

The seventh blind test highlights exciting developments in crystal structure prediction

Mihails Arhangelskis*

Faculty of Chemistry, University of Warsaw, 1 Pasteura Street, Warsaw, 02-093, Poland. *Correspondence e-mail: m.arhangelskis@uw.edu.pl

Keywords: crystal structure prediction; polymorphism; crystal forms; energy ranking.

The ability to predict possible crystal structures starting from the knowledge of molecular structure alone is a challenging, yet immensely attractive task, directly linked to materials design and crystal form screening. Being able to anticipate whether a given compound can potentially yield new, so far undiscovered polymorphs, is of paramount importance in the pharmaceutical industry, where the sudden occurrence of a new polymorph may require a complete redesign of the manufacturing process, as has been demonstrated in the famous case of a disappearing polymorph of Ritonavir (Chemburkar *et al.*, 2000).

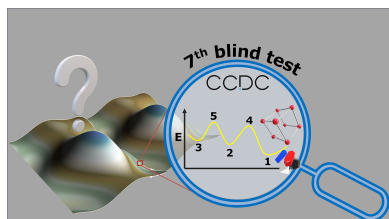
The blind tests of crystal structure prediction (CSP), organized by the Cambridge Crystallographic Data Centre (CCDC), have been instrumental in monitoring the progress in CSP method development, as well as gauging the limits of structural complexity that these methods can tackle. Since the inception in 1999, a total of six blind tests have been previously conducted (Lommerse *et al.*, 2000; Motherwell *et al.*, 2002; Day *et al.*, 2005; Day *et al.*, 2009; Bardwell *et al.*, 2011; Reilly *et al.*, 2016) with the results of the latest blind test being reported in the current issue of *Acta Crystallographica B, Structural Science, Crystal Engineering and Materials*, with two separate articles focusing on the key parts of the CSP process, namely on structure generation and energy ranking (Hunnisett, Nyman *et al.*, 2024; Hunnisett, Francia *et al.*, 2024).

The general organization of the blind test has remained unchanged: the participating research groups were provided with molecular diagrams of seven target systems, for which experimental crystal structures had been determined yet temporarily withheld from publication. The biggest change compared to the previous blind tests was the introduction of separate structure generation and energy ranking phases. In the newly introduced ranking phase, all participating groups were given identical sets of generated structures for each target system, and were tasked to rank the energies with their methods of choice.

Compared with the previous blind tests, the current edition presented several new challenges. In terms of the diversity of constituent elements, for the first time, the blind test exceeded the boundaries of purely organic materials, with an introduction of a copper coordination complex target (molecule XXVIII), as well as a silicon and iodine-containing molecule (molecule XXVII).

In addition to elemental complexity, participants of the blind test faced the challenge of greater crystallographic complexity of the target systems. This blind test placed particular emphasis on the ability to predict structures with multiple known polymorphs, including the particularly challenging case of system XXXII, represented by eight experimentally confirmed polymorphs. The complexity of this particular system was not limited to polymorphic diversity: the corresponding molecule contained 11 rotatable bonds, signifying great conformational flexibility.

The capabilities of modern CSP methods were also tested in the realm of multi-component solids. While structures of solvates, salts and cocrystals have been present in the previous blind tests, their stoichiometries had always been precisely defined. This time the participants were faced with a new challenge of discovering the appropriate cocrystal stoichiometry, having been told that the experimental crystal forms possess two of the possible three stoichiometric ratios of 1:1, 2:1 or 1:2. This added an additional degree of complexity to the structure search, which required an adaptation of various algorithms for the treatment of variable stoichiometry treatment, including convex hull search, consideration of lattice energies of individual crystal components, as well as consideration of the number of hydrogen-bond donor and acceptor groups in the



constituent molecules. Finally, four out of seven target systems contained disorder in their structures, with a potential implication for both structure generation and energy ranking parts of the search. In terms of structure generation, the presence of disorder may give rise to additional energy minima, which would not have been found during a fully ordered structure search. In terms of energy ranking, disorder is associated with increased entropy, and may lead to additional stabilization and re-ranking of the predicted structures in terms of their free energy.

Notable developments have been found in structure ranking methods. In addition to the force field and dispersion-corrected density-functional theory methods, for the first time we are seeing several participating groups adopting machine-learned potentials for energy ranking. More emphasis has been placed on going beyond static lattice energies, with the introduction of the vibrational contribution to the free energy, either using phonon or molecular dynamics simulations. Finally, a non-energetic ranking based on topological descriptors has been adopted by one of the groups.

Given the importance of CSP as an applied method for studying crystal energy landscapes and pre-emptively discovering yet-unknown crystal forms, the key emphasis in method development must be placed on maintaining the computational efficiency. With this in mind, for the first time, all participating groups were requested to report the computational cost of their CSP searches, allowing consideration of the time and energy cost associated with CSP calculations in relation to experimental crystal form screening. It is evident that the future of CSP, particularly for wide adaptation in industry, lies in finding the right balance between the accuracy of the calculations and their computational cost. No doubt, these considerations will affect the development of CSP methods, and will be reflected in the future iterations of the blind tests.

To conclude, the seventh blind test has given us an excellent demonstration of the capabilities of modern CSP methods and outlined the directions of future development for this exciting area of research. It has also showcased the amazing team effort of the CSP community, with a record-breaking 28 participating groups from around the world. I am sure that this blind test report will convince everyone that the future of CSP is bright!

Acknowledgements

The author would like to thank Lavanya Kumar and Dr Sivakumar Gurusamy for preparing the image for the table of contents.

References

Bardwell, D. A., Adjiman, C. S., Arnautova, Y. A., Bartashevich, E., Boerrigter, S. X. M., Braun, D. E., Cruz-Cabeza, A. J., Day, G. M., Della Valle, R. G., Desiraju, G. R., van Eijck, B. P., Facelli, J. C., Ferraro, M. B., Grillo, D., Habgood, M., Hofmann, D. W. M., Hofmann, F., Jose, K. V. J., Karamertzanis, P. G., Kazantsev, A. V.,

Kendrick, J., Kuleshova, L. N., Leusen, F. J. J., Maleev, A. V., Misquitta, A. J., Mohamed, S., Needs, R. J., Neumann, M. A., Nikylov, D., Orendt, A. M., Pal, R., Pantelides, C. C., Pickard, C. J., Price, L. S., Price, S. L., Scheraga, H. A., van de Streek, J., Thakur, T. S., Tiwari, S., Venuti, E. & Zhitkov, I. K. (2011). *Acta Cryst.* **B67**, 535–551.

Chemburkar, S. R., Bauer, J., Deming, K., Spiwek, H., Patel, K., Morris, J., Henry, R., Spanton, S., Dziki, W., Porter, W., Quick, J., Bauer, P., Donaubaue, J., Narayanan, B. A., Soldani, M., Riley, D. & McFarland, K. (2000). *Org. Process Res. Dev.* **4**, 413–417.

Day, G. M., Cooper, T. G., Cruz-Cabeza, A. J., Hejczyk, K. E., Ammon, H. L., Boerrigter, S. X. M., Tan, J. S., Della Valle, R. G., Venuti, E., Jose, J., Gadre, S. R., Desiraju, G. R., Thakur, T. S., van Eijck, B. P., Facelli, J. C., Bazterra, V. E., Ferraro, M. B., Hofmann, D. W. M., Neumann, M. A., Leusen, F. J. J., Kendrick, J., Price, S. L., Misquitta, A. J., Karamertzanis, P. G., Welch, G. W. A., Scheraga, H. A., Arnautova, Y. A., Schmidt, M. U., van de Streek, J., Wolf, A. K. & Schweizer, B. (2009). *Acta Cryst.* **B65**, 107–125.

Day, G. M., Motherwell, W. D. S., Ammon, H. L., Boerrigter, S. X. M., Della Valle, R. G., Venuti, E., Dzyabchenko, A., Dunitz, J. D., Schweizer, B., van Eijck, B. P., Erk, P., Facelli, J. C., Bazterra, V. E., Ferraro, M. B., Hofmann, D. W. M., Leusen, F. J. J., Liang, C., Pantelides, C. C., Karamertzanis, P. G., Price, S. L., Lewis, T. C., Nowell, H., Torrisi, A., Scheraga, H. A., Arnautova, Y. A., Schmidt, M. U. & Verwer, P. (2005). *Acta Cryst.* **B61**, 511–527.

Hunnisett, L. M., Francia, N., Nyman, J., Abraham, N. S., Aitipamula, S., Alkhalid, T., Almehairbi, M., Anelli, A., Anstine, D. M., Anthony, J. E., Arnold, J. E., Bahrami, F., Bellucci, M. A., Beran, G. J. O., Bhardwaj, R. M., Bianco, R., Bis, J. A., Boese, A. D., Bramley, J., Braun, D. E., Butler, P. W. V., Cadden, J., Carino, S., Cervinka, C., Chan, E. J., Chang, C., Clarke, S. M., Coles, S. J., Cook, C. J., Cooper, R. I., Darden, T., Day, G. M., Deng, W., Dietrich, H., DiPasquale, A., Dhokale, B., van Eijck, B. P., Elsegood, M. R. J., Firaha, D., Fu, W., Fukuzawa, K., Galanakis, N., Goto, H., Greenwell, C., Guo, R., Harter, J., Helfferich, J., Hoja, J., Hone, J., Hong, R., Hušák, M., Ikabata, Y., Isayev, O., Ishaque, O., Jain, V., Jin, Y., Jing, A., Johnson, E. R., Jones, I., Jose, K. V. J., Kabova, E. A., Keates, A., Kelly, P. F., Klimeš, J., Kostková, V., Li, H., Lin, X., List, A., Liu, C., Liu, Y. M., Liu, Z., Lončarić, I., Lubach, J. W., Ludík, J., Marom, N., Matsui, H., Mattei, A., Mayo, R. A., Melkumov, J. W., Mladineo, B., Mohamed, S., Momenzadeh Abardeh, Z., Muddana, H. S., Nakayama, N., Nayal, K. S., Neumann, M. A., Nikhar, R., Obata, S., O'Connor, D., Oganov, A. R., Okuwaki, K., Otero-de-la-Roza, A., Parkin, S., Parunov, A., Podeszwa, R., Price, A. J. A., Price, L. S., Price, S. L., Probert, M. R., Pulido, A., Ramteke, G. R., Rehman, A. U., Reutzel-Edens, S. M., Rogal, J., Ross, M. J., Rumson, A. F., Sadiq, G., Saeed, Z. M., Salimi, A., Sasikumar, K., Sekharan, S., Shankland, K., Shi, B., Shi, X., Shinohara, K., Skillman, A. G., Song, H., Strasser, N., van de Streek, J., Sugden, I. J., Sun, G., Szalewicz, K., Tan, L., Tang, K., Tarczynski, F., Taylor, C. R., Tkatchenko, A., Tom, R., Touš, P., Tuckerman, M. E., Unzueta, P. A., Utsumi, Y., Vogt-Maranto, L., Weatherston, J., Wilkinson, L. J., Willacy, R. D., Wojtas, L., Woollam, G. R., Yang, Y., Yang, Z., Yonemochi, E., Yue, X., Zeng, Q., Zhou, T., Zhou, Y., Zubatyuk, R. & Cole, J. C. (2024). *Acta Cryst.* **B80**, 548–574.

Hunnisett, L. M., Nyman, J., Francia, N., Abraham, N. S., Adjiman, C. S., Aitipamula, S., Alkhalid, T., Almehairbi, M., Anelli, A., Anstine, D. M., Anthony, J. E., Arnold, J. E., Bahrami, F., Bellucci, M. A., Bhardwaj, R. M., Bier, I., Bis, J. A., Boese, A. D., Bowskill, D. H., Bramley, J., Brandenburg, J. G., Braun, D. E., Butler, P. W. V., Cadden, J., Carino, S., Chan, E. J., Chang, C., Cheng, B., Clarke, S. M., Coles, S. J., Cooper, R. I., Couch, R., Cuadrado, R., Darden, T., Day, G. M., Dietrich, H., Ding, Y., DiPasquale, A., Dhokale, B., van Eijck, B. P., Elsegood, M. R. J., Firaha, D., Fu, W., Fukuzawa, K., Glover, J., Goto, H., Greenwell, C., Guo, R., Harter, J., Helfferich, J., Hofmann, D. W. M., Hoja, J., Hone, J., Hong, R.,

- Hutchison, G., Ikabata, Y., Isayev, O., Ishaque, O., Jain, V., Jin, Y., Jing, A., Johnson, E. R., Jones, I., Jose, K. V. J., Kabova, E. A., Keates, A., Kelly, P. F., Khakimov, D., Konstantinopoulos, S., Kuleshova, L. N., Li, H., Lin, X., List, A., Liu, C., Liu, Y. M., Liu, Z., Liu, Z.-P., Lubach, J. W., Marom, N., Maryewski, A. A., Matsui, H., Mattei, A., Mayo, R. A., Melkumov, J. W., Mohamed, S., Momenzadeh Abardeh, Z., Muddana, H. S., Nakayama, N., Nayal, K. S., Neumann, M. A., Nikhar, R., Obata, S., O'Connor, D., Oganov, A. R., Okuwaki, K., Otero-de-la-Roza, A., Pantelides, C. C., Parkin, S., Pickard, C. J., Pilia, L., Pivina, T., Podeszwa, R., Price, A. J. A., Price, L. S., Price, S. L., Probert, M. R., Pulido, A., Ramteke, G. R., Rehman, A. U., Reutzel-Edens, S. M., Rogal, J., Ross, M. J., Rumson, A. F., Sadiq, G., Saeed, Z. M., Salimi, A., Salvalaglio, M., Sanders de Almada, L., Sasikumar, K., Sekharan, S., Shang, C., Shankland, K., Shinohara, K., Shi, B., Shi, X., Skillman, A. G., Song, H., Strasser, N., van de Streek, J., Sugden, I. J., Sun, G., Szalewicz, K., Tan, B. I., Tan, L., Tarczynski, F., Taylor, C. R., Tkatchenko, A., Tom, R., Tuckerman, M. E., Utsumi, Y., Vogt-Maranto, L., Weatherston, J., Wilkinson, L. J., Willacy, R. D., Wojtas, L., Woollam, G. R., Yang, Z., Yonemochi, E., Yue, X., Zeng, Q., Zhang, Y., Zhou, T., Zhou, Y., Zubatyuk, R. & Cole, J. C. (2024). *Acta Cryst.* **B80**, 517–547.
- Lommerse, J. P. M., Motherwell, W. D. S., Ammon, H. L., Dunitz, J. D., Gavezzotti, A., Hofmann, D. W. M., Leusen, F. J. J., Mooij, W. T. M., Price, S. L., Schweizer, B., Schmidt, M. U., van Eijck, B. P., Verwer, P. & Williams, D. E. (2000). *Acta Cryst.* **B56**, 697–714.
- Motherwell, W. D. S., Ammon, H. L., Dunitz, J. D., Dzyabchenko, A., Erk, P., Gavezzotti, A., Hofmann, D. W. M., Leusen, F. J. J., Lommerse, J. P. M., Mooij, W. T. M., Price, S. L., Scheraga, H., Schweizer, B., Schmidt, M. U., van Eijck, B. P., Verwer, P. & Williams, D. E. (2002). *Acta Cryst.* **B58**, 647–661.
- Reilly, A. M., Cooper, R. I., Adjiman, C. S., Bhattacharya, S., Boese, A. D., Brandenburg, J. G., Bygrave, P. J., Bylsma, R., Campbell, J. E., Car, R., Case, D. H., Chadha, R., Cole, J. C., Cosburn, K., Cuppen, H. M., Curtis, F., Day, G. M., DiStasio, R. A. Jr, Dzyabchenko, A., van Eijck, B. P., Elking, D. M., van den Ende, J. A., Facelli, J. C., Ferraro, M. B., Fusti-Molnar, L., Gatsiou, C.-A., Gee, T. S., de Gelder, R., Ghiringhelli, L. M., Goto, H., Grimme, S., Guo, R., Hofmann, D. W. M., Hoja, J., Hylton, R. K., Iuzzolino, L., Jankiewicz, W., de Jong, D. T., Kendrick, J., de Klerk, N. J. J., Ko, H.-Y., Kuleshova, L. N., Li, X., Lohani, S., Leusen, F. J. J., Lund, A. M., Lv, J., Ma, Y., Marom, N., Masunov, A. E., McCabe, P., McMahon, D. P., Meekes, H., Metz, M. P., Misquitta, A. J., Mohamed, S., Monserrat, B., Needs, R. J., Neumann, M. A., Nyman, J., Obata, S., Oberhofer, H., Oganov, A. R., Orendt, A. M., Pagola, G. I., Pantelides, C. C., Pickard, C. J., Podeszwa, R., Price, L. S., Price, S. L., Pulido, A., Read, M. G., Reuter, K., Schneider, E., Schober, C., Shields, G. P., Singh, P., Sugden, I. J., Szalewicz, K., Taylor, C. R., Tkatchenko, A., Tuckerman, M. E., Vacarro, F., Vasileiadis, M., Vazquez-Mayagoitia, A., Vogt, L., Wang, Y., Watson, R. E., de Wijs, G. A., Yang, J., Zhu, Q. & Groom, C. R. (2016). *Acta Cryst.* **B72**, 439–459.