

## Recent Insights into Biological Regulation from Cell-Free Protein-Synthesizing Systems

VISHWANATH R. LINGAPPA, M.D., PH.D.<sup>1</sup>, AND JAISRI R. LINGAPPA, M.D., PH.D.<sup>2</sup>

### Abstract

We review the important role that cell-free protein-synthesizing systems (CFPSS) have played in the history of modern biology, and highlight two recent applications that illustrate their continued utility for the exploration of otherwise intractable aspects of gene expression and its regulation. Viral capsid assembly recreated in CFPSS reveals a catalyzed biochemical pathway involving transient, energy-dependent action of host proteins and discrete assembly intermediates, rather than the classical notion of self-assembly that was expected for capsid formation. Study of prion protein biogenesis reveals a new conformation critical for disease pathogenesis and advances the paradigm of *protein bioconformatics*, by which cells may productively regulate the folding of various proteins. In each example, the CFPSS made it easier to analyze biochemical mechanism than is possible in other currently available whole cell systems, illustrating why this approach is likely to be a continuing source of insight into important features of biological regulation.

**Key Words:** Protein synthesis, in vitro translation, biogenesis, protein folding, viral capsid, prion, wheat germ, reticulocyte lysate, translocation, endoplasmic reticulum, translocon.

---

### Cell-Free Protein-Synthesizing Systems as a Tool for Studying Cellular Function

PROTEIN SYNTHESIS IS THE primary biochemical process by which the information encoded in individual genes is expressed (1). Because of this central role in gene expression, the ability to carry out protein synthesis in a test tube, that is, to establish a cell-free protein-synthesizing system (CFPSS), can be a powerful first step towards understanding currently intractable features of biology. Many biological processes are difficult to study in humans. Investigating simpler model organisms, such as bacteria, yeast, worms, fruit flies, zebra fish, or

mice (2–7) or using mammalian cells in culture (8), often makes new insights possible, provided that the features observed in those organisms, or in cultured mammalian cells, are shared with humans. The development and study of such model systems has accounted for much of the rapid progress achieved in our understanding of gene expression over the past two decades. It is important to recognize, however, that there are limitations to model organisms: bacteria don't have the intracellular organelles found in eukaryotes; yeast lack a dimension of complexity in intercellular communication observed in metazoans; and even mice are different from humans in important aspects of both normal physiology and disease pathogenesis (9). These and other confounding features limit the extent, albeit often substantial, to which the study of model organisms illuminates human biology and diseases (10, 11). Even human cells in culture, especially immortalized lines, can be highly aberrant compared to their tissue of origin (12). Thus, a complete understanding of biology comes from integrating results obtained using multiple approaches, each with different strengths and weaknesses.

Even when simpler organisms or cultured cells *do* provide insight relevant to human biology, another difficulty persists: many biochemical path-

---

From the <sup>1</sup>Departments of Physiology and Medicine, University of California, San Francisco, San Francisco, CA; and <sup>2</sup>Department of Pathobiology, University of Washington, Seattle, WA.

This work was supported by grants from the NIH to VRL and to JRL.

A version of this manuscript was presented by VRL as a Dean's Lecture at the Mount Sinai School of Medicine, New York City, NY on October 30, 2002, and updated as of September 2004.

Address all correspondence to V.R. Lingappa, M.D., Ph.D., Department of Physiology, University of California, San Francisco, 513 Parnassus Avenue, San Francisco, CA 94143-0444; email: vrl@itsa.ucsf.edu

### Glossary of Abbreviations

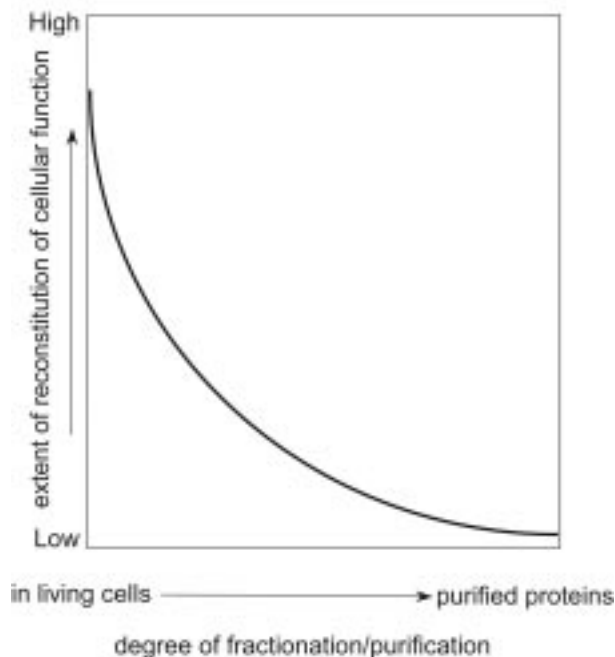
AR = autoradiography	PK = proteinase K
ATP = adenosine triphosphate	PrP = prion protein
C = carboxyterminal	RMJ = ribosome-membrane junction
CFPSS = cell-free protein-synthesizing system(s)	SARS = severe acute respiratory syndrome
CHO = Chinese hamster ovary	SDS-PAGE = sodium dodecyl sulfate-polyacrylamide gel electrophoresis
CoA = coenzyme A	Sec = secretory
EM = electron microscopy	SIV = simian immunodeficiency virus
ER = endoplasmic reticulum	SRP = signal recognition particle
GSS = Gerstmann-Straussler-Sheinker	StAR = steroidogenic acute regulatory
GTP = guanosine triphosphate	TrAF = translocation accessory factor
HBV = hepatitis B virus	TRAM = translocation-associating membrane
HMWC = high molecular weight complex	TRAP = translocon-associated protein
Mb = membranes	WG = wheat germ embryo extract
MDR-1 = multidrug resistance gene-1	wtPrP Tg = wild type PrP transgenic
N = amino terminal	

ways are simply difficult to study in the larger context of other events happening at the same time within the cell. This is because living eukaryotic cells are: (a) *extremely complex*, with a bewildering array of interacting gene products and subcellular components (13); (b) *extremely crowded*, in terms of protein concentrations within the cell (14); (c) *incredibly fast*, in terms of the speed with which they are able to carry out biochemical events; and (d) surrounded by a *plasma membrane* that separates their inner workings from the outside world, obscuring the former from the prying eyes of the scientist. There are several general ways by which these difficulties have been dealt with, in order to gain insight into how events are integrated in the complex environment of living cells.

First, molecular genetics can be brought to bear in model organisms, through techniques such as gene knockout (15), transgene expression (16, 17), microarray analysis (18), and proteomics (19). One limitation of many of these approaches is that they do not inform us directly about protein-protein (or protein-nucleic acid) interactions that are the basis for biochemical mechanism (the knowledge of which is critical for pharmaceutical drug development, for example). Thus, a pathway may have steps and interacting partners that are not easily detected (20). Similarly, detection of a change in gene expression (e.g., by microarrays) doesn't provide insight into the mechanism involved. Molecular genetic analysis can also be confounded by the tendency of important pathways to have evolved redundant, alternate routes that allow an organism to survive, should a primary pathway be inactivated (21–23). Thus, it may only be under special circumstances (e.g., stress states) that deficiency of the primary pathway is fully manifested (24).

Second, biochemical approaches can be used to study molecular interactions directly—but at a price

(Fig. 1; 25). Biochemical analysis is *reductionistic*: by isolating specific events away from the complex workings of the intact cell to varying degrees, biochemistry allows focus on a subset of events. Sometimes this is powerful, elucidating important but previously unappreciated molecular interactions. Sometimes it is misleading, because important interactions with other pathways within the cell are



**Fig. 1. Inverse relationship between the degree of purification of proteins and physiological relevance of the findings generated with those proteins.** The greater the degree of protein purification, the more detailed is the resulting knowledge of the protein. However, the knowledge generated, while detailed, is missing key aspects of physiological relevance in rough proportion to the degree of purification and isolation of the proteins. This is a variation on the theme of the “Heisenberg Uncertainty Principle” as it applies to biology.

lost, and irrelevant ones can come into play. Often it is both, advancing our thinking in some ways, while misleading us on other aspects (26, 27).

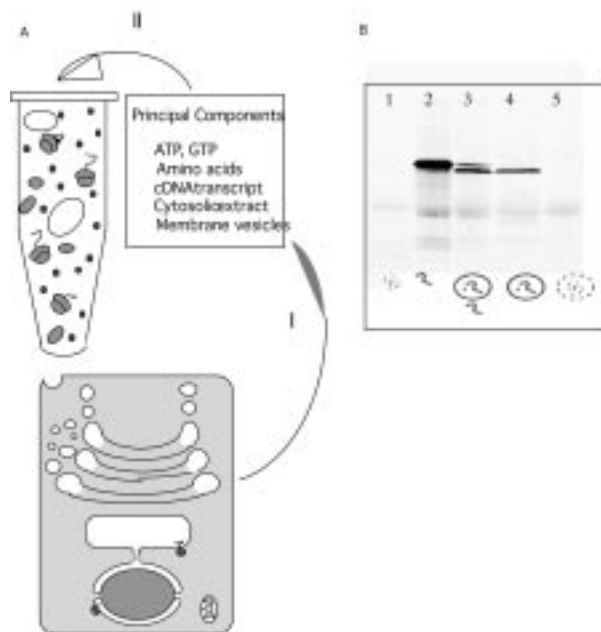
At its extreme, the biochemical approach involves the study of purified proteins, in order to facilitate elucidation of their sequence and structure, and some aspects of function: for example, purified proteins are necessary to grow protein crystals whose X-ray diffraction patterns provide the most precise structural information. However, this approach cannot easily be applied to proteins whose function is intrinsically linked to other complex processes within the cell. Protein synthesis is an example of a linked process that is frequently ignored when proteins are studied in isolation. Synthesis of a protein and related post-translational modifications are often treated as separate from the function of the protein, but in doing so intrinsic relationships important within the cell may be missed. This limitation is quite significant, because protein synthesis and post-translational events are central to gene expression. However, as will be described below for specific examples, studies from CFPSS allow many events that are linked to protein synthesis and post-translational events to be reconstituted and thus make possible their experimental investigation with far greater fidelity than is often possible with purified proteins.

CFPSS can be used in two very different ways. In the more commonly used practice, it makes possible expression of the protein encoded by a gene of interest (28). This is typically performed using various commercial cell-free translation kits, and represents the *end* of the investigation (e.g., of the cloning and expression of a gene of interest). A second use of these systems is based on their ability to reconstitute other events that have a critical relationship to protein synthesis. Examples of steps that are different from protein synthesis, but have a special relationship to that process, include post-translational modifications, trafficking, assembly events, and various downstream protein-protein interactions, including some related to a protein's ultimate function. Thus, CFPSS can be used to study the biological mechanisms involved in the linked events, *while maintaining the connection to protein synthesis*. In this approach, the synthesis of the protein is *just the beginning*, providing the context in which other more functional aspects of gene expression/regulation can be explored. A critical issue is whether protein synthesis and subsequent linked post-translational and regulatory events observed in the CFPSS occur in a manner that faithfully reflects what occurs *in vivo*. In multiple examples, some of which will be discussed here, the fidelity of CFPSS has been corroborated by findings *in vivo*. The power of this use

of CFPSS has not been fully appreciated by many biochemists and molecular biologists, and is the focus of the present review.

To prepare a CFPSS, plant or animal tissues are chosen that are specialized and/or primed for production of new proteins. For example, simple addition of water to seeds results in activation of dormant biochemical pathways in the embryo, including a burst of protein synthesis. This makes wheat germ embryo extracts (WG) a powerful starting material for preparation of a CFPSS (29, 30). Likewise, reticulocytes (immature red blood cells) are highly focused on production of a single protein, globin (the protein component of hemoglobin, which serves to carry oxygen in blood), making them an excellent starting material from which to prepare a CFPSS (rabbit reticulocyte lysate, RRL) (31, 32).

CFPSS start with a *crude extract* made from disrupted cells. Typically, initial disruption is carried out by homogenization, a process by which cells or tissue samples are subjected to shearing forces that mechanically break open cells. The resulting *homogenate* is then subject to centrifugation to remove unbroken cells and some large subcellular components that have been experimentally found to interfere with new protein synthesis. Disruption of the plasma membrane renders the cytoplasm and its organelles directly accessible to the experimenter. The resulting preparation contains all cytoplasmic proteins and many membrane-delimited organelles present in the cytoplasm, and is capable of synthesis of proteins *de novo*, when provided specific mRNA, amino acids, energy substrates and other components, and warmed to physiological temperatures. If one of the added amino acids is radiolabeled, it becomes possible to distinguish the newly synthesized protein encoded by the added mRNA, from the multitude of other proteins present in the crude extract, by analyzing the extract after completion of protein synthesis, by gel electrophoresis (Fig. 2). The ability to track the radiolabeled protein of interest as it interacts with unlabeled proteins of the cellular extract greatly simplifies analysis of various biochemical pathways that are in one way or another linked to protein synthesis. For example, the radiolabeled protein can be examined in the minutes following completion of synthesis to determine post-translational modifications or protein-protein interactions that may be necessary for a protein's function, measured separately in parallel aliquots. The extract itself can be subjected to varying degrees of manipulation prior to, during, and after protein synthesis, which can facilitate its use as an experimental system, as will be discussed below. Subcellular fractions can be removed or added before or



**Fig. 2. Cell-free protein-synthesizing systems (CFPSS).** (A) Schematic description of the components that comprise the CFPSS and their source. Cells are fractionated (I), and extracts prepared with or without various subcellular fractions, and (II) added to test tubes containing the energetic (e.g., adenosine triphosphate [ATP], guanosine triphosphate [GTP]) and substrate (e.g., amino acids, including one radiolabeled) requirements for protein synthesis and mRNA (e.g., cDNA transcript) encoding a particular protein of interest. When the mixture is warmed, specific protein synthesis proceeds in it, with the encoded (newly synthesized) protein radiolabeled. If microsomal membranes are present in the extract and the mRNA encodes a secretory or integral membrane protein, targeting to and translocation across the ER will occur. Newly synthesized proteins are distinguished from pre-existing proteins, because incorporation of the radiolabeled amino acid allows detection by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and autoradiography. (B) An autoradiograph of an SDS gel analyzing a reaction such as described in (A), where the translated polypeptide is a secretory protein (in this case, prolactin). Translation in the absence of added microsomal membranes is shown in lane 2. An aliquot of the material applied to lane 2 was first digested with proteinase K (PK) under standard (mild) conditions, resulting in complete degradation of the protein (see lane 1). Similar results are obtained when microsomal membranes are added after completion of protein synthesis, because only the *nascent* chain can target to the ER membrane. Lanes 3–5 show translation in the presence of added microsomal membranes. Lane 3 displays products after no other treatment. Note cleavage of the signal peptide, resulting in a small shift to lower molecular weight for most of the chains. Lane 4 displays the products after digestion with PK under mild conditions, with membranes intact. Note that the signal-cleaved form is protected from PK digestion because it is within the membrane vesicle, as a result of translocation. In Lane 5, the PK digestion takes place in the presence of nondeaturing detergent to abolish membrane integrity, rendering the protein fully accessible to PK. A schematic description of the consequences of each manipulation is shown at the bottom.

after carrying out the protein synthesis reaction, depending on the question of interest. For exam-

ple, when microsomal membranes derived from the endoplasmic reticulum (ER) are added prior to start of synthesis, they will engage secretory or membrane protein signal sequences and initiate transport of those nascent chains into the ER, the first stage of the secretory pathway (33–35). Thus, *temporal relationships* can also be appreciated. To continue the previous example, adding microsomal membranes more than a few minutes after initiation of protein synthesis results in failure of translocation of a newly synthesized chain, reflecting a need for precise timing of protein-protein interactions necessary for translocation of newly synthesized proteins across the ER membrane (36).

A number of characteristics of the CFPSS make it particularly useful for dissection of the molecular mechanisms involved in biological regulation. These features enable one to overcome problems that made analysis in living cells difficult in the first place (Table). First, empirically, the events that do occur in the CFPSS have been found to be remarkably *faithful* to the corresponding events in living cells. Second, while mechanistically faithful, the CFPSS tends generally to be relatively *slow and inefficient*, compared to the same process occurring in living cells. This proves to be a most fortuitous characteristic, in that it allows events that take place too rapidly to be detected and studied easily *in vivo*, to be slowed down enough for analysis in the CFPSS. Much as a slow-motion replay of a videotape would reveal the sleight-of-hand by which a magician carries out his/her tricks, so analysis in a CFPSS can reveal features previously undetected in living cells. In some cases, armed with knowledge of the characteristics of a biological process from study in CFPSS, it has been possible later to reinvestigate the process in cells—and detect the signature features of the pathway that were originally missed, as will be discussed in greater detail for two examples below (37, 38).

Third, CFPSS tend to reconstitute only a subset of the events observed in cells. As a result, that subset of reconstituted events is magnified and enhanced by the CFPSS—put, as it were, under a molecular spotlight. In some cases, a blockade that prevents a pathway from proceeding beyond a certain

**TABLE**

*Notable Features of CFPSS*

Fidelity to events <i>in vivo</i>
Slow-motion compared to events in living cells
Focused on a subset of events observed <i>in vivo</i>
Amenable to manipulation <i>in cis</i>
Fractionation and reconstitution <i>in trans</i>

CFPSS = cell-free protein-synthesizing system(s)

point can be used to further limit the progress of events in the CFPSS, for example, by adding or subtracting components, by changing reaction conditions, or through programming the CFPSS with mRNAs encoding mutant proteins. All of these manipulations make the study of molecular mechanism easier (39, 40). In this way, the *relationship* of early steps to later ones can be elucidated and explored.

Fourth, these systems are amenable to manipulations that cannot be done easily *in vivo*. For example, if the mRNA added to the CFPSS is prepared from a cDNA truncated upstream of the termination codon, the ribosome engaged in synthesizing the protein will be maintained at the point of truncation (due to the absence of a termination codon), thereby providing a prolonged *snapshot* of interactions occurring on the nascent chain at that point in chain growth (41, 42).

A fifth advantage of the CFPSS is its ability to be fractionated and subsequently reconstituted from relatively purified components, allowing proteins responsible for core biochemical activities to be distinguished from those involved in biological *regulation* (43–45).

### A Brief Intellectual History of CFPSS

Modern biology is the epitome of an inductive science: almost nothing in it can be safely concluded from deductive first principles. Hypotheses that may seem implausibly complex at one point in time are often supported by experimental studies later (46, 47). Thus, new experimental data frequently “trump” older theories that are based on more limited working hypotheses. Conversely, no matter how well we *think* we understand a biological process, we can never exclude the possibility that there are as-yet-unappreciated steps that intervene between any two currently known steps. Rather, we can anticipate that future experimental findings will enrich, revise, and even overturn our thinking about things that we feel quite confident of today.

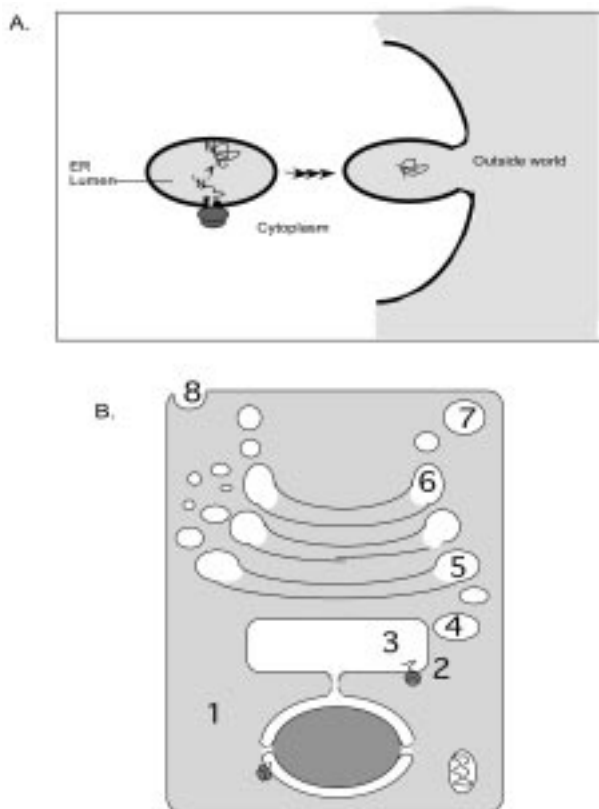
Classical biochemistry was born in the 1800s with the recognition that enzymes could be studied outside of cells (48), thus overturning the earlier physiological notion that the distinctive features of life could not be studied outside of living organisms (49). The power of this approach was demonstrated by the rapid progress made in analyzing biochemical pathways over the past century (50).

Similarly, cell biology, or *biochemical and biophysical cytology*, as it was once known (51), was born with the development of methods to look within cells. Previously, “protoplasm,” the stuff of living cells, was not seen as something that could

be rationally and systematically separated into distinctive components. However, the observation of obvious membrane-delimited compartments in cells, using the electron microscope (52), motivated the development of techniques for separating subcellular components by fractionation using ultracentrifugation of homogenates (53, 54).

Once organelles were defined by electron microscopy and correlated to specific enriched fractions generated by ultracentrifugation of cellular extracts/homogenates, the stage was set to use pulse-chase radiolabeling of cells, followed by subcellular fractionation and electron microscopic autoradiography, to define trafficking of proteins from their site of synthesis through their lifecycle (55–58). In this way, the *secretory pathway* by which proteins are transported out of the cell (secreted) was defined as a cell biological process (59; Fig. 3a).

The earliest step in protein secretion, namely translocation of the nascent chain across the membrane of the ER, proved to be a particularly delicate experimental challenge, because it occurs *while* the protein of interest is being made (Fig. 3b). This is a situation in which the focus of interest (i.e., the nascent chain), is constantly changing with the addition of amino acid residues, and thus, is not at equilibrium. For this reason, ferreting out the secrets of ER translocation involved some remarkable manipulations. First, systems were established in which protein chains that had been initiated *in vivo* were completed in cell-free extracts (60, 61). This set the stage for development of systems in which protein synthesis occurs *de novo*, i.e., the CFPSS. In those systems, the first step in the secretory pathway (translocation across the ER membrane), was reconstituted only if a particular subcellular fraction derived from the ER (termed “microsomal membranes”) was added (34, 62). Experiments in this system made it possible to appreciate a fundamental function of the ER: to initiate the sorting of *nascent* proteins destined for secretion from those destined to remain in the cytoplasm. These studies also elucidated the role of signal sequences in directing the nascent chain to receptor proteins that target the ER membrane, and explained why the ER membrane is typically studied with ribosomes engaged in the synthesis of (signal-sequence bearing) nascent chains. Finally, it was demonstrated that signal sequences were being removed *while the chain was still being made*, explaining why pulse-chase analysis of living cells generally failed to demonstrate a precursor-product relationship between completed, signal-sequence-containing, newly synthesized proteins and the mature, signal-sequence cleaved form



**Fig. 3. Early events in protein biogenesis at the endoplasmic reticulum (ER).** (A) Schematic representation of the ribosome-membrane junction (RMJ) at the ER and its relationship to the cell surface. Shown is the classical “closed” RMJ. Recent insights into protein bioconformatics demonstrate a newly appreciated level of complexity in this step of translocation. In the simplified example shown, the cytoplasm is white and the ER lumen is gray on the left. Ultimately, fusion of a vesicle derived from the ER with the plasma membrane exports products out of the cell. Note that the ER lumen is topologically contiguous with the outside world (also depicted as gray). (B) Outline of the major steps in the secretory pathway encompassed by the arrows between the vesicle and cell surface in A. Numbers refer to the major compartments traversed by a typical nascent secretory chain: (1) cytoplasm; (2) cytoplasmic face of the ER membrane; (3) ER lumen; (4) post-ER intermediate compartment; (5) cis-Golgi cisternae; (6) trans-Golgi cisternae; (7) secretory vesicles; and (8) exocytosis into the outside world.

(33). The importance of this lesson is often underappreciated: some biological processes are extremely difficult to “see” in living cells; the CFPSS offers opportunities to manipulate biochemical processes in ways that allow complex temporal and other special relationships between proteins to be analyzed and understood.

Much progress in understanding the function of the ER have been achieved since establishment of the new paradigm of protein targeting to and translocation across the ER membrane, as epitomized by the signal hypothesis (33). Most of these

new insights were generated first from studies using CFPSS (34). For example, the role of specific protein complexes, namely signal recognition particles (SRP) and SRP receptors was deciphered through modern variations on classical biochemistry: protein purification, manipulation and reconstitution (63–66). Some key insights into the mechanism of translocation from studies in the CFPSS included: (a) the role of SRP in translational pausing (67); (b) the role of guanosine triphosphate (GTP) as a regulatory mechanism for SRP/SRP receptor dynamics (68, 69); (c) the role of a special *protein-conducting channel* now termed the *translocon* (47); (d) the special role of SRP-receptor biogenesis in bypassing the general restrictions to protein traffic at the ER (70); and (e) the role of the ribosome-membrane junction (RMJ) as a novel point of regulation of chain translocation to the ER lumen (71). Most recently, studies of specialized nascent chains have revealed that the signal sequence plays an unexpected and still poorly understood role in distinguishing various pathways by which the chain can traverse the ER membrane (72).

While the focus of this discussion is the CFPSS, we do not mean to underplay the importance of other approaches to these difficult problems. For example, the Sec 61 complex of proteins that comprises the ER translocon was first identified as essential for secretion by a genetic approach in yeast (73). Likewise, the notion that not all secretory and membrane proteins are SRP-dependent also emerged from the study of yeast genetics (21). The point is that genetic and biochemical approaches complement one another. In particular, the CFPSS is often crucial for dissecting and understanding the details of protein-protein interactions (74, 75). The ER is not the only subcellular location whose function has been elucidated in significant measure by studies in CFPSS. Important post-translational protein-protein interactions in the cytoplasm (39, 40) (see below), the Golgi apparatus (76), and the mitochondria (77) have also been elucidated by studies of related cell-free systems.

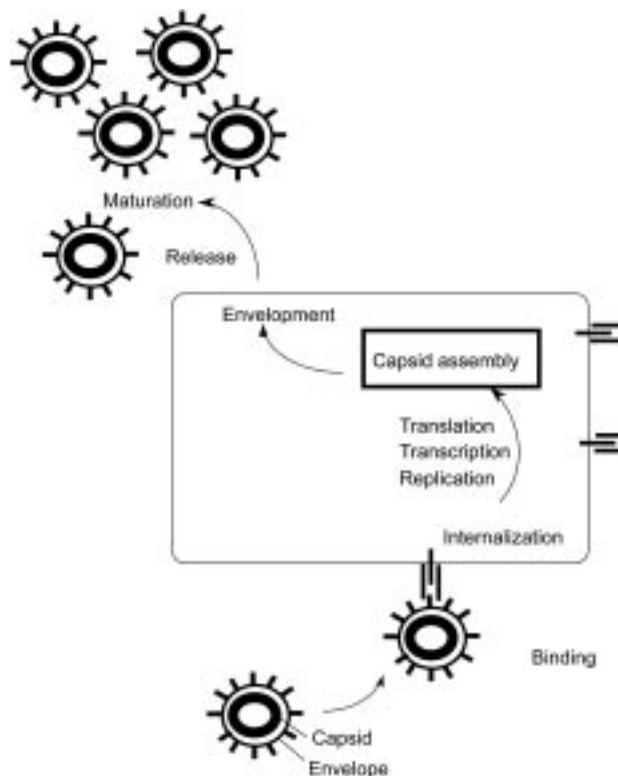
### Viral Capsid Assembly

In the last decade, CFPSS have provided provocative data for a new view of the mechanism of assembly of viral capsids. This process, by which a large, ordered, multimeric structure (the capsid) is assembled post-translationally from monomeric polypeptide units within the complex environment of the cytoplasm, has eluded biochemical dissection using other approaches, for many of the reasons discussed above. Here we will describe the rationale

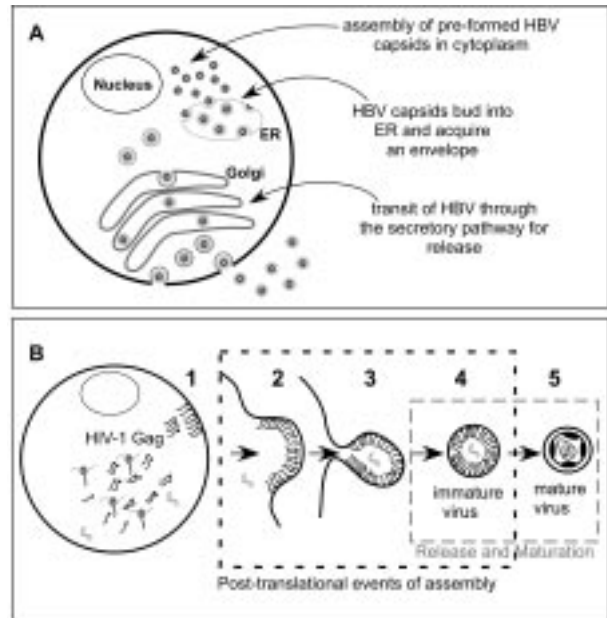
for using CFPSS for studying capsid assembly, the results obtained to date using such a system, and the implications of these findings for our understanding of virus cell biology.

While viral life cycles are extremely varied, some common themes emerge across many virus families (Fig. 4). The early events in the life of most viruses involve virion entry, virion disassembly, and replication of the viral genome, with different virus families using diverse strategies, cellular factors, and cellular compartments to achieve these goals (78). In contrast, the goal of the late events of the viral life cycle is production and release of new virions. Co-ordination to maximize virion production is critical for the success of the virus, and viruses have evolved diverse mechanisms for ensuring effective formation of infectious virions (Fig. 5).

The first step in virion formation involves assembly of the capsid, within which the viral genome and viral regulatory proteins are housed



**Fig. 4. The viral life cycle.** Schematic representation of the typical steps in the life cycle of an enveloped virus. Note that each viral family represents a unique variation on these themes; nevertheless, they must all bind to their unique receptor, internalize, replicate, transcribe and translate their proteins, assemble their capsids, and be released. In some cases, maturation occurs as a separate step or concomitant with release. Of these steps, capsid assembly has been a “black box” long ascribed to self-assembly.



**Fig. 5. Overview of hepatitis B virus (HBV) and HIV assembly.** (A) HBV capsids are composed of 180 or 240 copies of the HBV core protein that assemble into a 30 nm spherical capsid in the cytoplasm. These pre-formed capsids target to the endoplasmic reticulum (ER) membrane, where they are thought to exit the cell by traversing the secretory pathway. (B) The immature HIV-1 capsid is composed of close to 5,000 copies of the Gag polypeptide. Gag polypeptides are synthesized in the cytoplasm of the infected host cell and co-translationally myristoylated (1). Myristoylated Gag polypeptides target to the plasma membrane (2), which is where assembly of the spherical immature capsid occurs (3). As assembly of the 100 nm immature capsid is completed, it undergoes budding and envelopment by the plasma membrane, resulting in release of the immature virus particle (4). Subsequently, proteolytic cleavage of Gag mediated by the HIV-1 protease results in production of mature virus particles that have condensed, bullet-shaped capsids (5). Note that capsid formation of HBV and HIV involves coordination with encapsidation of the genome as well as packaging of viral and cellular proteins and factors that are critical for virus infectivity (not shown).

and protected. In the case of enveloped virions, the capsid subsequently acquires an envelope from host lipid bilayer membranes and undergoes release, directly or by a more circuitous process, into the extracellular space before, or in some cases after, maturation (79). For many viruses, the process of capsid formation within cells is less well understood than virion budding and release. Typically, viruses are described as having either spherical or helical capsids, based on electron microscopic images of capsid shape. Enveloped viruses from a number of unrelated families, including *Togaviridae*, *Flaviviridae*, *Retroviridae*, and *Hepadnaviridae*, form spherical capsids in the cytoplasm of eukaryotic host cells, often at the cytoplasmic face of a specific membrane compart-

ment such as the plasma membrane or ER. Examples of viral families with helical capsids include the Filoviridae such as Ebola virus and the Coronaviridae, including the virus causing severe acute respiratory syndrome (SARS). To date, only assembly of viruses with spherical capsids has been studied in CFPSS.

Viruses have evolved the means to assemble their capsids efficiently in the very crowded environment of the eukaryotic cytoplasm. The process of capsid assembly within the cell in the context of a viral infection may be extremely complex. Depending on the virus type, hundreds or even thousands of copies of a capsid protein are organized to form a single viral capsid. For example, a single hepadnavirus capsid is typically composed of either 180 or 240 copies of the viral core polypeptide, and a retroviral capsid is composed of thousands of viral Gag polypeptides. In addition, the viral genome is typically selected from a large pool of RNAs within the cytoplasm and packaged with appropriate stoichiometry during capsid formation (80). Furthermore, some viral capsids contain a large number of other viral proteins that also must be selectively packaged within the newly formed capsid. For example, human immunodeficiency virus (HIV-1) contains 15 viral proteins in addition to two copies of genomic RNA (81). Finally, capsid formation must be synchronized with the subsequent steps of envelopment and occur at the correct subcellular location to result in proper virion release. Thus, assembly of spherical capsids in the cytoplasm of an infected eukaryotic cell is most likely carefully regulated and highly coordinated.

From the vantage point of modern cell biology, it is clear that one must take into account the cytoplasmic environment where capsid formation occurs, to fully understand the mechanism of viral capsid formation. The cytoplasm is likely to be a source of obstacles and antiviral pathways that the virus must overcome or counteract for successful capsid formation. However, it could also be a source of factors and organelles that viruses have learned to commandeer in order to facilitate the process of capsid assembly. In fact, it is likely that the evolution of viral capsid assembly mechanisms was profoundly shaped by both obstacles and facilitators present in the ancient cytoplasm.

The question of how capsids assemble can be addressed on many levels. First, there is the question of how individual capsid proteins interact with each other and with other capsid components. On a larger level, there is the question of how these individual capsid protein interactions are influenced by post-translational events, cellular proteins, the cytoplasmic environment, and organelles. Yet an-

other level of complexity is introduced by studying sequential steps (discrete intermediates) in the formation of the capsid, and these need to be characterized in both time and space. Thus, capsid formation consists of sequentially occurring biochemical interactions with specific cellular proteins and organelles. The resulting complexes may be changing their subcellular location over time. Furthermore, these complicated events are probably subject to regulation, with molecular switches governing choices of pathways at each step in the temporal and spatial progression. Ironically, but not surprisingly, the question of how spherical capsids assemble was first addressed in an environment devoid of all cytoplasmic proteins or organelles during the 1970s. This was largely due to the growing awareness, during that time period, of the power of classical biochemical reductionism to provide a molecular view of a biologic process. Purified capsid proteins incubated in a buffer at extremely high concentrations (e.g., mg/mL) were found to have an intrinsic ability to "self-assemble" into spherical capsids. Such self-assembly was found to be typically promoted or "nucleated" non-specifically by oligonucleotides (Fig. 6a). This is perhaps best exemplified by seminal studies on tobacco mosaic virus (82), but has also been shown for retroviruses, flaviviruses, alphaviruses and other viruses, including phage and plant viruses.

While revealing an important property of capsid proteins, these self-assembly studies failed to take into account critical features of capsid assembly as it occurs in the context of the cell. These include the fact that, in cells, capsid assembly takes place in a cytoplasmic environment crowded with other proteins and processes, and occurs very, very quickly. A major contribution of CFPSS is their ability to allow assessment of such features, rendered difficult to assess by the conditions in living cells. In CFPSS, as in cells, the concentration of capsid protein is relatively low compared to the concentration of other (host) cytoplasmic proteins, which is as high as 300 mg/mL (83). This contrasts with the study of capsid proteins in isolation, in which the very high concentrations of capsid protein needed to drive the assembly reaction and cellular proteins are completely absent. A second critical feature common to cells and CFPSS is that capsid formation is temporally linked to translation of capsid proteins. In contrast, self-assembly systems use purified proteins, typically produced recombinantly from *Escherichia coli* and either purified in native form or denatured and renatured. Finally, cellular organelles, including membrane-bound compartments, are completely absent in recombinant self-assembly systems.

Nevertheless, despite these limitations, until the 1990s capsid assembly was studied largely using such self-assembly systems, combined with crystallography and mutational analysis of viral capsid proteins expressed in cells. Little was done to investigate the notions that cellular proteins could be involved in capsid formation, and that the pathway utilized by viruses in living cells might be different from that observed for self-assembly reactions. A notable exception to this trend was the demonstration that assembly of phage lambda heads fails to occur in the presence of mutations in the host bacterial GroE gene (84). Subsequent studies revealed that GroEL is a molecular chaperone involved in proper folding of newly synthesized proteins including phage lambda capsid proteins (85). In addition to ushering in the novel concept of molecular chaperones, these landmark studies introduced the idea that cellular factors may play a critical role in the viral life cycle by facilitating proper folding and assembly of viral structural proteins.

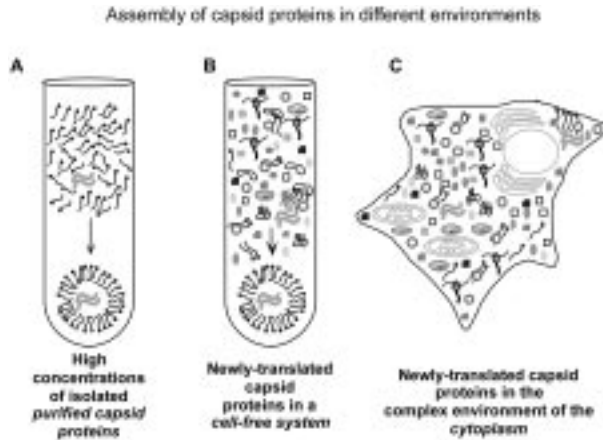
The first use of CFPSS for virion formation occurred in the early 1990s. In this system, infectious poliovirus was produced by programming a HeLa cell extract with poliovirus mRNA, demonstrating that CFPSS can faithfully reproduce all aspects of virion formation (86, 87). Subsequently, CFPSS was used to explore the mechanism of capsid assembly using hepatitis B virus (HBV) as a model system (40). The HBV capsid is relatively simple compared to other spherical capsids, and can be formed in both prokaryotic and eukaryotic cells simply by expression of the capsid protein. During infection, HBV capsids associate with the HBV envelope protein in the host ER, bud into the ER, and then traverse the secretory pathway for release outside the cell (88). Newly synthesized HBV capsids made in a CFPSS containing a wheat germ extract, closely resemble authentic HBV capsids by biochemical criteria (velocity sedimentation, buoyant density and protease resistance) and by the gold standard of electron microscopy (EM), demonstrating the fidelity of the system (40). Characterization of the steps involved in HBV capsid assembly in the CFPSS allowed the mechanism of HBV capsid assembly in a *cellular context* to be examined biochemically for the first time. Pulse-chase experiments in the CFPSS revealed that the HBV core progressed post-translationally through a high molecular weight complex (HMWC) with the characteristics of an *assembly intermediate*, before forming completed capsids (39). Characterization of the HMWC using immunoprecipitation revealed that it contained a cellular factor antigenically related to the molecular chaperone TCP-

1. Furthermore, when the HMWC-containing HBV core and the cellular factor were isolated from CFPSS, capsids could be generated by manipulation of this complex using energy substrates and other treatments (40). These findings suggest an alternate mechanism for capsid formation involving a *facilitated*, energy-dependent pathway, which proceeds via a discrete intermediate with cellular factors playing transient catalytic roles. This is very different from the previous notions of capsid formation through self-assembly.

The view of capsid assembly as facilitated, step-wise, and energy dependent was elucidated in more detail in a second and more challenging case. A CFPSS was established in which assembly of the immature capsid of HIV-1, a retrovirus, was achieved (39). Both the structure and the assembly pathway of the HIV-1 capsid are significantly more complex than that of HBV (89). The Gag polypeptide, the primary constituent of the HIV-1 capsid, undergoes a co-translational modification in which a 14-carbon fatty acid, myristic, is added to its N-terminus. Myristoylation in combination with specific N-terminal Gag domains targets Gag to the cytoplasmic face of the host cell plasma membrane, which is where the events of capsid assembly occur. Once the completed HIV-1 capsid is formed, it becomes enveloped by the plasma membrane, resulting in direct release to the extracellular space without traversing the secretory pathway (Fig. 6). Membrane targeting is thought to play an important role in concentrating Gag chains to a threshold required for assembly (90), although this remains to be demonstrated experimentally.

Thus, the challenge of establishing a CFPSS for HIV-1 capsid assembly was to reproduce the requirement for co-translational myristoylation and targeting to the cytoplasmic face of membranes. By adding purified myristoyl coenzyme A (CoA) and subfractionating wheat germ extracts, it was possible to produce a CFPSS that supported co-translational myristoylation of Gag and its targeting to endogenous membrane surfaces, and achieve *de novo* immature HIV-1 capsid assembly (39). Authenticity of immature capsids assembled in the CFPSS was supported by several criteria, including sedimentation value, buoyant density and EM appearance (compared to immature capsids produced in transfected mammalian cells). Furthermore, in this cell-free system, HIV-1 capsid formation failed to occur in the absence of co-translational myristoylation or if membrane integrity was abolished (39).

Having demonstrated formation of immature HIV-1 capsid in a manner that appears faithful to the product in cells, this system was utilized to ex-



**Fig. 6. Assembly of HIV-1 capsid proteins in different environments.** (A) Early studies of HIV-1 capsid assembly were performed using high concentrations of purified capsid proteins in isolation from cellular proteins and organelles. These studies revealed that capsid proteins have an intrinsic ability to self-assemble, but failed to give us insight into how capsid assembly occurs efficiently in the complex environment of the cell. (B) CFPSS were used to reconstitute HIV-1 capsid assembly by linking synthesis of capsid proteins to post-translational events of assembly in the presence of the high concentration of cellular protein as well as organelles. Thus, in contrast to self-assembly systems (A), the HIV-1 capsid assembly CFPSS (B) mimics many features of the HIV-1 assembly process that occurs in cells (C).

amine the mechanism of capsid formation. Pulse-chase experiments revealed that newly synthesized Gag polypeptides progress sequentially through a series of novel post-translational complexes of defined sizes (10S, 80S, 150S, and 500S) before culminating in formation of completed 750S immature capsids. Confirmation that these complexes are *bona fide* assembly intermediates came from the demonstration that assembly defective mutants in Gag, defined by others using cellular systems, were arrested at specific points in this assembly pathway, producing for each class of mutant, a subset of the same assembly intermediates that were observed for the wild type capsid protein in the CFPSS (39). In contrast, Gag mutants that were known to be assembly competent progressed through each of the assembly intermediates to the final 750S completed immature capsid (91).

Each of the key features of CFPSS indicated above (near the end of the first section) are illustrated in the studies analyzing viral capsid assembly. First, the fidelity of CFPSS in reproducing the capsid assembly pathway that occurs in cells has been confirmed by recent studies identifying capsid assembly intermediates in mammalian cells producing HIV-1 and other primate lentiviral capsids (36). Second, assembly intermediates are extremely transient, rapidly progressing, and are pre-

sent in very low steady-state quantities in mammalian cells. This explains why they had not been noticed before: by standard techniques; the signal-to-noise ratio was not sufficient to distinguish them from background. These assembly intermediates were first identified in a CFPSS because these systems reproduce events more slowly and inefficiently than do cells, as discussed above, allowing accumulation of these transient complexes to levels that were easily detectable. Third, since the CFPSS was programmed only with mRNA-encoding Gag polypeptide, resulting in synthesis and post-translational modification of Gag, the presence of the observed assembly intermediates could be attributed to capsid assembly. If they had been found first in the context and complexity of a cell, their relevance to capsid assembly would not have been obvious. Thus, identification of capsid assembly intermediates is an example of how the cell-free system can be a better model than living cells for initial dissection of complex biological systems. Of course, ultimately, these conclusions must be corroborated *in vivo*, as has been done for capsid assembly intermediates (37), but this is best done with a "heads up" as to what to look for, by virtue of studies of CFPSS.

The fourth advantage of CFPSS (as listed above) is that they provide the opportunity for manipulations that would be difficult to perform *in vivo*. This was illustrated by subsequent studies in which energy dependence was localized to the 80/150S step in the ordered pathway of assembly intermediates. The energy dependence was demonstrated by treating extracts *post-translationally* with an enzyme (apyrase) that hydrolyzes adenosine triphosphate (ATP) (39). While studies of ATP depletion can be performed in cells, its effects cannot easily be localized in time or space within a cell; this was achieved in the CFPSS. Since energy-utilization is a common hallmark of biological regulation, the observation of energy dependence during post-translational events in HIV-1 capsid assembly was crucial for overturning the notion that capsids self-assemble in cells, which was the dominant view at the time. Furthermore, this observation had important implications with respect to viral-host interactions. Since Gag does not contain ATP-binding domains, and Gag is the only viral protein programmed into the cell-free HIV-1 capsid assembly system, the requirement for ATP hydrolysis suggested that a cellular ATP-binding protein was needed for HIV-1 capsids to assemble.

The existence of a cellular factor required for HIV-1 capsid assembly had been indicated by previous fractionation and reconstitution data (39), illustrating the fifth advantage of CFPSS. Such an

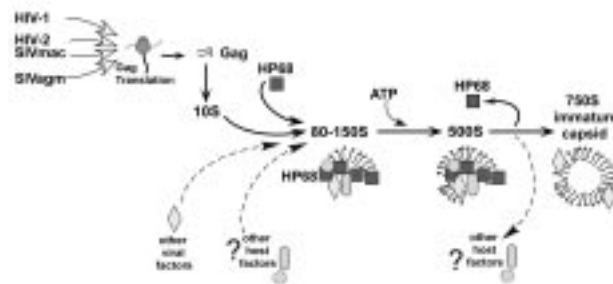
approach would not be possible in living cells. Together, these findings led to the search for a host factor critical for assembly, and resulted in identification of HP68 (also known as RNase L inhibitor), a cellular protein that contains two ATP-binding domains. HP68 was identified because it associates with 80S, 150S and 500S assembly intermediates, but not with completed 750S immature capsids in the CFPSS. In support of the fidelity of CFPSS, this association is also seen in transfected and infected primate cells (37, 92). The association of HP68 with Gag polypeptides in the assembly intermediates, but not the completed capsid, suggests that it plays a catalytic role in assembly analogous to that of enzymes in a biochemical pathway or of molecular chaperones in protein folding. It is puzzling why a protein that serves as a ribonuclease inhibitor should also facilitate building a viral capsid. Given this and other mysteries that surround the action of host factors involved in viral capsid assembly, the term “molecular facilitators” has been coined to distinguish them from conventional molecular chaperones, whose actions may or may not be similar.

Depletion and reconstitution in the CFPSS was used to determine whether HP68 is required for HIV capsid assembly. Immunodepletion of HP68 from the CFPSS resulted in an extract that supports near-normal levels of HIV-1 Gag synthesis but no longer allows efficient formation of 750S completed capsids. Presumably, the small amount of residual assembly observed in depleted extracts reflected the catalytic action of residual HP68. When recombinant purified HP68 was added back to the immunodepleted CFPSS extract, efficient 750S capsid assembly was reconstituted, supporting an essential role for HP68 in capsid formation. These findings were corroborated in living cells by demonstrating that dominant negative mutants of HP68 *blocked* release of HIV from transfected mammalian cells and HIV-infected human T cells (92).

Subsequently, other groups have confirmed the role of ATP in retrovirus assembly using cellular systems (93–95). However, since these studies did not distinguish between use of energy for capsid protein synthesis versus post-translation events in capsid assembly, we designed another approach to confirm the existence of an energy-sensitive step in capsid assembly in cells. We demonstrated that when extracts of primate cells expressing the nearly complete HIV-1 genome were harvested in the presence of ATP, the energy-dependent 80S and 150S complexes were present in minuscule quantities, suggesting that Gag polypeptides progress rapidly through these complexes when energy is not limiting. However, when harvested under ATP-depleted

conditions, all of the Gag-containing post-translational assembly intermediates we had identified in the cell-free system could be seen biochemically. The assembly intermediates that were found in the cell-free system fit the criteria previously defined for these complexes in the CFPSS. That is, they contained HP68 associated with Gag only in the 80S, 150S, and 500S complexes, and were not associated with Gag in unassembled 10S complexes or in the fully completed 750S capsids (37). Finally, these studies in cells allowed us to extend findings obtained for HIV-1 assembly to other members of the family Lentiviridae. Genetically divergent primate lentiviruses from 3 separate lineages all appear to utilize the HP68 and ATP-dependent, stepwise assembly pathway for immature capsid assembly (Fig. 7) when expressed in cells (37). Further cellular studies and, ultimately, animal studies will be important for confirming the accuracy of observations initially made in CFPSS.

Through studies of HBV and HIV capsid formation, the CFPSS has given us a radically different view of how viral capsids assemble. In this new view, the conceptual framework of self-assembly has been replaced by a cell-biological paradigm involving a stepwise, ordered biochemical *pathway* for capsid formation, characterized by discrete intermediates, energy and subcellular



**Fig. 7. Immature primate lentivirus capsids assemble in an energy-dependent manner by a pathway of assembly intermediates containing the cellular factor HP68.** Studies in CFPSS followed by confirmatory studies in cellular systems have revealed that HIV-1 assembles capsids by sending newly synthesized viral Gag polypeptides through a sequential pathway of assembly intermediate complexes identified by their sedimentation values (10S, 80–150S, and 500S). Other primate lentiviruses, including human immunodeficiency virus type 2 (HIV-2) and strains of simian immunodeficiency virus that infect African green monkeys (SIVagm) and macaques (SIVmac), also utilize this assembly pathway. In the 80–150S assembly intermediate, Gag polypeptides become associated with the host protein HP68. HP68 remains associated with the nascent virion until just before completion of the 750S immature capsid structure. HP68 has been shown by two methods to be essential for completion of capsid assembly. Other viral proteins present in completed virions are also found in assembly intermediates and recent studies suggest that other cellular factors may be present as well.

compartment-specific steps, and engagement of host proteins serving as catalysts or *molecular facilitators*. The role of such molecular facilitators in assembly may be similar to that of enzymes that serve to speed up chemical reactions that would occur spontaneously to a small extent in the absence of the enzymatic facilitator. There is now ample evidence for the existence of the same facilitated, energy-dependent assembly pathways in cells assembling primate lentiviruses. However, finding these assembly intermediates in cells would have been extremely difficult without the road map made available by studies in CFPSS, which facilitate detection of transient and complex post-translational regulatory events that are tightly linked to protein synthesis.

### The Biogenesis of Complex Proteins at the ER

Translocation across the ER membrane provided an important focus for the original development of the modern CFPSS. Therefore, it should not be surprising that this system has contributed significantly to our current understanding of complex protein biogenesis at the ER. Some of these insights are summarized below.

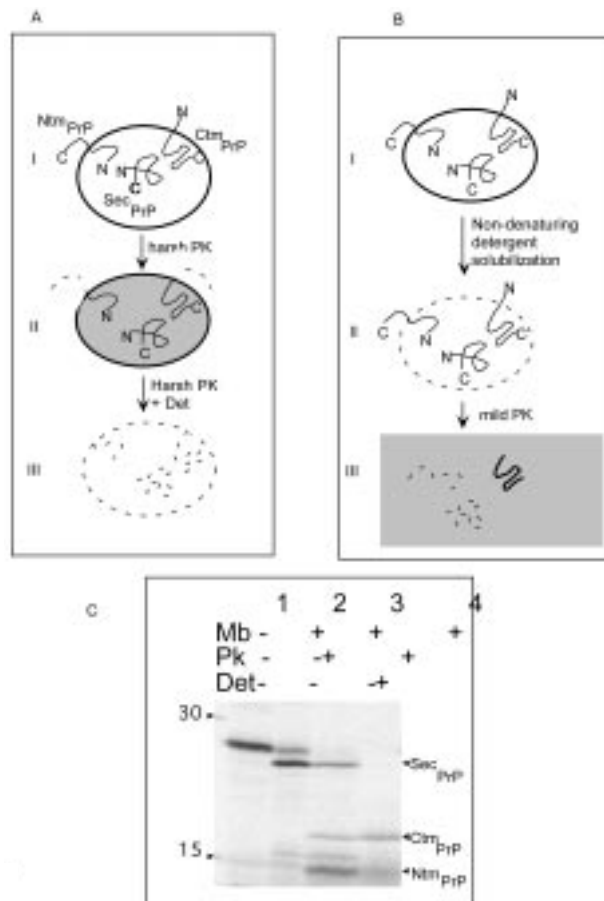
### Prion Protein

The prion protein (PrP) is a widely expressed glycoprotein which has been implicated in an extraordinary group of neurodegenerative disorders, some of which are transmissible, and others hereditary; still others occur sporadically in the absence of either family or exposure history (96). Initially, a distinction was made between “normal” PrP, i.e., that observed in uninfected cells (PrP<sup>C</sup>) and the form observed in infectious prion disease (PrP<sup>Sc</sup>) (97). Infectious prion disease involves conversion of PrP<sup>C</sup> into PrP<sup>Sc</sup>, an abnormally folded form that triggers neurodegeneration and, upon entry into cells, converts endogenous PrP<sup>C</sup> into more PrP<sup>Sc</sup>, thereby conferring the “infectious” character that is a hallmark of these disorders. Until recently, the normal function of PrP (PrP<sup>C</sup>) has been quite controversial. Some data has suggested that it is involved in apoptosis (98). Other data has argued for a neuroprotective function (99). Subsequent work suggests a far greater complexity and heterogeneity of what had been termed PrP<sup>C</sup>, which sets the stage to better understand previously mysterious features of prion disease (38, 44).

Studies of the CFPSS have contributed greatly to our current understanding of PrP physiology and pathophysiology. Upon expression in CFPSS, an initially *homogeneous* population of nascent PrP

chains is completed as a *heterogeneous* mix of three forms (termed “conformers”), identical in primary amino acid sequence, but distinctive in transmembrane topology and final folded state (38). That these forms are different in transmembrane topology can be demonstrated unambiguously by digestion with proteinase K (PK) under harsh conditions with membranes intact (Fig. 8A). However, upon digestion with PK under mild conditions, in the absence of membrane integrity, it is possible to distinguish intrinsic differences in PrP conformer folding (Fig. 8B). One form, termed C<sup>tm</sup>PrP, spans the membrane with its N terminal half in the cytoplasm and is intrinsically protease resistant. That is, upon solubilization of the ER membrane with nondenaturing detergents (i.e., which do not disturb protein folding), a “signature” mild PK digestion-resistant peptide fragment remains. Another form is termed SecPrP, because the region that spans the membrane in C<sup>tm</sup>PrP is fully translocated to the ER lumen in the CFPSS. *In vivo*, mature SecPrP is found anchored to glycolipids on the cell surface. SecPrP shows no protease resistance once membranes are solubilized with nondenaturing detergents. A third form, termed N<sup>tm</sup>PrP, spans the membrane with carboxy terminus in the cytoplasm, and, like SecPrP, shows no PK resistance, even with mild PK conditions, when membrane integrity is abolished. The significance of N<sup>tm</sup>PrP is unknown, although some have claimed to detect it *in vivo* (100). Fig. 8C illustrates how the CFPSS allows these forms to be detected (contrast this with Fig. 2B). Until these studies, it had not been possible to resolve PrP<sup>C</sup> into its three component conformers, and so, the possibility that different conformers are responsible for different subsets of the functions previously ascribed to PrP<sup>C</sup>, could not be appreciated. *In vivo* PrP heterogeneity is bewildering, with differentially glycosylated forms (e.g., utilizing either, both or none of the two available N-glycosylation acceptor sites), and endogenous proteolytic metabolic cleavage products (101), complicating the analysis and making it harder to resolve conformer distinctions that are seen clearly in the CFPSS.

The pathophysiological significance of PrP conformer heterogeneity was clarified in a series of studies (27, 38). First, the CFPSS was used to screen for mutations that skewed the distribution of PrP towards either C<sup>tm</sup>PrP or SecPrP. Such mutations were found and introduced as transgenes into a null background (i.e., into mice whose endogenous PrP gene had been deleted by homologous recombination). Those mice with transgenes favoring C<sup>tm</sup>PrP developed rapid spontaneous neurodegeneration with spongiform change and astroglio-



**Fig. 8. Assays of protein prion (PrP) biogenesis in CFPSS.** (A) Topological assay. Harsh proteolysis with PK with intact membrane vesicles removes domains outside the bilayer only (panel II). Membrane solubilization with nondenaturing detergent (Det) abolishes topological differences, so that all forms are degraded. (B) Conformational assay. Mild proteolysis with PK, with intact membranes vesicles, gives the same result as in the harsh protease topological assay. Solubilization of the membrane with nondenaturing detergent abolishes topological differences (panel II), but not differences in conformation. Thus, mild protease digestion under these conditions demonstrates intrinsic differences in folded state by protection of a signature fragment of C<sup>tm</sup>PrP (panel III). (C) Analysis of PrP biogenesis in cell-free translation systems with mild proteolysis, followed by SDS-PAGE and autoradiography (AR). Lanes 1–3 demonstrate the topological assay, displaying PrP synthesis in the absence of microsomal membranes (Lane 1), synthesis in their presence (lane 2) and after PK digestion (Lane 3). Lane 4, the conformational assay, involves solubilization of the protecting lipid bilayer so that the protein is fully accessible to PK. However, since PK digestion conditions used were “mild” conditions under which one conformer is fully digested (Sec<sup>PrP</sup>), while another conformer, C<sup>tm</sup>PrP, displays a signature protease-resistant fragment.

C = carboxyterminal; N = N terminal; Sec = secretory, and Mb = membrane.

sis, the pathological hallmarks of prion disease, and were found to have C<sup>tm</sup>PrP in their brains, but no PrP<sup>Sc</sup> (27). Furthermore, the brains of these sick

mice did *not* transmit prion disease when inoculated into healthy mice. Thus, infectivity and pathogenicity in prion disease appear distinct. Some of the mutations that favor C<sup>tm</sup>PrP are found in patients with Gerstmann-Straussler-Sheinker (GSS) syndrome, a hereditary dementia. Studies of patients with GSS revealed the presence of C<sup>tm</sup>PrP in their brain tissue. Mice with those transgenes developed spontaneous neurodegeneration, again without PrP<sup>Sc</sup>, and without transmissibility to other animals.

Those mice with transgenes favoring Sec<sup>PrP</sup> developed no spontaneous disease and were found to be protected from infectious prion disease. Upon inoculation with infectious prions, they developed disease very late and only after enormous levels of PrP<sup>Sc</sup> had accumulated. Sec<sup>PrP</sup> is the only mature conformer of PrP observed at steady state in normal adult brains of both humans and wild type PrP transgenic (wtPrP Tg) mice. A body of work (27) suggests that Sec<sup>PrP</sup> is the conformer of PrP that is the substrate for production of PrP<sup>Sc</sup>. A cell culture system that reproduces the conformer distribution achieved in CFPSS supports the notion that Sec<sup>PrP</sup> has a physiological function antagonistic to that of C<sup>tm</sup>PrP. Thus, Chinese hamster ovary (CHO) cells transfected with Sec<sup>PrP</sup>-favoring mutants appear *protected* from various forms of apoptosis (Saghafi et al., unpublished).

Finally, C<sup>tm</sup>PrP has been implicated in the signaling pathway by which PrP<sup>Sc</sup> causes neurodegeneration in infectious prion disease, by two independent lines of evidence. First, a striking correlation is observed between the propensity of a mutation to favor C<sup>tm</sup>PrP in the CFPSS, adjusted for the level of expression in a particular Tg line (termed the C<sup>tm</sup>-index), and the level of accumulation of PrP<sup>Sc</sup> at time of onset of clinical disease, after inoculation with scrapie (27). Together with other data, these findings suggest the following hypotheses:

- C<sup>tm</sup>PrP allows the cell to monitor the level of PrP<sup>Sc</sup> and responds to PrP<sup>Sc</sup> accumulation by triggering apoptosis;
- PrP<sup>Sc</sup> inactivates pathways that normally suppress C<sup>tm</sup>PrP expression. In this view, cells expressing C<sup>tm</sup>PrP-favoring PrP mutations have elevated levels of nascent C<sup>tm</sup>PrP and therefore are already closer to the threshold level of C<sup>tm</sup>PrP needed to exit the ER and trigger apoptosis, even prior to inoculation with PrP<sup>Sc</sup>. Upon inoculation, PrP<sup>Sc</sup> replication triggers C<sup>tm</sup>PrP export from the ER. Thus, a smaller amount of PrP<sup>Sc</sup> is needed to reach that threshold in

the case of Tg lines expressing  $C^{tm}PrP$ -favoring mutations.

These hypotheses and observations allowed formulation of a hypothesis to relate the different categories of prion disease and clarify the distinction between the events of *transmission* and the separate events of *disease pathogenesis* (see Fig. 9, stage I vs. stage II/III).

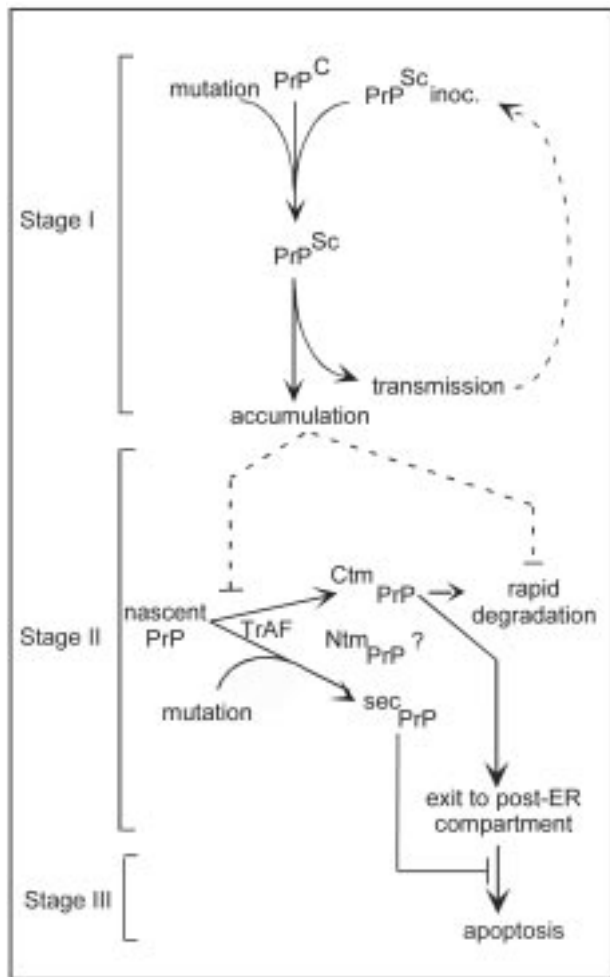
A body of work has clarified some of the molecular mechanisms by which different PrP con-

formers are made at the ER. First microsomal membranes were solubilized, fractionated and reconstituted with the few gene products absolutely needed for translocation across the ER (SRP receptor; Sec61p; and in some cases, TRAM). When wtPrP is synthesized in the presence of these membranes,  $C^{tm}PrP$  is the only conformer observed (44). Addition of a glycoprotein fraction reconstitutes formation of the  $SecPrP$  conformer, upon wtPrP translation (44). Subsequent work has identified translocon-associated protein (TRAP) alpha as a component of the glycoprotein complex needed to make  $SecPrP$  (103). Since  $SecPrP$  may have a role in protection from apoptosis and  $C^{tm}PrP$  appears to be a trigger of apoptosis (Saghafi et al., unpublished), this ER membrane glycoprotein complex is a "protective" factor that prevents apoptosis by redirecting the nascent PrP chain away from the default pathway leading to the formation of  $C^{tm}PrP$  (44). Other data from our lab suggest that at particular times in development,  $C^{tm}PrP$  is induced in order to trigger physiological apoptosis (V. Lingappa, unpublished). Thus, prion disease appears to be a case in which a physiological mechanism (triggering of apoptosis by  $C^{tm}PrP$ ) is activated at an inappropriate time and/or place.

When a mutation that is expressed only as  $SecPrP$  was translated in the presence of minimal membranes (which make only  $C^{tm}PrP$ ), the chains were found to target but not translocate across the ER membrane (44). These studies of PrP biogenesis suggested the existence of *multiple pathways* across the ER membrane, and that the correct fit between translocation substrate and accessible pathway is necessary, if the chain is to be faithfully translocated. The mutation in question appears to prevent the PrP chain from accessing the pathway utilized to form  $C^{tm}PrP$ , and the minimum membranes are missing the factor needed to establish the pathway utilized to form  $SecPrP$ . Subsequent studies have reinforced this conclusion (72, 104–106, and Ott and Lingappa unpublished).

The studies to date on PrP have highlighted the power of the CFPSS (Table):  $C^{tm}PrP$  was first detected in the CFPSS but later confirmed by studies in Tg mice; whereas  $C^{tm}PrP$  is difficult to detect *in vivo* at steady state in the adult, it is readily detected in the CFPSS, because only a subset of events occurring *in vivo* are reconstituted under standard conditions in the CFPSS (e.g., ER degradation is not). Mutations expressed in the CFPSS have clarified some aspects of PrP biogenesis (104–106), and manipulations in *trans* (fractionation and reconstitution) (44, 103).

CFPSS have not only allowed recognition of conformational heterogeneity in the initial prod-



**Fig. 9.** Steps in prion protein pathogenesis. Relationship of infectious cycle (stage I) to pathway of *de novo* PrP biogenesis (stage II) to pathway of apoptosis (stage III), copied with permission from Hegde RS, Tremblay P, Groth D, et al. Transmissible and genetic prion diseases share a common pathway of neurodegeneration. *Nature* 1999; 402(6763):822–826 (27). The three conformers that comprise what was previously known as  $PrP^C$  were first identified in CFPSS. This proposal also accounts for how a point mutation of PrP can result in spontaneous neurodegeneration (by favoring  $C^{tm}PrP$  production and activation of apoptosis), without generation of infectious  $PrP^{Sc}$ , which is necessary for spread of transmissible prion disease.

TrAF = translocation accessory factor.

ucts of PrP biogenesis, but have also allowed this phenomenon to be appreciated in a more measured fashion than would have been possible *in vivo*, where the bewildering plethora of diverse modified forms that constitutes heterogeneity of PrP is more daunting.

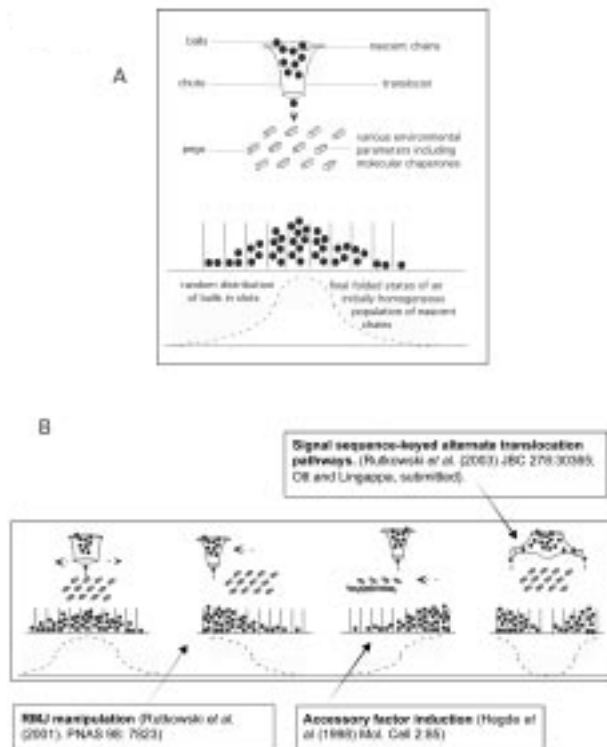
### Protein Bioconformatics

Initially, the evidence for alternate folded forms, or conformers, for PrP in the CFPSS was interpreted as revealing an unusual feature of PrP biology. However, a growing body of evidence suggests that the conformational heterogeneity observed for PrP is *not* unique or even unusual, and may well be a widespread property of secretory, membrane, and even cytosolic proteins. The seminal observation in favor of this radical new view was that replacing the signal sequence of PrP with that of a variety of different simple secretory proteins such as prolactin and immunoglobulin heavy chain, resulted in dramatic changes in the ratio of SecPrP to C<sup>tm</sup>PrP upon synthesis in the CFPSS (106). If the signal sequence's role is simply to target to the ER and translocate, then why should there be the observed correlation of specific signal sequences with the generation of a particular mix of conformers? The subsequent discovery that different PrP conformers have dramatically different functions (C<sup>tm</sup>PrP triggers apoptosis in transfected cells, SecPrP protects transfected cells from apoptosis and oxidative stress, Saghafi et al., unpublished) suggests that this observation is significant and raises some interesting possibilities.

One possibility is that the signal sequence plays a role beyond targeting to, and translocation across, the ER membrane *per se*. Specifically, the signal sequence appears to be a critical determinant of the *pathway* across the ER membrane. Hence the variation long noted between signal sequences might be *functional*, used as a means of fine-tuning the folding pathway to be taken by the nascent chain. That is, rather than variations between signal sequences reflecting the *lack* of specific information restricting their divergence, these variations may reflect the fact that different proteins have *selected* mutations of the signal sequence that access the particular mix of folding funnels desired by the cell for that particular protein. Such selection would be complex, occurring in concert with degradative and other mechanisms that allow the cell to “sculpt” the mix of conformers generated, in response to physiological or other cues.

The second implication of this observation is that perhaps what is unusual about PrP is not that it is made in multiple conformers, but rather that

those conformers are *easy to detect because they manifest as differences in topology*, which are far easier to score than subsets of chains of different conformation. A study of proteins other than PrP has now explored this hypothesis and suggests that it may be more generally correct, and that multiple mechanisms exist by which conformational hetero-



**Fig. 10. Protein bioconformatics.** (A) Analogy to protein folding through the translocon. The balls represent nascent chains, the chute represents the translocon, the pegs represent molecular chaperones and environmental parameters that alter protein-protein interactions between the nascent chain and translocon, and the slots at the bottom of the figure represent the diverse final folding states that achieve relative energy minima to which the chain can be directed, and into which the balls ultimately distribute. (B) The conformational distribution can be changed by manipulating one or more of several (separately regulated) parameters, including the accessible environment (e.g., cytosol vs. luminal), as a consequence of open vs. closed ribosome-membrane junctions, specific protein-protein interactions during the course of translocation, timing of events and modifications, and pathways out of the translocon. Over the course of evolution, the necessary “waste” of energy for synthesis of chains that achieve alternate conformations undergoes selection for particular purposes. In some cases, evolution may have closed off certain pathways by selecting signal sequences that do not allow the chain access to those folding funnels. In other cases, contingency was maintained perhaps by coupling to signaling cascades, so that one pathway or another was accessible at different times (e.g., in response to phosphorylation of particular trans-acting proteins). PrP biogenesis is probably an example of the latter. Note that evidence has been generated for three of the hypothetical examples of regulation indicated; see references 72 and 106.

geneity can be regulated in response to the physiological needs of the cell (Fig. 10) (72). One such mechanism appears to involve the RMJ. Long thought to form a tight seal between the secretory ribosome and the ER membrane, it has been observed that in some cases an open RMJ is established, while under other circumstances a closed RMJ is formed. When the RMJ is open, specific information later in the chain seems capable of effecting RMJ closure. When the RMJ is closed, specific information later in the chain can direct its opening. Opening and closing of the RMJ appears to be achievable either by discrete sequences in the mature chain, termed pause transfer sequences (42, 71), or by different classes of signal sequences themselves (106). Thus, while all signal sequences seem to direct their “passenger” nascent chains to the ER membrane, the precise *pathway* across the membrane (via the translocon) that is achieved depends, in part, on which class of signal sequence the protein has evolved.

Classes of signal sequences appear to differ in ways beyond the characteristics of the RMJ they establish. In a case where two signal sequence affect folding of two unrelated proteins (prolactin and TIGR) *without* opening the RMJ, a change in organization of the translocon (as demonstrated by chemical cross-linking to a truncated nascent chain in the CFPSS), has been correlated with each outcome (72). However, it remains to be determined whether these two forms of conformational regulation are truly different mechanisms, or whether they may represent different manifestations of a more limited set of mechanisms. Finally, factors such as TRAP alpha, an integral membrane glycoprotein, have been implicated in redirecting nascent chains in *trans* (103). For some of these mechanisms (e.g., regulation of the RMJ) there may be multiple sub-states (e.g., closed vs. partially open vs. largely open vs. completely open), rather than simply open vs. closed RMJs. In other cases, there may be a number of pathways through the translocon. In still other cases, there may be *trans*-acting factors in the cytoplasm, the ER membrane or the ER lumen. Each of these mechanisms may be keyed by different classes of sequences, encoded in different classes of proteins. In some cases, signal sequences play this role. Possibly, a hierarchy of such mechanisms is activated as a nascent chain progresses across the ER membrane, closing off some folding funnels while opening others, in the course of chain growth and translocation.

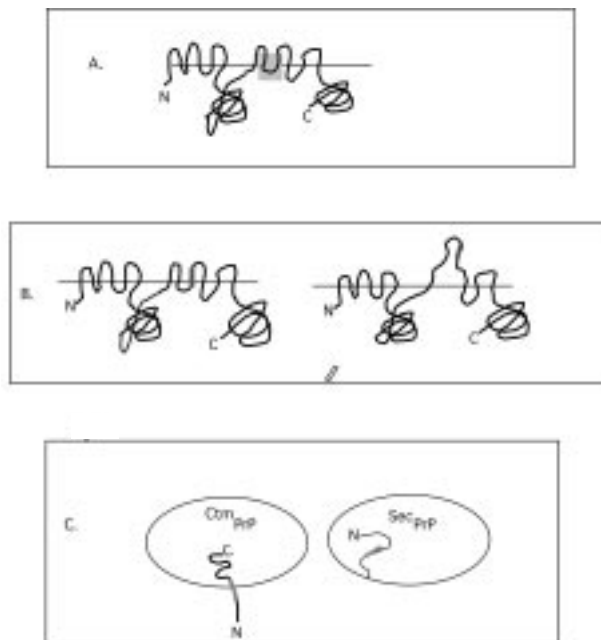
A third implication of the signal sequence swapping experiments is that seemingly subtle mutations in a signal sequence can have dra-

matic effects on the folding of the nascent chain “passenger.” Thus, the conventional view that signal sequences of different proteins display many small changes in precise amino acid sequence because they are relatively nonspecific in their function should be turned on its head: the precise sequence of the signal peptide has been tailored to direct the passenger to the precise mix of final folded states selected over the course of evolution. Further work is needed to corroborate or refute this new view of signal sequence heterogeneity.

### Multispanning Membrane Proteins

It is remarkable that swapping seemingly “simple” signal sequences from different secretory proteins affects the folding of the passenger nascent chain, even though the signal sequence is (eventually) cleaved. Whether the notion of protein bioconformatics growing out of this observation is generally applicable to most proteins or applies only to a small subset of secretory and membrane proteins, awaits further experimental investigation. Nevertheless, it is most tempting to apply the concept to biogenesis and folding of multispanning membrane proteins, where one might expect a great potential for heterogeneity in conformation/topology to expand the information content expressed in a gene product. However, these proteins are exceedingly difficult to study, and few studies have been carried out with precision to distinguish conformations/topological forms under conditions not confounded by rapid degradation in the ER. Here again lies an advantage of analysis in CFPSS.

In one case, that of multidrug resistance gene-1 (MDR-1), a striking finding highly analogous to the observations on PrP biogenesis has been made (107). It was found that this multi-spanning membrane protein, predicted to be made as two sets of six-membrane spanning loops, had a previously unanticipated degree of heterogeneity at loops 7–8. A little more than half of the chains had this sequence in the orientation expected (i.e., forming two transmembrane loops). However approximately 40% of the chains had the entire region localized to the ER lumen, in a manner reminiscent of PrP topological heterogeneity (see Fig. 11). This raises the possibility that certain membrane proteins utilize mechanisms similar to PrP to achieve different conformation/topology for different subsets of chains. Each of these subsets of chains of different conformation or topology may serve different functions in a manner that is extremely difficult to detect, let alone follow, *in vivo*.



**Fig. 11. Multispanning membrane protein heterogeneity. Summary of data from reference 107.** (A) Predicted topology of multidrug resistance gene-1 (MDR-1), a multispanning integral membrane protein of the ABC transporter superfamily. (B) Experimentally demonstrated mix of topological forms observed first in the CFPSS (107). (C) Similarity in local orientation of transmembrane loops of MDR-1 to topological heterogeneity of PrP, suggesting that common mechanisms may be at work.

### Anticipating Future Developments

It is possible that cytoplasmic proteins with no relationship to the ER also have sequences that, by association with *trans*-acting factors such as molecular facilitators, redirect the folding of nascent proteins. If activity of these *trans*-acting factors is entrained by signaling pathways, it becomes possible to induce a particular conformer in response to metabolic, developmental or cell type and tissue-specific circumstances. While some of these interactions may be disease, tissue, or cell-type specific, others may be more general and occur in CFPSS, as is true for redirecting nascent chains of PrP to SecPrP vs. C<sup>tm</sup>PrP, which will facilitate their detection and elucidation.

The complex cell-biology sleuthing in the CFPSS summarized above is still in progress, both for cytoplasmic proteins like HP68 and viral capsid proteins, and for secretory and membrane proteins like PrP, prolactin and MDR-1 and others. The full implications of the findings discussed here remain to be established. At the very least, they represent important variations on the theme of protein biogenesis, which need to be understood

for a full appreciation of human disease and biological diversity. In the broadest extrapolation, they represent a new way of thinking about protein folding and its regulation—protein bioconformatics. The studies of viral capsid assembly and protein bioconformatics discussed above amply illustrate the power of the CFPSS to jump start new lines of inquiry with respect to the mechanism by which these extraordinary biological regulatory events occur.

Recent findings suggest that CFPSS have a high potential for revealing other remarkable levels of biological regulation that were previously not anticipated. Thus, CFPSS have been instrumental in revealing that the steroidogenic acute regulatory (StAR) protein, whose activity is the rate-limiting step in steroidogenesis, and which is localized to the mitochondria matrix, carries out its biological function at the outer mitochondrial membrane. How it could work at a location other than that in which it was found was clarified by CFPSS studies that reconstituted StAR function and manipulated the signals for its biogenesis in a manner that revealed its novel mode of regulation (77). Likewise, the recent demonstration that the proximal and distal limbs of the secretory pathway probably “talk” to one another as part of a regulatory feedback loop engaged during specific protein synthesis (108) suggests still more aspects of biological regulation that the CFPSS may help decipher. Will CFPSS deliver fully on the promise suggested by these studies? Time will tell.

### Summary and Conclusions

We have presented a very different way of viewing CFPSS that is in continuity with the history of cell biology and biochemistry. We have showcased in detail two dramatic examples where use of CFPSS is ushering in new paradigms of thinking, first in viral capsid formation, and second, growing out of the study of prion protein biogenesis and leading to the theory of protein bioconformatics. We hope that the “take-home message” will be an appreciation of the power and still-to-be-tapped richness of this approach to complex issues in biology. In addition, we hope that the reader will be challenged to apply these systems to other topics in biology that are connected to protein synthesis but where the answers have been taken for granted. This is particularly true where some data already suggests a degree of complexity not addressed by simpler models. In such cases, the CFPSS may provide important and unexpected answers that must eventually be validated *in vivo*.

## Acknowledgments

We wish to thank Kevin Klein, Julia Dooher, and Lorne Walker for critical reading of the manuscript. VRL and JRL are scientific founders of Prosetta Corporation, whose focus is commercial application of the concepts described in this review.

## References

- Alberts B, Johnson A, Lewis J, et al. *Molecular biology of the cell*. 4th ed. New York: Garland Science; 2002.
- Davis RH. The age of model organisms. *Nat Rev Genet* 2004; 5(1):69–76.
- Castrillo J, Oliver S. Yeast as a touchstone in post-genomic research: strategies for integrative analysis in functional genomics. *J Biochem Mol Biol* 2004; 37:93–106.
- Putcha G, Johnson EJ. Men are but worms: neuronal cell death in *C. elegans* and vertebrates. *Cell Death Differ* 2004; 11:38–48.
- Guo S. Linking genes to brain, behavior and neurological diseases: what can we learn from zebrafish? *Genes Brain Behav* 2004; 3:63–74.
- Zhang Q, Calafat J, Janssen H, Greenberg S. ARF6 is required for growth factor- and rac-mediated membrane ruffling in macrophages at a stage distal to rac membrane targeting. *Mol Cell Biol* 1999; 19(12): 8158–8168.
- Lafaille J. T-cell receptor transgenic mice in the study of autoimmune diseases. *J Autoimmun* 2004; 22:95–106.
- Scott MR, Kohler R, Foster D, Prusiner SB. Chimeric prion protein expression in cultured cells and transgenic mice. *Protein Sci* 1992; 1(8):986–997.
- Winslow J, Insel T. The social deficits of the oxytocin knock-out mouse. *Neuropeptides* 2002; 36:221–229.
- Bogue C. Genetic models in applied physiology. *Functional genomics in the mouse: powerful techniques for unraveling the basis of human development and disease*. *J Appl Physiol* 2003; 94:2502–2509.
- Barr MM. Super models. *Physiol Genomics* 2003; 13(1):15–24.
- Dixon J, Ginsberg H. Regulation of hepatic secretion of apolipoprotein B-containing lipoproteins: information obtained from cultured liver cells. *J Lipid Res* 1993; 34:167–179.
- Soldati T, Geissler H, Schwarz E. How many is enough? Exploring the myosin repertoire in the model eukaryote *Dicystostelium discoideum*. *Cell Biochem Biophys* 1999; 30:389–411.
- Fulton A. How crowded is the cytoplasm? *Cell* 1982; 30:345–347.
- Zhao L, Bakke M, Hanley NA, et al. Tissue-specific knockouts of steroidogenic factor 1. *Mol Cell Endocrinol* 2004; 215(1–2):89–94.
- Ernst M, Jenkins BJ. Acquiring signalling specificity from the cytokine receptor gp130. *Trends Genet* 2004; 20(1):23–32.
- Bartke A, Chandrashekar V, Vailey B, et al. Consequences of growth hormone (GH) overexpression and GH resistance. *Neuropeptides* 2002; 36(2–3):201–208.
- Bryant PA, Venter B, Robins-Browne R, Curtis N. Chips with everything: DNA microarrays in infectious diseases. *Lancet Infect Dis* 2004; 4(2):100–111.
- Michnick S. Proteomics in living cells. *Drug Discov Today* 2004; 9:262–267.
- Higgin G, Jacobsen H. Transgenic mouse models of Alzheimer's disease: phenotype and application. *Behav Pharmacol* 2003; 14:19–38.
- Hann BC, Walter P. The signal recognition particle in *S. cerevisiae*. *Cell* 1991; 67(1):131–144.
- Ferreirinha F, Quattrini A, Pirozzi M, et al. Axonal degeneration in paraplegin-deficient mice is associated with abnormal mitochondria and impairment of axonal transport. *J Clin Invest* 2004; 113(2):231–242.
- Leiter E. Mice with targeted gene disruptions or gene insertions for diabetes research: problems, pitfalls, and potential solutions. *Diabetologia* 2002; 45:296–308.
- Thyagarajan T, Totey S, Danton MJ, Kulkarni AB. Genetically altered mouse models: the good, the bad, and the ugly. *Crit Rev Oral Biol Med* 2003; 14(3):154–174.
- Lingappa V, Farey K. *Physiological medicine*. New York: McGraw-Hill; 2000. p. 1006.
- Weissmann C, Flechsig E. PrP knock-out and PrP transgenic mice in prion research. *Br Med Bull* 2003; 66:43–60.
- Hegde RS, Tremblay P, Groth D, et al. Transmissible and genetic prion diseases share a common pathway of neurodegeneration. *Nature* 1999; 402(6763):822–826.
- Renesto P, Raoult D. From genes to proteins: in vitro expression of rickettsial proteins. *Ann N Y Acad Sci* 2003; 990:642–652.
- Herson D, Schmidt A, Seal S, et al. Competitive mRNA translation in an in vitro system from wheat germ. *J Biol Chem* 1979; 254(17):8245–8249.
- Endo Y, Sawasaki T. High-throughput, genome-scale protein production method based on the wheat germ cell-free expression system. *Biotechnol Adv* 2003; 21:695–713.
- Pelham H, Jackson R. An efficient mRNA-dependent translation system from reticulocyte lysates. *Eur J Biochem* 1976; 67:247–256.
- Shields D, Blobel G. Efficient cleavage and segregation of nascent presecretory proteins in a reticulocyte lysate supplemented with microsomal membranes. *J Biol Chem* 1978; 253(11):3753–3756.
- Blobel G, Dobberstein B. Transfer of proteins across membranes. I. Presence of proteolytically processed and unprocessed nascent immunoglobulin light chains on membrane-bound ribosomes of murine myeloma. *J Cell Biol* 1975; 67(3):835–851.
- Blobel G, Dobberstein B. Transfer to proteins across membranes. II. Reconstitution of functional rough microsomes from heterologous components. *J Cell Biol* 1975; 67(3):852–862.
- Lingappa VR, Devillers-Thierry A, Blobel G. Nascent pre-hormones are intermediates in the biosynthesis of authentic bovine pituitary growth hormone and prolactin. *Proc Natl Acad Sci U S A* 1977; 74(6):2432–2436.
- Rothman JE, Lodish HF. Synchronised transmembrane insertion and glycosylation of a nascent membrane protein. *Nature* 1977; 269(5631):775–780.
- Dooher JE, Lingappa JR. Conservation of a step-wise, energy-sensitive pathway involving HP68 for assembly of primate lentiviral capsids in cells. *J Virology* 2004; 78(4):1645–1656.
- Hegde RS, Mastriani JA, Scott MR, et al. A transmembrane form of the prion protein in neurodegenerative disease. *Science* 1998; 279(5352):827–834.
- Lingappa JR, Hill RL, Wong ML, Hegde RS. A multistep, ATP-dependent pathway for assembly of human immunodeficiency virus capsids in a cell-free system. *J Cell Biol* 1997; 136(3):567–581.
- Lingappa JR, Martin RL, Wong ML, et al. A eukaryotic cytosolic chaperonin is associated with a high molecular weight intermediate in the assembly of hepatitis B virus capsid, a multimeric particle. *J Cell Biol* 1994; 125(1):99–111.

41. Perara E, Rothman RE, Lingappa VR. Uncoupling translocation from translation: implications for transport of proteins across membranes. *Science* 1986; 232(4748):348–352.
42. Chuck SL, Lingappa VR. Pause transfer: a topogenic sequence in apolipoprotein B mediates stopping and restarting of translocation. *Cell* 1992; 68:9–21.
43. Gorlich D, Rapoport DA. Protein translocation into proteoliposomes reconstituted from purified components of the endoplasmic reticulum membrane. *Cell* 1993; 75(4):615–630.
44. Hegde RS, Voigt S, Lingappa VR. Regulation of protein topology by trans-acting factors at the endoplasmic reticulum. *Mol Cell* 1998; 2(1):85–91.
45. Hegde RS, Voigt S, Rapoport TA, Lingappa VR. TRAM regulates the exposure of nascent secretory proteins to the cytosol during translocation into the endoplasmic reticulum. *Cell* 1998; 92(5):621–631.
46. Engelman DM, Steitz TA. The spontaneous insertion of proteins into and across membranes: the helical hairpin hypothesis. *Cell* 1981; 23(2):411–422.
47. Simon S, Blobel G. A protein-conducting channel in the endoplasmic reticulum. *Cell* 1991; 65(3):371–380.
48. Barnett J, Lichtenthaler F A history of research on yeasts 3: Emil Fischer, Eduard Buchner and their contemporaries, 1880-1900. *Yeast* 2001; 18:363–388.
49. Barnett J. A history of research on yeasts 2: Louis Pasteur and his contemporaries, 1850-1880. *Yeast* 2000; 16:755–771.
50. Nirenberg M. Historical review: deciphering the genetic code—a personal account. *Trends Biochem Sci* 2004; 29:46–54.
51. Palade G. A small particulate component of the cytoplasm. *J Biophys Biochem Cytol* 1955; 1:59–68.
52. Palay S, Palade G. The fine structure of neurons. *J Biophys Biochem Cytol* 1955; 1:69–88.
53. Palade GE, Siekevitz P. Liver microsomes; an integrated morphological and biochemical study. *J Biophys Biochem Cytol* 1956; 2(2):171–200.
54. Siekevitz P. Protoplasm: endoplasmic reticulum and microsomes and their properties. *Annu Rev Physiol* 1963; 25:15–40.
55. Jamieson J, Palade G. Intracellular transport of secretory proteins in the pancreatic exocrine cell. I. Role of the peripheral elements of the Golgi complex. *J Cell Biol* 1967; 34:577–596.
56. Jamieson JD, Palade G. Intracellular transport of secretory proteins in the pancreatic exocrine cell. II. Transport to condensing vacuoles and zymogen granules. *J Cell Biol* 1967; 34:597–615.
57. Jamieson J, Palade G. Intracellular transport of secretory proteins in the pancreatic exocrine cell. 3. Dissociation of intracellular transport from protein synthesis. *J Cell Biol* 1968; 39:580–588.
58. Jamieson J, Palade G. Intracellular transport of secretory proteins in the pancreatic exocrine cell. IV. Metabolic requirements. *J Cell Biol* 1968; 39:589–603.
59. Palade G. Intracellular aspects of the process of protein synthesis. *Science* 1975; 189:347–358.
60. Redman C, Siekevitz P, Palade G. Synthesis and transfer of amylase in pigeon pancreatic microsomes. *J Biol Chem* 1966; 241:1050–1058.
61. Redman CM, Sabatini DD. Vectorial discharge of peptides released by puromycin from attached ribosomes. *Proc Natl Acad Sci U S A* 1966; 56(2):608–615.
62. Milstein C, Brownlee GG, Harrison TM, Mathews MS. A possible precursor of immunoglobulin light chains. *Nat New Biol* 1972; 239(91):117–120.
63. Walter P, Ibrahim I, Blobel G. Translocation of proteins across the endoplasmic reticulum. I. Signal recognition protein (SRP) binds to in-vitro-assembled polysomes synthesizing secretory protein. *J Cell Biol* 1981; 91:545–550.
64. Walter P, Blobel G. Translocation of proteins across the endoplasmic reticulum III. Signal recognition protein (SRP) causes signal sequence-dependent and site-specific arrest of chain elongation that is released by microsomal membranes. *J Cell Biol* 1981; 91:557–661.
65. Walter P, Blobel G. Translocation of proteins across the endoplasmic reticulum. II. Signal recognition protein (SRP) mediates the selective binding to microsomal membranes of in-vitro-assembled polysomes synthesizing secretory protein. *J Cell Biol* 1981; 91:551–558.
66. Siegel V, Walter P. Each of the activities of signal recognition particle (SRP) is contained within a distinct domain: analysis of biochemical mutants of SRP. *Cell* 1988; 52:39–49.
67. Wolin SL, Walter P. Signal recognition particle mediates a transient elongation arrest of preprolactin in reticulocyte lysate. *J Cell Biol* 1989; 109:2617–2622.
68. Connolly T, Rapiejko PJ, Gilmore R. Requirement of GTP hydrolysis for dissociation of the signal recognition particle from its receptor. *Science* 1991; 252(5010):1171–1173.
69. Miller JD, Wilhelm H, Gierasch L, et al. GTP binding and hydrolysis by the signal recognition particle during initiation of protein translocation. *Nature* 1993; 366(6453):351–354.
70. Andrews DW, Lauffer L, Walter P, Lingappa VR. Evidence for a two-step mechanism involved in assembly of functional signal recognition particle receptor. *J Cell Biol* 1989; 108(3):797–810.
71. Hegde RS, Lingappa VR. Sequence-specific alteration of the ribosome-membrane junction exposes nascent secretory proteins to the cytosol. *Cell* 1996; 85(2):217–228.
72. Rutkowski DT, Ott CM, Polansky JR, Lingappa VR. Signal sequences initiate the pathway of maturation in the endoplasmic reticulum lumen. *J Biol Chem* 2003; 278(32):30365–30372.
73. Novick P, Field C, Schekman R. Identification of 23 complementation groups required for post-translational events in the yeast secretory pathway. *Cell* 1980; 21:205–215.
74. Mandon E, Jiang Y, Gilmore R. Dual recognition of the ribosome and the signal recognition particle by the SRP receptor during protein targeting to the endoplasmic reticulum. *J Cell Biol* 2003; 162:575–585.
75. Song W, Raden D, Mandon EC, Gilmore R. Role of Sec61alpha in the regulated transfer of the ribosome-nascent chain complex from the signal recognition particle to the translocation channel. *Cell* 2000; 100(3):333–343.
76. Goda Y, Pfeffer SR. Cell-free systems to study vesicular transport along the secretory and endocytic pathways. *FASEB* 1989; 3:2488–2498.
77. Bose H, Lingappa VR, Miller WL. Rapid regulation of steroidogenesis by mitochondrial protein import. *Nature* 2002; 417(6884):87–91.
78. Knipe DM, Howley PM. *Fields virology*. 4th ed. Philadelphia: Lippincott Williams and Wilkins; 2001.
79. Hunter E. Virus assembly. In: Knipe DM, Howley PM, editors. *Fields virology*. 4th ed. Philadelphia: Lippincott, Williams and Wilkins; 2001. pp. 171–198.
80. Berkowitz R, Fisher J, Goff SP. RNA packaging. *Curr Top Microbiol Immunol* 1996; 214:177–218.
81. Frankel AD, Young JA. HIV-1: fifteen proteins and an RNA. *Annu Rev Biochem* 1998; 67:1–25.
82. Klug A. The tobacco mosaic virus particle: structure and assembly. *Philos Trans R Soc Lond B Biol Sci* 1999; 354(1383):531–535.
83. Ellis RJ. Macromolecular crowding: an important but neglected aspect of the intracellular environment. *Curr Opin Struct Biol* 2001; 11(1):114–119.

84. Georgopoulos CP, Hendrix RW, Casjens SR, Kaiser AD. Host participation in bacteriophage lambda head assembly. *J Mol Biol* 1973; 76(1):45–60.
85. Zeilstra-Ryalls J, Fayet O, Georgopoulos C. The universally conserved GroE (Hsp60) chaperonins. *Annu Rev Microbiol* 1991; 45:301–325.
86. Molla A, Paul AV, Wimmer E. Cell-free, de novo synthesis of poliovirus. *Science* 1991; 254(5038):1647–1651.
87. Molla A, Paul AV, Wimmer E. In vitro synthesis of poliovirus. *Dev Biol Stand* 1993; 78:39–53.
88. Nassal M. Hepatitis B virus morphogenesis. *Curr Top Microbiol Immunol* 1996; 214:297–337.
89. Wills JW, Craven RC. Form, function, and use of retroviral gag proteins. *AIDS* 1991; 5(6):639–654.
90. Freed EO. HIV-1 gag proteins: diverse functions in the virus life cycle. *Virology* 1998; 251(1):1–15.
91. Singh AR, Hill RL, Lingappa JR. Effect of mutations in Gag on assembly of immature human immunodeficiency virus type 1 capsids in a cell-free system. *Virology* 2001; 279(1):257–270.
92. Zimmerman C, Klein KC, Kiser PK, et al. Identification of a host protein essential for assembly of immature HIV-1 capsids. *Nature* 2002; 415(6867):88–92.
93. Tritel M, Resh MD. Kinetic analysis of human immunodeficiency virus type 1 assembly reveals the presence of sequential intermediates. *J Virol* 2000; 74(13):5845–5855.
94. Tritel M, Resh MD. The late stage of human immunodeficiency virus type 1 assembly is an energy-dependent process. *J Virol* 2001; 75(12):5473–5481.
95. Weldon RA Jr, Parker WB, Sakalian M, Hunter E. Type D retrovirus capsid assembly and release are active events requiring ATP. *J Virol* 1998; 72(4):3098–3106.
96. Mastrianni JA, Roos RP. The prion diseases. *Semin Neurol* 2000; 20:337–352.
97. Prusiner SB. Prions. In: T. Frängsmyr, editor. (Les Prix Nobel Lecture), in *Les Prix Nobel*. Stockholm, Sweden: Almqvist & Wiksell International; 1998. pp. 268–323.
98. Kretschmar HA, Giese A, Brown DR, et al. Cell death in prion disease. *J Neural Transm Suppl* 1997; 50:191–210.
99. Bounhar Y, Zhang Y, Goodyer CG, LeBlanc A. Prion protein protects human neurons against Bax-mediated apoptosis. *J Biol Chem* 2001; 276(42):39145–39149.
100. Holada K, Simak J, Risitano AM, et al. Activated platelets of patients with paroxysmal nocturnal hemoglobinuria express cellular prion protein. *Blood* 2002; 100(1):341–343.
101. Mishra RS, Gu Y, Bose S, et al. Cell surface accumulation of a truncated transmembrane prion protein in Gerstmann-Strausler-Scheinker disease P102L. *J Biol Chem* 2002; 277(27):24554–24561.
102. Hegde RS, Lingappa VR. Regulation of protein biogenesis at the endoplasmic reticulum membrane. *Trends Cell Biol* 1999; 9:132–137.
103. Fons RD, Bogert BA, Hegde RS. Substrate-specific function of the translocon-associated protein complex during translocation across the ER membrane. *J Cell Biol* 2003; 160(4):529–539.
104. Kim SJ, Mitra D, Salerno JR, Hegde RS. Signal sequences control gating of the protein translocation channel in a substrate-specific manner. *Dev Cell* 2002; 2:207–217.
105. Kim SJ, Rahbar R, Hegde RS. Combinatorial control of prion protein biogenesis by the signal sequence and transmembrane domain. *J Biol Chem* 2001; 276:26132–26140.
106. Rutkowski DT, Lingappa VR, Hegde RS. Substrate-specific regulation of the ribosome-translocon junction by N-terminal signal sequences. *Proc Natl Acad Sci U S A* 2001; 98(14):7823–7828.
107. Skach WR, Calayag MC, Lingappa VR. Evidence for an alternate model of human P-glycoprotein structure and biogenesis. *J Biol Chem* 1993; 268(10):6903–6908.
108. Lipschutz J, Lingappa V, Mostov K. The exocyst affects protein synthesis by acting on the translocation machinery of the endoplasmic reticulum. *J Biol Chem* 2003; 278:20954–20960.