A Journal of the Gesellschaft Deutscher Chemiker
Angewandte
GDCh
International Edition
Chemie
www.angewandte.org

Accepted Article

Title: Fabrication of Pascal-triangle Lattice of Proteins by Inducing Ligand Strategy

Authors: Guosong Chen

To be cited as: *Angew. Chem. Int. Ed.* 10.1002/anie.202000771
Angew. Chem. 10.1002/ange.202000771

Link to VoR: http://dx.doi.org/10.1002/anie.202000771
http://dx.doi.org/10.1002/ange.202000771

WILEY-VCH

10.1002/anie.202000771

WILEY-VCH

RESEARCH ARTICLE

Fabrication of Pascal-triangle Lattice of Proteins by Inducing Ligand Strategy

Rongying Liu[a], Zdravko Kochovski[b], Long Li[a], Yue-wen Yin[c], Jing Yang[a], Guang Yang[a], Guoqing Tao[a], Anqiu Xu[a], Ensong Zhang[a], Hong-ming Ding*[c], Yan Lu*[b,d], Guosong Chen*[a,e], Ming Jiang[a]

Abstract: In this paper, a new type of supramolecular architecture, protein Pascal triangle has been constructed by using the inducing ligand strategy that we developed for protein assembly previously. Although mathematical studies on this famous geometry has a long history, no work on such Pascal triangle fabricated from native proteins appeared so far due to their structural complexity. In this work, by carefully tuning the specific interactions between the native protein building block WGA and the inducing ligand **R-SL**, Pascal-triangle 2D lattice with three types of triangular voids has been assembled. Moreover, 3D crystal structure was obtained based on the 2D Pascal triangles. The distinctive carbohydrate binding sites of WGA and intralayer & interlayer dimerization of RhB was the key to make such nanofabrication in solution possible. This strategy may be applied to prepare and explore various sophisticated assemblies based on native proteins.

Introduction

Fractal geometries such as Sierpinski triangle,[1] tree-like architectures[2] and other self-similar entities[3] have attracted increasing attention of chemists due to their importance in mathematics, engineering and life science.[4] However, Pascal triangle, as another significant type of pattern which appears similar to Sierpinski triangle but is not fractal, has been rarely favored by chemists. Virtually, Pascal triangle, to some extent, can be regarded as the origin of Sierpinski triangle.[5] In-depth investigating on Pascal triangle not only provides access to further ascertain the origin of fractals but also holds great potential for gaining a coherent description of the design principles underlying living organisms. [6] Whereas the study of Pascal triangle so far is still limited to the field of mathemathics as a consequence of high complexity of itself; i.e. it still remains a great challenge to fabricate Pascal-triangle entities via self-assembly strategy.

Recently, through protein self-assembly, lots of nanostructures with high complexity have been obtained,[7] including hollow tubes,[8] ring-like structures,[9] multiscale layers and crystals,[10] which fully exhibited the high capability of protein self-assembly to construct sophisticated nano-sized objects. We proposed and developed a strategy to construct protein assemblies via the dual non-covalent interactions participated by the small molecular "inducing ligand", with which different lectins have been piled up to form architectures at nanoscale via carbohydrate-protein interactions and dimerization of rhodamine B (RhB).[11] However, as far as we know, biomacromolecular Pascal triangles have not been achieved by using native proteins as building blocks. In our previous work,[11] the identical carbohydrate-binding sites tended to result in isotropic protein self-assemblies. While the key to Pascal triangle construction seems to be distinctive multiple interactions among building blocks. Hence, the proteins with distinctive carbohydrate binding sites rather than the identical carbohydrate-binding sites we previously employed, are probably capable of constructing Pascal triangle. In this paper, we would like to probe the feasibility of utilizing self-assembly of native proteins to fabricate Pascal triangle. Herein, we sought to leverage wheat germ agglutinin (WGA) to assemble with sialyllactoside-linked RhB as the small molecular ligand; the unique nature of eight carbohydrate-binding sites of WGA made the specific binding among WGA accessible. Through controlling the ratio of inducing ligand to WGA, Pascal-triangle 2D lattices with three types of triangular voids, 3D crystals of WGA were constructed. To the best of our knowledge, such fabrication of Pascal-triangle 2D lattices from native proteins has not been reported in the literature. The results reported here demonstrate great potential of the inducing ligand strategy for construction of protein Pascal triangle pattern with high complexity and provide in-depth understanding on the design principles of sophisticated geometries in living organisms.

Results and Discussion

Choices of protein building block and inducing ligand. Given Pascal triangle shares similar construction rules with Sierpinski triangle, the two critical rules for construction of the latter are available to consider:[1] (1) The building blocks with anisotropic shapes are favored, i.e. V-shape or K-shape. (2) The optimal selection of driving force is a combination of two or more interactions with different strength, detailly, the strong interaction is likely to be responsible for initially assembling the small molecules into basic building blocks, and then the weak one tends to enable the basic building blocks to further assemble into Sierpinski triangle. According to these design rules, both the

[a] R.Y. Liu, L. Li, J. Yang, G.Yang, G.Q. Tao, A.Q. Xu, E.S. Zhang, Prof. Dr. G.S. Chen, Prof. M. Jiang.
The State Key Laboratory of Molecular Engineering of Polymers and Department of Macromolecular Science
Fudan University, Shanghai 200433, China
E-mail: guosong@fudan.edu.cn (G. C.)
[b] Dr. Z. Kochovski, Prof. Dr. Y. Lu.
Soft Matter and Functional Materials, Helmholtz-Zentrum Berlin für Materialien und Energie, 14109 Berlin, Germany
E-mail: yan.lu@helmholtz-berlin.de (Y. L.)
[c] Y. W. Yin, Prof. Dr. H. M. Ding.
Center for Soft Condensed Matter Physics and Interdisciplinary Research, School of Physical Science and Technology
Soochow University, Suzhou 215006, China
E-mail: dinghm@suda.edu.cn (H. D.)
[d] Prof. Dr. Y. Lu.
Institute of Chemistry, University of Potsdam, 14476 Potsdam, Germany
[e] Prof. Dr. G.S. Chen
Multiscale Research Institute of Complex Systems, Fudan University, Shanghai 200433, China

Supporting information for this article is given via a link at the end of the document.

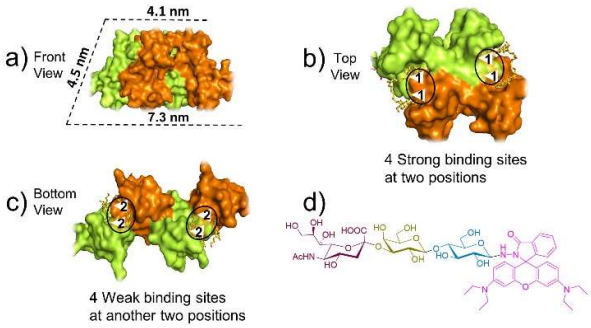


Figure 1. a) The shape of a WGA dimer, one monomer was colored with orange and another monomer was shown as yellow green (structure adapted from reported crystal structure, PDB code: 2X52). b) The 4 strong carbohydrate-binding sites at two positions (white triangle, four strong carbohydrate-binding sites were labeled as '1' on the top of WGA). c) The 4 weak carbohydrate-binding sites at another two positions (white triangle, labeled as '2' at the bottom of WGA). The four black circles in b and c represent the four distribution positions of binding sites. d) The molecular structure of inducing ligand (**R-SL**).

shape and carbohydrate-binding sites of the protein used to construct Pascal triangle are preferred to be anisotropic.

After careful selection, it was found that wheat germ agglutinin (WGA) appeares to be an ideal candidate for construction of a Pascal triangle. WGA is a homodimer, with each monomer containing four domains and the two monomers dimerize in a "head-to-tail" fashion forming a horseshoe-shaped complex[12a] (Fig. 1a). This anisotropic shape seems to favor the Pascal triangle construction. Meanwhile, eight independent carbohydrate-binding sites are concentrated at four positions on the monomer interface of WGA. The four binding sites on the top of WGA (Fig. 1b) exhibit a higher affinity than those at the bottom (Fig. 1c), with a stronger binding constant of at least a factor of two.[12] In terms of interactions, the four high-affinity carbohydrate-binding sites on the top of WGA are roughly identical, and the same for the four low-affinity ones. Therefore, there are two sets of carbohydrate-binding sites: four strong binding sites distributed at two positions present on the top of WGA (labeled as 1 in Fig. 1b) and four weak binding sites distributed at another two positions locate at the bottom of WGA (labeled as 2 in Fig. 1c).[12a, 12c] The isothermal titration calorimetry (ITC) experimently indicated the presence of two sets of binding sites on a WGA (Fig. S1).

To fabricate Pascal-triangle 2D lattice, WGA is supposed to be capable of initially forming a triangular basic building block after addition of inducing ligand. Based on our previous strategy,[11] here we synthesized a inducing ligand composed of two parts: α2,3-sialyllactose, which is responsible for binding with WGA; RhB, whose dimerization connects the ligand-attached WGA. Here this inducing ligand was denoted as **R-SL** (Fig. 1d). The all-atom molecular dynamics (MD) simulation was firstly employed to examine whether WGA can form triangular complex after binding with **R-SL**. Given that the strong carbohydrate-binding sites distributed at two positions would be firstly occupied and the steric effect (detailed discussion will be given in the next part) could arise from pre-binding **R-SL** at one position, the WGA attached by two **R-SL** was firstly used in simulation. After a 20-ns equilibrium, the final orientation of two **R-SL** was not parallel and their relative orientation to the protein was also different (Fig. 2a). Such property may not favor the formation of closed dimer but make triangular trimer construction possible (Fig. 2b, 2c).

Aiming at further estimating the probability of WGA forming triangular trimers, the coarse-grained (CG) Brownian dynamics (BD) simulation was applied, where each amino acid was treated as one CG bead in the simulation (Fig. 2d). Initially, 25 WGA proteins containing two **R-SL** (on the strong sites) were placed in the simulation box (Fig. S2). As time proceeded, the triangular trimer firstly appeared at 43800 τ (Fig. S3c), then more and more

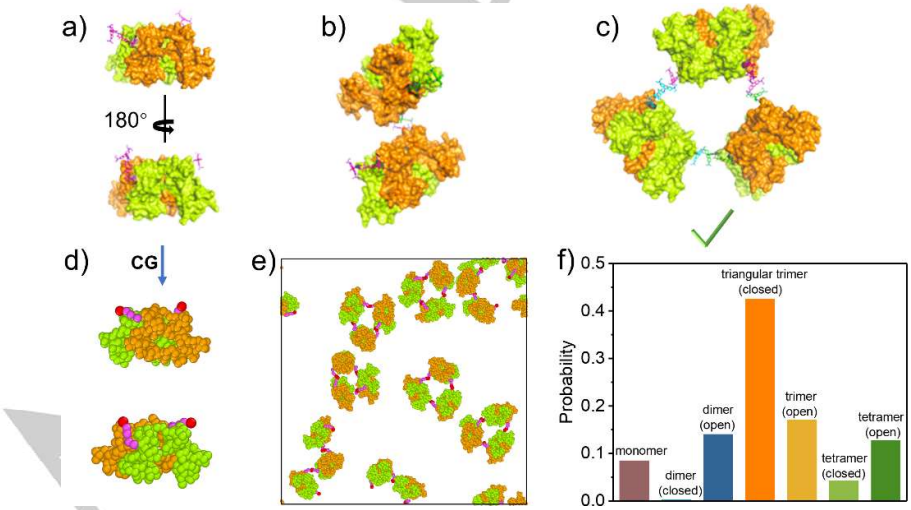


Figure 2. The results of molecular dynamics (MD) simulations and Brownian dynamics (BD) simulations. (a) The all-atom MD result of two ligands binding to the strong sites of WGA protein. The possible packing way of (b) two WGA proteins and (c) three WGA proteins via RhB dimerization. (d) Schematic illustration of the coarse-grained (CG) model for the WGA proteins (with two ligands at the strong binding sites). (e) The final snapshot for the assembly of twenty-five proteins in BD simulation. (f) The probability of different states of proteins in the simulations (averaged by seven independent runs, i.e., Fig. 2e and Fig. S7a-f). Open trimer refers to trimer with non-closed geometry, while closed trimer means those trimers with closed-loop geometry. For dimer and tetramer, details are shown in Fig S7.

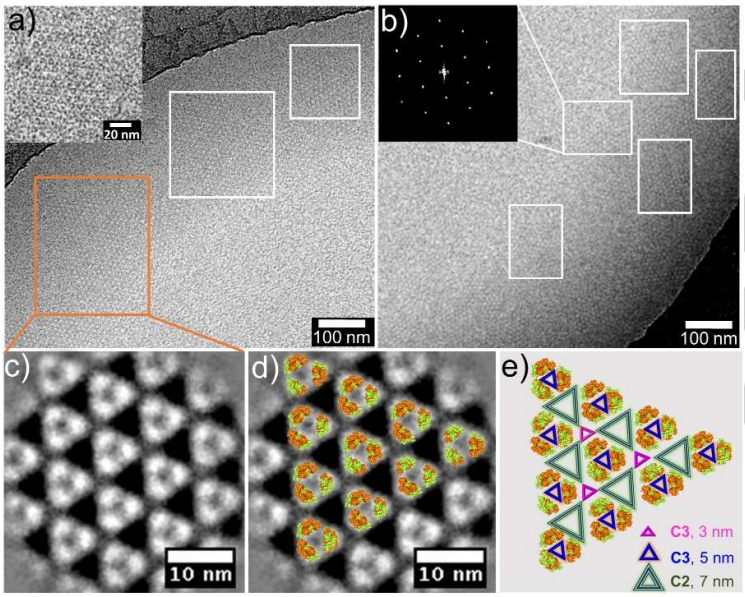


Figure 3. a, b) Cryo-EM image of the WGA 2D lattice (inset of a: enlarged Cryo-EM image of the WGA 2D lattice; inset of b: Fourier transform of selected area). c) 2D class average of 414 sub-areas from Cryo-EM images of the 2D lattice (one white dot represents one WGA protein dimer). d) Overlay of WGA crystal structures (PDB:2X52) with the 2D class average. e) Pascal- triangle pattern model (three types Pascal triangle pattern coexist, containing: small triangle void (pink, C3 symmetry, 3 nm length side), medium triangle void (blue, C3 symmetry, 5 nm length side), big triangle void (green, C2 symmetry, 7 nm length side)).

triangular trimers can be found in the simulation box (Fig. 2e, 2f, S3e). However, here the closed dimer was not observed during the whole simulation period (Fig. 2e, 2f), which was consistent with the result from all-atom MD simulations. Besides, we also observed closed tetramers in the simulation, but it was a rare event, because this type of closing loop would sharply decrease with increased monomers (Fig. 2f, S7). In addition, due to the limited time scale of the simulation, some non-closed oligomers such as open dimers, trimers and tetramers still remained (Fig. S7). Whereas due to their flexibility, these oligomers were not able to further assemble into large and ordered structures. In general, these simulation results suggested the feasibility of our design, namely, WGA could initially form triangular trimers, which would be used as basic building block in the subsequent Pascal-triangle 2D lattice construction.

Preparation and characterization of the Pascal-triangle 2D lattice of WGA. The corresponding synthesis details of **R-SL** were shown in supporting information (Scheme S1). The procedure of generating a 2D protein lattice is as follows: equal molar concentrations of **R-SL** and WGA were mixed in 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) buffer ([HEPES] =20 mM, [NaCl] = 40 mM, [$CaCl_2$] = 5 mM); the final concentration of **R-SL** and WGA was 2.0 × 10^{-4} M. The mixture was incubated at 4 °C for 48 h, then the Pascal-triangle 2D lattice was obtained as observed under Cryo-EM. The results shown in Fig. 3a and Fig. S4 demonstrate a clear lattice structure, with representative areas marked with white boxes. It can be clearly seen that a triangle structure, rather than the trapezoid shape of WGA, constituted the 2D lattice (Fig. 3a inset). The self-assembly of WGA induced by **R-SL** was also confirmed by dynamic light scattering (DLS) (Fig. S5). The height of this 2D lattice was about 4.1 nm as revealed by atomic force microscopy (AFM) (Fig. S6), which was consistent with the width of WGA, indicating that the 2D lattice was mono-layered.

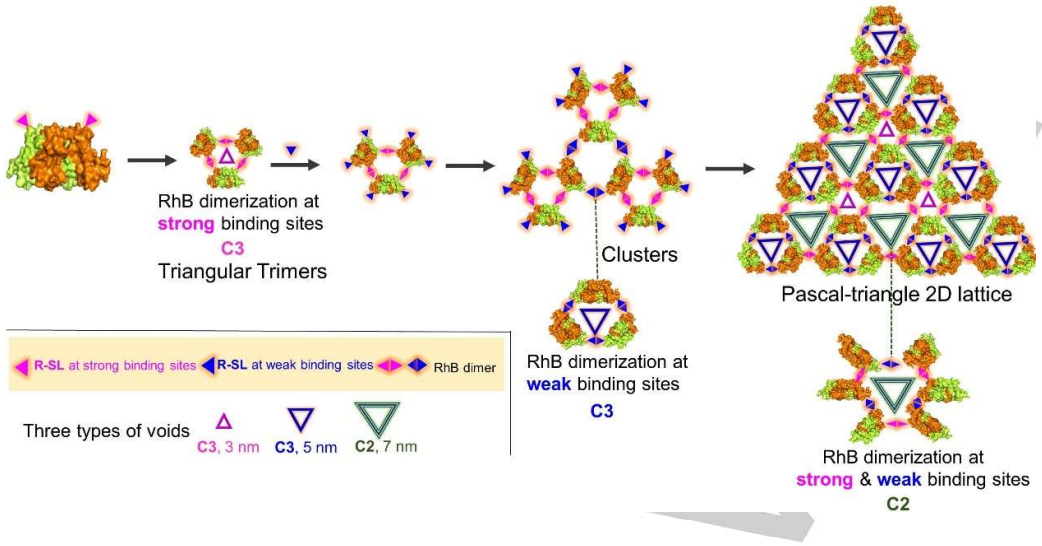
The successful construction of Pascal-triangle 2D lattice was further confirmed by reference-free 2D class averaging shown in Fig. 3c, 3d. Fourier transformation suggested the crystal nature of this 2D lattice (Fig. 3b inset). In addition, it was notable that planar crystals of such large size could grow without support within the confines and be dispersed in solution, which was supported by DLS data (Fig. S8).

To identify the driving force of the formation of Pascal-triangle 2D lattice, several control experiments have been conducted. Firstly, DLS measurement showed that the WGA was not able to self-assemble into such 2D lattice in the absence of **R-SL** (Fig. S9), which was consistent with previous literatures.[13] Secondly, the UV-vis and circular dichroism (CD) spectra results confirmed the RhB dimerization between **R-SL** (Fig. S10). The red-shift in UV-vis absorption spectrum could be ascribed to the dimerization of RhB. Meanwhile, the obvious CD signal agreed with the peak found in UV-vis, further confirmed the occurrence of RhB dimerization in the chiral environment provided by proteins. Moreover, the addition of either free α2,3-sialyllactose, which competitively bound with WGA; or β-CD, which inhibited the dimerization of RhB, leads to the dissociation of the assemblies (Fig. S11). This again confirmed that the formation of such Pascal-triangle 2D lattice was mainly driven by the dimerization of RhB.

We used cryo-EM to trace the formation process of the Pascal-triangle 2D lattice. Some triangular aggregates with about 7 nm side length were firstly observed after incubation for 8 h (Fig. S12a). When increasing the incubation time to 16 h,

10.1002/anie.202000771
WILEY-VCH

RESEARCH ARTICLE



Scheme 1. The proposed self-assembly mechanism underlying the formation of Pascal-triangle 2D lattice with p3 symmetry. It includes two sets of RhB dimerization (RhB dimerization at strong binding sites results in triangular trimer with C3 symmetry, RhB dimerization at weak binding sites gives rise to trimer with C3 symmetry, combination of two sets of RhB dimerization leads to hexamer with C2 symmetry.

clusters composed of three triangular aggregates were also detected (Fig. S12b). In addition, the DLS analysis confirmed the sequential formation from triangular trimers to clusters (Fig. S8). In short, the self-assembly process observed here was consistent to our design.

How did the triangular trimers form and further assemble into a Pascal-triangle 2D lattice? Given the presence of two sets of binding sites on a WGA (Fig. S1) and the proximity of two binding sites at each position (Figure 1b, 1c), the steered all-atom molecular dynamics (MD) simulation was employed to test the probability of the ligand binding to the different sites, where one free ligand was pulled to the strong binding site of the WGA with a very slow relative velocity (~0.2 nm/ns) in the absence (case 1) and presence (case 2) of the adjacent pre-binding ligand, and to the weak binding site in the absence (case 3) and presence (case 4) of the adjacent pre-binding ligand respectively (Fig. S13a). As shown in Fig. S13b, the pulling forces exhibited a sharp decrease when the ligand began to pack into the binding site in case 1 and case 3, but the decrease in case 1 was more sharp than that in case 3, suggesting that it was easier for the ligand to be occupied to the strong binding site. More importantly, the pulling forces both increased monotonously in case 2 and case 4, possibly since the steric effect between the free ligand and the pre-binding ligand (Figure S13c, d). As a result, the probability of the ligand binding to the strong site with adjacent pre-binding ligand (i.e., case 2) became lower than that of the ligand binding to the weak site without adjacent pre-binding ligand. Moreover, we also compared the binding energy of the WGA and the ligand, and the binding energy in case 1 was the lowest among the four cases (Figure S14). On the basis of above discussion, the order of binding priority appears to be as follows: case 1 > case 3 > case 2 > case 4, namely, two strong binding sites distributed at different positions tended to be firstly occupied, and then two weak binding sites distributed at different positions subsequently were occupied, the remaining two strong binding sites and two weak binding sites having adjacent pre-binding ligands were finally occupied.

Notably, the experiment and theoretical calculation results suggested that about four ligands bound to each protein as the ratio of inducing ligand **R-SL**/WGA was set at 1:1 (Fig. S15, Table S1). Consequently, based on above discussion, for the four ligands, two of them would firstly bind to the strong sites distributed at different positions, allowing the formation of triangular trimers, then another two ligands may occupy two weak binding sites located at different positions, permitting the triangular trimers to assemble into clusters and further into Pascal-triangle 2D lattice. On the contrary, when the ratio of **R-SL** to WGA was decreased to 0.5:1 (in this case, the average number of ligands binding to each protein was about two, see Table S1 and Fig. S15), only triangular trimers but no 2D lattice appeared (Fig. S16). In other words, the dimerization of RhB at the strong binding sites can only induce the formation of triangular trimers but that at the weak binding sites is essential for the further assembly of WGA into Pascal-triangle 2D lattice.

Based on the simulation and experimental results, the mechanism underlying the formation of the Pascal-triangle 2D lattice was proposed as illustrated in Scheme 1. At first, the RhB dimerization at strong binding sites gave rise to triangular trimers; then the RhB dimerization at weak binding sites tended to trigger the assembly of the triangular trimers into clusters and further into a Pascal-triangle 2D lattice.

More interestingly, along with WGA assembled into a Pascal-triangle pattern with p3 symmetry, three types equilateral triangle voids were also formed. Detailly, three WGAs made up an equilateral triangle voids with 3 nm side-length and C3 symmetry by RhB dimerization at strong binding sites. Similarly, equilateral triangle voids with 5 nm side-length, C3 symmetry by RhB dimerization at weak binding sites, and six WGAs made up an equilateral triangle voids with 7 nm side-length, C2 symmetry

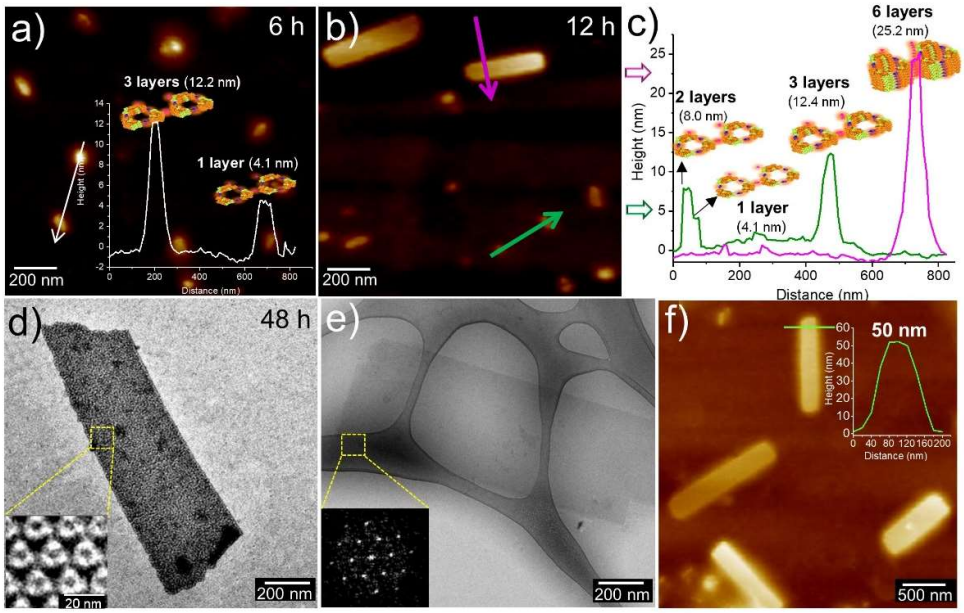


Figure 4. a) the AFM height image of **R-SL**/WGA (1.5:1) after incubation 6 h. b) The AFM height image and c) corresponding height profiles of (**R-SL**/WGA (1.5:1) after incubation 12 h. d) Negative stain TEM (inset: 2D class average image of selected area), e) Cryo-EM (inset: Fourier transform of selected area) and f) AFM height image of 3D crystals (**R-SL**/WGA (1.5:1) after incubation 48h.

results from RhB dimerization at both strong and weak binding sites (Fig. 3e, Scheme 1).

Although some transmembrane proteins can self-assemble into 2D lattice in a membrane environment, they exist as the free-monomer state in solution.[14] Additionally, some researchers found that a 2D lattice could be generated by designed protein interacting interface.[7h, 15] However, we checked all the 200 types (up to now) of 2D lattices from Reticular Chemistry Structure Resource (RCSR),[16] no similar Pascal-triangle 2D lattice with three types of voids could be found.

Fabrication and characterization of 3D crystals. After elucidating the self-assembly mechanism of the Pascal-triangle 2D lattice, we next investigated whether the 2D lattice could be further packed into 3D crystals. As an important driving force of protein self-assembly, protein-protein interaction was firstly considered to promote the stacking of 2D lattices. Thus, we attempted to extend incubation time. However, as the ratio of **R-SL**/WGA was set at 1:1, no obvious 3D crystals came to emerge even incubated for 120 h (Fig. S17). Further, more NaCl was added into the assembly solution to enhance the protein aggregation tendency, whereas only a few small 3D crystals can be observed (Fig. S18). This suggested that here the protein-protein interaction could indeed, to some extent, promote the stacking of 2D lattice but was not strong enough to generate large 3D crystals. To strengthen the interaction between layers, given the carbohydrate-binding sites of WGA were still unsaturated as the ratio of **R-SL** to WGA was set at 1:1, increasing this ratio may be a practical way to permit more binding sites to be occupied (or even become saturated).[11a,11d,17] In addition, it was reported that the interlayer dimerization of RhB tended to promote the 2D lattice stacking and 3D crystal growth.[18] Therefore, it can be envisioned that more inducing ligands may give rise to the stacking of 2D lattice and 3D crystal growth.

To test this possibility, the ratio of **R-SL** to WGA was directly increased to 1.5:1. CD and Cryo-EM were firstly employed to trace the assembly process. During the initial 6 h of incubation, a pronounced RhB dimerization signal increase was observed in CD spectra (Fig. S19), which often appeared as the RhB dimerization transformed from a free state into a confined state.[18] Indeed, as observed by Cryo-EM, some small 3D aggregates appeared after 6 h incubation, despite many Pascal triangles remained (Fig. S20a). This indicated that more inducing ligands participated in interlayer RhB dimerization, promoting the growth of 3D aggregates[18]. When increasing the incubation time to 12 h, the CD signal assigned to RhB dimerization increased significantly (Fig. S19), indicating that most of the Pascal triangles assembled into larger 3D aggregates (Fig. S20b). Therefore, it can be deduced that more inducing ligands gave rise to interlayer RhB dimerization compared to the case of 2D lattice, which enabled the formation of 3D aggregates.

AFM allowed further detection of the assembly process. After incubation of 6 h, both single-layer 2D lattice and multi-layer small aggregates were observed (Fig. 4a). Upon increasing the incubation time to 12 h, some larger aggregates appeared (Fig. 4b). A common feature shared by these aggregates was that their thicknesses were integral multiple of that of a single-layer 2D lattice (4.1 nm). Typically, even in the same aggregate, there were two different thicknesses: 4.1 nm and 8.0 nm (Fig. 4c), implying that certain parts of this aggregate already grown in the third dimension, while the other areas still retained a single-layer

structure. In other words, a hierarchical self-assembly occured in two and three dimensions simultaneously.

Notably, regular 3D crystals appeared after 48 h incubation (Fig. 4d-e). The AFM height image suggested that the thickness of 3D crystals was about 50 nm (Fig. 4f), which corresponded to about twelve layers of WGA. Significantly, both the results from 2D class average and Fourier-transform for this 3D crystal (Fig. 4d inset, 4e inset) were similar to those of 2D lattices (Fig. 3c, 3b inset), suggesting that the 3D crystals shared analogous protein packing with 2D lattices.

To summarize, a possible self-assembly mechanism of 3D crystals was proposed as illustrated in Fig. S21. Initially, the WGA still assembled into triangular trimers through RhB dimerization at strong binding sites. Then, the RhB dimerization at weak binding sites enabled the triangular trimers assemble into 2D clusters; at the same time, the interlayer RhB dimerization contributed by more inducing ligands allowed the clusters to further grow in the third dimension, namely, more inducing ligands permitted the simultaneous occurrence of self-assembly in two dimensions and three dimensions. As a result, these clusters gradually evolved into 3D crystals through self-assembly.

Conclusion

In this work, we constructed Pascal-triangle 2D lattices using well-selected protein building blocks with anisotropic shapes and two sets of carbohydrate binding sites. Moreover, the dynamic and exchangeable nature of non-covalent interactions make the further manipulation on 2D lattices to 3D crystals possible. Such results not only demonstrate the first construction of Pascal-triangle 2D lattice from native protein, but also drive the protein lattices fabricated by our inducing ligand strategy to an unprecedented level. We believe that this work opens a new avenue on fabrication of novel protein assemblies with desired shapes, sizes and functionalities, which might help us to understand the design principles underlying living organism better.

Acknowledgements

G.C. thanks NSFC/China (No. 51721002, 21861132012 and 91527305) for financial support. Y.L. thanks the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - Project number 410871749 for financial support. H.D. thanks NSFC/China (No. 21604060 and 11874045) for financial support. The authors would like to thank the Joint Lab for Structural Research at the Integrative Research Institute for the Sciences (IRIS Adlershof, Berlin) for Cryo-EM imaging. We thank Dr. Xiao Xu and Dr. Qidi Ran for helpful discussions. This work is supported by Shanghai Municipal Science and Technology Major Project (No.2018SHZDZX01) and ZJ Lab.

Keywords: inducing ligand • pascal triangle • self-assembly • WGA

[1] (a) Newkome, G. R.; Wang, P.; Moorefield, C. N.; Cho, T. J.; Mohapatra, P. P.; Li, S.; Iancu, V. *Science*. **2006**, 312(5781), 1782-1785. (b) Sarkar, R.; Guo, K.; Moorefield, C. N.; Saunders, M. J.; Wesdemiotis, C.; Newkome, G. R. *Angew. Chem. Int. Ed*. **2014**, 126(45), 12378-12381. (c) Newkome, G. R.; Moorefield, C. N. *Chem. Soc. Rev*. **2015**, 44(12), 3954-3967. (d) Zhang, X.; Li, N.; Gu, G. C.; Wang, H.; Nieckarz, D.; Szabelski, P.; Lu, J. T; Wu, K; Wang,Y,F. *ACS Nano*. **2015**, 9(12), 11909-11915. (e) Li, N.; Zhang, X.; Gu, G. C.; Wang, H.; Nieckarz, D.; Szabelski, P.; Peng, L. M. *Chinese. Chem. Lett*. **2015**, 26(10), 1198-1202. (f) Zhang, X.; Li, N.; Liu, L.; Gu, G.; Li, C.; Tang, H.; Wang, Y. *Chem. Commun*. **2016**, 52(69), 10578-10581. (g) Xing, L.; Peng, Z.; Li, W.; Wu, K. *Acc. Chem. Res*, **2019**, 52(4): 1048-1058. (h) Sun, Q.; Cai, L.; Ma, H.; Yuan, C.; Xu, W. *Chem.Commun*. **2015**, 51(75), 14164-14166. (i) Jiang, Z.; Li, Y.; Wang, M.; Liu, D.; Yuan, J.; Chen, M.; Li, X; Wang, P. *Angew. Chem. Int. Ed*. **2017**, 129(38), 11608-11613.

[2] (a) Khire, T. S.; Kundu, J.; Kundu, S. C.; Yadavalli, V. K. *Soft Matter*. **2010**, 6(9), 2066-2071. (b) Shin, S.; Gu, M. L.; Yu, C. Y.; Jeon, J.; Lee, E.; Choi, T. L. *J. Am. Chem. Soc*. **2018**, 140(1), 475-482. (c) Hernández, N. E.; Hansen, W. A.; Zhu, D.; Shea, M. E; Khalid. M.; Manichev, V.; Putnins, M.; Chen, M. Y.; Dodge, A. G.; Yang, L.; Berríos, I. M.; Banal, M.; Rechani, P.; Gustafsson, T.; Feldman, L. C.; Lee, S. H.; Wackett, L. P.; Dai, W.; Khare, S. D. *Nat. Chem*. **2019**, 11(3): 605-614.

[3] (a) Wang, M.; Wang, C.; Hao, X. Q.; Liu, J.; Li, X.; Xu, C.; Li, X. *J. Am. Chem. Soc*. **2014**, 136(18), 6664-6671. (b) Jiang, Z.; Li, Y.; Wang, M.; Song, B.; Wang, K.; Sun, M.; Li, X; Wang, P. *Nat. Commun*. **2017**, 8, 15476-15484. (c) Wang, L.; Liu, R.; Gu, J.; Song, B.; Wang, H.; Jiang, X.; Wang, M; Li, X. *J. Am. Chem. Soc*. **2018**, 140(43), 14087-14096. (d) Yin, G. Q.; Wang, H.; Wang, X. Q.; Song, B.; Chen, L. J.; Wang, L.; Li, X. *Nat. Commun*. **2018**, 9(1), 567-576.

[4] Barnsley, M. F. Fractals Everywhere. Morgan Kaufmann Publishers: Burlington, MA, **1993**.

[5] (a) Lakhtakia, A.; Messier, R.; Varadan, V. K.; Varadan, V. V. *J. Phys. A*: Math. Gen. **1987**, 20(11), 735-738. (b) Reiter, A. M. *Fibonacci. Quart*. **1991**, 31(2), 112-120.

[6] (a) Weibel, E. R. Am. *J. Physiol-Lung.C*. **1991**, 261(6): 361-369. (b) Losa, G. A. Fractals in biology and medicine. **2006**. (c) Reznikov, N; Bilton, M.; Lari, L.; Stevens, M. M.; Kröger, R. *Science*. **2018**, 360, No. eaao2189.

[7] (a) Salgado, E. N.; Radford, R. J.; Tezcan, F. A. *Acc. Chem. Res*. **2010**, 43, 661 – 672. (b) Koji, O.; Akira, O.; Takashi, H. *Chem.Commun*. **2012**, 48, 11714 – 11726; (c) Yeates, T. O.; Liu, Y.; Laniado, J. *Curr. Opin. Struct. Biol*. **2016**, 39, 134 – 143. (d) Luo, Q.; Hou, C.; Bai, Y.; Wang, R.; Liu, J. *Chem. Rev*. **2016**, 116(22), 13571-13632. (e) Bai, Y.; Luo, Q.; Liu, J. *Chem. Soc. Rev*. **2016**, 45(10), 2756-2767. (f) Sun, H.; Luo, Q.; Hou, C.; Liu, J. *Nano Today*. **2017**, 14, 16-41. (g) Baneyx, F.; Matthaei, J. F. *Curr. Opin. Biotech*. **2014**, 28, 39-45. (h) Norn, C. H.; André, I. *Curr. Opin. Struc. Biol*. **2016**, 39, 39-45. (i) King, N. P.; Bale, J. B.; Sheffler, W.; McNamara, D. E.; Gonen, S., Gonen; Baker, D. *Nature*. **2014**, 510, 103-108. (j) Huang, P. S.; Boyken, S. E.; Baker, D. *Nature*. **2016**, 537, 320-327. (k) Klinker, K; Schafer, O; Barz, M. *Angew. Chem. Int. Ed*. **2017**, 56(32), 9608-9613.

[8] (a) Brodin, J. D.; Smith, S. J.; Carr, J. R.; Tezcan, F. A. J. Am. Chem.Soc. **2015**, 137, 10468 – 10471. (b) Biswas, S.; Kinbara, K.; Oya, N. I.; Taguchi, H.; Aida, T. J. Am. Chem. Soc. **2009**, 131, 7556 –7557. (c) Miao, L.; Fan, Q.; Zhao, L.; Qiao, Q.; Zhang, X.; Hou, C.; Liu, J. *Chem. Commun*. **2016**, 52(21), 4092-4095.

[9] (a) Chou, T. F.; So, C.; White, B. R.; Carlson, J. C.; Sarikaya, M.; Wagner, C. R. *ACS Nano*. **2008**, 2(12), 2519 – 2525. (b) Bai, Y.; Luo, Q.; Zhang, W.; Miao, L.; Xu, J.; Li, H.; Liu, J. *J. Am. Chem. Soc*. **2013**,135(30), 10966 – 10969.

[10] (a) Li, X.; Qiao, S.; Zhao, L.; Liu, S.; Li, F.; Yang, F.; Liu, J. *ACS Nano*. **2019**, 13(2), 1861-1869. (b) Der, B. S.; Kuhlman, B. *Curr. Opin. Struc. Biol*. **2013**, 23(4), 639-646.

[11] (a) Sakai, F.; Yang, G.; Weiss, M. S.; Liu, Y.; Chen, G.; Jiang, M. *Nat. Commun*. **2014**, 5, 4634-4641. (b) Yang, G.; Zhang, X.; Kochovski, Z.; Zhang, Y.; Dai, B.; Sakai, F.; Liu, C; Chen, G.; Jiang, M. *J. Am. Chem. Soc*. **2016**, 138(6), 1932-1937. (c) Yang, G.; Ding, H. M.; Kochovski, Z.;

10.1002/anie.202000771

Hu, R.; Lu, Y.; Ma, Y. Q.; Chen, G.; Jiang, M. *Angew. Chem. Int. Ed.* **2017**, 56(36), 10691-10695. (d) Qi, W.; Zhang, Y.; Kochovski, Z.; Wang, J.; Lu, Y.; Chen, G.; Jiang, M. *Nano Research*. **2018**, 11(10), 5566-5572.

[12] (a) Wright, C. S.; Kellogg, G. E. *Protein. Sci*. **1996**, 5(8), 1466-1476. (b) Lis, H.; Sharon, N. *Chem. Rev*. **1998**, 98(2), 637-674. (c) Wright, C. S. *J. Biol. Chem.* 1992, 267(20): 14345-14352.

[13] (a) Muraki, M.; Ishimura, M.; & Harata, K. *BBA. Gen. Subjects*. **2002**, 1569(1-3), 10-20. (b) Wright, C. S.; Jaeger, J. *J. Mol. Biol.* **1993**, 232(2): 620-638.

[14] (a) Reviakine, I.; Bergsma, S. W.; Mazeres, D. C.; Govorukhina, N.; Brisson, A. *J. Struc. Biol*. **2000**, 131(3), 234-239. (b) Richter, R. P.; Him, J. L. K.; Tessier, B.; Tessier, C.; Brisson, A. R. *Biophys. J*. **2005**, 89(5), 3372-3385.

[15] (a) Gonen, S.; DiMaio, F.; Gonen, T.; Baker, D. *Science*. **2015**, 348, 1365-1368. (b) Suzuki, Y.; Cardone, G.; Restrepo, D.; Zavattieri, P. D.; Baker, T. S.; Tezcan, F. A. *Nature*. **2016**, 533, 369-373.

[16] O'Keeffe, M.; Peskov, M. A.; Ramsden, S. J.; Yaghi, O. M. *Acc. Chem. Res.* **2008**, 41(12), 1782-1789.

[17] Dotan, N.; Arad, D., Frolow; F.; Freeman.A. *Angew. Chem. Int. Ed.* **1999**, 38(16), 2363-2366.

[18] Brodin, J. D.; Ambroggio, X. I.; Tang, C.; Parent, K. N.; Baker, T. S.; Tezcan, F. A. *Nat. Chem*. **2012**, 4(5), 375-382.

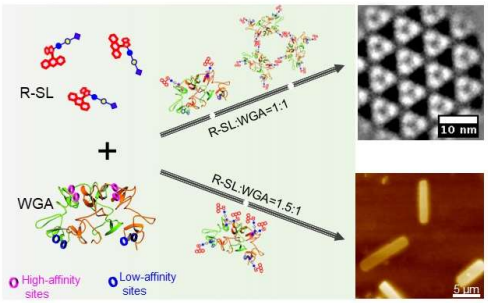
10.1002/anie.202000771

WILEY-VCH

Entry for the Table of Contents

RESEARCH ARTICLE

We constructed Pascal-triangle lattice by using a careful selection of protein building blocks with anisotropic shapes and two sets of carbohydrate binding sites. Moreover, the dynamic and exchangeable nature of non-covalent interactions make the further manipulation on 2D lattices to 3D crystals possible.



Rongying Liu, Zdravko Kochovski, Long Li, Yue-wen Yin, Jing Yang, Guang Yang, Guoqing Tao, Anqiu Xu, Ensong Zhang, Hong-ming Ding, Yan Lu*, Guosong Chen*, Ming Jiang*

Page No. – Page No.

Fabrication of Pascal-triangle Lattice of Proteins by Inducing Ligand Strategy