STRUCTURAL BIOLOGY COMMUNICATIONS

ISSN 2053-230X

A note on the appearance of PEG in macromolecular crystals

Alexander McPherson*‡

Department of Molecular Biology and Biochemistry, University of California, Irvine, Irvine, California, USA. *Correspondence e-mail: amcphers@uci.edu

Received 30 September 2025
Accepted 19 November 2025

Edited by J. Agirre, University of York, United Kingdom

‡ Currently Visiting Professor at Scripps Research Institute, Jupiter, Florida, USA.

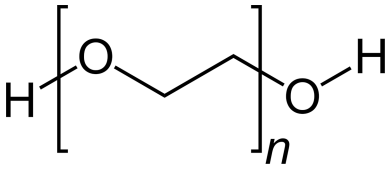
Keywords: polyethylene glycol; mechanism; crystallization; refinement; modeling.

In protein crystal structures that achieve high resolution, 1.5 Å or higher, what we commonly refer to as the 'solvent regions' are distinctively different depending on whether the crystals were grown using polyethylene glycol (PEG) as the precipitating agent or whether the crystals were grown from a salt solution or using some other PEG-free mother liquor, such as 2-methyl-2,4-pentanediol (MPD), or low ionic strength. $2mF_o - F_c$ and other difference Fourier syntheses intended to show ordered features present in the 'solvent regions' of crystals from 'non-PEG mother liquors' can generally be satisfactorily interpreted in terms of well defined water molecules and salt ions against a background of disordered solvent. Ordered water molecules are usually discrete and separate entities (neglecting water molecules with alternate locations, which tend to elongate densities). Even when clusters or chains of water molecules present themselves, the individual waters are spherical, resolved and make sense in chemical terms (distances between the water molecules, electrostatic environment *etc.*). At lower resolution water molecules may run together and obscure their discrete nature, but for well refined high-resolution structures this is not usually the case.

This is especially not the case for crystals grown from PEG. Discrete water molecules are seen, as in salt- or alcohol-grown crystals, but intriguingly, strands, chains, clusters, tangles and odd, usually extended, electron-density constellations also appear on the surfaces of proteins, as well as distant from protein. These densities are inconsistent with chains or clusters of water molecules. They are, however, consistent with ordered, or partially ordered, PEG molecules of lengths 2–8, or greater. I contend that is indeed how they should be interpreted.

I am not the first to remark on these suggestive strands and chains. They have been reported many times in the structures of protein crystals (Hašek, 2006; Hašek *et al.*, 2011; Skálová *et al.*, 2010), and the Protein Data Bank (Berman *et al.*, 2000) is replete with entries that include PEG molecules of various lengths (PEG, PG4, PG5, PG6 and probably in some cases even greater lengths). PG6 is about the same length as PEG 400. Using physical approaches that deal with phase separations in polymer–water solutions containing PEG and proteins (Tanaka *et al.*, 2003; Abe *et al.*, 2022) also led to the conclusion that 'a part of the PEG molecules may be taken into protein crystals'.

It has been noted elsewhere that PEG molecules of average molecular weight 1000 would have a hydrodynamic radius (for radii of gyration and exclusion-zone measures, see Tanaka *et al.*, 2003) so large that the probability of them entering, through channels and intersticies, into existent protein crystals would be very low (Ray & Bracker, 1986; Ray & Puvathingal, 1986). Thus, we might conclude that the observation of PEG molecules in protein crystals is, in most cases, and certainly for larger lengths, not due to diffusion into the crystals, but due to the incorporation of PEG molecules into the lattice at the time of formation. That is, they are inherent components of the crystal, and the interactions that they make with the protein, with water molecules and with each other are determinants in defining the overall crystal structure.



This is an important distinction. It suggests that the reason that PEG has been successful in the crystallization of macromolecules and viruses is not strictly due to 'volume-exclusion' effects and the sequestration of water by PEG (Atha & Ingham, 1981; Ingham, 1990; Israelachvili, 1997; Sheth & Leckband, 1997), but is also due to the propensity of PEG to create order in the crystal through intermolecular protein contacts, the local ordering of solvent molecules and the self-organization of the PEG molecules in

the crystal itself. Any factor that promotes order of any kind in a crystal lattice is energetically favorable and encourages crystallization success.

It has been pointed out by physical and polymer chemists (Israelachvili, 1997; Banquy *et al.*, 2016; Gaballa *et al.*, 2023) that PEG is a unique, highly unusual polymer. Its properties are quite different from those of polypropylene glycol and methylene glycol, from which it differs by only one carbon in the monomeric unit. It is, furthermore, quite different from otherwise chemically similar polymers such as polyvinyl alcohol or Jeffamine, which use different linkages. PEG is extremely hydrophilic, as demonstrated by its exceptional solubility in water. 50%(*w*/*v*) stock solutions of PEG 3350 are commonly maintained in laboratories. It is estimated that 2.3–2.6 molecules of H_2O are bound per PEG monomer (Sheth & Leckband, 1997; Israelachvili, 1997; Banquy *et al.*, 2016). Given the number of water molecules bound per monomer and the molecular weights of a PEG monomer (58 Da) and water (18 Da), it can be calculated that a given weight of PEG in solution will sequester about an equivalent weight of water, for example 50 g of PEG will bind about 50 g of water. Yet PEG, paradoxically, also displays significant hydrophobic characteristics (Israelachvili, 1997; Sheth & Leckband, 1997; Banquy *et al.*, 2016; Kuhl *et al.*, 1996; Herrmann *et al.*, 1983).

It has, furthermore, been reported that PEG may promote protein association, and therefore a reduction in solubility, by essentially 'pushing' the macromolecules together (Sheth & Leckband, 1997). The idea is similar to that which explains the 'crystallization' of otherwise non-interacting colloidal particles, such as polystyrene spheres. Hašek, who has studied numerous aspects of the role of PEG in the crystallization of proteins, has hypothesized that PEG provides a surface-shielding effect for proteins that is dependent on PEG length and that PEG is included in protein crystals (Hašek, 2006). He suggests that this leads to a reduction in stacking faults and other dislocations in protein crystals (Malkin *et al.*, 1996), which in turn reduces mosaicity and leads to protein crystals that yield higher resolution diffraction. His conclusions are echoed by Abe and coworkers using light scattering and other physical techniques (Abe *et al.*, 2022). The conceptual basis for the overall interactions of PEG with biological macromolecules remains elusive, however, and it is still not entirely clear which effects or combination of effects best explain the value of PEG for protein crystallization.

There are data to support this hypothesis, other than the numerous observations in the literature of incorporated PEG; among them, the distinctly different appearance of electron-density maps and the otherwise inexplicable success of PEG precipitants in protein crystal growth. Our own experiments (McPherson & Larson, 2018; McPherson, 2019*a*,*b*, 2026) have suggested that protein crystals grown from salts or low ionic strength readily allow the diffusion of dyes into their interior. Crystals grown from PEG, in general, do not. The popular dye reagent 'IZIT' (methylene blue; Hampton Research, Aliso Viejo, California, USA) used to discriminate protein crystals from small-molecule crystals often gives false negatives for crystals grown from PEG (Bob Cudney, Hampton Research, personal communication).

An organizing principle attributed to PEG may also help to explain other puzzles associated with its use in crystallization. For example, why, in some cases, low-molecular-weight PEGs produce crystals, while longer length PEGs having much greater 'volume-exclusion' and solvent-sequestration effects are not effective. At the extreme, crystals of many proteins have been grown from very low concentrations of PEG: 1–2%. This is below the level necessary to effect significant volume exclusion and water sequestration. It seems apparent that some other feature of low-molecular-weight PEG must be in play. Other cases have been reported in which the addition of a low concentration of low-molecular-weight PEG to a salt, or to a low ionic strength mother liquor, produces protein crystallization, while an equivalent PEG-free mother liquor does not (see, for example, Kuznetsov *et al.*, 2000; Ray & Bracker, 1986; Ray & Puvathingal, 1986; Harris *et al.*, 1995). It might be noted that in the popular commercial Crystal Screen (Hampton Research, Aliso Viejo, California, USA), of the 48 crystallization reagents, nine contain low-molecular-weight PEGs, and five of these are salt solutions doped with low-molecular-weight PEG.

It should be noted that the majority of successful protein crystallizations from PEG are from high-molecular-weight PEGs such as molecular weight 3350 or 8000. These numbers, however, only define the mean molecular weight of the PEG, which in reality is a broad distribution of weights and includes the low-molecular-weight PEG under discussion here (Shen *et al.*, 2024; Gaballa *et al.*, 2023). In addition, high-molecular-weight PEG, depending on storage, preparation, purification and other factors, degrades spontaneously into the low-molecular-weight PEG fragments that can be observed in crystals (Wilson *et al.*, 2020). Thus, even if protein crystals are grown from concentrated, high-molecular-weight PEG, there is ample opportunity for low-molecular-weight PEG to be present at significant concentrations.

For protein (and nucleic acid and virus) crystals grown from, or in the presence of, PEG, evidence indicates that the 'solvent regions' separating protein molecules in the lattice are not occupied exclusively by ordered water molecules along with disordered bulk solvent. The spaces are occupied by a mixture of water and PEG molecules of varying lengths. The content of these regions should be thought of more as a paste, a plaster, a mortar or perhaps even a fibrous gel that serves to create order and to bind the macromolecules together. Tanaka and coworkers reached similar conclusions: 'the stable liquid phase formed a sponge-like structure in both PEG 8000 and 20 000 solutions (Tanaka *et al.*, 2003).

Re-evaluation of the nature of the solvent regions in crystals grown from PEG has implications for protein refinement. PEG molecules in the solvent regions are usually associated with protein surfaces, but also extend away from the protein into distal regions. This undoubtedly constrains their dynamic propensity and contributes to their ordering. Nevertheless, PEG molecules exhibit flexibility and this is reflected in their generally low occupancies, 0.2–0.6 in general, and their

high atomic displacement parameters, often 50–80 Å^2. These features of partial order, however, do not prevent electron-density maps from betraying their presence.

PEG molecules, despite their degree of disorder, avidly bind water molecules, as noted above, which may in turn be distinctive in electron-density maps. Such PEG-bound water molecules may appear in maps at much greater distances from protein atoms or other waters than commonly allowed [3.5 Å is the default limit in *Coot* (Emsley & Cowtan, 2004) and *phenix.refine* (Adams *et al.*, 2010; Afonine *et al.*, 2012) for elimination]. Such PEG-bound water molecules will not be included in a crystal structure, though valid, if an automated water-placement procedure is utilized.

PEG molecules form curls, tangles, twisted threads and other aggregates which give rise to electron-density shapes and forms that cannot be explained by clusters or chains of discrete water molecules. Water molecules have a hydrogen-bonding distance of about 2.5–3.0 A, and in high-resolution crystal structures are incompatible with the PEG strands. Because of partial occupancies and self-association of PEG molecules, clashes and clashscores designed for macromolecular structure are highly questionable. For example, in one trypsin–ligand complex that we recently refined at 1.2 Å resolution (McPherson *et al.*, 2025) the clashscore of the protein alone in the unit cell with all PEG molecules removed was 3. With the PEG molecules used in the refinement in place the clashscore was 22. The clashscore appears to be irrelevant when applied to the entire crystal structure for crystals grown from PEG.

Finally, it should be emphasized that the behavior and structure of PEG is almost entirely derived from investigations of PEG in bulk solvent, usually water solutions. There is little data on the behavior or structure of PEG molecules in constrained spaces, such as in the cavities, channels and intersticies that characterize protein crystals, although this issue has not been completely ignored (Rodrigues *et al.*, 2011). Nor is there data on those features of PEG in the presence of other macromolecules. In the lattice spaces, PEG must certainly be more restricted in motion and more ordered in form.

References

Abe, S., Tanaka, J., Kojima, M., Kanamaru, S., Hirata, K., Yamashita, K., Kobayashi, A. & Ueno, T. (2022). *Sci. Rep.* **12**, 16031.

Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L.-W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C. & Zwart, P. H. (2010). *Acta Cryst.* D**66**, 213–221.

Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., Terwilliger, T. C., Urzhumtsev, A., Zwart, P. H. & Adams, P. D. (2012). *Acta Cryst.* D**68**, 352–367.

Atha, D. H. & Ingham, K. C. (1981). *J. Biol. Chem.* **256**, 12108–12117.

Banquy, X., Lee, D. W., Kristiansen, K., Gebbie, M. A. & Israelachvili, J. N. (2016). *Biomacromolecules*, **17**, 88–97.

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. & Bourne, P. E. (2000). *Nucleic Acids Res.* **28**, 235–242.

Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.

Gaballa, S. A., Naguib, Y., Mady, F. & Khaled, K. (2023). *J. Adv. Biomed. Pharm. Sci.* **7**, 26–36.

Harris, L. J., Skaletsky, E. & McPherson, A. (1995). *Proteins*, **23**, 285–289.

Hašek, J. (2006). *Z. Kristallogr. Suppl.* **2006**, 613–618.

Hašek, J., Labský, J., Skálová, T., Kolenko, P., Dohnálek, J., Dušková, J., Štepánková, A. & Koval, T. (2011). *Z. Kristallogr.* **28**, 475–480.

Herrmann, A., Pratsch, L., Arnold, K. & Lassmann, G. (1983). *Biochim. Biophys. Acta*, **733**, 87–94.

Ingham, K. C. (1990). *Methods Enzymol.* **182**, 301–306.

Israelachvili, J. (1997). *Proc. Natl Acad. Sci. USA*, **94**, 8378–8379.

Kuhl, T., Guo, Y., Alderfer, J. L., Berman, A. D., Leckband, D., Israelachvili, J. & Hui, S. W. (1996). *Langmuir*, **12**, 3003–3014.

Kuznetsov, Y. G., Day, J., Newman, R. & McPherson, A. (2000). *J. Struct. Biol.* **131**, 108–115.

Malkin, A. J., Kuznetsov, Y. G. & McPherson, A. (1996). *J. Struct. Biol.* **117**, 124–137.

McPherson, A. (2019*a*). *Acta Cryst.* F**75**, 132–140.

McPherson, A. (2019*b*). *Crystals*, **9**, 126–132.

McPherson, A. (2025). *Cryst. Growth Des.* In the press.

McPherson, A. (2026). In preparation.

McPherson, A. & Larson, S. B. (2018). *Acta Cryst.* F**74**, 593–602.

Ray, W. & Bracker, C. J. (1986). *J. Cryst. Growth* **76**, 562–576.

Ray, W. & Puvathingal, J. M. (1986). *J. Biol. Chem.* **261**, 11544–11549.

Rodrigues, C. G., Machado, D. C., da Silva, A. M. B., Júnior, J. J. S. & Krasilnikov, O. V. (2011). *Biophys. J.* **100**, 2929–2935.

Shen, Y., Son, J. & Yu, X.-Y. (2024). *Sci. Rep.* **14**, 14980–14996.

Sheth, S. R. & Leckband, D. (1997). *Proc. Natl Acad. Sci. USA*, **94**, 8399–8404.

Skálová, T., Dušková, J., Hašek, J., Kolenko, P., Štěpánková, A. & Dohnálek, J. (2010). *J. Appl. Cryst.* **43**, 737–742.

Tanaka, S., Ataka, M., Onuma, K. & Kubota, T. (2003). *Biophys. J.* **84**, 3299–3306.

Wilson, J., Ristic, M., Kirkwood, J., Hargreaves, D. & Newman, J. (2020). *iScience*, **23**, 101219.