A Personal Account of Chaperonin History

George H. Lorimer*  
Center for Biomolecular Structure and Organization and Department of Chemistry and Biochemistry, University of Maryland, College Park, Maryland 20742

“Oh wad some Power the giftie gie us  
To see oursels as ithers see us!  
It wad frae monie a blunder free us.”  
Robert Burns, 1786

The history of the chaperonins involves two diverse, seemingly unrelated observations dating back to the 1970s—the genetics of bacteriophage morphogenesis and the synthesis of Rubisco during the biogenesis of chloroplasts. They coalesce in the late 1980s with the demonstration by the groups of Georgopoulos and Ellis that GroEL and the Rubisco-binding protein are homologous to one another (9) and with our demonstration in 1989 of what it was the chaperonin proteins really do (7).

GENETICS OF BACTERIOPHAGE MORPHOGENESIS

In the early 1970s Costa Georgopoulos isolated temperature-sensitive (ts), mutant strains of Escherichia coli that were unable to support the growth of phage λ and a number of other bacteriophages (6 and references therein). The E. coli gene responsible for this phenotype was eventually given the name groE. By dint of some further genetics Costa, together with Barbara Hohn, identified the groE gene product, a protein with a subunit \( M \) of approximately 65 that we now know as GroEL. This enabled the groups of Roger Hendrix and Barbara Hohn to purify and record the first, now familiar, electron micrographs of GroEL. It was a 14-mer consisting of two heptameric rings stacked back to back. It also hydrolyzed ATP. Quite what GroEL did in the cell and how it was involved in phage morphogenesis was not known at this stage. Nevertheless, Hohn et al. (10) wrote “In groE mutants, bacteriophage T4 capsid protein aggregates in lumps at the cell membrane instead of forming normal capsids; bacteriophage \( \lambda \) head protein aggregates to form polyheads, spirals and inactive prehead-like particles and bacteriophage T5 tail assembly is affected. These phenomena could be explained by assuming an assembly aiding function present in normal cells but missing or malfunctioning in the mutant.”

In retrospect you can see that they were describing a malfunction in protein folding that leads to aggregation instead of the native state. But recall also that these observations were made immediately after Anfinsen’s seminal work (1) that provided the dominating intellectual framework in the 1970s and 1980s in all matters related to protein folding. By then the prevailing dogma was that protein folding was a spontaneous event and suggesting that the folding of some proteins was assisted by other proteins was tantamount to heresy.

With his bank of groE mutants Costa next designed a screen to look for suppressor mutations. This led to the discovery of GroES (6) and to evidence that the two groE gene products, GroEL and GroES, interacted with one another in vivo. Next Costa’s group purified GroES, showed that in vitro it formed a 1:1 complex with GroEL in the presence of ATP, that GroES inhibits the ATPase activity of GroEL, and that it had a ring-like oligomeric structure. Still the function of GroEL and GroES remained elusive.

Costa also reported two further genetic observations relating to GroEL and GroES the significance of which was not then immediately obvious but which, with the benefit of hindsight, would later become more significant. The first was that a temperature-sensitive mutation could be suppressed by over expressing GroEL and GroES (4). It is now apparent that many (not all) ts mutations are the result of a malfunction in protein folding at the restrictive temperature. Later, when we knew what GroEL and GroES really did, my colleagues at DuPont demonstrated that one could suppress ts-mutations in a wide variety of structurally unrelated proteins simply by over-expressing both GroEL and GroES in the cells harboring the ts-mutant genes (15).

The second important observation relates to the indispensable nature of the chaperonins. By the late 1980s it had become apparent that GroEL and GroES were an important part of the heat shock response. Using a genetic approach Costa’s group demonstrated that both GroEL and GroES were indispensable proteins for bacterial growth at all temperatures between 20°C and 40°C (5). Again, the full significance of this result does not really become apparent until one knows the function of GroEL and GroES within the cell, assisting other proteins to fold. There are few events more central, more fundamental to the cell.

BIOSYNTHESIS OF RUBISCO

In the late 1970s, John Ellis and his colleagues found that one could study protein synthesis in chlor-
roplasts by feeding radiolabeled amino acids to intact, isolated chloroplasts. The major product of such a synthesis is the large subunit of Rubisco. After incubating intact, isolated chloroplasts with [35S]Met for varying lengths of time, Barraclough and Ellis (2) quenched the reaction by osmotically lysing the chloroplasts and removed the green membranous material. They first subjected the soluble material to non-denaturing PAGE. After electrophoresis they either stained the gel with Coomassie or subjected it to autoradiography. They observed that the radioactivity was initially associated with a slower migrating band than the holoenzyme of Rubisco. Only later did radioactivity appear together with the holo-enzyme, implying a precursor-product relationship. When the slower migrating radioactive band was excised and subjected to SDS-PAGE, the radioactive band comigrated with the large subunit of Rubisco, whereas the Coomassie stain was associated with a protein with a subunit mass of about 60 kD. They concluded that the synthesis of the holo-enzyme of Rubisco involved the transient formation of a complex with another large, oligomeric protein, the Rubisco subunit binding protein (RsuBP). Because at that time there was no intellectual framework into which the observation could be fitted, so pervasive was the Anfinsen view (1), this observation had little impact beyond those interested in chloroplast biogenesis or Rubisco. Ellis’s group subsequently purified RsuBP and noted that it was an oligomer of approximately 750 kD with a subunit mass of approximately 60 kD and that it bound ATP (11).

About then Harry Roy’s group did a series of pulse-chase experiments following the fate of the Rubisco large subunits (13). They showed that the “hot” large subunits first sedimented as a 29S complex; this was the binary RsuBP•large subunit complex seen in the Ellis’s non-denaturing gels. Next, in the chase the “hot” large subunits sedimented as a broad 7S species. With the benefit of hindsight and the crystal structure, I believe that these 7S species were a mixture of L2 dimers and (L2)2 tetramers en route to forming the (L2)4 octomeric core. They also showed that ATP was needed to chase the material out of the 29S complex into the 7S species. Finally, the 7S species was chased into the holoenzyme, which sediments at 18S. In retrospect these papers were very insightful, although at the time they were unappreciated given the prevailing dogma.

In 1987 at a NATO meeting John Ellis first used the term “molecular chaperone” in referring to RsuBP. Later (3) he speculated on the existence of a class of proteins (molecular chaperones) “whose function is to ensure that the folding of certain other polypeptide chains and their assembly into oligomeric structures occur correctly.” He went on to suggest that molecular chaperones “do not form part of the final structure nor do they necessarily possess steric information specifying assembly.”

Shortly after that Tony Gatenby and I decided to take a serious look at the RsuBP. By “serious” I meant mechanistic studies with purified materials. So I purified some RsuBP while Tony developed an E. coli-based in vitro transcription-translation system for making radiolabeled large subunits. However, our first experiment yielded a surprise. While the autoradiogram of the non-denaturing-PAGE gel clearly showed the presence of RsuBP-large subunit binary complex, the control, to which no RsuBP had been added, also showed the very same complex! I rudely suggested to Tony that he must have added RsuBP to both samples. Tony was not amused! But we did not wait long for an explanation, for within a short time John Ellis informed us that RsuBP was homologous to GroEL, which obviously (in hindsight) was present in the E. coli transcription-translation system. Our response was “GroEL? What’s that?”

### CONNECTING PHAGE MORPHOGENESIS AND RUBISCO SYNTHESIS

A protein in plants similar to GroEL was first observed by Tsuprun and Pushkin in 1981 (12). They reported electron micrographs and other properties of a protein from pea leaves indistinguishable from GroEL. However, in the absence of a function for this protein the significance of these results was not obvious or appreciated. A further 6 years were to pass before the connection between phage morphogenesis and Rubisco synthesis was established (9). This came about by the cloning and sequencing of RsuBP and of the GroE operon. RsuBP and GroEL were clearly homologous proteins. The gene for GroES was located immediately upstream of GroEL. The authors begin the discussion as follows: “We have described a ubiquitous, conserved, abundant protein that is associated with the post-translational assembly of at least two structurally distinct oligomeric protein complexes. We conclude that the role of this protein is to assist other polypeptides to maintain or assume conformations which permit their correct assembly into oligomeric structures.”

The function of RsuBP in assisting the assembly of Rubisco in the chloroplast is immediately obvious. Less obvious was the function of GroEL and GroES in E. coli. It is clearly not there to assist the folding of Rubisco, a protein not normally found in E. coli; nor can the sole function of GroEL and GroES be the assembly of new phage particles. It has to be doing something of benefit to E. coli. So we (Pierre Goloubinoff, Tony Gatenby, and I) reasoned that if the synthesis of Rubisco in chloroplasts involves RsuBP, the chloroplast version of GroEL, then the synthesis of recombinant Rubisco in E. coli would also involve GroEL and perhaps also GroES. We had two plasmids, one encoding the cyanobacterial Rubisco operon for both large and small subunits and the other encoding the dimeric L2 Rubisco. While expres-
sion of these proteins in E. coli did yield some soluble, biologically active Rubisco, most of the Rubisco was an insoluble, inactive inclusion body. Pierre engineered a plasmid encoding the GroE operon so that one could over-express GroEL and GroES and Rubisco in the same E. coli cell. The total amount of Rubisco protein expressed by E. coli was approximately the same whether or not one also over-expressed GroE. What was dramatically different, however, was the fate of Rubisco. In the cells that over-expressed GroE nearly all of the Rubisco was soluble and biologically active. For this to occur required both GroEL and GroES post-translationally at some point between the formation of the nascent polypeptide and the formation of the L2 dimer (8).

At this point I felt we had a great opportunity to figure out exactly what GroEL and GroES did. We knew the identity of one of its in vivo substrates, the large subunit of Rubisco. We knew how to assay Rubisco. We had also narrowed the involvement of GroE to two possibilities: either the folding of the monomer and/or the association of the folded monomers to give the biologically active dimer. Some have described our next step as "very bold." We purified GroEL and GroES to homogeneity and set out to do some neo-Anfinsen experiments using chemically denatured L2 Rubisco as the substrate.

The results of the very first such experiment sent Pierre over the moon. "It works, it works! We’ve got it! I was less impressed. It was true that no active Rubisco was formed at all if one omitted any of the components (GroEL, GroES, Mg2+, or ATP). But compared with the control with native Rubisco, the recovery was miserable, approximately 1% as I recall. However, I spotted a potential flaw in the experimental protocol. Thinking that GroEL/ES would act as an enzyme, Pierre had used substrate (unfolded Rubisco) in molar excess. In short there was too much Rubisco in the pot. So Pierre was sent back to the bench with instructions to use equimolar amounts of unfolded Rubisco and GroEL 14-mers. Wow! Now we recovered >80%. It was almost too good to be true and so the very next day I repeated (successfully) the experiment with my own hands. Now I was over the moon! Now we could do the kind of experiments that enzymologists like to do with purified components instead of with messy, ill-defined mitochondrial or chloroplast extracts. It did not take us very long to work out some of the essential features, and we reported this in a paper to Nature (7).

The Nature paper contains many mechanistic insights. We begin by defining three species of Rubisco on the basis of circular dichroism: native Rubisco N, unfolded (in urea or GdnHCl) Rubisco U, and acid denatured Rubisco A. Rubisco A had some secondary structure, and we suggested that it might correspond to a molten globule, a term that was then fashionable but which I no longer consider very useful. We were lucky on two counts. First, we arbitrarily chose conditions for refolding that were non-permissive; i.e. no reconstitution of enzymatic activity occurred in the absence of the complete system (GroEL, GroES, and MgATP). Second, there was a fourth component needed that we luckily included, a monovalent cation K+ or NH4+. Our preparations of GroEL and GroES were stored as ammonium sulfate precipitates, and enough NH4+ was carried over into the reaction mix to satisfy that particular requirement.

We showed that ATP hydrolysis was required by quenching the reaction with hexokinase plus Glc (to convert the ATP to ADP). The yield of refolded Rubisco depended on whether one started with Rubisco U or Rubisco A, but the rate constant was the same regardless. We concluded that "reconstitution involves a common intermediate, the formation of which precedes the same chaperonin-dependent rate-determining step. We propose that on removal of the denaturant by dilution, a rapid chaperonin-independent folding of Rubisco U to a state rich in secondary structure and resembling Rubisco A (the Rubisco I state), is followed by reaction with GroEL, GroES and Mg-ATP."

We noted that one required a slight molar excess of GroEL oligomers over Rubisco monomers. We attributed this to the instability of the Rubisco-I state and established that, absent a molar equivalent of GroEL oligomers, Rubisco-I underwent irreversible aggregation. We wrote "when either Rubisco U or Rubisco A was diluted into solutions containing the chaperonins, two competitive and mutually exclusive processes occurred. The unfolded or partly folded protomers either formed biologically unproductive aggregates or they formed a binary complex with GroEL, which prevented aggregation and directed the protomers along a biologically productive pathway." Later, we amplified this in the discussion "It is clear that the concept of a stable substrate ‘patiently’ awaiting interaction with substoichiometric quantities of a catalyst does not apply here. Instead, there is an urgent need to stabilize this fickle intermediate by formation of a binary complex with GroEL."

We next demonstrated the existence of this binary complex by electrophoretic and immunological methods commenting that "the need to sequester unstable folding intermediates evidently exists in vivo as well as in vitro. The formation of a binary complex between the Rubisco large subunit and chloroplast GroEL was demonstrated long ago." Here we refer to the Barraclough-Ellis experiment of a decade before (2).

We next showed that the discharge of this binary GroEL-Rubisco-I complex depended on GroES and MgATP. We concluded that "the chaperonin-dependent reconstitution of Rubisco involves a strictly ordered set of reactions—first, the formation
of a stable binary complex of GroEL•Rubisco-I, which at least in vitro can be demonstrated as an MgATP- and GroES-independent event, followed by the MgATP- and GroES-dependent discharge of folded and stable, but catalytically inactive monomers, which subsequently assemble into active dimers.”

In our concluding paragraph we also noted that “the E. coli chaperonins clearly did not evolve to facilitate the folding or assembly, or both, of Rubisco, a protein not normally found in E. coli. It follows that the interaction between GroEL and the partly folded Rubisco cannot be specific for Rubisco alone. Instead, the specificity of the interaction must lie in some so far unidentified structural element of the partly folded protein. Because the native protein shows no detectable interaction with GroEL, this structural element must be missing from, or inaccessible in, the native protein.”

It was not very long before purified, chemically denatured proteins were being thrown at every molecular chaperone and heat shock protein known to man. Gone were the messy experiments with organelles and other variants of chicken soup.

In the ensuing 10 years much progress has been made unraveling the mechanistic and structural details of the chaperonin nano-machine (for reviews, see 16 and 14). Paradoxically, however, the problem that Tony Gatenby and I set out to investigate 12 years ago, how to refold hexadecameric, higher plant Rubisco $S_4 (L_2)_4 S_4$ from its denatured subunits, remains unsolved. To date, despite several forlorn attempts both with or without chaperonins, not so much as a whiff of Rubisco activity has been reconstituted. This only goes to show that the grand old protein of plant biology still has a few secrets to divulge.

LITERATURE CITED