on the scissile phosphate. The Mg-bound water functions as a general acid to provide a proton to the cleaved DNA product. The bound DNA substrate is thereby hydrolyzed in a sequence-independent manner to produce cleaved products with a 5′-end phosphate and a 3′-end OH group.

Why EndoG has a diminished nuclease activity when it is dissociated into monomers is intriguing, as each protomer can bind and degrade DNA separately. To answer this question, they further constructed two obligatory monomeric CPS-6 mutants (P207E and K131D/F132N), which degrade DNA with a diminished activity because of a DNA-binding affinity poorer than that of wild-type CPS-6. Unexpectedly, the P207E mutant exhibited predominantly a 3′-to-5′ exonuclease activity, indicating a possible activity change from endonuclease to exonuclease. The dimer conformation of CPS-6 is thus essential to maintain its optimal DNA-binding and endonuclease activity. Compared to other non-specific endonucleases, which are typically monomeric enzymes, EndoG is a unique dimeric endonuclease, of which the activity can be modulated with oxidation to induce a dimer-to-monomer conformational change.

In summary, these results provide a molecular basis to explain how EndoG degrades DNA substrates without a sequence preference and why EndoG exhibits optimal endonuclease activity as a homodimer. The authors suggest that stabilizing the dimeric conformation of EndoG might provide a way to promote its endonuclease activity and to combat diseases induced by oxidative stress. (Reported by Chun-Jung Chen)

This report features the work of Hanna S. Yuan and her co-workers published in Nucleic Acid Res. ***44****, 10480 (2016).*

TLS 15A1 Biopharmaceuticals Protein Crystallography

- Protein Crystallography
- Biological Macromolecules, Protein Structures, Life Sciences

| References |

1. J. L. J. Lin, C. L. Lin, Y. Y. Hsiao, L. G. Doudeva, W.-Z. Yang, Y.-T. Wang, and H. S. Yuan, J. Biol. Chem. **287**, 7110 (2012).
2. Q. Zhou, H. Li, H. Li, A. Nakagawa, J. L. Lin, J. L. Li, E.-S. Lee, B. L. Harry, R. Skleen-Gaar, Y. Shehiro, D. William, S. Mitani, H. S. Yuan, B.-H. Kang, and D. Xue, Science **353**, 394 (2016).
3. J. L. J. Lin, A. Nakagawa, R. Skeen-Gaar, W.-Z. Yang, X. Ge, S. Mitani, D. Xue, and H. S. Yuan, Cell Rep. **16**, 279 (2016).
4. J. L. J. Lin, C.-C. Wu, W.-Z. Yang, and H. S. Yuan, Nucleic Acid Res. **44**, 10480 (2016).

Molecular Averaging Is Powerful for Crystal Structures

A new method, molecular averaging in real space, is developed to evaluate effectively the phasing power and to enhance the success of determining new protein structures.

X-ray crystallography of proteins remains a predominant method to determine three-dimensional structures of biological macromolecules. Despite great progress towards its automation and efficiency, phasing massive diffraction reflections remains a critical step in the determination of structures. *Ab initio* phasing, which requires only one native data set, is eagerly expected but is still a challenging method. One *ab initio* phasing is a method to utilize many equivalent molecules in a cell unit. Supposing electron densities of each equivalent molecule to be the same, the phases of amplitude data are greatly restricted; a correct molecular density is obtained. It is called molecular averaging. This technique is generally common for phase improvement after molecular replacement to eliminate the initial model biases. *Ab initio* phasing by molecular phasing is a challenging topic. Many trial calculations and much discussion of *ab initio* phasing with molecular averaging have been reported. In particular, a viral particle composed of many well ordered capsid proteins is an effective target for application with *ab initio* molecular averaging, but there has been no successful case of *ab initio* trials by molecular averaging to a novel protein structure.

In 2015, a research team of Chun-Jung Chen from the Life Science Group, NSRRC, applied this method through *ab initio* molecular averaging to crystal data of a novel grouper nervous necrosis virus-like particle (GNNV-LP) and obtained the first crystal structure.[1] In 2016, they reported the detailed description of the method applied to the GNNV-LP structure, a quantitative discussion of the feasibility of the method on introducing a novel conventional index.[2] This challenging trial and its success of *ab initio* phasing push the boundary of structure determination of biological macromolecules, which will benefit the community of structural biology.

The GNNV-LP data and crystal data of its protrusion domain were recorded at Protein Crystallography beamlines, including **TLS 15A1**, **SP 12B2** and **SP 44XU**. The GNNV-LP data were collected at resolution 3.6 Å, and the protrusion domain at 1.2 Å. The first *ab initio* method is to use non-crystallographic symmetry averaging (NCSA). From the GNNV-LP data, it assumed 30 redundant copies of a capsid protein molecule that formed viral icosahedral particles in the crystallographic asymmetric unit. This icosahedral orientation was derived from the self-rotation function of the amplitude data. Capsid protein particle positions and the solvent region in the unit cell were guessed from the form of related virus of the same genera: alpha-nodavirus. Beginning from a naive shell model, the electron density of the molecular copies was averaged; the density at the solvent region was flattened. In reciprocal space, amplitude data were added little by little. Between the density maps of real space and reciprocal space was linked with Fourier transform (FT) or inverse-FT. The basic cycle was calculated iteratively from low to high resolution. **Figure 1** shows a schematic diagram of the basic cycles. The final density map for a novel GNNV-LP was obtained (**Fig. 2**).

In the obtained map of electron density, there was still a region of unclear density that was for the protrusion domain located on the surface of the viral particle. The protrusion domain has a flexible structure in its viral states. They crystallized this protrusion domain solely and obtained high-resolution crystal data of four kinds of the space group and cell sizes. Secondly, they applied the molecular averaging procedure to the equivalent molecular densities among the three crystal data as shown in **Fig. 3**. The procedure was called cross-crystal averaging (CCA).

As an initial density, they used the envelope obtained from the unclear protrusion region of the GNNV-LP map. After an iterative calculation of averaging as the first NCSA case, the phases were much improved to produce high-resolution structures for the protrusion domain and the entire GNNV-LP structure (**Fig. 3**).

From this experience, the authors introduced a new index named a free fraction (*ff*) to show the phasing power of the molecular averaging. When the copy number of NCS is *n*, *ff* is expressed as

$$ff = \frac{1 - S}{n},$$

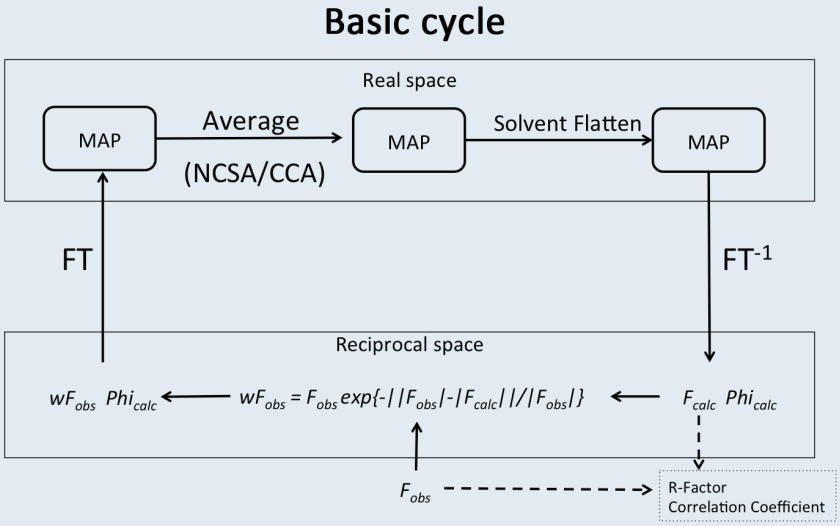

Fig. 1: Schematic diagram of a basic iteration cycle of the averaging method. [Reproduced from Ref. 2]

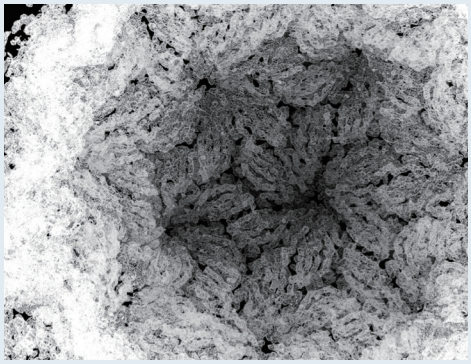

Fig. 2: The traceable density map (white mesh) in a view from the inner surface of GNNV-LP was obtained with *ab initio* averaging. [Reproduced from Ref. 2]

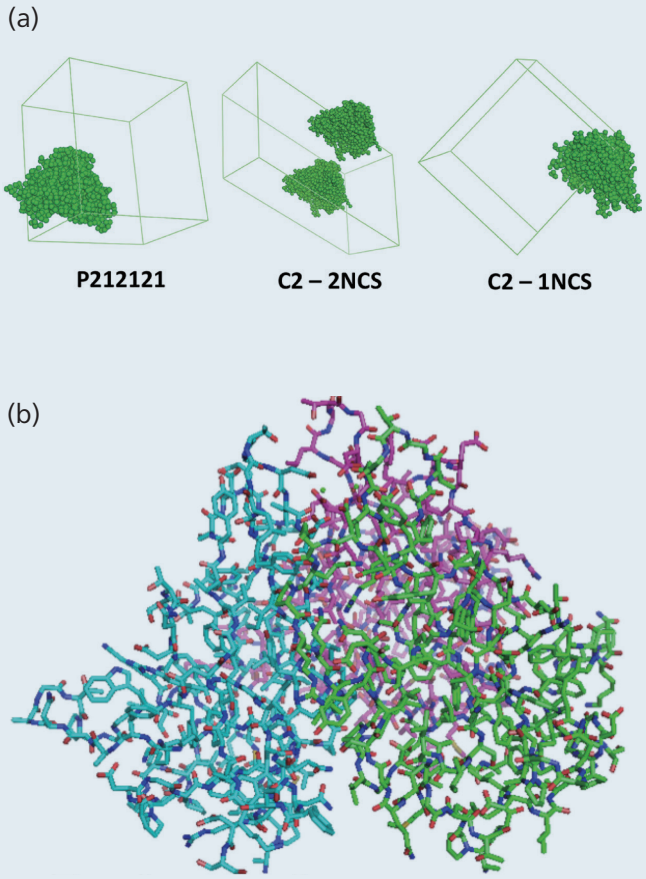

Fig. 3: Protrusion domain positions among the obtained crystals of three kinds (a) and the final structure (b). [Reproduced from Ref. 2]

in which S is the fraction of the solvent region. As *ff* shows the fraction of the unconstrained region in the total cell, (*n* - 1) copies and the solvent region are constrained in the averaging. Taking the benefit of the simplicity of this formula, *ff* is readily expanded to the multi-crystal situation of CCA.

If the equation is generalized to *m* pieces of crystal, crystal *k* has n_k NCS, S_k and P_k are fractions of the solvent and protein regions, respectively; *ff* becomes

$$ff = 1 \Big/ \left(\sum_{k=1}^{m} n_k + n_1 \frac{S_1}{P_1} + n_2 \frac{S_2}{P_2} + \ldots + n_m \frac{S_m}{P_m} \right).$$

The first successful NCSA case was *ff* = 0.024, the second CCA case was *ff* = 0.038. Adding test calculations with varying *ff*, a value less than about 0.1 was found to have a great phasing power to show the interpretable clear density map.

The authors discussed the meaning of *ff*. Given amplitudes and their phases, one can obtain an entire density map in the crystal cell. With only amplitude data in a typical case, the number of the data is decreased by half. From a point of view of information content, to obtain the entire density in the cell more than half part of the density should be known and *ff* should be less than 0.5. The region 0.5 < *ff* < 1 is considered to be an over-fitting condition in which the number of data is less than the number of parameters to be determined. The authors noted that *ff* is another expression of a data parameter ratio. *ff* is more useful than a data parameter ratio because the values of *ff* can be used without considering the resolution limit or the grid spacing of the density map. They obtained a criterion *ff* < 0.1; it corresponds to a data parameter ratio > 5, which is a reasonable value with another *ab initio* method for a successful condition such as sub-atomic resolution.

Their new index *ff* can show a quantitative phasing power. By its simple calculation, it indicates not only the feasibility of *ab initio* phasing trials but also how much density modification work is required to remove the initial model biases. This method of molecular averaging in real space can effectively evaluate the phasing power to enhance the success of determining new structures.

This report features the work of Chun-Jung Chen, Masato Yoshimura and their co-workers published in Acta Cryst. D ***72****, 830 (2016).*

TLS 15A1 Biopharmaceuticals Protein Crystallography
SP 12B2 BM – Protein X-ray Crystallography
SP 44XU Macromolecular Assemblies

- Protein Crystallography
- Biological Macromolecules, Protein Structures, Life Sciences

| References |

1. N.-C. Chen, M. Yoshimura, H.-H. Guan, T.-Y. Wang, Y. Misumi, C.-C. Lin, P. Chuankhayan, A. Nakagawa, S. L. Chan, T. Tsukihara, T.-Y. Chen, and C.-J. Chen, PLoS Pathog. **11**, e1005203 (2015).
2. M. Yoshimura, N.-C. Chen, H.-H. Guan, P. Chuankhayan, C.-C. Lin, A. Nakagawa, and C.-J. Chen, Acta Cryst. D **72**, 830 (2016).