STRUCTURAL BIOLOGY

ISSN 2059-7983

Perspective on a large-scale ligand structure characterization

Elspeth F. Garman*

Department of Biochemistry, Dorothy Crowfoot Hodgkin Building, University of Oxford, South Parks Road, Oxford OX1 3QU, United Kingdom. *Correspondence e-mail: elspeth.garman@bioch.ox.ac.uk

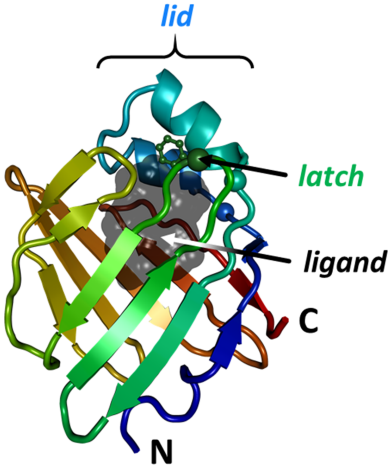
Keywords: fatty acid-binding protein; ligand structure characterization.

In this issue, a triumvirate of papers (Ehler, Bartelmus *et al.*, 2025; Ehler, Benz & Rudolph, 2025; Casagrande *et al.*, 2025) from the Lead Discovery Department at F. Hoffmann-La Roche Ltd in Basel, Switzerland, presents a detailed multi-faceted study of fatty acid-binding protein (FABP) structures. FABPs are a class of important proteins that act in the intracellular transport and metabolism of fatty acids. They occur in a number of isoforms, denoted FABP1, 2, 3 *etc.*, and isoforms 4 and 5 are potential targets for the treatment of diabetes and atherosclerosis, the latter being a disease caused by deposition of fatty acids on the inner walls of arteries.

The work reported in these three papers has resulted in the submission of a total of 229 structures to the Protein Data Bank, of which 13 are of apo protein and 216 are ligand-bound. The investigation has highlighted a number of issues requiring the attention of the macromolecular crystallography community, which are perhaps deserving of an airing in a longer future editorial.

The first paper in the series addresses the dynamics of the FABP4 isoform and ligand binding to it using NMR to look at how ligands and inhibitors change the conformational flexibility of the protein, especially that of the entrance to the binding site, the so-called portal region. Pivotally, a preparation of human FABP4 (hFABP4), shown to have no endogenous ligands bound before the experiments, was used to probe the binding site. hFABP4 was found to become more rigid when a ligand was bound. Moreover, the protein's biophysical behaviour was affected by the nature of the ligand, in that it either remained associated with bicelles via the portal region, or detached from them into solution, a prerequisite for FABP4 to achieve nuclear access *in vivo*.

The second paper describes the crystallographic determination at 100 K of 229 high-resolution (median 1.2 Å) ligand- and inhibitor-bound structures of human FABPs 1, 3, 4, 5 and 9, and mouse FABPs 4 and 5, along with much relevant affinity data. The ligands explored a wide chemical space in which natural fatty acids had their carboxylate and aliphatic groups replaced by other chemical substituents. Some structures were (serendipitously) phased using single-wavelength anomalous dispersion from the signal arising from a bromine in the ligand. A detailed analysis of the inhibitor/ligand-binding poses is provided, made possible by the consistently high-resolution structural information and showing the different effects on the protein conformation as a function of acidic head groups and overall ligand conformation.



In the last paper, the authors present some unexpected results that arose directly from the availability of the high-resolution crystallographic structures described in the second paper. A significant number (approximately 15%) of the ligands found in the electron density deviated substantially from those that had nominally been soaked into the crystals. Possible reasons for this unwelcome proportion of structures with aberrant ligands bound are well articulated in a detailed discussion, with examples of each pathology drawn from the 216 liganded FABP structures. The issues range from possible human error in the initial registration of the ligand's chemical structure to unexpected products from the chemical synthesis that led to initial building of the wrong compounds into the electron-density maps (*e.g.* the wrong heterocycle or the wrong assignment of chiral centres in enantiomers and diastereomers). Further possible complications are also enumerated in which covalent modification by ring opening occurs, or conversely intramolecular condensation reactions lead to ring closures. Chemical side reactions can result in the presence of different substituents than those expected, or the shift of a chemical group from one atom to another. Sometimes a substituent is unobserved in the

final map due to reasons including ligand decomposition during storage or at the crystal-soaking stage, and traces of the synthesis starting material being present. Racemic mixtures also may cause challenges in identification of the chiral centres during model building.

Unexpected ligand binding can arise from impure preparations where the impurity binds preferentially and outcompetes the intended ligand. There are a number of such examples in the literature, and I, myself, battled with such a case in the 1990s while refining structures of *Salmonella typhimurium* sialidase with different *N*-acetylneuraminic acid phosphonate derivatives bound (Crennell *et al.*, 1996). Two forms were synthesized, one with an equatorial phosphonate group (ePANA) and the other with an axial phosphonate group (aPANA), and the structures solved to 1.9 Å resolution. However, after much head scratching about the poor fits of the aPANA model to the electron density, we realized that a small impurity of ePANA in the aPANA sample had resulted in 40% of the active sites having ePANA bound, and 60% aPANA (PDB entry 1dil). The correlation of atomic *B* factor and occupancy can make refinement of such mixed-ligand structures challenging.

In summary, there are serious implications for lower resolution ligand-bound structures where the possible errors enumerated in the third paper remain undetected, and thses structures are submitted to the PDB. In particular, if machine-learning algorithms are using these structures in training sets, predictions made for drug discovery will be adversely affected.

References

Casagrande, F., Ehler, A., Burger, D., Benz, J., Ross, A. & Rudolph, M. G. (2025). *Acta Cryst.* D**81**, 423–435.
Crennell, S. J., Garman, E. F., Philippon, C., Vasella, A., Laver, W. G., Vimr, E. R. & Taylor, G. L. (1996). *J. Mol. Biol.* **259**, 264–280.
Ehler, A., Bartelmus, C., Benz, J., Plitzko, I. & Rudolph, M. G. (2025). *Acta Cryst.* D**81**, 451–464.
Ehler, A., Benz, J. & Rudolph, M. G. (2025). *Acta Cryst.* D**81**, 436–450.