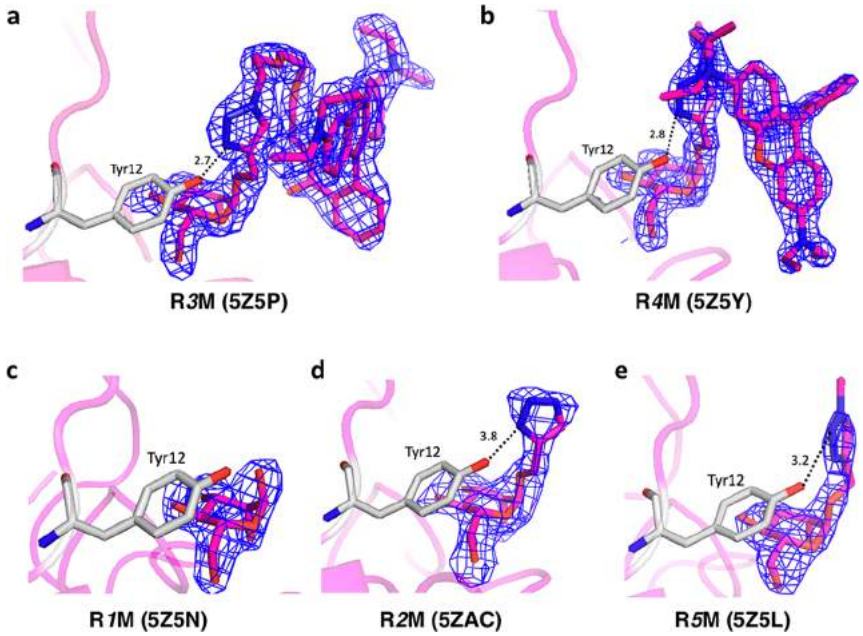


Figure 5. Detailed conformation of the ligands (PDB codes in parentheses). Complete **R3M** (a) and **R4M** (b) structures were modeled in the $2f_o-f_c$ electron density, while only mannose group of **R1M** (c) and mannose group with triazole of **R2M** (d) and **R5M** (e) were modeled in the $2f_o-f_c$ electron density maps. The maps are contoured at 1.0 sigma level. This figure was prepared using PyMOL.

similar to the structure of **ConA-R4M** (Figure S6a) obtained via supramolecular ways which has been reported in our previous paper. The rapid crystallization of **ConA-R3M** by liquid diffusion method, leads to the rapidly porous packing of ConAs via the dimerization of **R3M**, while **ConA-R3M** crystals prepared from vapor diffusion induce interpenetrating frameworks. However, in the **R4M** case, whether in the liquid diffusion method or in the vapor diffusion method, ConAs in the **ConA-R4M** crystal are both closely packed via the dimerization of **R4M** with the low solvent fraction 39.1% (Figure S6a,c). Thus, in the comparison for both methods in **R4M** case, they both resulted in interpenetrating frameworks. Again, in the new crystal of **ConA-R3M** or **ConA-R4M**, the two supramolecular interactions, that is, protein−ligand specific binding between ConA-Man and the dimerization of RhB played the important role (Figures 3d and S6b). In short, although there is no obvious difference on the binding of **R3M** to ConA observed between the two crystals, the two different crystallization methods (liquid diffusion vs vapor diffusion), could result in either interpenetrating or noninterpenetrating frameworks (Figure 3a,c), which could be attributed to the promotion of protein−protein interactions in the vapor diffusion method.

3.3. All-Atom MD Simulation Results of ConA-R3M and ConA-R4M. To obtain more physical insights into the above experimental results, we applied all-atom MD simulations to investigate the dimerization behavior of **ConA-R3M**. As shown in Figure 4a, the mannose-ConA (sugar-lectin) interactions and $\pi-\pi$ stacking in the **ConA-R3M** system were very stable, and two **R3M** molecules were packed due to the dimerization of RhB (nearly parallel to each other). More importantly, we found that not only the chain A of **ConA-R3M** in MD simulation (color in cyan) was aligned pretty well to the same chain from **ConA-R3M** crystal (color in magenta) but also the other chain (i.e., chain C′) in MD simulation was aligned well to that in the adjacent ConA (Figure 4b), which indicates that the simulation methods used here can catch most of the atomic/molecular details in the experiments. Then, we checked the dimerization behaviors of **ConA-R4M**, and found that the $\pi-\pi$ stacking between two **R4M** molecules can also occur, in agreement with the experimental results. However, with the increase of length, the flexibility of the molecule increased (see Figure S7), which has some negative impacts on the packing. As a result, the $\pi-\pi$ stacking of RhB part in the case of **ConA-R4M** is not perfectly parallel, and there exists the small angle between the benzene rings of the two **R4M** molecules, leading to the weaker interaction energy between the benzene rings of the two ligands compared to that in the case of **ConA-R3M** (see the caption of Figure 4). On the basis of above discussions, we can conclude that **R3M** may have better dimerization capacity than **R4M.** This might be used to explain the UV−vis results in the experiments. Dimerization of RhB group in only **ConA-R3M** precipitates was characterized by UV−vis spectra in Figure S8,[27,28] whereas no distinct characteristic peaks of RhB's dimerization in other precipitates from **ConA-R*n*M**. Obviously, the characteristics in UV−vis spectra is in accordance with the MD simulation results, that is, the ligand **R3M** has the best dimerization capacity than all the other ligands.

3.4. Crystallographic Results from ConA-R*n*M (*n* = 1, 2, 5). To elucidate the contribution of the tether length of **R*n*M**, in this work, we cocrystallized ConA with not only **ConA-R3M** and **ConA-R4M**, but also **ConA-R*n*M** (n = 1, 2, 5) by hanging-drop vapor diffusion. As mentioned, both **ConA-R3M** and **ConA-R4M** structures were refined in about 2 Å resolution, reveled well-defined electron density for **R3M** and **R4M** (Figure 5a, b). The triazole ring of **R3M** or **R4M** interacts with Tyr12 (shown in sticks in Figure 5a) of ConA via hydrogen bond (2.7 and 2.8 Å, respectively) between N5 and O atom of phenol group (dashed line in Figure 5a), stabilizing the configuration from mannose to triazole ring. However, in the binding pockets of **R2M** and **R5M**, only the mannose moieties with triazole ring portions could be fitted well in the electron density maps, and the distances between triazole ring and

DOI: 10.1021/acs.iecr.8b00657
Ind. Eng. Chem. Res. XXXX, XXX, XXX−XXX

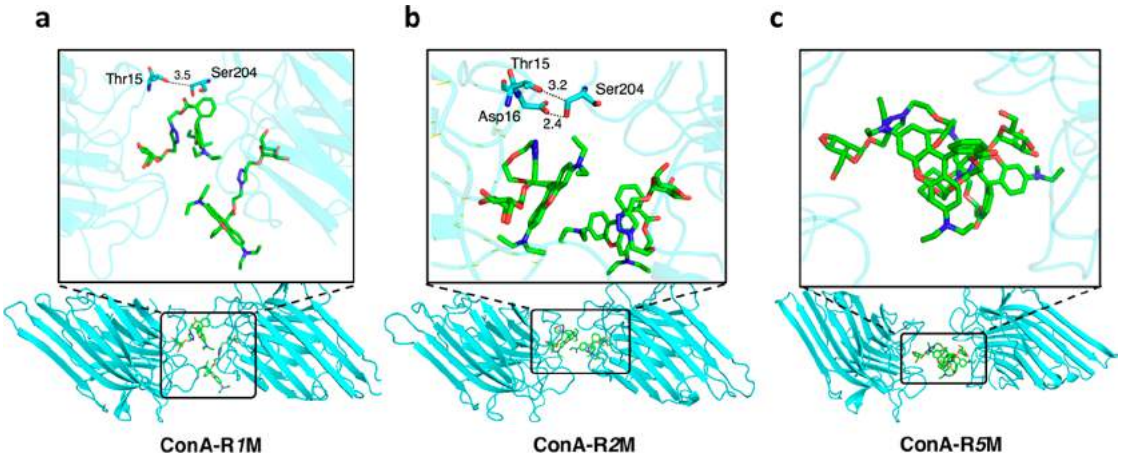


Figure 6. Final snapshots of the packing of (a) **ConA-R*1*M**, (b) **ConA-R2M**, and (c) **ConA-R5M** in all-atom MD simulations. Some hydrogen bonds are established by the interaction of residues Thr15 and Ser204 (of neighboring ConA) in **ConA-R*1*M** and Thr15, Asp16, and Ser204 (of neighboring ConA) in **ConA-R2M**, which prevent the approach of the two **R*1*M** or **R2M** molecules.

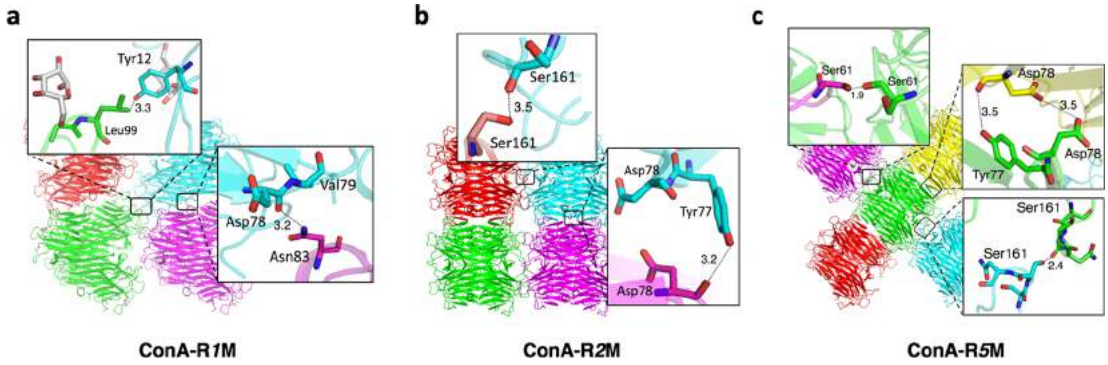


Figure 7. Protein−protein interactions observed in the crystal lattice of **ConA-R*n*M** (n = 1, 2, 5) (a) Three hydrogen bonds obtained in **ConA-R*1*M** with space group $P2_12_12_1$, (b) five hydrogen bonds obtained in **ConA-R2M** with space group $P2_12_12_1$, and (c) two hydrogen bonds obtained in **ConA-R5M** with space group $P2_1$. The zoom-in images show the contacts of crystal lattice for all three **ConA-R*n*M** (n = 1, 2, 5).

Tyr12 of ConA are 3.8 Å in **R2M** and 3.2 Å in **R5M** (dashed line in Figure 5d,e). As indicated by the long distances (3.8 and 3.2 Å), no stable interaction forms between triazole ring and Tyr12 of ConA in the **ConA-R2M** and **ConA-R5M** structures (Figure 5d,e), leading to the disordering of RhB "tails" in **ConA-R2M** and **ConA-R5M** complexes. In the **ConA-R1M** structure, only mannose moiety of **R1M** was fitted in density maps of each of four binding sites of ConA (Figure 5c). All the RhB groups at the other end of flexible ethylene glycol spacer were disordered and not modeled in the structures of the **ConA-R*n*M** (n = 1, 2, 5) complexes. One possible explanation for these findings is that the linker length of ligands in 1, 2, and 5 repeating oligoethylene glycol units are not suitable for RhB's dimerization inducing protein frameworks.

3.5. All-Atom MD Simulation Results of ConA-R*n*M (n = 1, 2, 5). To better answer the question why dimerization did not occur in **ConA-R*n*M** (n = 1, 2, 5) systems, the all-atom MD simulation was again employed. When the length (oligoethylene glycol units) of **R*n*M** is short (n = 1 or 2), the side chain of the proteins could prevent the approaching of the two **R*n*M** molecules, thus the $\pi-\pi$ stacking was not observed (see Figure 6a,b). On the contrary, when the tether length was very long, as aforementioned, the molecule became more flexible (see Figure S7). Even though the two **R5M** molecules can contact with each other closely, the $\pi-\pi$ stacking was unstable due to the thermal fluctuation of the long molecules. As a result, the efficient dimerization in the case of **R5M** cannot be observed either (see Figure 6c). In general, since the ligands **R*1*M**, **R2M**, and **R5M** were not of suitable length, the dimerization of the RhB moieties on neighboring **ConA-R*n*M** complexes was excluded.

3.6. Protein−Protein Interactions in ConA-R*n*M (n = 1, 2, 5). ConA forms a tetramer of four identical ∼26 kDa subunits at or above neutral pH; however, when the pH value drops below 5.6, ConA dissociates into active dimers of ∼52 kDa. In this work, all the crystals were grown at pH 7.2; as expected, ConA exists as tetramer in all the crystals. According to the structural results of all the **ConA-R*n*M** (n = 1, 2, 5), there are hydrogen bonds established by the interaction of residues Tyr12, Asp78, Asn83, and Leu99 in **ConA-R*1*M**, and Tyr77, Asp78, and Ser161 in **ConA-R2M**, and Ser61, Tyr77, Asp78, and Ser161 in **ConA-R5M** (Figure 7; more details in Table S1). Obviously, the interactions between neighboring tetramers in **ConA-R*n*M** (n = 1, 2, 5) are more compact than those in **ConA-R3M** or **ConA-R*4*M**. The crystallographic packing structure (Figure 7) showed that the crystallization was induced by protein−protein interactions rather than ligand−ligand interactions (i.e., $\pi-\pi$ stacking between the RhB) in the cases of **R*1*M**, **R2M**, and **R5M**. The zoom-in images showing the detailed interactions between neighboring tetramers within the crystal lattice of all three **ConA-R*n*M** (n = 1, 2, 5) complexes are depicted in Figure 7.

DOI: 10.1021/acs.iecr.8b00657
Ind. Eng. Chem. Res. XXXX, XXX, XXX−XXX

4. CONCLUSION

In summary, we demonstrated different protein crystalline frameworks of **ConA-R3M** crystals, either interpenetrating or noninterpenetrating frameworks, which can be achieved by different crystallization methods. Noninterpenetrating frameworks formed in **ConA-R3M** crystal via liquid−liquid crystallization, whereas the same protein (ConA) binding with the same ligand (**R3M**) crystallized from hanging-drop vapor diffusion method, arranged into two interpenetrating frameworks. These observations suggested that variant packing frameworks with ligands might be controlled by different crystallization methods. Furthermore, these results also indicated that the different protein packing models were controllable via slightly changes in the structure of the inducing ligands, such as tiny differences in length of the spacer that links mannose and RhB together. Only when the ligand **R*n*M** has a the suitable spacer length (n = 3, 4) can the protein form desired and stable crystalline frameworks. The completed structures of other **R*n*M**s with shorter or longer spacer lengths (n = 1, 2, 5) could not be observed in the crystals, which was verified by all-atom MD simulations.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.iecr.8b00657.

Synthesis scheme of **R*n*M** (n = 1, 2, 5); MALDI-TOF of **R*n*M** (n = 1, 2, 5); liquid−liquid diffusion method; packing structure of **ConA-R4M** crystal via different methods; root mean square deviations (RMSD) of five ligands as a function of time in MD simulations; UV−vis spectrum of precipitants and supernate of **ConA-R*n*M**; hydrogen bonds in **ConA-R*n*M** (n = 1, 2, 5) (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: dinghm@suda.edu.cn.
*E-mail: ganjhh@fudan.edu.cn.
*E-mail: guosong@fudan.edu.cn.

ORCID

Guosong Chen: 0000-0001-7089-911X

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank National Natural Science Foundation of China (No. 21504016, 91427302, 91527305, and GZ962), the State High-Tech Development Plan (the "863 Program", No. 2015AA020914) and Shanghai Synchrotron Radiation Facility (synchrotron BL17U).

■ REFERENCES

(1) Sinclair, J. C.; Davies, K. M.; Venien-Bryan, C.; Noble, M. E. Generation of protein lattices by fusing proteins with matching rotational symmetry. *Nat. Nanotechnol.* **2011**, *6*, 558−62.

(2) Gonen, S.; DiMaio, F.; Gonen, T.; Baker, D. Design of ordered two-dimensional arrays mediated by noncovalent protein-protein interfaces. *Science* **2015**, *348*, 1365−8.

(3) Lanci, C. J.; MacDermaid, C. M.; Kang, S. G.; Acharya, R.; North, B.; Yang, X.; Qiu, X. J.; DeGrado, W. F.; Saven, J. G. Computational design of a protein crystal. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 7304−9.

(4) Yang, L. L.; Liu, A. J.; Cao, S. Q.; Putri, R. M.; Jonkheijm, P.; Cornelissen, J. J. L. M. Self-Assembly of Proteins: Towards Supramolecular Materials. *Chem. - Eur. J.* **2016**, *22*, 15570−15582.

(5) Yeates, T. O.; Liu, Y. X.; Laniado, J. The design of symmetric protein nanomaterials comes of age in theory and practice. *Curr. Opin. Struct. Biol.* **2016**, *39*, 134−143.

(6) Brodin, J. D.; Ambroggio, X. I.; Tang, C.; Parent, K. N.; Baker, T. S.; Tezcan, F. A. Metal-directed, chemically tunable assembly of one-, two- and three-dimensional crystalline protein arrays. *Nat. Chem.* **2012**, *4*, 375−82.

(7) Sleytr, U. B.; Messner, P.; Pum, D.; Sara, M. Crystalline bacterial cell surface layers (s layers): from supramolecular cell structure to biomimetics and nanotechnology. *Angew. Chem., Int. Ed.* **1999**, *38*, 1034−54.

(8) Ringler, P.; Schulz, G. E. Self-assembly of proteins into designed networks. *Science* **2003**, *302*, 106−109.

(9) Sontz, P. A.; Bailey, J. B.; Ahn, S.; Tezcan, F. A. A Metal Organic Framework with Spherical Protein Nodes: Rational Chemical Design of 3D Protein Crystals. *J. Am. Chem. Soc.* **2015**, *137*, 11598−601.

(10) Sakai, F.; Yang, G.; Weiss, M. S.; Liu, Y.; Chen, G.; Jiang, M. Protein crystalline frameworks with controllable interpenetration directed by dual supramolecular interactions. *Nat. Commun.* **2014**, *5*, 4634.

(11) Luo, Q.; Hou, C.; Bai, Y.; Wang, R.; Liu, J. Protein Assembly: Versatile Approaches to Construct Highly Ordered Nanostructures. *Chem. Rev.* **2016**, *116*, 13571−13632.

(12) Liljestrom, V.; Mikkila, J.; Kostiainen, M. A. Self-assembly and modular functionalization of three-dimensional crystals from oppositely charged proteins. *Nat. Commun.* **2014**, *5*, 4445.

(13) Hendrickson, W. A. Stereochemically restrained refinement of macromolecular structures. *Methods Enzymol.* **1985**, *115*, 252−70.

(14) Chayen, N. E.; Saridakis, E. Protein crystallization: from purified protein to diffraction-quality crystal. *Nat. Methods* **2008**, *5*, 147−53.

(15) Matthews, B. W. Solvent content of protein crystals. *J. Mol. Biol.* **1968**, *33*, 491−7.

(16) Naismith, J. H.; Emmerich, C.; Habash, J.; Harrop, S. J.; Helliwell, J. R.; Hunter, W. N.; Raftery, J.; Kalb, A. J.; Yariv, J. Refined structure of concanavalin A complexed with methyl alpha-D-mannopyranoside at 2.0 A resolution and comparison with the saccharide-free structure. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1994**, *50*, 847−58.

(17) Yang, G.; Ding, H. M.; Kochovski, Z.; Hu, R. T.; Lu, Y.; Ma, Y. Q.; Chen, G. S.; Jiang, M. Highly Ordered Self-Assembly of Native Proteins into 1D, 2D, and 3D Structures Modulated by the Tether Length of Assembly-Inducing Ligands. *Angew. Chem., Int. Ed.* **2017**, *56*, 10691−10695.

(18) Mueller-Dieckmann, C.; Kauffmann, B.; Weiss, M. S. Trimethylamine N-oxide as a versatile cryoprotective agent in macromolecular crystallography. *J. Appl. Crystallogr.* **2011**, *44*, 433−436.

(19) 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.

(20) Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. The Amber biomolecular simulation programs. *J. Comput. Chem.* **2005**, *26*, 1668−1688.

(21) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926−935.

(22) Sousa da Silva, A. W.; Vranken, W. F. ACPYPE - AnteChamber PYthon Parser interfacE. *BMC Res. Notes* **2012**, *5*, 367.

(23) Nose, S. A Unified Formulation of the Constant Temperature Molecular-Dynamics Methods. *J. Chem. Phys.* **1984**, *81*, 511−519.

(24) Hoover, W. G. Canonical Dynamics - Equilibrium Phase-Space Distributions. *Phys. Rev. A: At., Mol., Opt. Phys.* **1985**, *31*, 1695−1697.

(25) Parrinello, M.; Rahman, A. Polymorphic Transitions in Single-Crystals - a New Molecular-Dynamics Method. *J. Appl. Phys.* **1981**, *52*, 7182−7190.

DOI: 10.1021/acs.iecr.8b00657
Ind. Eng. Chem. Res. XXXX, XXX, XXX−XXX

(26) 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*, 8577−8593.
(27) Nishikiori, H.; Fujii, T. Molecular forms of rhodamine B in dip-coated thin films. *J. Phys. Chem. B* **1997**, *101*, 3680−3687.
(28) Alig, A. R.; Gourdon, D.; Israelachvili, J. Properties of confined and sheared rhodamine B films studied by SFA-FECO spectroscopy. *J. Phys. Chem. B* **2007**, *111*, 95−106.

DOI: 10.1021/acs.iecr.8b00657
Ind. Eng. Chem. Res. XXXX, XXX, XXX−XXX