Crystallography, Evolution, and the Structure of Viruses

Published, JBC Papers in Press, February 8, 2012, DOI 10.1074/jbc.X112.348961

Michael G. Rossmann
From the Department of Biological Sciences, Hockmeyer Hall of Structural Biology, Purdue University, West Lafayette, Indiana 47907

My undergraduate education in mathematics and physics was a good grounding for graduate studies in crystallographic studies of small organic molecules. As a postdoctoral fellow in Minnesota, I learned how to program an early electronic computer for crystallographic calculations. I then joined Max Perutz, excited to use my skills in the determination of the first protein structures. The results were even more fascinating than the development of techniques and provided inspiration for starting my own laboratory at Purdue University. My first studies on dehydrogenases established the conservation of nucleotide-binding structures. Having thus established myself as an independent scientist, I could start on my most cherished ambition of studying the structure of viruses. About a decade later, my laboratory had produced the structure of a small RNA plant virus and then, in another six years, the first structure of a human common cold virus. Many more virus structures followed, but soon it became essential to supplement crystallography with electron microscopy to investigate viral assembly, viral infection of cells, and neutralization of viruses by antibodies. A major guide in all these studies was the discovery of evolution at the molecular level. The conservation of three-dimensional structure has been a recurring theme, from my experiences with Max Perutz in the study of hemoglobin to the recognition of the conserved nucleotide-binding fold and to the recognition of the jelly roll fold in the capsid protein of a large variety of viruses.

Early Education

I was born in Frankfurt, Germany, on July 30, 1930. Grandfather Rossmann was a high school teacher of French who had written a well known textbook. My mother’s family members were merchants and academic historians with expertise in classical Greece and Italy. My mother had studied art at the famous Bauhaus Art School in Weimar after the end of the First World War. By the time of my birth, she was a correspondent for the Frankfurter Zeitung, a local newspaper with a national readership similar in nature to the former Manchester Guardian in England. She illustrated her articles about local events with her sketches.

During my first few school years in Germany, I was under constant threat of beatings by other boys and some of the teachers on account of the Jewish ancestry of my mother’s family. We immigrated to England in July of 1939. Although I could not speak or understand English when we first arrived, school soon became, for the first time, a pleasure, unlike in Germany. With the help of kind teachers, I became fascinated by geometry and enjoyed the grammatical analyses in Latin classes. My mother had joined the Society of Friends (Quakers) as a young person in Germany soon after the end of the First World War. After I had successfully taken an entrance examination that qualified me for a bursary, it was financially possible for my mother to enter me as a pupil into the Friends’ School, Saffron Walden, Essex, in 1942. Here, I was happy while discovering my interest in science. I was an undergraduate at the Regent Street Polytechnic from 1948 to 1951.
studying physics and mathematics. I stayed on to work on a master’s degree, measuring the vapor pressures of metals. In 1952, I obtained a position as a lecturer in “Natural Philosophy” (physics) at the Royal Technical College (now, the University of Strathclyde) in Glasgow, Scotland (1952–1956). After arriving in Glasgow, I completed the M.Sc. degree in 1953. However, I was dissatisfied with my intellectual progress and was able to arrange to simultaneously study for a Ph.D. degree under J. Monteath Robertson at the University of Glasgow (1953–1956) while teaching at the technical college, about a one-mile bicycle ride away. At the University, I studied the crystal structures of aromatic hydrocarbons, doing all calculations by hand. During this time, I married Audrey Pearson. Our wedding was at the Adel Friends meeting house on a beautiful summer’s day in July of 1954 in Leeds. On completing my Ph.D. studies in 1956, I was accepted as a postdoctoral fellow by Bill Lipscomb at the University of Minnesota. I was thankful for a Fulbright scholarship that paid not only my travel expenses, but also those of my family. As a postdoctoral fellow, I worked on the structures of some plant natural products using, for the first time, an electronic computer and writing some early crystallographic computer programs.

Cambridge (1958–1964)

In 1958, my wife (Audrey, pregnant with Heather) and our two children (Martin and Alice) returned to England, where I had been accepted by Max Perutz to work in the Medical Research Council’s laboratory in Cambridge (later, the Laboratory of Molecular Biology) (Fig. 1). Max had collected three-dimensional data on horse oxyhemoglobin. The new EDSAC 2 computer had just started to become functional. My first task was to find the relative coordinates of the heavy atoms in the C2 space group of the hemoglobin crystals. A number of methods had been previously proposed by Perutz, Crick, Bragg, and Wyckoff, but none were entirely satisfactory. Using the three-dimensional Fourier program I had written for the new
EDSAC 2 computer, I invented a Patterson-like technique (1) for finding the position and refining the occupancies of the heavy atom markers. We were able to determine the 5.5 Å resolution structure of hemoglobin (Fig. 2) in the summer of 1959 (2, 3) and recognize the similarity to Kendrew’s 6 Å resolution myoglobin structure, determined a year or so earlier. These were the first protein structures to be solved. The evolutionary relationship between these structures, confirming the evolution of living organisms at a basic molecular level, has been a major guide to my research direction ever since.

The work on the hemoglobin structure determination was also a stimulus for the development, in collaboration with David Blow, of crystallographic techniques that formed the technical foundations of structural biology (4–7). These included the use of anomalous dispersion (4–7).
Dehydrogenases and Evolution of Protein Domains (1964–1980)

In 1964, my family and I moved to Lafayette, Indiana, where I had an opportunity to develop my own laboratory at Purdue University. I decided to work on lactate dehydrogenase (LDH) based on a vague suspicion that there might be a common structural motif among NAD-dependent dehydrogenases, much as there was among the oxygen carriers myoglobin and hemoglobin. The structure of LDH (12) was the first structure of an enzyme with a small metabolite as a substrate. It was also by far the largest structure solved to date. Three years later, we also solved the structure of glyceraldehyde-3-phosphate dehydrogenase (GAPDH) (13). The striking similarity of the NAD-binding domain in these two structures, as well as in alcohol and malate dehydrogenases, determined by Carl Brändén and Leonard Banaszak, respectively, confirmed the earlier expectations. Furthermore, I recognized (14) that flavodoxin, a FMN- and FAD-binding protein whose structure had been determined both by Lyle Jensen and by Martha Ludwig, and adenylate kinase, an ATP-binding protein whose structure had been determined by Georg Schulz, as well as other structures that also bound nucleotides, all had a similar fold, giving rise to the recognition of a common nucleotide-binding fold (13–16). I suspected that this fold might be of central importance to life because of its ability to establish a functional relationship between a protein and a nucleotide. Indeed, it is now clear that this fold is one of the most common protein folds. The early structures of the dehydrogenases also showed that, in general, the building blocks of proteins were structural domains, each with primitive functions and each having an independent evolutionary history. Gene duplication and fusion produced more sophisticated enzymes where the substrate bound between domains, with each domain providing an essential function (16).

X-ray Diffraction Data Processing (1979–2000)

In 1970, David Haas and I demonstrated the use of frozen crystals to minimize radiation damage (17), a technique that was subsequently popularized by Ada Yonath in her studies of ribosomes. Today, frozen crystals are used for almost all protein crystal x-ray diffraction data collection.

A few years later, I developed data processing procedures for oscillation photography as an essential component to our work on virus structure determination (18, 19). Many of these procedures are now incorporated into the popular HKL and MOSFLM processing techniques. During the early days of our use of synchrotron radiation, we realized the value of avoiding the damaging and time-consuming traditional crystal setting procedures by inventing the “American method” of shooting first and thinking (computing) later to find the crystal orientation relative to the camera axes (20). This required the development of algorithms to determine the crystal orientation (21). All of these procedures are standard practice today. Indeed, the earlier technique of “setting” a crystal with its axes in a known relationship to the axes of the x-ray camera is now mostly a forgotten skill.

Small Icosahedral Viruses (1971 to Present)

It had been my intention to study virus structures even before leaving Cambridge. The title of my first National Science Foundation grant was “The Structure of Proteins and Viruses.” It was submitted in 1963, even before my actual arrival at Purdue. The vagueness of the title shows that solving the three-dimensional structure of any new protein to a resolution sufficient for the rough recognition of amino acids was, at that time, a reasonable ambition but likely to take many years of exploratory work. The structure of viruses, however, was a yet unattainable dream.
Nevertheless, I was funded, and that same grant continues today after almost fifty years and more than about ten competitive renewals.

After success with the dehydrogenase studies and a half-year sabbatical leave during 1971 with Bror Strandberg in Uppsala, Sweden, working on the structure determination of satellite tobacco necrosis virus (STNV), I started work on viruses in earnest. Some small RNA plant viruses, such as STNV, could be readily propagated, purified in gram quantities, and crystallized. Eventually, in 1980, this led to the structure of southern bean mosaic virus (22). The structure of tomato bushy stunt virus had been determined by Steve Harrison a year or so earlier. To everybody’s great surprise, the capsids of these viruses consisted of 180 copies of a viral protein subunit that had a similar tertiary “jelly roll” fold assembled into a similar \( T=3 \) quaternary structure, demonstrating once again the conservation of tertiary structure to retain function.

We next turned our attention to animal viruses in collaboration with Roland Rueckert, the leading expert on picornaviruses and working at the University of Wisconsin. This led to the structure of human rhinovirus serotype 14 in 1985 (11), which provided broad insights on assembly, neutralization by antibodies, and receptor recognition. The “canyon hypothesis” proposed that the receptor would bind into a depression on the viral surface (the canyon) that was inaccessible to larger antibodies, thus escaping from host immune surveillance. This site was confirmed in 1993 for the major group of rhinoviruses that use ICAM1 (intercellular adhesion molecule 1) as their cellular receptor (23) and later for other viruses as well (24, 25). We also discovered that certain anti-rhinovirus drugs bound to a pocket in the capsid (26), a discovery that led to recognizing that the stable infectious virions were destabilized on binding to a receptor by ejecting a bound “pocket factor” molecule, thus initiating infection. Extensive work, first with Sterling-Winthrop, Inc. and later with ViroPharma Inc., led to the “pleconaril” drug, which scored well in phase III clinical trials, but was not licensed by the Food and Drug Administration primarily because of undesirable side effects for women on birth control hormones.

The accumulation of crystallographic techniques now opened the door for the determination of many other icosahedral viruses in my laboratory and elsewhere. Among the virus structures we published were Mengo virus (27), canine parvovirus (28), bacteriophage \( \phi X174 \) (29), cosackievirus B3 (30), human parvovirus B19 (31), and shrimp and silkworm parvoviruses. Other aspects such as viral assembly intermediates could also be investigated now (32). All of these viruses were found to have the same jelly roll structure for their capsid proteins, indicating that at least a part of their viral genomes had a common origin.

### Electron Microscopy of Icosahedral Enveloped Viruses (1995 to Present)

In 1981, I took my second sabbatical leave, this time back in Cambridge, learning some electron microscopy from Richard Henderson at the Laboratory of Molecular Biology, my home of 20 years earlier. On my return to Purdue, it was not difficult to persuade my colleagues that we should hire an expert in the use of electron microscopy for three-dimensional reconstructions. This led to the hiring of Tim Baker, who quickly established himself as a major contributor to the study of viruses. My first collaborative project with him was the confirmation of the rhinovirus canyon as being the site of binding for the cellular receptor ICAM1 molecule (23), mentioned above.

In 1991, we determined the crystal structure of the nucleocapsid protein of Sindbis virus (33), a member of the alphavirus family. In contrast to our earlier work, alphavirions have a lipid envelope around their nucleocapsid, making it difficult to crystallize such viruses. Fortunately, Richard Kuhn, a virologist, joined the Purdue faculty. Furthermore, Tim Baker was now also a member of our faculty. With Richard producing the virus, Tim producing the cryo-electron microscopy (cryo-EM) structure, and myself developing techniques of combining the crystal structure of the capsid protein (33) with the electron microscopy results, we were able to publish the structure of an alphavirus (34).

The alphavirus investigation led to the development of hybrid technology in combining crystallography with electron microscopy (35). In this way, we obtained the pseudoatomic structure of Sindbis virus (36, 37) and of flaviviruses such as dengue and West Nile viruses (38, 39). Similarly, we were able to determine the structure of immature flaviviruses (40), establishing, together with the insightful work of my colleague Jue Chen, the maturation process leading to infectious virus (41, 42).

### Tailed Bacteriophages (1998 to Present)

We employed the combination of electron microscopy and crystallography in the study of tailed bacteriophages. These viruses are incredibly efficient, requiring usually only one particle to infect their host, whereas other viruses would take tens or hundreds of particles to be successful. The tail organelle is the weapon by which these viruses have established their evolutionary success and their enormous abundance in water. In these studies, we and others
developed hybrid techniques for combining the crystal structures of individual proteins with cryo-EM structures of the virus or virus fragments to obtain pseudo-atomic resolution structures (43). In collaboration with Dwight Anderson of the University of Minnesota, we determined the structure of assembly intermediates of the small tailed φ29 phage (44) and the machine, located at one of the twelve icosahedral vertices, that packages the genomic DNA into the empty procapsid of both φ29 (45) and, in collaboration with Venigalla Rao of the Catholic University of America, of the very much larger T4 bacteriophage (46). In collaboration with Vadim Mesyanzhinov of the Lomonosov Moscow State University, we also determined the structure of the T4 tail base plate before and after ejecting its genome into the host (47, 48), thus providing some detail on how these viruses efficiently infect their hosts.

Large dsDNA Icosahedral Viruses

The occurrence of accurate icosahedral symmetry diminishes as the virus being examined becomes larger and more complex, making it progressively more difficult to use the techniques that have been especially developed to study icosahedral virus structures. Indeed, it was the development of these techniques that was among my motivations for the study of viruses! In particular, we have been studying Mimivirus (49, 50) in collaboration with Didier Raoult of the University of the Mediterranean in Marseille, France. Until recently, Mimivirus was the biggest known virus both in its physical dimensions and in its genome. This virus straddles the definition of a “dead” virus and a simple “living” cell in terms of the types of genes that are included in its genome. It has a diameter of ~5000 Å, a genome of 1.2 million bp, and a special “stargate” vertex from which the dsDNA genome can exit while infecting a host. The major capsid protein consists of two consecutive jelly roll domains, as is also the case for adenovirus and many other large dsDNA viruses (51) studied by us, Roger Burnett (University of Pennsylvania), Dave Stuart (University of Oxford), Dennis Bamford (University of Helsinki), and others using a combination of crystallography and cryo-EM. Of particular interest is Paramecium bursaria chlorella virus 1 (51), which, like Mimivirus, we showed, in collaboration with James Van Etten (University of Nebraska), has a special vertex (52, 53).

Epilogue

The challenge to structural virology now is to study progressively less symmetric and more complex viruses. Although investigations of crystallizable components of pleomorphic viruses is by no means new, the recent progress in recording high quality cryo-EM tomograms is making it possible to put the structural fragments into the context of the whole virus. For instance, in my laboratory, we are now studying Newcastle disease virus, a member of the paramyxovirus family, which includes the more commonly known measles and mumps viruses.1

On looking back, I realize that I have traveled far from my original motivation, which was based primarily on mathematical solutions of the crystallographic phase problem, the central problem of any crystallographic structural determination. I am greatly indebted to Max Perutz, who opened my eyes to the basic puzzles of biology and made me realize that good science is much more than the fun of puzzle solving, but is a study of Nature. Nevertheless, mathematics and crystallography have remained central to my analytical processes. I felt especially honored when the International Union of Crystallography asked me to contribute a volume describing the techniques that constitute the science of structural biology, which I then attempted to do in collaboration with Eddy Arnold with the first edition of Volume F of the International Tables for Crystallography. It has been particularly satisfying to see the success of the molecular replacement method, the universal adoption of the American method for collecting data, and the rapidly expanding use of hybrid methods. However, in the end, none of this would be worthwhile were it not for the enormous increase in knowledge of the structures and evolution of viruses and their implication for life on Earth.

Acknowledgments—I apologize to the many friends and colleagues whose work I have mentioned but have not referenced. Wherever possible, I have named the person, but because of the limitation of the number of references, I have cited only my own work. I also wish to thank the many postdoctoral fellows, graduate students, collaborators, friends, colleagues, and technicians who have made the work described here possible. I also wish to thank Cheryl Kelly, who helped to prepare this article for print. I have been very fortunate that my wife, Audrey, understood, as she often pointed out, that marriage to a scientist requires a very special kind of wife. Furthermore, Audrey welcomed everybody who came to my laboratory, helping all to settle into life in Lafayette. She insisted on knowing about every new arrival and made sure she knew all of his or her specific needs and interests. I am very grateful for the many years of generous support by the National Institutes of Health and the National Science Foundation, for industrial support especially from the Sterling-Winthrop Co., and for help from Purdue University.

*Math Author’s Choice—Final version full access. Address correspondence to: mr@purdue.edu.

1 A. J. Battisti, G. Meng, D. C. Winkler, L. W. McGinnes, P. Plevka, A. C. Steven, T. G. Morrison, and M. G. Rossmann, unpublished data.
REFERENCES

1. Rossmann, M. G. (1960) The accurate determination of the position and shape of large replacement groups in proteins. Acta Crystallogr. 13, 221–226.

2. Perutz, M. F., Rossmann, M. G., Cullis, A. F., Muirhead, H., Will, G., and North, A. C. T. (1960) Structure of hemoglobin: a three-dimensional Fourier synthesis at 5.5 Å resolution, obtained by x-ray analysis. Nature 185, 416–422.

3. Cullis, A. F., Muirhead, H., Perutz, M. F., Rossmann, M. G., and North, A. C. T. (1961) The structure of hemoglobin. VIII. A three-dimensional Fourier synthesis at 5.5 Å resolution: determination of the phase angles. Proc. R. Soc. A 265, 15–38.

4. Rossmann, M. G. (1961) The position of anomalous scatterers in protein crystals. Acta Crystallogr. 14, 383–388.

5. Rossmann, M. G., and Blow, D. M. (1961) The single isomorphous replacement method. Acta Crystallogr. 14, 1195–1202.

6. Blow, D. M., and Rossmann, M. G. (1962) The detection of subunits within the crystallographic asymmetric unit. Acta Crystallogr. 15, 24–31.

7. Rossmann, M. G., and Blow, D. M. (1962) The single isomorphous replacement method. Acta Crystallogr. 14, 265–417.

8. Blow, D. M., Rossmann, M. G., and Jeffery, B. A. (1964) The arrangement of alpha-chymotrypsinogen molecules in the monoclinic crystal form. J. Mol. Biol. 8, 65–78.

9. Dodson, E., Harding, M. M., Hodgkin, D. C., and Rossmann, M. G. (1966) The crystal structure of insulin. III. Evidence for a 2-fold axis in rhombohedral zinc insulin. J. Mol. Biol. 16, 227–241.

10. Main, P., and Rossmann, M. G. (1966) Relationship among structure factors due to identical molecules in different crystallographic environments. Acta Crystallogr. 21, 67–72.

11. Rossmann, M. G., Arnold, E., Erickson, J. W., Frankenberger, E. A., Griffith, J. P., Hecht, H. J., Johnson, J. E., Kamer, G., Luo, M., Moser, A. G., Rueckert, R. R., Sherry, B., and Vriend, G. (1985) Structure of a human cold virus and functional relationship to other picornaviruses. Nature 317, 145–153.

12. Adams, M. J., Ford, G. C., Kabsch, W., Liljas, A., Brünger, A. T., and Babu, R. A. (1975) The complete description of the 1H, 15N, and 13C chemical shifts of human insulin. J. Mol. Biol. 95, 285–293.

13. Rossmann, M. G., and Blow, D. M. (1963) The refinement of structures partially determined by the isomorphous replacement method. Acta Crystallogr. 16, 319–326.

14. Rossmann, M. G., and Blow, D. M. (1963) Structure and assembly of icosahedral virus particles. Adv. Virus Res. 8, 41–137.

15. Rao, S. T., and Rossmann, M. G. (1973) Comparison of super-secondary structures in proteins. Adv. Protein Chem. 28, 1–103.

16. Rossmann, M. G., Arnold, E., Erickson, J. W., Frankenberger, E. A., Griffith, J. P., Hecht, H. J., Johnson, J. E., Kamer, G., Luo, M., Moser, A. G., Rueckert, R. R., Sherry, B., and Vriend, G. (1985) Structure of a human cold virus and functional relationship to other picornaviruses. Nature 317, 145–153.

17. Haas, D. J., and Rossmann, M. G. (1970) Crystallographic studies on lactate dehydrogenase. J. Mol. Biol. 50, 65–78.

18. Rossmann, M. G., and Blow, D. M. (1962) The detection of subunits within the crystallographic asymmetric unit. Acta Crystallogr. 15, 24–31.

19. Rossmann, M. G., Arnold, E., Erickson, J. W., Frankenberger, E. A., Griffith, J. P., Hecht, H. J., Johnson, J. E., Kamer, G., Luo, M., Moser, A. G., Rueckert, R. R., Sherry, B., and Vriend, G. (1985) Structure of a human cold virus and functional relationship to other picornaviruses. Nature 317, 145–153.

20. Rossmann, M. G., and Blow, D. M. (1962) The single isomorphous replacement method. Acta Crystallogr. 14, 265–417.

21. Rossmann, M. G., Arnold, E., Erickson, J. W., Frankenberger, E. A., Griffith, J. P., Hecht, H. J., Johnson, J. E., Kamer, G., Luo, M., Moser, A. G., Rueckert, R. R., Sherry, B., and Vriend, G. (1985) Structure of a human cold virus and functional relationship to other picornaviruses. Nature 317, 145–153.

22. Rossmann, M. G., and Blow, D. M. (1963) The refinement of structures partially determined by the isomorphous replacement method. Acta Crystallogr. 16, 319–326.
Rossmann, M. G. (2004) Three-dimensional rearrangement of proteins in the tail of bacteriophage T4 on infection of its host. Cell 118, 419–429

Xiao, C., Chipman, P. R., Battisti, A. J., Bowman, V. D., Renesto, P., Raoult, D., and Rossmann, M. G. (2005) Cryo-electron microscopy of the giant Mimivirus. J. Mol. Biol. 353, 493–496

Xiao, C., Kuznetsov, Y. G., Sun, S., Hafenstein, S. L., Kostyuchenko, V. A., Chipman, P. R., Suzan-Monti, M., Raoult, D., McPherson, A., and Rossmann, M. G. (2009) Structural studies of the giant Mimivirus. PLoS Biol. 7, e92

Nandhagopal, N., Simpson, A. A., Gurnon, J. R., Yan, X., Baker, T. S., Graves, M. V., Van Etten, J. L., and Rossmann, M. G. (2002) The structure and evolution of the major capsid protein of a large, lipid-containing DNA virus. Proc. Natl. Acad. Sci. U.S.A. 99, 14758–14763

Cherrier, M. V., Kostyuchenko, V. A., Xiao, C., Bowman, V. D., Battisti, A. J., Yan, X., Chipman, P. R., Baker, T. S., Van Etten, J. L., and Rossmann, M. G. (2009) An icosahedral algal virus has a complex unique vertex decorated by a spike. Proc. Natl. Acad. Sci. U.S.A. 106, 11085–11089

Zhang, X., Xiang, Y., Dunigan, D. D., Klose, T., Chipman, P. R., Van Etten, J. L., and Rossmann, M. G. (2011) Three-dimensional structure and function of the Paramecium bursaria chlorella virus capsid. Proc. Natl. Acad. Sci. U.S.A. 108, 14837–14842