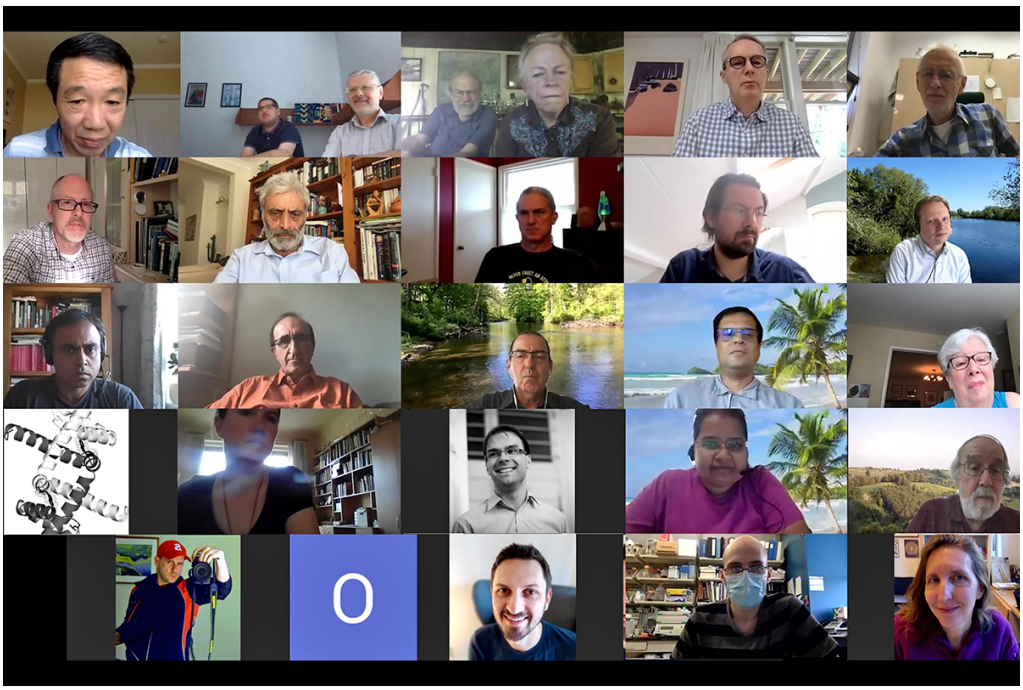
IUCr Newsletter (2020) Volume 28, number 4

International Union of Crystallography
NEWSLETTER

Special report

Nucleic Acid Valence Geometry Working Group

Bohdan Schneider, Ian Bruno, Stephen K. Burley, David A. Case, Jiří Černý, Rhiju Das, Martin Egli, Paul Emsley, Zukang Feng, Mariusz Jaskolski, Robbie P. Joosten, Catherine L. Lawson, Nigel W. Moriarty, Abhik Mukhopadhyay, Garib Murshudov, Gary Parkinson, Jane S. Richardson, Oleg V. Sobolev, Brinda Vallat, Sameer Velankar, Evan Worden, Ivan Nikolay Zheludev and Helen M. Berman



Participants in the virtual workshop to establish the Nucleic Acid Valence Geometry Working Group.

This communication reports on an initiative to establish the Nucleic Acid Valence Geometry Working Group, whose task will be to define and implement a uniform dictionary for nucleic acid valence geometry parameters for use in modeling, refinement and validation systems.

Biomolecular structure refinement depends both on experimental data and on chemical restraints that encode prior knowledge regarding acceptable geometries. Chemical restraints used for nucleic acid (NA) valence geometries (bond lengths and angles) in the major refinement and validation software packages are divergent, especially in the standard deviations used and in whether they are dependent on sugar pucker, RNA versus DNA, etc. This impacts many aspects of NA-containing structure studies including model building, refinement and validation. We propose to review those dictionaries and protocols, first seeking agreement on a common compromise.

This effort is an initiative of the ELIXIR Community on Structural Bioinformatics, 3D-Bioinfo (Orengo *et al.*, 2020) "to encourage development and use of software tools to describe, analyze, annotate, and predict nucleic acid (NA) structures. In particular, standards are needed for initial model building and refinement of nucleic acid molecular structures." This task has become urgent as techniques such as cryo-EM (Afonine *et al*., 2018; Lawson *et al.*, 2020) and high-resolution neutron diffraction (Liebschner *et al*., 2019) are generating experimental data on 3D structures containing RNA and DNA molecules.

To initiate the process, we held a virtual meeting on 24–25 August 2020. More than 20 subject-matter experts in the field discussed the status of standards for valence geometry in nucleic acids. We heard 15 talks on the topic from various perspectives. We learned that the most commonly used software tools for refinement and validation use different sets of values for the valence geometries. Of particular concern is the fact that the refinement programs *REFMAC* (Murshudov *et al.*, 2011) and *Phenix* (Liebschner *et al.*, 2019), which includes X-ray, neutron and cryo-EM, use somewhat different values. The wwPDB (Berman *et al*., 2003) validation pipeline (Gore *et al*., 2017) uses the values published on the current NDB site (Berman *et al*., 1992) while *Phenix* uses the Molprobity (Williams *et al*., 2018) values. Lack of consistency between refinement and validation dictionaries leads to validation reports that may identify unforeseen or even spurious outliers.

Valence geometry standards have traditionally been determined separately for the nitrogenous bases and the sugar-phosphate backbone. The first set of values was determined by the Kennard/CSD group in 1982 (Taylor & Kennard, 1982*a*) (Taylor & Kennard, 1982*b*). Using a larger number of crystal structures, somewhat more detailed parameters were determined for the bases (Clowney *et al.*, 1996) and the sugar backbone (Gelbin *et al.*, 1996; Parkinson *et al*., 1996). Current refinement and

validation programs use some combination of those 1996 values. Recently an updated set of CSD-based values was developed again separately for the bases (Gilski *et al*., 2019), phosphates (Kowiel *et al.*, 2016) and sugars (Kowiel *et al*., 2020). The latter two analyses show the correlation between the valence geometry and conformation. These values have been tested for a limited number of refinements. The complicated geometry of the phosphodiester linkage was also investigated as a function of charge (Schneider *et al*., 1996).

Extensive discussion at the workshop resulted in two conclusions. One is that the valence geometry standards cannot be isolated from the other geometric and conformational descriptors of nucleic acids, including sugar puckers, phosphate conformations, base pair geometry and hydrogen bonding. These all need updates of standards using state-of-the-art statistical tools and well-curated sets of experimental data before adoption, especially by the rightly conservative wwPDB validation. Ways to find correlated geometry values for larger building blocks such as the dinucleotide unit were suggested and need to be further discussed. A second conclusion is that in the short term, while that research is being done and evaluated by the community, it is very important for all stakeholders to agree on a uniform, optimal compromise for the simpler restraint values currently being used.

Outcomes and recommendations of this Workshop are the following:

- To establish the Nucleic Acid Valence Geometry Working Group, whose task will be to make these changes happen.
- The short-term goal of the Working Group will be to define, obtain agreement on, and implement a conservative and uniform compromise on simple parameters for current use in all the major modeling, refinement and validation systems.
- More complex definitions that take into account local identities and conformations should be compared, tested and tried out by the community, to reach a more accurate, next-generation set of valence parameters.
- Further research and testing should eventually be expanded to conformationally (primarily dihedral-angle) dependent restraints.

Acknowledgements

This project has been supported by the implementation plan of the ELIXIR Community on Structural Bioinformatics, 3D-Bioinfo.

References

Afonine, P. V., Poon, B. K., Read, R. J., Sobolev, O. V., Terwilliger, T. C., Urzhumtsev, A. & Adams, P. D. (2018). *Acta Cryst*. D**74**, 531–544.

Berman, H. M., Olson, W. K., Beveridge, D. L., Westbrook, J., Gelbin, A., Demeny, T., Hsieh, S. H., Srinivasan, A. R. & Schneider, B. (1992). *Biophys. J.* **63**, 751–759.

Berman, H. M., Henrick, K. & Nakamura, H. (2003). *Nat. Struct. Biol.* **10**, 980.

Clowney, L., Jain, S. C., Srinivasan, A. R., Westbrook, J., Olson, W. K. & Berman, H. M. (1996). *J. Am. Chem. Soc.* **118**, 509–518.

Gelbin, A., Schneider, B., Clowney, L., Hsieh, S.-H., Olson, W. K. & Berman, H. M. (1996). *J. Am. Chem. Soc.* **118**, 519–528.

Gilski, M., Zhao, J., Kowiel, M., Brzezinski, D., Turner, D. H. & Jaskolski, M. (2019). *Acta Cryst*. B**75**, 235–245.

Gore, S., Sanz Garcia, E., Hendrickx, P. M. S., Gutmanas, A., Westbrook, J. D., Yang, H., Feng, Z., Baskaran, K., Berrisford, J. M., Hudson, B. P., Ikegawa, Y., Kobayashi, N., Lawson, C. L., Mading, S., Mak, L., Mukhopadhyay, A., Oldfield, T. J., Patwardhan, A., Peisach, E., Sahni, G., Sekharan, M. R., Sen, S., Shao, C., Smart, O. S., Ulrich, E. L., Yamashita, R., Quesada, M., Young, J. Y., Nakamura, H., Markley, J. L., Berman, H. M., Burley, S. K., Velankar, S. & Kleywegt, G. J. (2017). *Structure*, **25**, 1916–1927.

Kowiel, M., Brzezinski, D., Gilski, M. & Jaskolski, M. (2020). *Nucleic Acids Res*. **48**, 962–973.

Kowiel, M., Brzezinski, D. & Jaskolski, M. (2016). *Nucleic Acids Res*. **44**, 8479–8489.

Lawson, C. L., Berman, H. M. & Chiu, W (2020). *Struct. Dynam*. **7**, 014701.

Liebschner, D., Afonine, P. V., Baker, M. L., Bunkóczi, G., Chen, V. B., Croll, T. I., Hintze, B., Hung, L.-W., Jain, S., McCoy, A. J., Moriarty, N. W., Oeffner, R. D., Poon, B. K., Prisant, M. G., Read, R. J., Richardson, J. S., Richardson, D. C., Sammito, M. D., Sobolev, O. V., Stockwell, D. H., Terwilliger, T. C., Urzhumtsev, A. G., Videau, L. L., Williams, C. J. & Adams, P. D. (2019). *Acta Cryst*. D**75**, 861–877.

Murshudov, G. N., Skubak, P., Lebedev A. A., Pannu N. S., Steiner R. A., Nicholls R. A, Winn M. D., Long F. & Vagin, A. A. (2011). *Acta Cryst*. D**67**, 355–367.

Orengo, C., Velankar, S., Wodak, S., Zoete, V., Bonvin, A., Elofsson, A., Feenstra, K. A., Gerloff, D. L., Hamelryck, T., Hancock, J. M., Helmer-Citterich, M., Hospital, A., Orozco, M., Perrakis, A., Rarey, M., Soares, C., Sussman, J. L., Thornton, J. M., Tuffery, P., Tusnady, G., Wierenga, R., Salminen, T. & Schneider, B. (2020). F1000Res 9.

Parkinson, G., Vojtechovsky, J., Clowney, L., Brunger, A. T. & Berman, H. M. (1996). *Acta Cryst.* D**52**, 57–64.

Schneider, B., Kabelac, M. & Hobza, P. (1996). *J. Am. Chem. Soc.* **118**, 12207–12217.

Taylor, R. & Kennard, O. (1982*a*). *J. Am. Chem. Soc.* **104**, 3209–3212.

Taylor, R. & Kennard, O. (1982*b*). *J. Mol. Struct.* **78**, 1–28.

Williams, C. J., Headd, J. J., Moriarty, N. W., Prisant, M. G., Videau, L. L., Deis, L. N., Verma, V., Keedy, D. A., Hintze, B. J., Chen, V. B., Jain, S., Lewis, S. M., Arendall, W. B., 3rd, Snoeyink, J., Adams, P. D., Lovell, S. C., Richardson, J. S. & Richardson, D. C. (2018). *Protein Sci*. **27**, 293–315.

22 November 2020

Copyright © - All Rights Reserved - International Union of Crystallography

SUBSCRIBE SUBMIT ADVERTISE

EDITORIAL BOARD

DECTRIS
detecting the future

EIGER2 R CdTe

HPC for any wavelength!

- Fast data acquisition
with high efficiency from Cu to In

- True photon counting for superior dynamic range

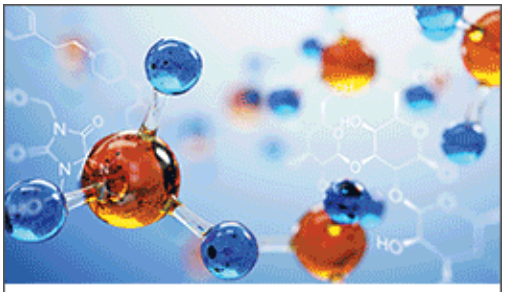
- Better spot separation
thanks to small pixels and direct detection

XtaLAB Synergy Flow

A Revolution in High-Throughput Crystallography

LEARN MORE

Rigaku



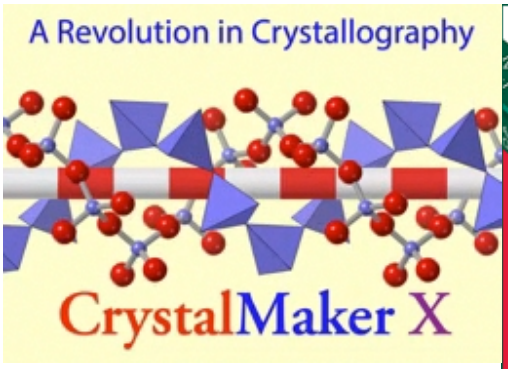
- Real-time XRD
- Revolutionary Detector
- Industry & Research
- 21 CFR 11 Compliance

Find out more

ThermoFisher
SCIENTIFIC

PHOTON III Mixed-mode Detectors –
A better way to count photons

BRUKER


A Revolution in Crystallography
CrystalMaker X

50% discount*
for IUCr Associates
PLEASE CLICK HERE FOR MORE INFORMATION
*Off the purchase price of printed volumes for personal use.

IUCrData
The IUCr's peer-reviewed open-access data publication