A Synopsis of Bioconformatics

1/15/02
Vishwanath R. Lingappa, M.D., Ph.D.
Professor of Physiology and Medicine
University of California, San Francisco
Email: **vrl@itsa.ucsf.edu**

List of topics covered

Summary ……..……Page 1

The cell biology of gene expression……..…….…… Page 2

Importance of faithful trafficking to intracellular compartments …. Page 2

The problem of protein folding……Page 3

Prion protein and the origins of the bioconformatic hypothesis……Page 5

Bioconformatics: a new view of protein biogenesis and folding……Page 8

Implications for the philosophy and sociology of science… …….Page 11

References…… …..Page 11

Brief figure legends ……Page 13

Summary

Studies on protein biogenesis suggest a new view that greatly expands the information content of the genome – and reveals a new degree of uncertainty in the conventional proteome-based understanding of gene expression. These findings lead to the bioconformatic hypothesis, which proposes that many proteins have multiple folded states, each with different functions. Here I provide an overview of these findings with emphasis on the thought process that leads to these conclusions. Then I indicate briefly some epistemological implications of these findings that pose a challenge for the future.

The cell biology of gene expression

Gene expression is the set of events by which the information in the genome is decoded in the form of **proteins** (1). As the structures that carry out the work of a cell, proteins provide the molecular basis for what an organism does. But just making a protein isn't enough for gene expression to mean anything: the protein has not only to be made, but also correctly **folded** (2), **modified** (3), and **localized** (4) to where it's work will do some good. The process of faithful protein **biogenesis** includes all of these processes and occurs via intricate, highly regulated, and poorly understood mechanisms, carried out by complex subcellular machines, with myriad variations on the fundamental themes (5). Thus, much like the conversion of blueprints for a city into an actual thriving metropolis, the process of gene expression culminating in proper protein biogenesis and localization is incredibly complex and involves numerous steps***, many of which relate to the final goal in a far from obvious manner*** (see Fig. 1). To extend the analogy, squabbling between the mayor and the head of the sanitation workers union, resulting in a garbage strike, could paralyze city operations in a way hard to anticipate in the abstract.

Importance of faithful trafficking to intracellular compartments

Different forms of life can be classified according to an evolutionary "tree" that illustrates their phylogenetic relationship (6). Over the past 3 billion years, this process has resulted in a series of leaps in complexity of organization of life, which is reflected in key specializations found in different phyla. (see Fig. 2). Each of these developments must have demanded a correspondingly great increase in the information expressed from the genome.

Perhaps most fundamentally, procaryotes (such as bacteria) are distinguished from eukaryotes (including both plants and animals, single-celled yeasts to ourselves), by the presence in the latter of a multiplicity of intracellular **membrane-delimited compartments**, including the nucleus, mitochondria, endoplasmic reticulum (ER), Golgi apparatus, and lysosome (see Fig. 3). These membrane-delimited compartments make possible additional layers of complexity in organization of both gene expression and protein function. However, because most proteins are made in the cytoplasm, the benefits of intracellular compartmentalization can be maintained only if cells have also evolved a means of faithful **intracellular trafficking**, by which newly synthesized proteins end up in their correct final location and orientation. Mislocalization to the wrong compartment is as bad as not expressing a protein at all – or worse. By analogy, an air traffic controller can't do his or her job properly if they are sent to the cafeteria. In fact, if enough people are erroneously sent there, it will get too crowded with people just "hanging out" for the people who want to buy food to get in!

For secretory proteins, proper localization involves complex machines to carry out various tasks. Thus there is a machine to make the protein (the **ribosome**), one to target the nascent chain to the ER (**signal recognition particle**), another to allow the nascent chain to cross the ER membrane bilayer (the **translocon**, see Fig. 4. A discrete length of the nascent chain, termed the signal sequence, activates the targeting and translocation machines. During and immediately after the process of synthesis, targeting and translocation, proteins fold. Also during synthesis, many of them are modified by addition of sugars and other structures, and some associate with one another to form larger, multi-protein complexes. From there, yet other machines take over to move the protein to its correct final residence. Because the events of targeting and initial localization take place while the protein is still being made, investigators face an interesting problem in their study: the growing polypeptide chain is ever-changing – a veritable "moving target". The biogenesis of proteins embedded in membranes, and proteins targeted to other compartments of the cell involve variations on these same themes of signal sequences and receptors for trafficking, as exemplified in the biogenesis of secretory proteins described above.

Most workers have emphasized the "constitutive" aspects of this process (Fig 4A), e.g. proteins that don't have signal sequences don't translocate to the ER lumen; those that do are secretory proteins unless a stop transfer sequence anchors them to the membrane, etc. However, over the years, we and others have found evidence suggesting that translocation across the ER membrane may be a far more regulated process than generally recognized (2, see Fig. 4B). For example, **translocational pausing** is a remarkable phenomenon in which literally hundreds of amino acids of a nascent chain are transiently exposed to the cytoplasm prior to translocation into the ER lumen. This appears to be a consequence of regulation of the ribosome-membrane junction (r-mj), during the translocation of specialized secretory proteins (5,7,8,9, see Fig. 4B).

The problem of protein folding

But how do these events involving proteins relate to the larger process of gene expression? There are two notable relationships. First, the precise sequence that make up these polymers of amino acid residues is encoded in the precise sequence of the mRNA transcribed from segments of DNA within individual genes. Second, each individual protein folds in a distinctive way determined by thermodynamics operating on the set of interactions within and between proteins, for example, between "side groups" that distinguish one amino acid from another. The number of ***possible*** folded states for a protein is astronomical – about 10 to the 50^{th} power for a typical sized protein – but most of these are thermodynamically unstable, i.e. would require input of energy to maintain, and therefore are extremely unlikely to be long-lived end-point folded states. The conventional assumption is that only ***one*** of these end-point folded states

represents “correct” folding (i.e. a folding state in which each aspect of the protein occupies a local minimum in the energy landscape, from which it would be functional), and is achieved by most newly synthesized proteins, upon completion of biogenesis, most of the time. While a lot is known about the first process (i.e. the decoding of information in the DNA sequence into amino acids of the protein sequence), far less is known about the process of how a newly synthesized protein “chooses” its final folded state. Even less is known about the possible relationship between the genome and protein folding states.

The problem for understanding protein folding is that, like a golf ball meandering slowly towards the cup, there are many infinitesimally small “bumps and gullies” that can influence it’s path (e.g. for the golf ball, as a function of variables such as speed, direction, wind temperature, etc.). An early notion from the 1960’s, based on test tube experiments, was that protein folding was “spontaneous” i.e. dictated solely by the thermodynamics of protein-protein interactions (10). In the 1980’s it was recognized that protein folding occurs faithfully in living cells in an environment orders of magnitude more crowded, and at rates far faster than in the test tube, hardly the expectations for a spontaneous process. Families of proteins collectively known as “**molecular chaperones**” were discovered to play roles in folding of newly synthesized proteins (11). In general, they are believed to work by preventing a protein in an unfolded state from engaging in inappropriate or undesired interactions, thereby allowing the opportunity for the correct interactions to take place, resulting in proper folding.

To deal with the distinct problem of how a protein reaches its final folded state, given such a large number of alternates to choose from, the notion of “folding funnels” was advanced (12). The idea here is that, as a protein starts to fold, it precludes some pathways of further folding and favors others, much like a marble in a pinball machine. Thus, as a protein starts to fold, early events in folding direct the chain towards some later events to the exclusion of others, allowing “correct” folding to proceed without sampling all possible folded states.

The view that any given protein has only one “proper” final folded state, with all others representing “misfolding”, is a crucial feature of current thinking. This view is reinforced by the fact that such alternately folded states are, typically, rapidly degraded by the cell. Indeed, a stringent and quite unforgiving quality control machinery has been discovered and shown to prevent alternately folded copies of proteins from leaving the ER, directing them instead to a pathway of degradation (13). On the other hand, it is possible that this assumption (i.e. that there is one “proper” folded state for each protein and that the others represent “misfolding”) could be wrong, as will be discussed below.

Before considering this possibility, however, let us consider its implications. Literally thousands of scientists work on the assumption that measurement of the level of a given protein or its mRNA tells you what you need to know about expression of that form of the gene, both to understand its function in a healthy

cell and its dysfunction in disease. , Hundreds of millions of dollars of research money each year are staked on this assumption – which is not true to the extent that the bioconformatic hypothesis is valid. The current paradigm could be stated as : “Each protein has one final folded form and performs one function in the cell”. The alternative, that a protein could achieve more than one conformation that is allowed by quality control machinery, would mean that simply measuring the total level of a protein could be completely misleading. Cells with identical levels of a protein could have those molecules folded in different ways -- ***in which case they might as well be different proteins and from your measurements, you wouldn’t know it!*** Thus, a lot of conclusions on normal function and disease would need to be revisited. What a mess!

Furthermore, if one protein has multiple conformations, with the different conformations having different functions, a means of distinguishing the specific conformational isoforms is needed if they are to be studied. Currently, ***the tools to do this do not exist for any protein, and to generate them would be time consuming, expensive and need to be done separately for each of the tens of thousands of gene products expressed!*** Thus, there is a huge and entirely understandable reluctance to consider a possibility that would wreak havoc on what we think we know and how we currently go about getting new knowledge on gene expression. Until recently, such skepticism was appropriate, as a precedent did not exist for this alternative. Recently however, that has started to change.

Prion protein and the origins of the bioconformatic hypothesis

The prion protein (PrP) is a gene product expressed in normal brain and a number of other tissues, most notably in cells of the immune system. It is a relatively “new” gene in the course of evolution, being present in birds and mammals, but not insects. It came to prominence because of its involvement in an unusual set of neurodegenerative diseases termed prion diseases (14). The most notorious of these disorders are transmissible (i.e. “infectious”) diseases such as scrapie and mad cow disease in animals, and kuru and new variant Creutzfeld-Jacob disease in humans. A discussion of the current views of the pathogenesis of these disorders is outside the scope of this discussion. Suffice it to say that a large body of evidence leads to the following conclusions:

i) normal PrP is non-infectious;

ii) “infectious” PrP (termed PrP^{Sc}) appears to be folded in a different manner from normal PrP;

iii) somehow, the abnormally folded PrP^{Sc} appears capable of inducing normal PrP to fold into the abnormal form (which is the basis for its “infectious” character);

iv) some prion diseases are genetic diseases, occurring spontaneously over time, due to mutations in PrP rather than altered folding of PrP whose sequence is normal;

v) some of these genetic prion diseases are non-infectious, and have no accumulation of PrP^{Sc} despite producing classical prion disease histopathology in brain (spongiform change and astrogliosis) and clinical signs in animals (including humans).

vi) PrP^{Sc} while transmissible (i.e. "infectious") is not intrinsically pathological, because animals do not get prion disease if their endogenous PrP gene is missing. Presumably, there are backup genes that replace the normal function of PrP since these animals are grossly normal. Furthermore, when given transplants of normal brain that ***are*** infected with prions, the animals with no endogenous PrP gene develop high titers of infectious PrP^{Sc} in their brains, ***but have no pathological findings or clinical signs of disease*** (15).

Because we were interested in the pathogenesis of prion diseases, we decided to try and learn more about normal PrP itself, and therefore studied its biogenesis. Those studies, starting some 15 years ago, revealed a number of remarkable findings (16-22, and unpublished studies). In particular:

i) Whereas all other proteins studied up to that time were classifiable uniquely as one topological type (e.g. as either cytosolic proteins, secretory proteins, integral membrane proteins, or proteins destined to other organelles), the identical full length PrP, made from a homogeneous population of nascent chains, encoded in a single mRNA, was found to be made as ***either*** a secretory protein or in one of two different integral membrane forms of different transmembrane orientation, ***depending on factors present in the cytosol, ER membrane, and ER lumen***(see Fig. 5).

ii) Mutations of PrP that cause spontaneous neurodegenerative disease in mice and humans skewed the distribution of topological forms towards one particular transmembrane form, termed ^{Ctm}PrP.

iii) The different ***topological*** forms are also different in ***conformation*** because mild protease digestion degraded some forms but not others in detergent solutions that abolish all topological distinction (see Fig 5)***.***

iv) Furthermore, the ^{Ctm}PrP form while normally rapidly degraded in the ER, was allowed by quality control machinery to exit the ER under certain circumstances.

v) The relationship between the different ***topological*** forms of PrP and the infectious PrP^{Sc} was clarified by studies demonstrating an inverse

relationship ship between the propensity of a transgenic mouse to make ^{Ctm}PrP (e.g. severity of the mutation x level of expression) and the level of PrP^{Sc} accumulated upon onset of clinical disease. Thus, PrP^{Sc} appears to signal production of ^{Ctm}PrP.

vi) In the course of scrapie infection, ^{Ctm}PrP is induced just prior to onset of clinical disease.

vii) Normal PrP could be made in the ^{Ctm}PrP conformation in response to changes in intracellular signaling, during development.

viii) The normal function of PrP varies according to its topological type: ^{Ctm}PrP appears to be a trigger of apoptosis (programmed cell death). Studies from other workers can be interpreted to suggest that secretory PrP is an anti-apoptotic factor that protects cells from apoptosis (23, 24).

ix) Swapping the PrP signal sequence for that of other simple secretory proteins resulted in a dramatic change in distribution among the PrP forms, even where the authentic PrP chain was the wild-type protein.

x) The signal sequences that favor one versus another topological form of PrP could be classified on the basis of the nature of the ribosome-membrane junction (rm-j) established for the nascent chain early in PrP biogenesis. Those establishing an “open” rm-j strongly favored ^{Ctm}PrP, those establishing a “closed” rm-j strongly favored secretory PrP, etc.

xi) Swapping of signal sequences can alter folding of other proteins, unrelated to PrP. A protein whose native signal sequence formed a “closed” rm-j could have its folding altered (as shown by a reporter glycosylation site, chemical crosslinking, and protease digestion) by simply swapping its signal sequence to that of an open rm-j.

xii) Point mutations within the signal sequence could reproduce the changes observed by gross swapping of signal sequences.

xiii) Open and closed rm-j represent only one of several dimensions of regulation as some signal sequences that establish closed junctions were also capable of altering folding.

The important take home messages from these studies have been two fold. First, a pathophysiological model for prion disease has been established with three stages that accounts for available data (see Fig. 6).

Second, perhaps the alternate conformational forms of PrP (for example, ^{Ctm}PrP) were easily detected in the case of PrP ***because*** they manifested as topologically

different forms. For most other proteins, such **conformational heterogeneity** may well exist (as demonstrated, for example with the glycosylation reporter indicated in point x above) , but is simply not be easily detected because it is not associated with differences in topology.

Bioconformatics: an alternative view of protein biogenesis and folding

Bioconformatics refers to a new view of protein biogenesis and folding. The key tenets of the bioconformatic hypothesis are that:

i) For some, and perhaps many, proteins the outcome of synthesis is heterogeneous, generating multiple forms of identical amino acid sequence, termed **conformers**, that may often be extremely difficult to detect or distinguish, given currently available tools. (Fig. 7A).

ii) The different conformers are generated by the effects of transient differences in the environment seen by the growing nascent chain (e.g. including protein-protein interactions, redox differences and others).

iii) Some of these alternate forms are likely not “mis-folded” in that they have distinct functions and are allowed by “quality control” machinery to leave the ER at particular times.

iv) Functional heterogeneity may physically manifest in many different ways, for example, as alternate topological forms, final folded states, stable or transient associations with other proteins, or due to trafficking to one compartment versus another.

v) The cell has mechanisms, not currently easily recognized, by which synthesis of one form and its corresponding function can be favored at one time, while a different form and its corresponding function are favored at another time. This may be due to simply altering the stochastic probabilities of particular protein-protein interactions or folding funnel choice, or through more directed mechanisms (see Fig. 7B).

vi) The machinery involved in translocation to the ER lumen is an example of how proteins designed for one set of functions (targeting and translocation) have been elaborated to take on other functions (regulation of protein folding). Quality control and other pathways may serve to further modify the mix of forms evident at steady state, by suppression of particular forms at certain times through degradation. These mechanisms may be controlled by signal transduction pathways (ways that one part of the cell can talk to another), directly or indirectly.

vii) Dysregulation of these normally tightly regulated pathways of biogenesis, resulting in the wrong form being in the wrong place, or made at the wrong time, can result in disease.

viii) For secretory and integral membrane proteins, the signal sequence is one crucial, regulated, determinant of bioconformatic heterogeneity. At least in part, its action occurs via the nature of the r-mj it establishes (see Fig.4B).

Given these tenets, let us summarize their implications for the natural history of protein evolution. Early in evolution primitive protein biogenesis may have resulted in a certain heterogeneity of folded forms, perhaps simply because of stochastic variation at key "branch points" of the folding funnel. Initially, this heterogeneity was a problem. Hence the evolution of quality control mechanisms which converted heterogeneity into simply a waste. Later, however, regulation of quality control, and elaboration of means to switch from a program favoring one versus another folding funnel (e.g. open vs closed rm-j), allowed what had been a waste to become a huge reservoir of new information content for gene expression. Selection and expansion of a particular alternate conformation occurred in response to the need for some different function that the alternate conformation could carry out. Finally, the regulation was fine-tuned with respect to the timing of expression of one form versus another and the tissue specificity of expression.

The next steps for advancement of this perspective are:

i) define the extent of bioconformatics (does it occur during biogenesis of only a few proteins, or many, or most?).

ii) develop the tools with which to study it rapidly and easily

iii) uncover the regulatory mechanisms by which it is controlled and with which it interfaces to be entrained by other dimensions of cellular and organismal homeostasis.

There are three primary reasons why bioconformatics has been overlooked in studies of other proteins in which, unlike PrP, conformational heterogeneity is not accompanied by topological differences (see Fig. 8).

First, because cells have powerful mechanisms, including ER degradation, by which heterogeneity can be masked. These mechanisms can be turned "on" and "off" over a time scale that challenges the investigator's abilities at detection. It is difficult to detect heterogeneity in living cells below the 10% level. Thus, if a major conformation of a protein comprises 40% of chains, and many minor forms comprise the rest, with no individual form at greater than say, 5% of chains, a complex traffic of minor forms could easily be lost as "background" to standard

analysis. For example, ^{Ctm}PrP, now believed to be central to disease pathophysiology, is normally almost undetectable, and yet, becomes readily apparent in dysregulation syndromes. It is notable that despite years of intensive investigation of prion diseases in brain, ^{Ctm}PrP was first detected in cell-free translation systems (15). Only much later, with the development of new tools, was it finally detected in diseased brain (18). Indeed, from unpublished studies ^{Ctm}PrP now appears to have a developmentally regulated physiological role. Generalization of these observations would suggest that rapidly degraded proteins could represent alternately folded forms with specialized functions rather than simply misfolded "waste".

Recent studies with other proteins are provocative in this regard. The extent of presumptive misfolding and degradation in some cases , strains credulity. For example, as much as 90% of newly synthesized cystic fibrosis transmembrane regulator (CFTR), the gene product whose defect causes cystic fibrosis, is rapidly degraded upon synthesis -- even in the case of the wild-type, non-mutant form found in people without the disease! The conventional view of this extraordinary phenomenon would be that CFTR is a difficult protein to synthesize, and the cell makes mistakes most of the time (27).

The bioconformatic hypothesis proposes an alternative explanation for CFTR degradation. According to this view, it is the consequence of many alternate conformational forms, most of which the cell does not need at any given time. Any one of these forms is a small percentage of the total and therefore not easily detected and followed, in contrast to the major "constitutive" form which is responsible for the conventional function. Consider the analogy of a Martian watching a football game: they might conclude it was nothing more than a ritual in which two groups of large individuals in red and blue uniforms meet in the middle of a field and beat each other up. However that interpretation would completely miss nuances such as whether the offense or defense is on the field, whether the team is within field goal range, whether a team is "running out the clock" or following a scripted "two minute drill", or changing its strategy inside the "red zone", etc. These are the features that add richness and meaning to the process. Such subtle yet profound details are what makes the bioconformatic hypothesis difficult to appreciate for most proteins.

Second, because many of the currently available tools for high resolution study of protein conformation (e.g. crystallization and x-ray diffraction) require highly purified proteins, often of recombinant origin, where the native heterogeneity in their physiological tissues, in real time, is not maintained. Likewise, other techniques (e.g. nuclear magnetic resonance) might easily lose the signal from a minor conformational form in the noise of background protein denaturation or masked by the signal of the major conformation present.

Third, because there are a paucity of tools developed for rapid and easy examination of conformational heterogeneity as it occurs ***in vivo***. For example,

conformation-specific monoclonal antibodies (mAbs) have been used to study only a few extraordinary structural transitions (24), and are not generally sought for most proteins, in part because of their relatively low affinity, and also because their initial identification depends upon the ability to distinguish those that recognize one conformation from another, which is generally not possible unless the conformations were already recognized and purified in the first place!

Implications of the bioconformatic hypothesis for philosophy and sociology of science

i) To the extent that the bioconformatic hypothesis proves to be an accurate reflection of reality, it undermines current thinking – and much of the promise of existing genomics and proteomics – by introducing a new dimension of uncertainty in our ability to understand gene expression.

ii) The findings giving rise to bioconformatics reveal some of the limitations of induction and reductionist science. It's the "only game in town" and yet the propensity for misdirection of the experimenter is profound. A jaundiced view suggests that our methodology amounts to "looking under the lamp post for the wallet we lost in the dark alley, because at least we can see there". Hopefully, not all of our efforts are as nonsensical, but historically, they may amount to as much.

iii) The interconnection of technology, ideology and other dimensions of society on the way "facts" are stated and interpreted raises important questions about their meaning.

iv) The power of homeostasis and evolution, with the nervous and immune systems (28) as notable showcases for its achievements, is quite breathtaking. A billion years is a long time. Selection can achieve a lot – and hide a lot – in that time frame (30).

v) Finally, the "conflicts of interest" manifest when an investigator has a high personal stake in the success or failure of a paradigm, suggest that "hard science" is far less objective than is generally believed.

References

1. Lewin, B. Genes VII, Oxford University Press 7th ed. 1999.

2. Richards, F.M. (1991). The protein folding problem. Sci. Am. 264:54-57.

3. Yan, S.C.B. et al. (1989). Post-translational modifications of proteins: some problems left to solve. Trends Biochem Sci. 14:264-8..

4. Blobel, G. (1980). Intracellular protein topogenesis. Proc. Natl. Acad. Sci. USA 77:1496-1500.

5. Hegde, R. and Lingappa, V.R. (1999) Regulation of protein biogenesis at the endoplasmic reticulum. (perspectives) Trends Cell Biol. 9:132-137.

6. Futuyama, D. Evolutionary Biology, Sinauer Assoc. 3rd ed. 1997

7. Chuck, S.L. and Lingappa, V. R. (1992) Pause Transfer: A topogenic sequence in apolipoprotein B mediates stopping and restarting of translocation. Cell 68, 9-21.

8. Hegde, R. S., and Lingappa, V. R. (1996). Sequence-specific alteration of the ribosome-membrane junction exposes nascent secretory proteins to the cytosol. Cell. 85: 217-228.

9. Hegde, R.S. et al. (1998).· TRAM regulates the Exposure of Nascent Secretory Proteins to the Cytosol During Translocation into the Endoplasmic Reticulum. Cell 92:621-31.

10. Anfinsen, C. B. (1973). Principles that govern the folding of protein chains. Science. 181:223-230.

11. Ellis, R.J., and F.U. Hartl (1996). Protein folding in the cell: competing models of chaperonin function. Faseb Journal 10:20-6.

12. Dill, K.A. and H.S. Chan (1997). From Levinthal to pathways to funnels. Nat. Struc.Biol 4:10.

13. Ellgaard, L., et al. (1999). Setting the standards: quality control in the secretory pathway. Science. 286: 1882-1888.

14. Aguzzi A, et al (2001). Prions: health scare and biological challenge. Nat Rev Mol Cell Biol. 2:118-26. Review.

15. Brandner S, et al. (1996). Normal host prion protein necessary for scrapie-induced neurotoxicity. Nature.379:339-43.

16. Hay, B. et al. (1987). Biogenesis and transmembrane orientation of the scrapie prion protein. Mol. Cell Biol. 7:914-919.

17. Hay, B. et al. (1987). Evidence for a secretory form of the cellular prion protein. Biochemistry 26:8110-8115.

18. Lopez, C.D. et al. (1990) Unusual topogenic sequence directs prion protein biogenesis. Science, 248, 226-229.

19. Hegde, R. S., et al. (1998). A transmembrane form of the prion protein in neurodegenerative disease. Science. 279: 827-834.

20. Hegde, R. S.et al. (1998). Regulation of protein topology by trans-acting factors at the endoplasmic reticulum. Molecular Cell 2:85-9.

21. Hegde, R. S., et al. (1999). Transmissible and genetic prion diseases share a common pathway of neurodegeneration. Nature. 402: 822-826.

22. Rutkowski, D. T., et. al.. (2001). Substrate-specific regulation of the ribosome-translocon junction by N-terminal signal sequences. Proceedings of the National Academy of Sciences of the United States of America. 98: 7823-7828.

23. Kuwahara C. et al. (1999). Prions prevent neuronal cell-line death. Nature. 1999 Jul 15;400(6741):225-6.

24. Bounhar, Y. et al. (2001). Prion protein protects human neurons against Bax-mediated apoptosis.J Biol Chem. 2001 Oct 19;276(42):39145-9.

25. Lingappa, V.R., et al. Conformational control through translocational regulation: a new view of secretory and membrane protein biogenesis. BioEssays submitted.

26. Copeland, CS; et al. (1986). Assembly of influenza hemagglutinin trimers and its in intracellular transport. Journal of Cell Biology 103:1179-91.

27. Ward, C. L., et al. (1995). Degradation of CFTR by the ubiquitin-proteasome pathway. Cell. 83: 121-127.

28. Bonifacino, J.S., et al. (1990). A. Novel post-translational regulation of TCR expression in CD4+CD8+ thymocytes influenced by CD4. Nature 344:247-51.

29. J. Dewey, J. The quest for certainty. (1923.)

30. Lingappa, V.R. and Farey, K. Physiological Medicine. McGraw-Hill 1st ed. (2000).

Brief figure legends:

Figure 1. Summary of the key regulatory steps in gene expression.

Figure 2. A crude evolutionary tree indicating some notable quantum leaps in complexity of organization of life.

Figure 3. Membrane delimited intracellular compartments of the typical eukaryotic cell. 1-6 depict the temporal order of compartments traversed by newly synthesized secretory proteins. 1= cytoplasm; 2=cytoplasmic face of the ER; 3= ER lumen; 4= Golgi apparatus; 5= vesicle trafficking from trans Golgi network to cell surface; 6= release of vesicle contents into the outside world upon fusion of the transport vesicle or granule to the plasma membrane.

Figure 4 A. Conventional view of protein translocation across the ER. Note the constitutive (non-regulated with respect to contingency) nature of the process, the closed ribosome-membrane junction, and a single final functional folded state for the protein. B. Bioconformatic view of protein translocation across the ER membrane. Note the regulation of the ribosome-membrane junction, allowing distinct differences in protein-protein interactions in the cytosol, membrane and lumen, and different final folded states of the protein.

Figure 5. PrP biogenesis: A. Topological assay. Harsh proteolysis (PK) with intact membrane vesicles removes domains outside the bilayer only (panel II). Membrane solubilization with non-denaturing detergent abolishes topological differences, so all forms are degraded. B. Conformational assay. Mild protease digestion with membranes intact, gives the same result as in the harsh protease topological assay. Solubilization of the membrane with non-denaturing detergent abolishes topological differences (panel II), but not differences in conformation. Thus, mild protease under these conditions demonstrates intrinsic differences in folded state by protection of a signature fragment of ^{Ctm}PrP (panel III).

Figure 6. Pathophysiology of prion diseases. Summary of findings from ref 19.TRAF refers to translocation accessory factor that directs PrP folding funnel selection during translocation.

Figure 7. A. Analogy to protein folding through the translocon. A standard demonstration of Gaussian distribution applied to protein folding in the abstract. B. Regulation of protein folding by altering chute characteristics, relationship to environmental factors, or pathways out of the translocon.

Figure 8. How cell biological complexity confounds analysis of gene expression. Multiple conformations without tools to detect or distinguish them easily in real time; different fates, half-lives, associations, etc. for different conformations.