Check for updates

Received: 7 November 2020 | Accepted: 15 November 2020

DOI: 10.1002/jccs.202000498

SPECIAL ISSUE PAPER

JOURNAL OF THE CHINESE CHEMICAL SOCIETY

Molecular dynamics simulation combined with small-angle X-ray/neutron scattering defining solution-state protein structures

Shang-Wei Lin[1] | **Kuan-Hsuan Su**[1] | **Yi-Qi Yeh**[2] | **U-Ser Jeng**[2,3] | **Chun-Ming Wu**[2] | **Hsiao-Ching Yang**[1]

[1]Department of Chemistry, Fu Jen Catholic University, New Taipei City, Taiwan

[2]National Synchrotron Radiation Research Center, Hsinchu Science Park, Hsinchu, Taiwan

[3]Department of Chemical Engineering, National Tsing Hua University, Hsinchu, Taiwan

Correspondence
Hsiao-Ching Yang, Department of Chemistry, Fu Jen Catholic University, New Taipei City 24205, Taiwan.
Email: hcyang_chem@mail.fju.edu.tw

Funding information
Ministry of Science and Technology of Taiwan

Abstract

Protein crystallography is frequently used to obtain the atomic-resolution structure of a particular protein by x-ray diffraction of its crystallized form. Nevertheless, a single static structure cannot represent the sequence structural dynamics of a protein's function. Based on our previous study (*J. Phys. Chem. B* 2017, *121* (50), 11229–11240 and *ACS Catal.* 2018, *8* (3), 2534–2545), we herein introduce a strategy of determining the protein hydration structures by small-angle X-ray and neutron scattering (SAXS and SANS) with contrast variation techniques, in combination with molecular dynamics simulation (MD), to describe hydrodynamics conformation transitions of biomacromolecules. Water is a partner that conducts the dynamics of proteins, and hydration interactions with proteins affect their dynamics. To shed light on the way of the chemical constitution of a protein, the water interactions, and the dynamics of its structure underlie the specific mechanisms of enzyme functions, we review an integrated approach using small-angle X-ray and neutron scattering (SAXS and SANS) in combination with molecular dynamics (MD) simulation to probe the structural dynamics of a given protein, thus shedding light on its function.

KEYWORDS

molecular dynamics (MD) simulations, protein water hydration dynamics, small angle X-ray/neutron scattering (SAXS and SANS)

1 | INTRODUCTION

The protein dynamical transition is triggered by its structure coupling with water hydration to function biologically.[1–3] A single static structure cannot represent the solution structures of a protein to the dynamical properties in function.[4,5] Solution nuclear magnetic resonance (NMR) is usually performed for studying the dynamical structures of a protein, but it depends on the protein molecular weight with residue identification and labelling techniques. Despite the many advantages in cryo-electron microscopy (EM) in covering a wide molecular weight range of specimens for structural studies of dynamic biological macromolecules, until recently, cryo-EM could not deliver high-resolution structures of membrane proteins of interest. To date, the protein data bank currently contains more than 170,000 crystal structures,[6] which can provide detailed information about structures. Nevertheless, a significant number of structures of non-crystallizable large proteins remain either unsolved or only partially solved.

Water is a partner that conducts the dynamics of proteins, and hydration interactions with proteins affect

 © 2020 The Chemical Society Located in Taipei & Wiley-VCH GmbH

their dynamics.[7–10] To shed light on how the chemical constitution of the protein, the water interactions, and the dynamics of its structure underlie the specific mechanisms of enzyme functions, in this paper, we review an integrated approach using small-angle X-ray and neutron scattering (SAXS and SANS) in combination with molecular dynamics (MD) simulation (Figure 1**)** to investigate the structural dynamics of a protein, thus shedding light on its function. [11,12]

Synchrotron SAXS provides a means to probe the time-preserved structural state(s) of bio-macromolecules in solution, characterizing the size and shape evolution with resolution of 1–100 nm,[8,13] complementary to X-ray crystallography with angstrom atomic resolution. Figure 1(a) illustrates the measurement and preliminary analysis of SAXS, the scattering vector q is defined as $q = 4\pi\lambda^{-1}\sin\theta$, with 2θ for the scattering degree and λ the wavelength, to determine the size and shape of a given protein, which is assumed to be derived from a number of coexisting conformational states in solution.[11,14] Due to the complementary relationship that binds SAXS and SANS, SANS provides unique and valuable information in the dynamical transition behavior of hydration water on biomolecules and materials.[3,15,16] Because of the elastic scattering feature, SANS and SAXS (Figure 1c) provide complementary structure information to each other. It should be noted that the intense X-ray beams of synchotrons may damage protein samples. The findings of the study indicate that the scanning HPLC–SAXS approach effectively excludes scattering containment from coexisting oligomer species while collecting mono-dispersion monomer protein structures (see Figure 1c) and limits damage from radiation via an on-line flow system. [11,17,18]

We demonstrated an approach combining simulations of all-atom molecular dynamics and small-angle X-ray scattering (SAXS)/ultraviolet–visible spectroscopy (UV–vis) absorption measurements, which allows probing of the structures and dynamical behaviors of two counter-functioning proteins, thromboxane synthase (TXAS) and prostacyclin synthase (PGIS), in solution. [11] A large number of MD structures can be used to simulate

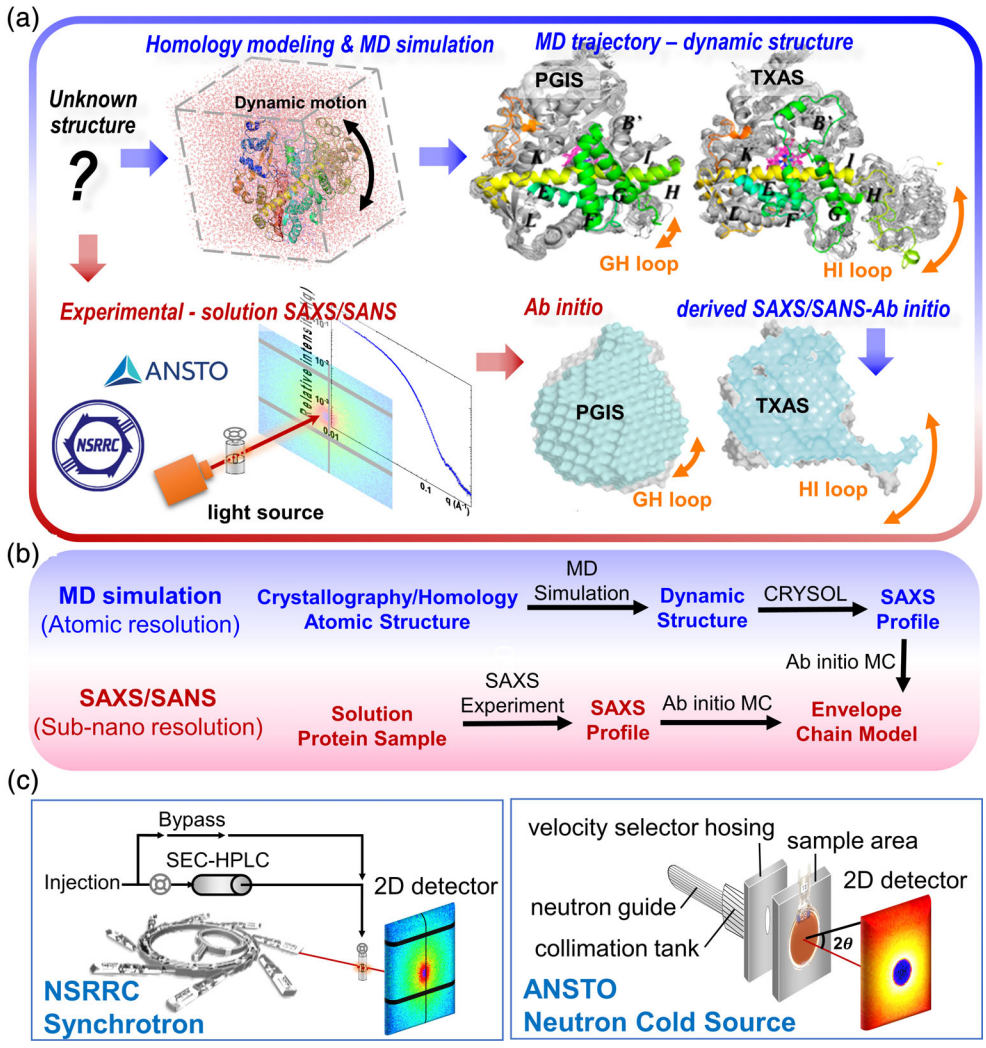

FIGURE 1 (a),(b) Defining solution state protein structures and enzyme catalysis by with small-angle x-ray/neutron scattering combined molecular dynamics simulation.
Source: (*J. Phys. Chem. B* **2017**, *121* (50), 11,229–11,240)
(c) Illustration of NSRRC HPLC-SAXS and ANSTO SANS beamline[29]

SAXS and SANS profiles (Figure 1b) in the software CRYSOL and CRYSON.[19,20] The results of MD and experimental-solution SAXS had excellent agreement. We also examine the significant internal motions within proteins, structures of the whole molecules, and possible problems with protein folding, crystallization, and functionality.

2 | PROTEIN SIZE AND FLEXIBILITY

Here we choose two antagonistic proteins, PGIS and TXAS, for demonstration.[11] Currently, the X-ray crystallographic structures of PGIS are available, but lack of TXAS structures. Given in Figure 2a, we used an online size exclusion high-performance liquid chromatographic (SE-HPLC) system to collect individual monodisperse from oligomer.[17,18,21] Because the proteins aggregate slowly enough after purification, we could capture monomer protein structures by SE-HPLC different elution time for timely SAXS data collection before aggregation.[17,18,21] After detected using 280 nm and 420 nm UV–VIS, the sample solution was slowed down for SAXS data collection with an exposure frame rate of 30 s; the SAXS data frames of the same radius of gyration (R_g) confirms a single species corresponding to the protein monomers;[11] these data were averaged for improved statistics as shown in Figure 2b, which inset shows the data are fitted using the Guinier approximation $\ln[I(q)] = \ln[I(0)] - (R_g^2/3)q^2$ for the R_g value and globular protein hydrodynamic radius $R = (5/3)^{0.5}R_g$.[14] Figure 2c shows the globular protein size estimated from the R_g value. PGIS and TXAS are purified to homogeneity as 56.63 and 59.02 kDa.[11] The R_g of PGIS and TXAS solution is estimated 25.4 Å and 31.5 Å, and protein radius is 32.8 and 40.7. Their molecular weights are very similar, but their protein radius differ by 7.9 Å. This significant difference is consistent with the simulated dynamic structures with MD simulation. Both consistently indicate that TXAS is more flexible and extended than PGIS. Further, we examined the folding shapes of the proteins using the Kratky plots, which show that both PGIS and TXAS are globular in structures far from that of unfolded proteins. The relatively asymmetric profile of TXAS, however, suggests consistently a relatively extended conformation. Figure 2d shows the different profiles corresponding to a specific protein folding. Homology modeling of the TXAS structure with detailed MD simulations in water allows estimation of the configuration space with loop motion and the origins of the protein behaviors in solution. In contrast to the stability of the conserved PGIS structure

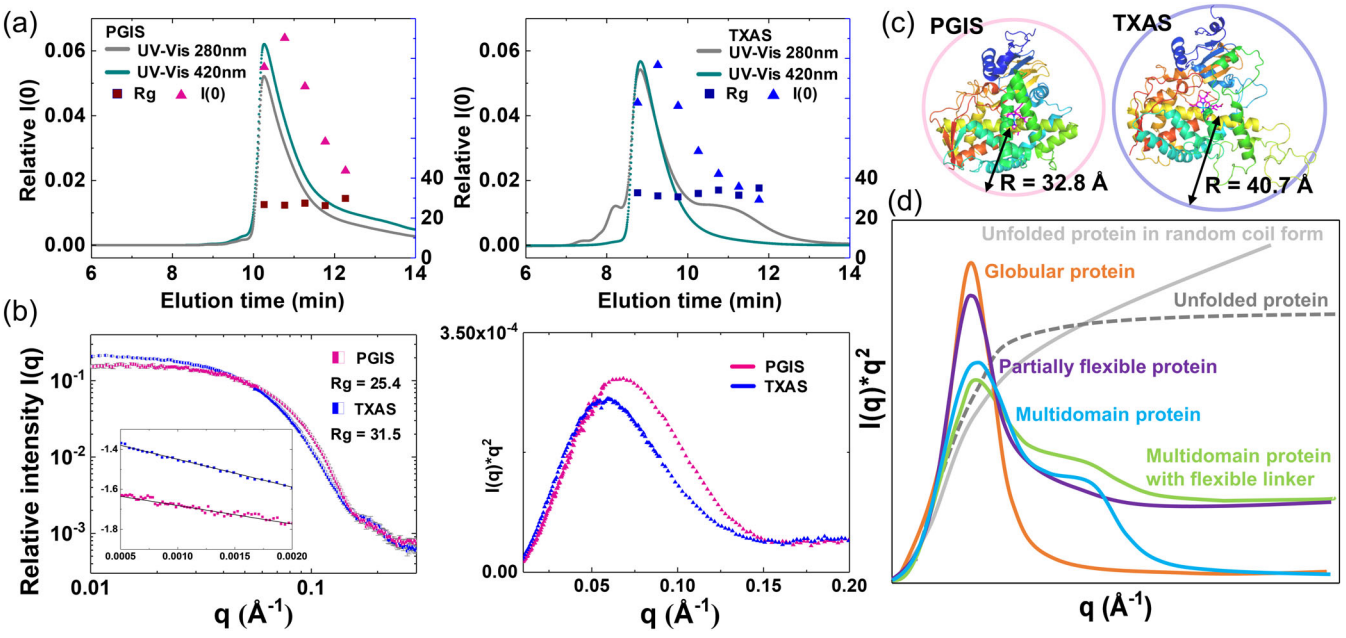

FIGURE 2 The protein sample quality evaluation by the (a) SEC-SAXS synchronization system based on UV–Vis (280/420 nm) verification, and the corresponding radius of gyration (R_g) and I(0) extracted from the SAXS measurements along the elution time. (b) The substrate-free protein SAXS profile and that in the Kratky presentation of $I(q)q^2$ versus q.[11] (c) Protein dimensions estimated using $R = (5/3)^{0.5}R_g$ for spheres or globular protein. (d) The Kratky presentation of typical SAXS profiles for protein conformations of featured flexibility and/or degree of unfolding[9]

in solution, the pronounced TXAS flexibility has unstructured loop regions associated with the characteristic P450 structural elements.[11]

3 | SHAPE AND ENVELOPE STRUCTURE

The ATSAS software in EMBL covers great needs in SAXS data analysis in Figure 3a.[22] But before analyzing the SAXS profile, one must examine carefully SAXS data quality.[14] In practice, SAXS data can be merged from two separated measurements of different q-ranges or concentrations, for one integrated data set that covers a wider q-range with better low-q (contributed by lower concentration samples with less aggregation effect) and high-q data (contributed by higher concentration samples of better data statistics).[11,14,23] Thus, we conduct two experiments separately as necessary. With the merged data, one can employ the code Primus for the R_g value.[23] It is useful to averaged data measured in different frames for improved data statistics. In this regard, data of overlapped profiles are required; profile overlapping can be conveniently examined using the Kratky presentation in Figure 3a;[9] in which, overlapping of the peak maxima of the profiles provides a convenient check for data frames of a same monomeric protein; overlapping of the data in the high-q region can be amplified in the Kratky plots for easier identification.[9]

After a good set of merged SAXS data, we use GNOM to convert the SAXS data into a distance distribution function $p(r)$ in Figure 3a,[24] in which maximum value reveals critically a possible longest dimension of the protein, D_{max}. When protein solution and crystal structures are the same, a corresponding D_{max} can also be derived from the known protein crystal structure for consistency check. Even for proteins of unknown structure, one can estimate the size of the protein through its molecular weight, assuming a globule shape.[14,22] Given an appropriate size, the ATSAS software can provide an estimated value to suggest whether the distance distribution function is reliable,[22] through comparison of the D_{max} value. For proteins of a same mass, however, D_{max} of a globular shape tends to be smaller than that of an extended shape.

One can use GASBOR, DAMMIN, or DAMMIF,[25–27] of different constraints on the building blocks (Figure 3a)

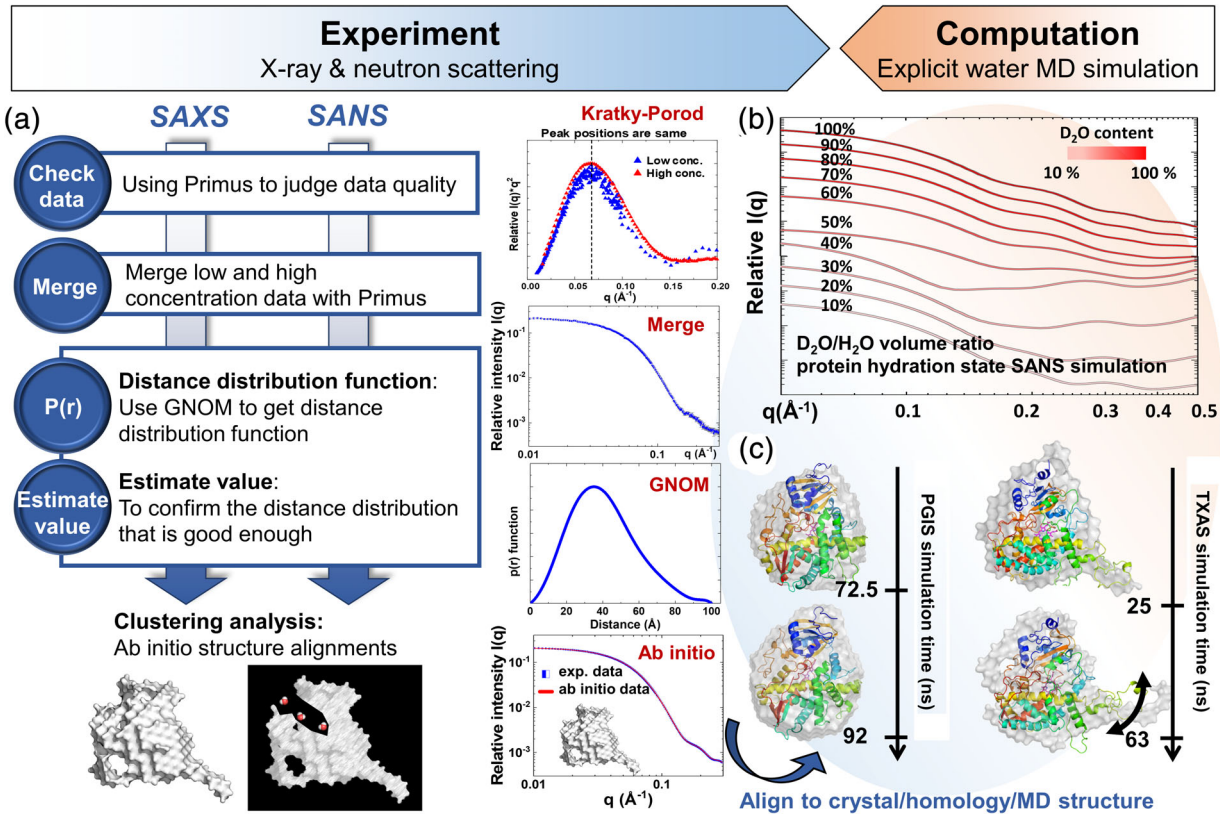

FIGURE 3 Complementarity of X-ray/neutron scattering monuments and explicit water MD simulation. (a) Steps of SAXS and SANS to derive the protein structure (gray envelope structure). (b) The SANS simulation profiles generated by the crystallography/MD structures with CRYSON[20] (c) The solutions SAXS-derived structures overlay the crystallography/MD structure (rainbow ribbon structure) displaying the conformation difference between PGIS and TXAS originated from the protein's flexible region[11,22]

used (for different purposes), to reconstruct an envelope structure with $p(r)$, on the basis of ab initio calculations.[22] We use MD structures to simulated SAXS and SANS profiles (Figure 3b) using in the CRYSOL and CRYSON to quickly select the MD structures that match our experimental data.[19,20] For an evaluation of convergence of the simulated structures generated in repeated runs with the same $p(r)$, normalized spatial discrepancy (NSD) of the simulated structures can be calculated; below unity is required.[22,28] NSD values below unity indicate similar (hence converged) structures; a null NSD value corresponds to identical structures. Furthermore, we can use a low-resolution solution structure of the protein obtained with SANS and SAXS, to confine a high-resolution atomic structure built from the sequence of amid acids upon minimization of free energy. The SAXS-confined protein structure of atomic information hence provides a good basis in revealing the protein dynamic behaviors from the spatial changes of the MD trajectories in Figure 3c;[11] which paves a way to understand the mechanisms in the enzymatic reactions.

4 | CONCLUSION

Currently, it still remains a challenging task in the target membrane protein crystal growth. SAXS are significantly suitable for those structures of non-crystallizable either unsolved or only partially solved, and especially the proteins with flexible linkers. Moreover, biological processes of signal transduction with molecule transport and binding often involve a sequence dynamic structures of protein-water interactions that are difficult to probe. SANS with the contrast variation technique via prepared solvent H-D substation to proteins, together with SAXS, provide complementary information for studying protein solution structures, including topographies of protein hydration structures and protein external/internal water interfaces. Such an integration approach of SAXS and SANS combining MD simulation brings potential impacts on the structural biology research, of particular interest, the topological origins of protein hydration surfaces in aqueous solution and their dynamics to the molecular principles.

ACKNOWLEDGMENTS

This work was supported by grant MOST 109-2113-M-030-002 from the Ministry of Science and Technology of Taiwan. We also thank the TLS BL23A1 and TPS 13A National Synchrotron Radiation Research Center of Taiwan for providing beam time and technical support.

ORCID

Hsiao-Ching Yang https://orcid.org/0000-0002-7970-1357

REFERENCES

[1] L. Liu, S.-H. Chen, A. Faraone, C.-W. Yen, C.-Y. Mou, *Phys. Rev. Lett.* **2005**, *95*, 117802.
[2] A. Faraone, L. Liu, C. Y. Mou, C. W. Yen, S. H. Chen, *J. Chem. Phys.* **2004**, *121*, 10843.
[3] S. H. Chen, L. Liu, E. Fratini, P. Baglioni, A. Faraone, E. Mamontov, *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 9012.
[4] H.-C. Yang, Y.-C. Ge, C.-H. Yang, W.-C. Chao, *ACS Catal.* **2018**, *8*, 2534.
[5] W.-C. Chao, J.-F. Lu, J.-S. Wang, H.-C. Yang, H.-H. Chen, Y.-K. Lan, Y.-C. Yu, P.-T. Chou, L.-H. Wang, *J. Am. Chem. Soc.* **2011**, *133*, 18870.
[6] P. W. Rose, A. Prlić, A. Altunkaya, C. Bi, A. R. Bradley, C. H. Christie, L. D. Costanzo, J. M. Duarte, S. Dutta, Z. Feng, R. K. Green, D. S. Goodsell, B. Hudson, T. Kalro, R. Lowe, E. Peisach, C. Randle, A. S. Rose, C. Shao, Y.-P. Tao, Y. Valasatava, M. Voigt, J. D. Westbrook, J. Woo, H. Yang, J. Y. Young, C. Zardecki, H. M. Berman, S. K. Burley, *Nucleic Acids Res.* **2016**, *45*, D271.
[7] P. Bernadó, E. Mylonas, M. V. Petoukhov, M. Blackledge, D. I. Svergun, *J. Am. Chem. Soc.* **2007**, *129*, 5656.
[8] A. G. Kikhney, D. I. Svergun, *FEBS Lett.* **2015**, *589*, 2570.
[9] V. Receveur-Brechot, D. Durand, *Curr. Protein Peptide Sci.* **2012**, *13*, 55.
[10] Y.-H. Cheng, H.-C. Yang, P.-T. Chou, *J. Phys. Chem. Lett.* **2020**, *11*, 9468.
[11] H.-C. Yang, C.-H. Yang, M.-Y. Huang, J.-F. Lu, J.-S. Wang, Y.-Q. Yeh, U. S. Jeng, J. Phys, *Chem. B* **2017**, *121*, 11229.
[12] M. Sonntag, P. K. A. Jagtap, B. Simon, M. S. Appavou, A. Geerlof, R. Stehle, F. Gabel, J. Hennig, M. Sattler, *Angew. Chem. Int. Ed.* **2017**, *56*, 9322.
[13] L. Boldon, F. Laliberte, L. Liu, *Nano Rev.* **2015**, *6*, 25661.
[14] J. Trewhella, A. P. Duff, D. Durand, F. Gabel, J. M. Guss, W. A. Hendrickson, G. L. Hura, D. A. Jacques, N. M. Kirby, A. H. Kwan, J. Perez, L. Pollack, T. M. Ryan, A. Sali, D. Schneidman-Duhovny, T. Schwede, D. I. Svergun, M. Sugiyama, J. A. Tainer, P. Vachette, J. Westbrook, A. E. Whitten, *Acta Crystallogr. D* **2017**, *73*, 710.
[15] D. Liu, Y. Zhang, Y. Liu, J. Wu, C.-C. Chen, C.-Y. Mou, S.-H. Chen, *J. Phys. Chem. B* **2008**, *112*, 4309.
[16] D. Liu, Y. Zhang, C.-C. Chen, C.-Y. Mou, P. H. Poole, S.-H. Chen, *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 9570.
[17] Y.-Q. Yeh, K.-F. Liao, O. Shih, Y.-J. Shiu, W.-R. Wu, C.-J. Su, P.-C. Lin, U. S. Jeng, *J. Phys. Chem. Lett.* **2017**, *8*, 470.
[18] O. Shih, Y.-Q. Yeh, K.-F. Liao, C.-J. Su, P.-H. Wu, R. K. Heenan, T.-Y. Yu, U. S. Jeng, *J. Phys. Chem. Lett.* **2018**, *9*, 4287.
[19] D. Svergun, C. Barberato, M. H. J. Koch, *J. Appl. Crystallogr.* **1995**, *28*, 768.
[20] D. I. Svergun, S. Richard, M. H. J. Koch, Z. Sayers, S. Kuprin, G. Zaccai, *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 2267.
[21] E. Brookes, P. Vachette, M. Rocco, J. Perez, *J. Appl. Crystallogr.* **2016**, *49*, 1827.
[22] D. Franke, M. V. Petoukhov, P. V. Konarev, A. Panjkovich, A. Tuukkanen, H. D. T. Mertens, A. G. Kikhney, N. R.

Hajizadeh, J. M. Franklin, C. M. Jeffries, D. I. Svergun, *J. Appl. Crystallogr.* **2017**, *50*, 1212.
[23] P. V. Konarev, V. V. Volkov, A. V. Sokolova, M. H. J. Koch, D. I. Svergun, *J. Appl. Crystallogr.* **2003**, *36*, 1277.
[24] A. V. Semenyuk, D. I. Svergun, *J. Appl. Crystallogr.* **1991**, *24*, 537.
[25] D. I. Svergun, M. V. Petoukhov, M. H. Koch, *Biophys. J.* **2001**, *80*, 2946.
[26] D. I. Svergun, *Biophys. J.* **1999**, *76*, 2879.
[27] D. Franke, D. I. Svergun, *J. Appl. Crystallogr.* **2009**, *42*, 342.
[28] M. B. Kozin, D. I. Svergun, *J. Appl. Crystallogr.* **2001**, *34*, 33.
[29] K. Wood, J. P. Mata, C. J. Garvey, C.-M. Wu, W. A. Hamilton, P. Abbeywick, D. Bartlett, F. Bartsch, P. Baxter, N. Booth, W. Brown, J. Christoforidis, D. Clowes, T. d'Adam, F. Darmann, M. Deura, S. Harrison, N. Hauser, G. Horton, D. Federici, F. Franceschini, P. Hanson, E. Imamovic, P. Imperia, M. Jones, S. Kennedy, S. Kim, T. Lam, W. T. Lee, M. Lesha, D. Mannicke, T. Noakes, S. R. Olsen, J. C. Osborn, D. Penny, M. Perry, S. A. Pullen, R. A. Robinson, J. C. Schulz, N. Xiong, E. P. Gilbert, *J. Appl. Crystallogr.* **2018**, *51*, 294.

AUTHOR BIOGRAPHIES

Shang-Wei Lin Shang-Wei Lin is a Ph.D. student studying under Prof. Hsiao-Ching Yang at the department of chemistry, Fu Jen Catholic University. His research is to develop molecular dynamics (MD) simulation combined small-angle X-ray scattering (SAXS) and small-angle neutron scattering (SANS) for determining solution-state protein structures and characterizing the enzyme dynamics of cytochrome P450 proteins. He graduated with his M. Sc. in Fu Jen Catholic University in 2020.

Kuan-Hsuan Su Kuan-Hsuan Su is a masters student studying under Prof. Hsiao-Ching Yang at the department of chemistry, Fu Jen Catholic University. Her interests are CYP450 enzyme catalysis reactions through molecular dynamics simulation, exploiting the driving forces and stereoselectivity mechanisms of enzyme kinetics, animating the static crystal structure to explore the three-dimensional structure characteristics between molecules through molecular-level visualization platforms, and using various energy calculations to analyze the interactions among substrates, amino acids, and waters. She received her B.S. in chemistry from Fu Jen Catholic University in 2019.

Chun-Ming Wu Dr. Chun-Ming Wu is an instrument scientist for Taiwan's first cold-neutron triple-axis spectrometer—SIKA and two SANS instruments employed by National Synchrotron Radiation Research Center (NSRRC) station at Australian Neutron Science and Technology Organization (ANSTO). He is nominated as the NSRRC representative at ANSTO since 2018. Dr. Wu was the SIKA project officer to construct SIKA at ANSTO for the 8 million Australian dollar project in 2007–2013. He received his Ph.D. degree of Chemistry in 2005 from the Fu-Jen Catholic University (FJCU), Taiwan. His expertise is in neutron scattering techniques, and more recent work is focused on developing the neutron technology for the advanced GI-SANS instrument designs and data analysis.

Hsiao-Ching Yang Hsiao-Ching Yang, Professor, department of chemistry, Fu Jen Catholic University. Prof. Yang has focused on the SAXS/SANS data analysis by incorporating innovated methodologies of computational chemical dynamics that allow one to meaningfully interpret the scattering data from monodisperse systems, from transient complexes as well as flexible and heterogeneous systems in terms of structural models. Her group's current interest has focused on the hybrid approaches utilizing SAXS/SANS with high-resolution molecular dynamics simulation techniques, but also with biochemical, biophysical, and numerical programming methods by a series of breakthrough experiments with the synchrotron/neutron beamline (TLS 23A and TPS 13A at NSRRC/QUOKKA at ANSTO) and developed computational techniques for solving the inverse mapping problems of X-ray/neutron scattering.

How to cite this article: Lin S-W, Su K-H, Yeh Y-Q, Jeng U-S, Wu C-M, Yang H-C. Molecular dynamics simulation combined with small-angle X-ray/neutron scattering defining solution-state protein structures. *J Chin Chem Soc*. 2021;68: 403–408. https://doi.org/10.1002/jccs.202000498