Bragg centennial

Early days in drug discovery by crystallography – personal recollections

Peter M. Colman

The Walter and Eliza Hall Institute of Medical Research, 1G Royal Parade, Parkville, Victoria, Australia. Correspondence e-mail: pcolman@wehi.edu.au

The influences of Lawrence Bragg and Max Perutz are evident in the contemporary emphasis on 'structural enablement' in drug discovery. On this occasion of the centenary of Bragg's equation, his role in supporting the earliest structural studies of biological materials at the Cavendish Laboratory is remembered. The 1962 Nobel Prizes for the structures of DNA and proteins marked the golden anniversary of the von Laue and Bragg discoveries.

Max Perutz's idea for an 80th birthday event was to gather together people interested in the relevance of protein crystallography to medicine. His Medical Research Council colleagues duly organized a memorable meeting at the Royal Institution on 23 September 1994, 'Protein Crystallography and Medicine – a meeting in honour of Max Perutz' (Fig. 1). Of course Perutz himself (Perutz, 1994) stole the show with his discourse on polar zippers in haemoglobin and Huntington's disease. Almost twenty years earlier he had written (Perutz, 1976) 'I started to study haemoglobin by X-ray crystallography in 1937 because at that time the structure of proteins seemed the most important unsolved problem in biochemistry, but I never dreamt that its solution would one day throw light on the nature of inherited diseases.'

Today it is rare to see a publication on the molecular structure of a protein that just might be a drug target without reading concluding remarks about the implications of the work for new medicines. Evidently it was not always so. Our

Figure 1
Max Perutz on the occasion of his 80th birthday celebration at the Royal Institution. Left to right: Tony Kossiakoff, Wim Hol, Michael Rossmann, Aaron Klug, Max Perutz, David Blow, Peter Colman, Don Wiley and Wayne Hendrickson.

own work on influenza viruses was about to begin when the above lines were written. When I moved to CSIRO in 1978 it was very convenient to quote Max Perutz to my masters. My real interest at that time was in understanding antigenic variation in influenza viruses and its consequences for vaccination, but the lines I used drew on Perutz's predictions on the utility of structure, serine proteases in his example, in drug discovery (Perutz, 1976). 'Knowledge of the three-dimensional structure of these enzymes may open the way to the synthesis of tailor-made inhibitors for the treatment of such diseases.' My enthusiasm aside, that our work would ever affect medical practice seemed most unlikely, a view that many people shared and some continued to hold even after we had discovered zanamivir.

The first crystals of influenza virus neuraminidase (Laver, 1978) held great promise for realising a structural analysis, but it was necessary to resort to screening multiple strains and subtypes for suitable crystals to eventually determine its structure. The electron-density maps were only interpretable after averaging two copies of a 1957 strain in one crystal form with a single copy of a 1967 strain in another (Varghese et al., 1983). The 21 amino-acid sequence differences between these strains are no impediment to this approach. The first example of the now common β-propeller fold was revealed. It was also immediately evident from the structure that the active centre of the enzyme was invariant across all then known strains of influenza (Colman et al., 1983).

Averaging independent images is akin to knowing the value of the molecular transform at non-integral values of the reciprocal lattice and the use of information of this type was raised by Bernal (Bernal et al., 1938), and first seriously contemplated as a phasing tool by Bragg and Perutz (Bragg & Perutz, 1952). Crick was right about the use of this technique for *ab initio* phasing (Crick, 1988), at least in the early 1950s, but in this work we see the origins of algorithms for solvent flattening, non-crystallographic symmetry averaging (Bricogne, 1974; Colman, 1974) and, most recently, image recovery in diffractive imaging (Fienup, 1987, Miao et al., 1999).

The neuraminidase antigen lives in the shadow of its partner antigen, the viral haemagglutin, a molecule to really excite a structural biologist with its capacity to facilitate acid-catalysed fusion of the viral and endosomal membranes (Skehel & Wiley, 2000). It came as a shocking but welcome surprise that crystallography could inform a process as dynamic as membrane fusion. Like drug discovery, it is now commonplace to learn about such biological events from crystal structure studies.

By comparison to the haemagglutinin, the neuraminidase, another glycosidase in a lineage going back to lysozyme, held little interest. Furthermore, even the anti-neuraminidase antibodies raised during influenza infection were of secondary value to the host compared to those raised against the haemagglutinin. And then there was the question around the role for this enzyme activity on the surface of the virus. It was thought that the main function of neuraminidase in the life cycle was to liberate progeny virions from the infected cell surface (Palese & Compans, 1976), surely a rather late stage for therapeutic intervention. On the other hand, here was a truly unusual example of a viral enzyme whose function was not intracellular, and drugs targeting it would not need to cross the infected cell membrane to do their work.
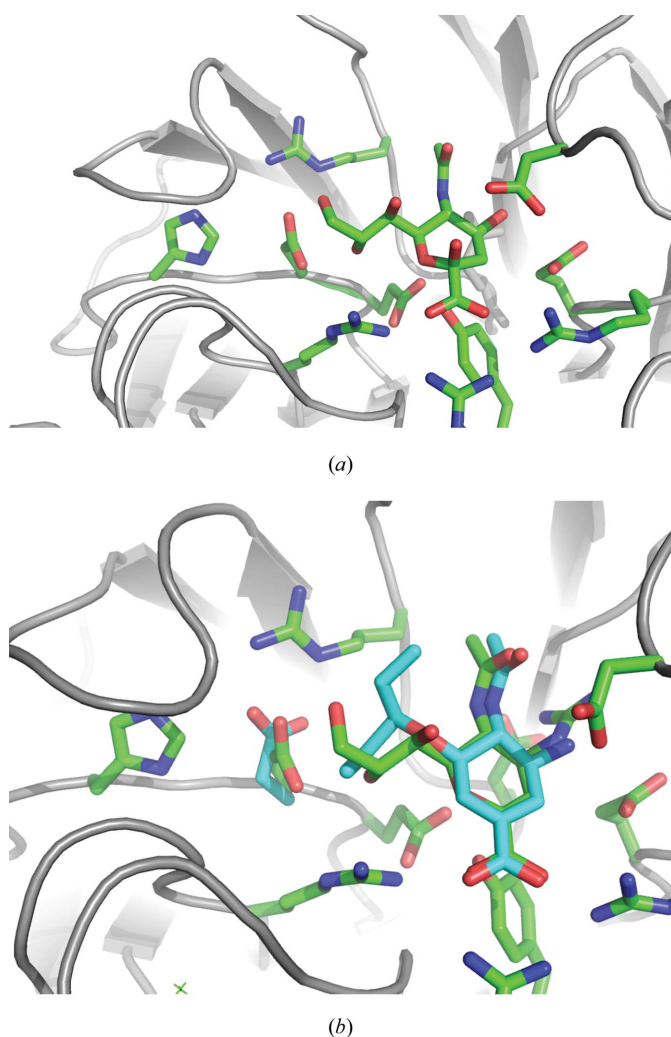
The seminal crystallographic observation that led to zanamivir was the structure of the complex between neuraminidase and the product of the enzyme reaction, sialic acid (Varghese et al., 1992) (Fig. 2). The bound sugar is twisted, its carboxylate lying coplanar with the pyranose ring. The 4-hydroxyl entity points to a water-occupied pocket on the enzyme where strain-invariant negatively charged amino acids are located. Three modifications to sialic acid each provide an increase of some two orders of magnitude to its affinity with the enzyme. The first is the establishment of an sp^2 carbon at C2, placing the carboxylate equatorial and not axial, to the sugar. The second and third come from replacing the 4-hydroxyl with 4-ammonium and then 4-guanidium (zanamivir), thereby engaging protein carboxylate residues in the binding pocket for the 4-hydroxyl and displacing bound water molecules there (Varghese et al., 1995; von Itzstein et al., 1993).

A compound as hydrophilic as zanamivir does not cross membranes, neither does it need to unless the drug is to be delivered orally. The upper airways are readily accessible to inhaled drugs and Relenza is a formulation of zanamivir for oral inhalation. Our own attempts to find an orally available 'version' of zanamivir were unsuccessful (Taylor et al., 1998), but others succeeded (Kim et al., 1997). The requisite properties found in oseltamivir carboxylate (the active substance in Tamiflu) derive from its carbocyclic core and the substitution of the C6-glycerol moiety with a pentyl ether. Unlike the glycerol, the pentyl ether is not found in the natural substrate, sialic acid, and the consequences of this departure are that the neuraminidase must undergo a small conformation change to bind oseltamivir carboxylate that is not required for its binding to either zanamivir (Fig. 2) or the substrate sialic acid (Varghese et al., 1998). This structural difference is a small opening through which drug-resistant viruses are unleashed (Colman, 2009).

If we take the public remarks of the drug industry at face value, it is now unusual for a drug-discovery program to be initiated if the target is not 'structurally enabled'. Some contemporary work is quite remarkable, such as the discovery of ABT-737 (Oltersdorf et al., 2005) and its clinical analogue ABT-263. These molecules mimic the interaction between a helical peptide and the pro-survival Bcl-2 proteins, an event which triggers cell death by apoptosis and is frequently dysregulated in tumour cells (Lessene et al., 2008). Such advances hold great promise that ever more-challenging drug targets will be hit in the future.

Figure 2
(a) The product of the neuraminidase-catalysed reaction, sialic acid, bound to the enzyme. (b) Overlay of zanamivir (green) and oseltamivir carboxylate (cyan) bound to neuraminidase.

References

Bernal, J. D., Fankuchen, I. & Perutz, M. (1938). *Nature (London)*, **141**, 523–524.
Bragg, L. & Perutz, M. F. (1952). *Proc. R. Soc. London Ser. A*, **213**, 425–435.
Bricogne, G. (1974). *Acta Cryst.* A**30**, 395–405.
Colman, P. M. (1974). *Z. Kristallogr.* **140**, 344–349.
Colman, P. M. (2009). *Annu. Rev. Biochem.* **78**, 95–118.

Bragg centennial

Colman, P. M., Varghese, J. N. & Laver, W. G. (1983). *Nature (London)*, **303**, 41–44.

Crick, F. (1988). *What Mad Pursuit: a Personal View of Scientific Discovery*. New York: Basic Books.

Fienup, J. (1987). *J. Opt. Soc. Am. A*, **4**, 118–123.

Itzstein, M. von, *et al.* (1993). *Nature (London)*, **363**, 418–423.

Kim, C. U., Lew, W., Williams, M. A., Liu, H., Zhang, L., Swaminathan, S., Bischofberger, N., Chen, M. S., Mendel, D. B., Tai, C. Y., Laver, W. G. & Stevens, R. C. (1997). *J. Am. Chem. Soc.* **119**, 681–690.

Laver, W. G. (1978). *Virology*, **86**, 78–87.

Lessene, G., Czabotar, P. E. & Colman, P. M. (2008). *Nat. Rev. Drug Discov.* **7**, 989–1000.

Miao, J., Charalambous, P., Kirz, J. & Sayre, D. (1999). *Nature (London)*, **400**, 342–344.

Oltersdorf, T., *et al.* (2005). *Nature (London)*, **435**, 677–681.

Palese, P. & Compans, R. W. (1976). *J. Gen. Virol.* **33**, 159–163.

Perutz, M. F. (1976). *Nature (London)*, **262**, 449–453.

Perutz, M. (1994). *Protein Sci.* **3**, 1629–1637.

Skehel, J. J. & Wiley, D. C. (2000). *Annu. Rev. Biochem.* **69**, 531–569.

Taylor, N. R., Cleasby, A., Singh, O., Skarzynski, T., Wonacott, A. J., Smith, P. W., Sollis, S. L., Howes, P. D., Cherry, P. C., Bethell, R., Colman, P. & Varghese, J. (1998). *J. Med. Chem.* **41**, 798–807.

Varghese, J. N., Epa, V. C. & Colman, P. M. (1995). *Protein Sci.* **4**, 1081–1087.

Varghese, J. N., Laver, W. G. & Colman, P. M. (1983). *Nature (London)*, **303**, 35–40.

Varghese, J. N., McKimm-Breschkin, J. L., Caldwell, J. B., Kortt, A. A. & Colman, P. M. (1992). *Proteins*, **14**, 327–332.

Varghese, J. N., Smith, P. W., Sollis, S. L., Blick, T. J., Sahasrabudhe, A., McKimm-Breschkin, J. L. & Colman, P. M. (1998). *Structure*, **6**, 735–746.