

COMMENTARY

Open Access

Ways and means of coping with uncertainties of the relationship of the genetic blue print to protein structure and function in the cell

Ernst JM Helmreich

Abstract

As one of the disciplines of systems biology, proteomics is central to enabling the elucidation of protein function within the cell; furthermore, the question of how to deduce protein structure and function from the genetic readout has gained new significance. This problem is of particular relevance for proteins engaged in cell signalling. In dealing with this question, I shall critically comment on the reliability and predictability of transmission and translation of the genetic blue print into the phenotype, the protein. Based on this information, I will then evaluate the intentions and goals of today's proteomics and gene-networking and appraise their chances of success. Some of the themes commented on in this publication are explored in greater detail with particular emphasis on the historical roots of concepts and techniques in my forthcoming book, published in German: *Von Molekülen zu Zellen. 100 Jahre experimentelle Biologie. Betrachtungen eines Biochemikers.*

Introduction

The formation of functional proteins in the cell depends on the reliable transcription and translation of the genetic information. Today, genetics aims at describing the functional correlate of the genome in the context of the living cell, in other words, it tries to build a bridge from genotype to phenotype. In classical genetics however, a gene was an abstract concept, a materially undefined entity that ferried a characteristic property from parent to offspring. Later, genes became chemically, structurally and functionally defined RNA or DNA polymers, [1], and chromosomes, where the genes are located in the cell, were shown to be tightly packed protein-ribonucleic acid complexes, containing a linear array of genes [2]. Finally, the syntax, the genetic message, was deciphered and shown to consist of words consisting of three letter syllables. Today, the question, what is a gene and how does it function [3] must be raised anew, a question that in the not too distant past would have seemed silly, because every schoolchild believed to know what a gene is. The root cause for this confidence came from the organisation of the genome in prokaryotes, where the gene and the product of its

transcription, the mRNA, and the product of its translation, the resulting protein's amino acid sequence, were seen to be co-linear. This assumption was experimentally verified [4]. It also led to euphoric expectations in human genetics. After the human genome had been deciphered, the hope was to replace a diseased gene with a functional one, bringing, at the very least, the expectation of curing monogenetically-caused diseases. Co-linearity in prokaryotes is possible, because transcription and translation are not separated by cell compartmentalisation, as is the case for eukaryotes. In prokaryotes, the nascent genetic readout, the mRNA, binds directly to ribosomes where it is translated into protein. However, due to the spatial separation of genes and ribosomes, direct coupling of gene transcription, translation and protein biosynthesis is no longer possible in Eukarya. Whereas gene expression is confined to the nucleus, translation is an exclusively cytoplasmic event. Although separation of both processes makes gene expression much more complicated than in bacteria, it enabled the evolution of split genes and pre-messenger RNA splicing in Eukarya, which in turn led to a massive diversification of genetic information [5,6]. This, however, is only one of the key aspects that have profoundly changed our views. Most recently, advances in genetics have led to the comprehensive knowledge of the

Correspondence: e.helmreich@mail.uni-wuerzburg.de
Weindl Lenzstrasse 19, 83727 Schliersee, Germany

genome's output: the full extent of transcription and the regulatory role of the products of the non-coding DNA, the micro RNAs [7,8].

After the gene was shown to be composed of DNA, [9,10], its structure solved by Watson and Crick, [11,12], and its replication verified by Meselson and Stahl, [13], the next question that had to be answered was, how is the information in DNA transcribed and translated into the language of proteins? The direction of the search was set by Francis Crick's brilliant idea of a *unidirectional path, DNA→RNA→Protein*. Crick had suggested that '*...the main function of the genetic material, DNA, in a cell is to control, but not necessarily to direct the synthesis of proteins...*' thus hinting at a role of molecules, other than DNA, in protein synthesis. Moreover, on the same occasion [14], he pronounced his central dogma: '*... once information has passed from the nucleic acids into protein, it can not get out again, only transfer from nucleic acid to nucleic acid or from nucleic acid to protein is possible...*' (Citation from: Horace Freeland Judson's book, *The eight day of creation. Makers of the revolution in Biology*" [15]). Crick's dogma was a modern version of August Weismann's [16] in which the '*Keimplasma*', containing the genes, was described as being separate from the cell soma, which contained everything else. A unidirectional transfer of information from DNA to protein would also have forestalled any kind of environmental influence on the gene, as Lamarck had envisioned, because proteins are the agents that communicate with the environment. While Crick's idea laid the foundations for the identification and characterization of the instruments of gene transcription and translation, tRNAs and mRNAs, and while he was right on transfer only from nucleic acid to nucleic acid, the idea of an unidirectional transfer of information, like any other dogma in biochemistry, was short lived. After Howard Temin [17] and David Baltimore [18] independently discovered reverse transcription from RNA to DNA, Crick's paradigm had to be abandoned.

A focal point of my discussion of the route from gene to protein is the plasticity of structure and function of some proteins, notably of antibodies and receptors [19] which allows one and the same protein to assume different structures and functions in the same cellular environment. I shall give examples of pliable and adaptable proteins and shall cite possible causes of ligand-induced malleability of proteins. Finally, I shall comment briefly on the revolutionary progress made in recent years in the design and improvement of new and old techniques respectively *in vivo* observations, advances that give me confidence that the ambitious goals of the *in vivo* 21st century biochemistry might be, at least in part, realized.

Comments

The genetics of today

The existence of split genes [20] was demonstrated by Richard J. Roberts und Philip A. Sharp [21]: Within exons, genomic regions containing coding sequences which are transcribed and translated into defined protein products, are embedded what were initially thought to be non-coding regions or introns. The splicing process leads to removal of the introns, and the remaining exons are fused and transcribed. Such a spliced gene may have contained many introns ranging from 40 to 100,000 nucleotides in length. Transcription of spliced genes results in a split RNA transcript. In the 1980s Thomas Cech and his colleagues observed that the processing and splicing of a 400 nucleotide long intron of the ribosomal RNA of *Tetrahymena thermophila* did not require any added proteins [22], leading to the conclusion that RNA itself can carry out splicing of its own RNA. The RNA of *Tetrahymena* is a ribozyme [23].

Splicing of ribosomal RNA, or rRNA, and splicing of pre-mRNA, are two different processes [1]. Removal of introns of pre-mRNAs and ligation of the remaining exons is executed by a nuclear machinery, the spliceosome. The spliceosome assembles itself anew on each intron of a mRNA that requires splicing. Consisting of four different, small nuclear ribonucleoproteins (snRNPs), and 100-150 additional proteins, the spliceosome constitutes one of the most complex cellular macromolecular machines identified to date.

Most mRNAs encoding more than one different protein can be spliced by the spliceosome in a manner, also known as alternative splicing, that ensures that different proteins can be generated from a single pre-mRNA. In this way, splicing contributes to the efficient and exhaustive use of the genetic repertoire and has thus been instrumental in facilitating the evolutionary process within organisms.

Alternative splicing also enables transcription to start at a gene and then continue transcribing another gene with another function or encoding a different protein, located remotely or even on another chromosome [24]. 4-5% of DNA, assigned to identified proteins is transcribed in this way.

Mobile or 'jumping' genes, were discovered in the 1950s by Barbara McClintock in corn, have also gained new relevance, one of the reasons being that they can come under different promoter control resulting in differential expression; a process shown to be responsible for cancerous cell growth [25]. One of the reasons why genes must be mobile is the compartmentalisation of gene transcription and the discrete location of genomic transcription factories, although the exact location of the latter remains uncertain. This makes movement of

genes mandatory, because the number of transcription factories in each cell nucleus is limited. Consequently, genes located tens of megabases apart or even on different chromosomes must move to a transcription factory, in order to be transcribed [26].

Compartmentalisation of gene transcription, transcription-mediated gene fusion and genome-wide transcription, [27,28], and the implications of disorder in proteins [29] for regulation and protein interactions, will make it difficult for today's gene networkers to assign a coding region to a gene in a genome map. However, without such an assignment, it is problematic to link a protein to the gene that encodes it.

A further layer of complexity is contributed by mechanisms, some old, some new, that modify genes and chromosomes epigenetically [30], without permanently altering the inheritable sequence of the DNA. Moreover, not only is transcription more complex in eukaryotes and metazoa than in prokaryotes, the same holds for the translation of the message and its regulation. In recent years, we have learned that non-coding DNA [7] encodes small RNAs or micro RNAs [8] which control gene expression at the level of transcription and translation in metazoa and humans.

This short summary of the state of current genetics clearly illustrates that co-linearity of genotype and phenotype, of gene and protein, represents a simplistic interpretation of the situation in eukarya, and is only one side of the coin. The other problem for systems biology and proteomics, is the inherent intricacy of the path from the genetically-determined primary structure of a protein to its functional structure in the cell.

From Gene to Protein

The optimistic expectation, that the genetically determined primary structure of a protein contains all the information required for formation of a unique, three-dimensional, secondary and tertiary, spatial, three dimensional structure was widely shared by my generation of biochemists. This is also what Jacques Monod had expected [31]. He thought that a protein has only '*...one defined structure and one unique communiqué...*' and indeed there are examples of where this expectation holds true.

1. Protein folding, the other half of the genetic code

Pauling's, [32] perception of the ways and means by which higher order structures of a protein are formed from the backbone structure of primary amino acids, guided Christian B. Anfinsen, [33] in his attempts to decipher the other half of the genetic code. He could show that de-natured ribonuclease, with its four -S-S- bonds reduced, spontaneously refolded into its native structure, once the denaturing agent, 8 M urea, was removed by dialysis and the scrambled enzyme exposed

to a small amount of an -S-S- reducing agent, mercaptoethanol. Under these conditions, disulfide interchange eventually formed out of a mixture of 105 incorrectly folded isoforms, a single homogenous product which was indistinguishable from native ribonuclease. (Renaturation is driven by the free energy, ΔG , which is gained when the unfolded polypeptide chain refolds into a stable, native structure. The gain in enthalpy offsets the loss of entropy, when the more ordered, folded structure is formed).

Further supporting evidence came from the successful recombination experiments of Frederic M. Richards, [34], with ribonuclease S, a 20 amino acid fragment, cut off from the N-terminal half of ribonuclease, and the remaining core of the enzyme with its 104 residues, containing all four disulfide bridges of the native, parent molecule. The small and large fragment recombined correctly and the structure and function of the native ribonuclease were completely restored. The experiments of Frederic M. Richards proved that cooperative, non-covalent forces are also responsible for protein-protein recognition and interaction, just as Linus Pauling had foreseen [32].

Another way to solve the problem of protein folding rests on the expectation that it should be theoretically possible to predict the three-dimensional structure of a protein from the chemical and physicochemical properties of the constituent amino acid side chains and their behaviour in aqueous solution. However, although estimates of the forces involved in folding a protein have been made and continuously refined [35], all these efforts have been insufficient to permit reliable prediction of the higher order structures of larger proteins from their amino acid sequences.

Ribonuclease therefore turned out to be more of an exception rather than the rule. In many cases, one and the same protein can assume different conformations with different functions in the same cellular environment and location.

In the decades since Anfinsen's experiments, the instruments that help a cell to fold its proteins have been identified. Anfinsen and his colleagues [33] were already aware that folding may require assistance, because renaturation of ribonuclease *in vitro* took hours, much longer than the time needed for folding of a polypeptide chain of this length *in vivo*, where it is known to occur in about a minute or less. However, when Anfinsen and his colleagues [36] added an extract containing an enzyme which catalyzes disulfide interchange, renaturation occurred *in vitro* as fast as it did *in vivo*. In the meantime, a number of folding partners were discovered, including enzymes which help to form disulfide bonds and enzymes such as peptidyl prolyl isomerases which accelerate the cis-trans isomerizations of peptide

bonds featuring a preceding proline residue [37,38]. The correct placement of disulfide bonds is often rate-limiting in protein folding, notably of proteins which are secreted. Some secreted proteins, for example trypsin, chymotrypsin [39] and insulin [40] are formed as inactive precursors, which are converted to the active form, with the right arrangement of disulfide bonds, by limited proteolysis. Moreover, cells that secrete proteins harbour disulfide isomerases that help to promote disulfide interchange and rearrangement of proteins to be secreted. Proteins which are secreted, or those targeted to the cell surface first enter the luminal space of the endoplasmic reticulum (ER) where they are properly folded and in some cases modified, for example, by glycosylation. Only properly folded and processed proteins can exit from the ER. Peter Walter and his colleagues, in San Francisco have studied the regulatory mechanisms in eukaryotic cells that control folding and secretion of proteins in the ER [41]. They concluded that in the event of misfolding, the unfolded protein response of the cell (UPR) comes into play, which can order the death of a cell with misfolded proteins by apoptosis. Folding malfunctions seem to play a role not only in familial protein folding diseases, but also in diabetes, cancer, viral infections and prion-based diseases [42].

The discovery and characterization of the heat shock proteins, such as HSP 60 and HSP 70, helped to gain new insights into the fate of proteins in the living cell. These and other chaperones accompany proteins in the cell from their formation to their degradation and removal. While without their help, newly folded proteins would form aggregates, this is only one of their many essential functions. By 1989, Walter Neupert's laboratory in Munich had described a role for a heat shock protein in the assembly of proteins in mitochondria [43]. Subsequently, chaperones were shown to help folding polypeptide chains in the course of their synthesis, to protect already folded proteins and to mark those which are incorrectly folded for degradation [44,45]. Chaperones may even participate in the folding process itself, like a folding enzyme [46].

Remarkably, chaperones are not only required for protein folding, they also help in the assembly of macromolecular complexes such as proteasomes, ribosomes and spliceosomes. For example, an assembly chaperone has recently been described that assembles the protein building blocks of the spliceosome and helps to build this complex machine in the cell [47,48]. This kind of chaperones probably assists in the biogenesis of all macromolecular complexes in the cell. (For more information see the review of Hartl, F. U. und Hayer-Hartl, M. in *Science*. (2002) [44], and the publication dealing with *cellular functions of molecular chaperones*, in: *Mechanisms of Protein folding* [49]).

Despite the significant body of knowledge that has accumulated relating to the mechanism of protein folding in the living cell, reliable folding of larger complex proteins in the laboratory remains to be achieved. As we learn more about ligand-induced conformational changes of proteins [50], it is becoming clear that the cause of our failure might be the ease with which some proteins are capable of changing their structure and function. Thus, the genetically-determined primary, backbone structure of a protein can in some cases form different higher order structures. This process is not only induced by binding of ligands, which regulate and determine the functionally-relevant actions and interactions of such proteins in the cell, because there are also intrinsically disordered proteins or parts of proteins.

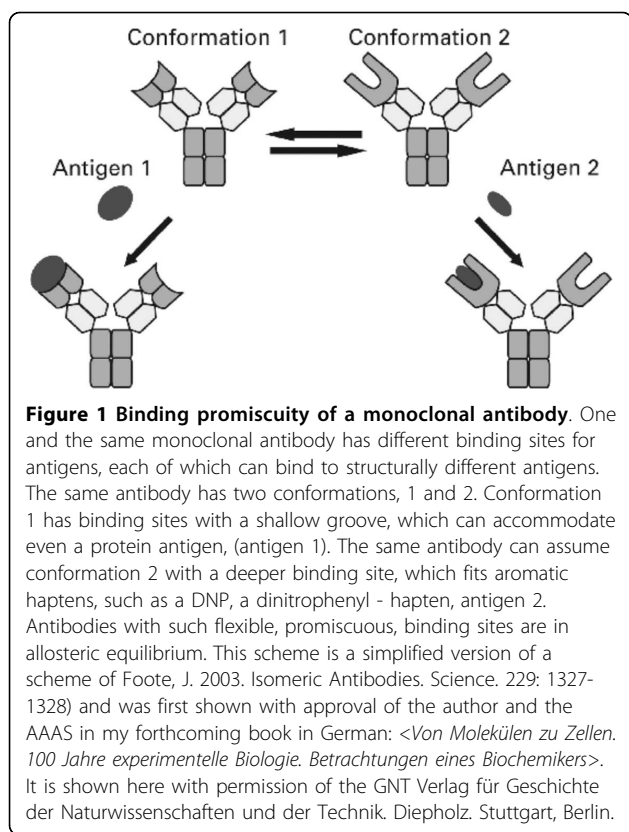
Flexible Proteins

Although Linus Pauling [51] had to abandon his instructive theory of antibody formation, his notion of a flexible secondary and tertiary structure of antibody binding sites in particular, and of proteins in general, has been vindicated.

1. The antibody: an example of a flexible protein

Today, there is convincing evidence that antibody binding sites can accommodate structurally different haptens or antigens [52-54]. For example, a monoclonal antibody raised against the HIV-1p24 protein can bind to several unrelated peptides [55]. James, Roversi and Tawfik [56], from the Weizmann Institute in Rehovot have demonstrated that a monoclonal antibody, specific for a 2,4 dinitrophenyl (DNP) hapten also bound, with a broad range of affinities, to several, unrelated aromatic compounds and even to a protein. Adaptability of the antibody binding sites is also in agreement with different crystal structures of the same antibody, unbound and bound to a hapten. A fragment, Fv, of the variable binding site of this monoclonal antibody did crystallize, in the absence of the ligand, in two different, stable conformations and in complexes with aromatic haptens and with a proteinaceous antigen, it adopted a third and a fourth conformation [57]. Fig:1.

An antibody clearly illustrates the flexibility of a protein and the unpredictability of the structure-function relationship [58]. Thus, a monoclonal antibody exists in solution as an ensemble of isomers, each capable of binding a structurally-different ligand. The plasticity of antibodies is exploited in current medicine, notably of camelid antibodies (antibodies of camels), which are homologous to human VH, variable heavy antibody chains, and possess a simpler structure than other antibodies. Whether structural changes are limited to a spatial rearrangement of contact residues for the ligand in the antibody binding site, or are propagated allosterically to other parts of the antibody molecule remains to be



determined. Israel Pecht [59], considering the high activation enthalpies for binding hapten or antigen, is inclined to assume that hapten-binding leads to a large, long-range, conformational transition of the whole antibody.

II. Receptors are examples of flexible proteins, demonstrating the advantages of structural plasticity and functional versatility for regulatory proteins

In the first sentence of the preface of his unsurpassed treatment of allosterism, Max Perutz praised the versatility of proteins [60]: *'In the popular view the structure of DNA has told us all about the molecular basis of life, but in fact DNAs and most RNAs are chemically inert, whereas proteins are the workhorses of the living cell. They function as catalysts and genetic regulators, pumps and motors, receptors and transducers, stores and transporters, scaffolds and walls, toxins and antitoxins, conductors and insulators, and much more'*.

Aside from antibodies, receptors are prime examples of ligand-induced adaptability of structure and functional versatility. However, receptors also exemplify the functional advantages of structural plasticity and functional versatility. Some receptors can bind different ligands and perform different actions. Promiscuity of ligand coupling is biologically important, because it enables, for example, one and the same cytokine to

cross-talk to different receptors, thus allowing one and the same cytokine to control different biological pathways. This considerably enlarges the signalling repertoire of a given cytokine.

Receptors are proteins engaged mostly in signal transfer from the outside to the inside of and between cells [61]. Aside from receptors in the plasma membrane of cells, there are receptors bound to intracellular boundaries, to mitochondrial and nuclear membranes. Nuclear receptors transmit signals from steroid and non-steroidal hormones such as thyroid hormones and retinoic acids, which can pass through the plasma membrane. Finally, pathogenic microbes [62] may transfer their receptors into cells, making them susceptible to infection [63].

Like receptors, receptor ligands have many diverse structures and functions acting as sensory signals, odorants and photons, hormones, growth factors, cytokines and chemoattractants, and much more. Ligands may be as small as a photon or as large as a protein, such as insulin. In some cases, the interaction with ligands causes a structural transition of the receptor, which triggers tyrosine phosphorylation. Some receptors have intrinsic tyrosine kinase activity and phosphorylation sites on the cytoplasmic side of the receptor chains, others recruit tyrosine kinases when activated and, in other cases, activation of a receptor and signalling requires the recruitment of a co-receptor.

Robert M. Stroud and James A. Wells [64] draw the distinction between two kinds of receptors. In one kind, ligand-induced structural changes bring about lateral association of separated subunits in the membrane, leading to dimerization or oligomerization of the receptor. In the other kind, structural changes are confined to a re-arrangement of a pre-existing polymeric assembly. Stroud and Wells have named receptors belonging to the first group, horizontal receptors and those belonging to the latter group, vertical receptors. Though I have adopted their classification here, there are some receptors that do not fit into their scheme. Fig. 2 and Fig:3.

Horizontal receptors bind mono- or multimeric protein ligands. This kind of receptors control gene transcription and translation and regulate cell replication, growth and cell death. The three-dimensional structures of almost a dozen complexes of the ligand-binding extracellular parts of these receptors with their ligands have been solved [65]. Although the structural rearrangements on binding of the ligand differ considerably from one receptor to the next, as Walter Sebald et al. have shown [66-68], all horizontal receptors eventually form dimers or oligomers on ligand binding and activation. A consequence of receptor oligomerization is the formation of new interactive surfaces on the inside of the membrane for coupling partners, mostly components of intracellular signalling chains. The best-characterized

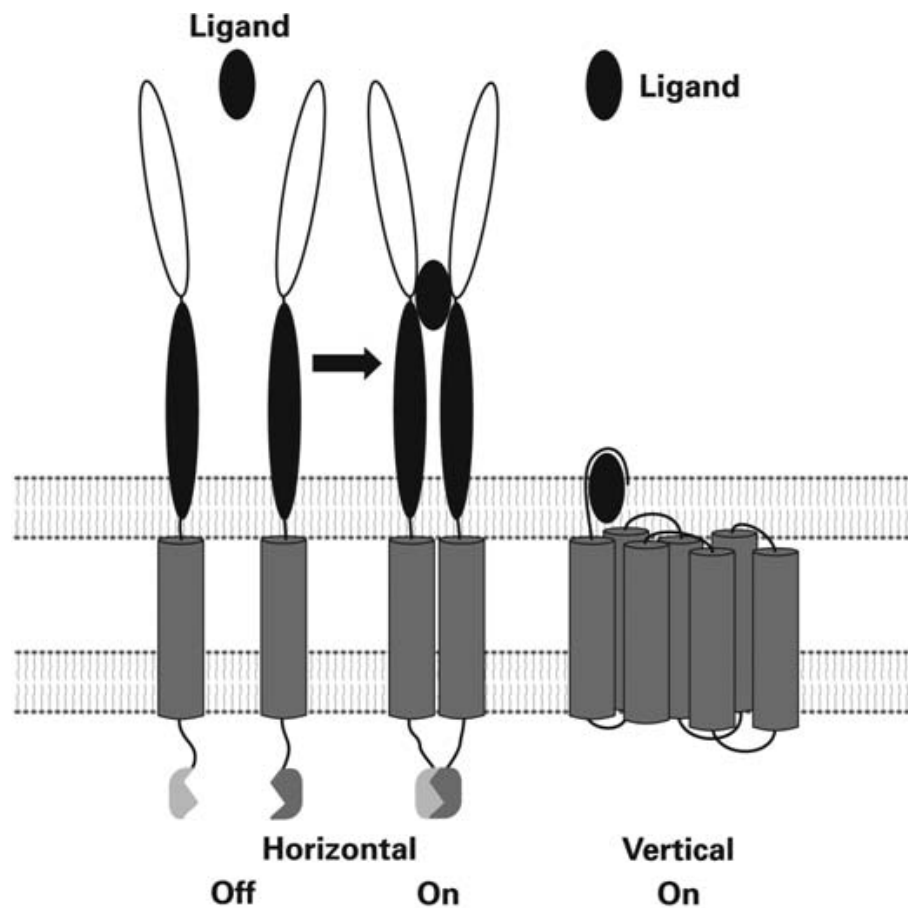


Figure 2 Key characteristics of horizontal and vertical receptors. The differences in structural changes on activation by binding the ligand are indicated schematically. The striking feature of horizontal receptors is that they associate and form dimers or multimers upon activation, whereas activation of vertical receptors does not involve quaternary structure changes of the already associated, preformed oligomer. This figure was originally reproduced from R.M. Stroud and J. A. Wells. 2004. Mechanistic Diversity of Cytokine Receptor Signalling across Cell Membranes, and the journal, *Science STKE*.2004.Re: 7, pps: 1-17, with approval of the authors and the AAAS in my forthcoming book in German: *<Von Molekülen zu Zellen. 100 Jahre experimentelle Biologie. Betrachtungen eines Biochemikers>*. It is shown here with permission of the GNT Verlag für Geschichte der Naturwissenschaften und der Technik. Diepholz. Stuttgart, Berlin.

examples of horizontal receptors are members of the family of human growth hormone receptors (hGHRs). The structural flexibility and coupling versatility of both receptor and ligand is startling. The same structures in human growth hormone can bind to different receptors [69], and, vice versa, hGH receptors can use the same structural elements for binding structurally different ligands.

Consisting of a modular structure and accepting both intra- and extra-cellular signals, the integrins constitute another group of horizontal receptors with remarkable structural pliability and functional versatility. Upon phosphorylation by tyrosine kinases, the tails of integrins bind specifically to muscle proteins such as talin, and transmit signals to the submembranous cytoskeleton [70,71], whereas the extracellular domain of integrins binds to constituents of the extracellular matrix. In this

way, integrins modulate cell-cell interactions [72] and play a central role in cell-cell adhesion.

The largest family of receptors is the species of seven transmembrane, heptahelical, receptors, coupled to G proteins, (G = GTP-binding). One cannot assign these receptors unequivocally to the vertical receptor class, because there is increasing evidence of oligomerization of G-protein-coupled receptors [73]. These receptors are ubiquitous: they are found in the cell membranes of archaea, halobacteria and metazoa, as well as plants, mammals and humans. Rhodopsin, located in the retina of the eye, which is activated by light [74], bacterio-rhodopsin [75], opsin [76] and the β_2 -adrenergic receptor [77] also belong to this group. These are currently the best characterized receptors at the structural level and are also targets of modifiers and inhibitors with important therapeutic value in medicine, particularly

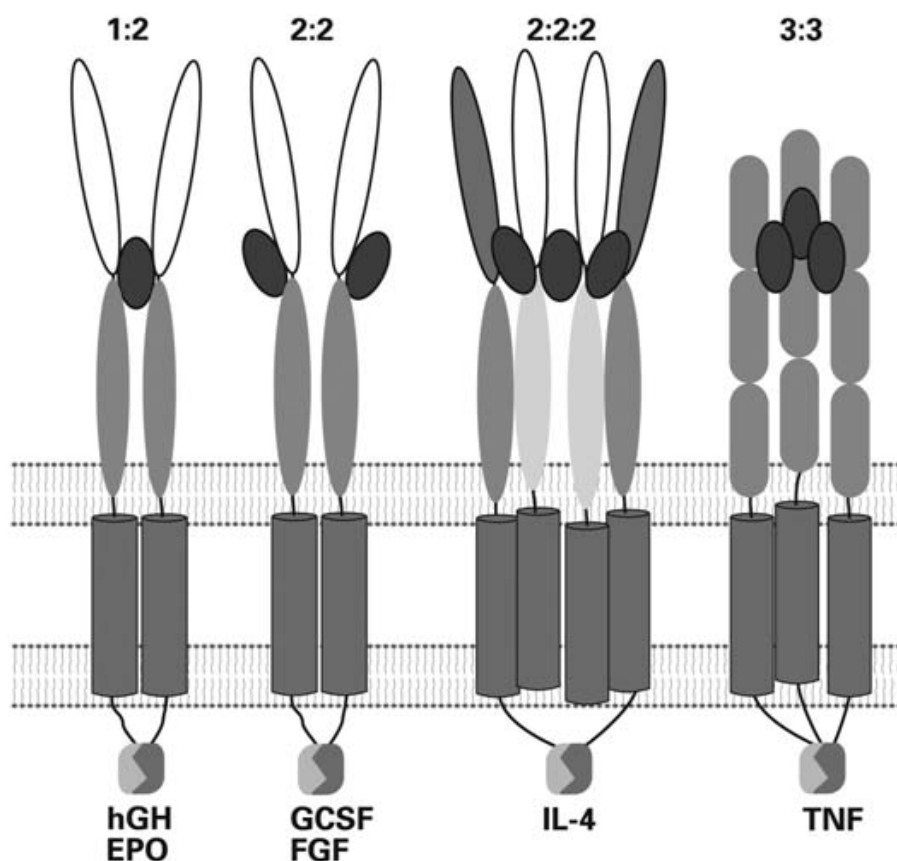


Figure 3 Illustration of the variety of different multimeric assemblies of horizontal cytokine-receptors on ligand binding and activation. hGH = human growth hormone. EPO = erythropoetin. GCSF = Granulocyte Colony-stimulating factor. FGF = Fibroblast growth factor. IL-4 = Interleukin 4. TNF = Tumour necrosis factor. This figure is originally reproduced with permission of the authors R. M. Stroud and J. A. Wells (2004) and the AAAS from *<Mechanistic Diversity of Cytokine Receptor Signalling Across Cell Membranes>*, and the journal, *Science* STKE.2004.Re: 7, pps: 1-17, and is shown in my forthcoming book in German: *<Von Molekülen zu Zellen. 100 Jahre experimentelle Biologie. Betrachtungen eines Biochemikers>*. It is reproduced here with permission of the GNT Verlag für Geschichte der Naturwissenschaften und der Technik. Diepholz. Stuttgart, Berlin.

cardiology. These receptors control ion pumps, phosphorylation-cascades and cell metabolism. One way, but not the only way, to transmit signals from these receptors after activation is through G proteins. For more information, see [19,78-80] and the recent excellent review by Ilka Böhme and Annette G. Beck-Sickingner in this journal [81].

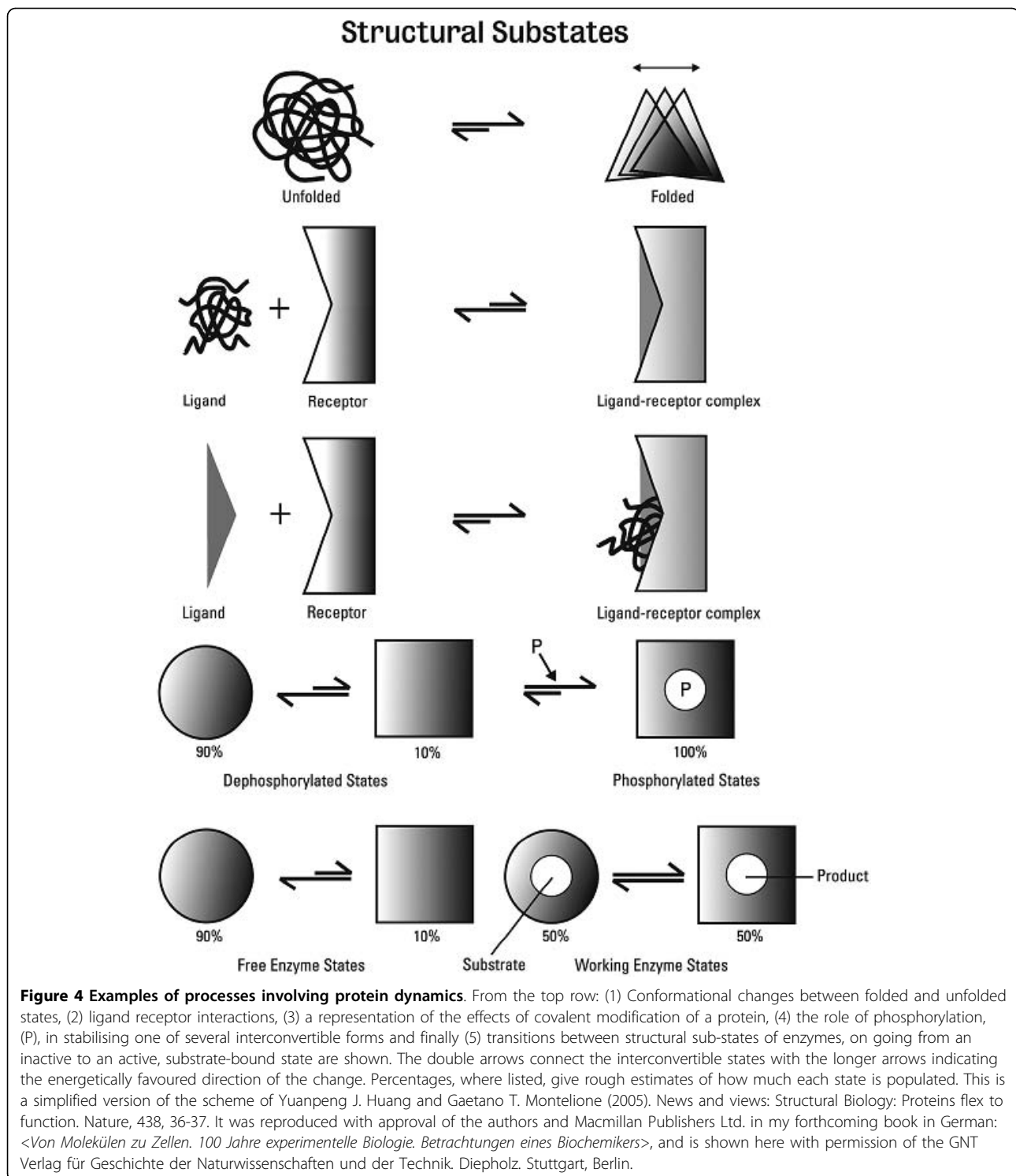
The flexibility and functional versatility of antibodies and receptors raise the question of how a protein must be designed in order to be able to drastically change its structure and function on binding ligands. New insights into the dynamics and physicochemical properties of some proteins may help to answer this question. The dynamics of proteins may also explain why protein-folding *in vitro* has turned out to be such an intractable problem.

Internal Dynamics of Proteins

Figure 4 lists examples of processes in which protein dynamics play a role. Hans Frauenfelder [82], a pioneer

in the field of protein physico-chemistry, described the dynamics of a protein, myoglobin, as early as the 1970s. According to Hans Frauenfelder et al., [83], some proteins are an assemblage of conformational variants which move in an intricate, multidimensional energy landscape and interconvert from one structure to another in tens of femtoseconds (A femtosecond is 10^{-15} sec) or hours. A detailed analysis of the temporal order of the conformational changes may eventually shed more light on the topography of the energy landscape and explain the properties of these proteins [84]. Since individual molecules may follow different paths, single-molecule methods are now increasingly used. Analysis of this intricate problem is being actively pursued and progress in numerical simulations is already permitting increasing complexity to be handled [85,86].

NMR spectroscopy, has also provided structural information on proteins exhibiting internal dynamics [87-89]. (For a review see: [[90]). For example, Eisenmesser et al.



[91,92] studied the structural changes in the enzyme cyclophilin A, on going from a non-active to a catalytically active state. (Cyclophilins are prolyl isomerases, that are inhibited by cyclosporin [93]). Eisenmesser et al. observed that during catalysis some amino acids can change their position as fast as the elementary steps

of the catalytic reaction, suggesting that conformational changes of the enzyme are synchronized with catalysis. However, surprisingly, identical structural changes also occur in the non-active enzyme. The only difference between the active and inactive enzyme is that the sub-states are differentially populated. One might speculate

that proteins with an intrinsic plasticity such as cyclophilin A may have been earmarked by evolution to become enzymatic catalysts, because all the structural prerequisites for catalysis are already present in the non-active protein.

Another protein, for which a comprehensive range of NMR data are available is human ubiquitin, a relatively small protein of 76 amino acids. The native state of ubiquitin was described [94] as an ensemble of different conformers, interconverting on a picosecond to nanosecond scale. The mobility of the structure of this protein is not restricted to its surface loops and side chains, it also includes its core, which was described as liquid-like and fluctuating [95,96]. It is remarkable that changes including mutational changes, leading to a complete redesign of its hydrophobic core do not compromise the structural integrity of this protein. According to Martin Karplus and his colleagues [97], such properties may even be representative of proteins in general.

The bewildering variety of conformational sub-states of some proteins is an indication of their structural plasticity and versatility of function. We are just beginning to understand what makes some proteins flexible and what kind of forces drive structural transitions. The extraordinary malleability of this kind of proteins is also of great interest for protein engineers trying to redesign proteins in the laboratory.

The progress in structural analysis and, particularly, NMR spectroscopy, has significantly extended our insight into co- and post-translational modifications in the last decade, a theme which has already been reviewed exhaustively in recent years [29,98,99]. In this context, it suffices to say that the diversity and versatility of structural disorder of proteins and of protein modifications, in addition to the consequences of rRNA and mRNA splicing, has greatly increased the repertoire of possible protein structures and functions, way beyond the information deducible from the genome. Disorder in proteins has important consequences, notably for the function of regulatory proteins, but also for protein interactions in the cell. It remains to be seen to what extent systems biology and proteomics will be able to handle the structural and functional versatility of proteins, notably of proteins carrying out regulatory functions.

In conclusion, structural flexibility seems to be an inherent property of some, if not all, proteins. It allows versatility of function and is essential for regulation. Such properties can not be deduced from the genetic readout, from the primary back bone structure of these proteins. Moreover, in many cases, regulated proteins can change their structure and function in the cell as a consequence of reversible covalent modifications, such as by phosphorylation and/or other changes. About 2%

of human genes encode protein kinases and phosphatases. Such reversible modifications also leave no traces in genetic readouts.

Another aspect that cellular biochemistry must deal with is the role of the environment. Compared with hemoglobin and other water soluble, allosterically, ligand-regulated proteins, the information on structural transformations of membranous receptors on ligand-binding and activation is limited. The reason is that receptors are embedded in a water-lipid interface. This makes it difficult to isolate them in a native conformation and identify the structures of the active conformers at high resolution. Therefore, the structural information at hand often encompasses only the water soluble extracellular part of the receptor, and does not allow assignment of specific roles in signal transmission to reactive interfaces on the cytosolic side, which are exposed to the cell's interior, where coupling with signal transmitters occurs. (Notable exceptions are the heptahelical G-protein coupled receptors).

The role of the environment

Although I shall concentrate here on the role of the membranous environment for receptor structure and function, the influence of the environment on actions and interactions of proteins, deserves, of course, serious consideration in all cases, because environmental influences on structurally variable proteins are not restricted to the membrane and are most likely also responsible for configurational rearrangements of proteins in other cellular environments.

How can the membranous environment affect receptor structure and function? [100]. The membranous environment may affect receptor activation and de-activation, signal transduction and receptor re-cycling in many ways [101]. Membrane lipids could either influence receptors through global changes in the physical state of the water lipid bilayer, for example through microviscosity changes, or individually through discrete and specific interactions. While, the former is unlikely in homoiothermic animals, the latter is to be expected. Although, the central importance of this problem is widely recognized, we still do not have enough data on the physicochemical restraints that individual membrane lipids may exert on the mobility and reactivity of proteins embedded in the membrane. An example is the role that is being attributed to complexes, or transducisomes, containing whole signal transmission chains, arranged on lipid scaffolds in the membrane [102,103]. Although, this is undoubtedly an attractive perspective, the role of lipids in forming and maintaining such transducisomes in membranes is not yet clear. We neither know the composition, size or the lifetime of such lipid-protein complexes [104].

Although, it is evident that the behaviour of proteins at the water-lipid interface is difficult to study, I also recall that for the older generation of biochemists, the lipid bilayer was mostly a undefined mixture that is better circumvented. For example, Carl F. Cori, [105], one of the great biochemists of the 20th century, with whom I had the privilege to work with in the 1950/60s, although finally accepting a role of insulin in making the membrane permeable for sugars [106], insisted that an intracellular enzyme, glucokinase or hexokinase, rather than a membrane-bound insulin receptor activating a membranous glucose carrier, controls the uptake of glucose into muscle.

The increasing awareness of the important role of the membranous environment in reactions of pivotal importance for the life of cells, brings hope that in the future more biochemists will try to decipher the influence of this particular solvent milieu on the many important reactions that occur in or at water-lipid interfaces in the cell. I expect that an important incentive for such greatly needed studies will be the availability of new methods for *in vivo* observations, methods that allow visualization of proteins in their natural environment and study of their movements and interactions *in situ*.

New and old ways of *in vivo* observation

Already in the 1960s, biochemists of my generation, notably those who were interested in regulation, were keen to determine the cellular location of the enzyme or the protein they were studying. However, cellular compartmentalisation made any extrapolation of the *in vitro* to the *in vivo* situation speculative and tentative. I described the situation in 1969 [107]: '*Equally as important as the actual molecular mechanism of allosteric transitions which may differ from enzyme to enzyme, is the in vivo expression of the regulatory potentiality of enzymes which control multienzyme sequences. The regulatory behaviour of an enzyme in the isolated state can only serve as a guide which may hopefully lead us to its actual control properties in the living cell*'. At that time, one tried to localize metabolites in cells by using radioactively labelled 2-deoxyglucose which, after being phosphorylated to 2-deoxyglucose-6-phosphate, accumulates in the cell [108]. Because it can not be metabolized further, this compound could be localized by a combination of radioautography and electron microscopy and from its position in the cell, one could deduce the location of the kinase which had phosphorylated the sugar analogue [109]. Later, receptors were labelled with fluorescent dyes [107] and their movements in cells studied [110]. Today, the introduction of a gene for fluorescent aquaporin into a living cell and its fusion to a gene expressing the protein to be

localized, has led to new and improved methodologies to monitor a protein and visualize its position in the intact cell [111,112]. Roger Tsien's methods in conjunction with other physical methods, such as fluorescence energy transfer, FRET, fluorescence recovery after photobleaching, FRAP, and fluorescence correlation spectroscopy with single molecule detection, [113], today allow the identification and tracking of an increasingly wide range of molecules, both large and small, in their natural environment. The introduction of new and refinement of older methods has already brought us a wealth of new and valuable information on receptor function in living cells [114-116]. In the future, the aim is to enlarge the number of proteins that are visualized until one gets a *movie*, showing movements and interactions of all kinds of proteins in a living cell [117].

Spectacular advances have also been made in optical microscopy by Stefan W. Hell in Göttingen [118,119]. Until now, the Abbe diffraction barrier of optical microscopy limited the resolution of conventional light microscopy to 200 to 300 nm in lateral dimensions, leaving many intracellular organelles and molecular structures unresolvable, thus limiting the use of microscopy for three-dimensional imaging of biological structures. Now however, the *Abbe* diffraction limit has been transcended, and resolutions of 20-30 nm and 50-60 nm in the lateral and axial dimensions respectively have been achieved. With new techniques [120], such as stochastic three-dimensional optical reconstruction microscopy, STORM, a resolution can be obtained which enables the visualization of cellular structures by light microscopy in three dimensions.

Equally impressive are the advances in recent years in the development of cryoelectron microscopy and cryoelectron tomography. Following the pioneering work in the 1960s of De Rosier and Aaron Klug, cryoelectron tomography is now beginning to allow to make visible complex molecular assemblies within a living cell, such as actin bundles, intermediate filaments, which build the scaffold of a cell. Moreover, one can monitor proteins and nucleic acids on ribosomes, the cellular sites of protein synthesis [121,122]. In the future, all will depend on how far one will be able to increase the resolution, which is still in the tenth nm range.

Attempts are also being made today to study the structure of proteins in the living cell by in cell NMR spectroscopy [123,124]. This might be the beginning of a new era of structural analysis of proteins *in situ*. Another technical development that deserves attention, is the application of high resolution scanning X ray diffraction microscopy, or SXDM, to *in situ* objects [125,126].

Conclusions

The chances of today's functional genetics and proteomics

Today, it is agreed that a functional assignment of the genome is needed [127]. For that purpose, the ENCODE project, ENCyclopaedia Of DNA Elements) [128], which should provide functional assignments to all genes in the human genome has been organized and implemented. The most advanced sequencing techniques, combined with the most sophisticated computational methods are being applied. A pilot study, covering 1%, or 30 megabases, of the human genome, showed that nearly all of the DNA in a human chromosome is transcribed into RNA, as well as revealing the existence of many, still unrecognized, stretches of DNA with as yet unknown functions.

Another objective lies in finding out how the expression of DNA, packed in the chromatin of the chromosome, is regulated and epigenetically controlled. This is a sign that in the future, in addition to all the efforts aiming at a functional definition of the genome, one will also take into account the bewildering and still growing number of ways and means of how the genetic readout is modified, controlled and regulated. Depending on how cell-specific gene expression is, one might even have to identify the functional role of genome sequences in individual, differentiated cells.

Stefan Bornholdt, [129], has pointed out that going from measurements of all the factors, including all regulatory factors, that are responsible for the expression of a single gene, to measurements of the same parameters controlling the expression of hundreds and thousands of genes in a gene network, would make the number of differential equations that have to be solved to describe such a realistic situation intractably large, at least for the time being. But, mathematical integration is necessary to handle large amounts of data and model complex biological systems. Such concerns are realistic, and have also been raised by the physicist and evolutionary biologist, Alfred Gierer [130]. Thus, it seems that today's systems biology must, like Odysseus, guide its ship with care and imagination, avoiding the perils of Scylla and Charybdis. On the one hand, large amounts of data are necessary to describe complex biological processes in living systems and on the other, it may turn out that the use of this plethora of data is limited by the availability of algorithms capable of integrating them. This will be a difficult problem to solve for mathematicians and computer specialists.

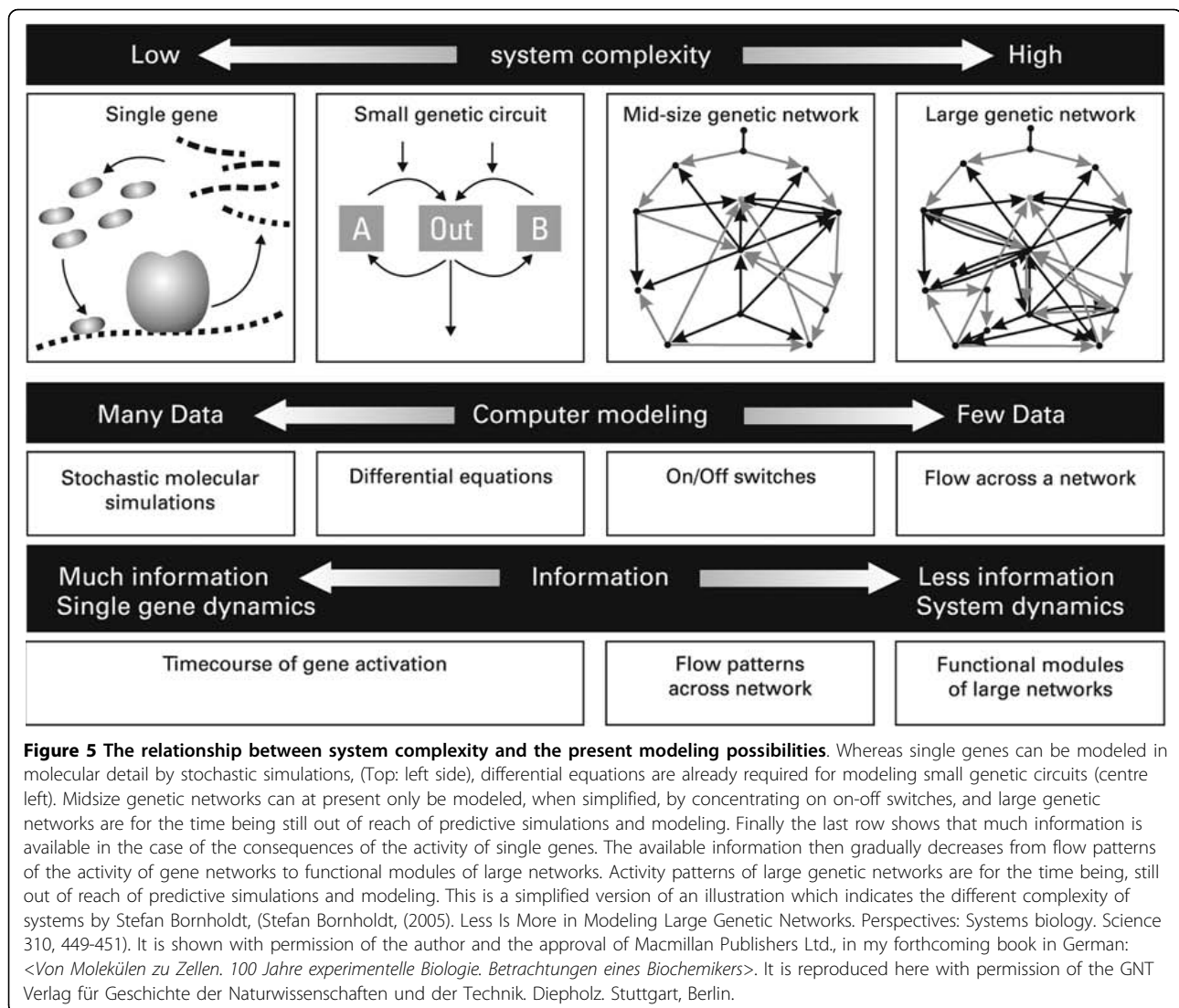
The relationship between system complexity and the present modelling possibilities is shown in Fig: 5.

The obvious way to deal with such immense complexity, is simplification and reduction, although there are

cogent reasons for which one has had to turn, for example, from single genes to gene networks, because deletions or mutations of single genes often have no discernible effect on the cell. For the time being however, I do not see any other way. However, here again a word of caution is warranted: for any kind of reduction, the question is of course how far can one reduce a complex biological process, without losing characteristic features? There are different ways to reduce complexity. One possibility is to choose a simple object. This is exemplified by Sidney Brenner's [131] introduction and successful use of *Caenorhabditis elegans* in genetics. The recent discovery of RNA-interference in *Caenorhabditis* by Craig Mello und Andrew Fire [132] demonstrates the value of a simple experimental object. Although, these experiments show the value of a suitable model organism for experiments in biology, one should not forget that the success of using simple biological objects, depends on the kind of questions asked.

Another possibility to simplify metabolic and genetic networks is to focus only on on/off switches and feedback loops that control whole systems [133]. Rather than studying complex sequences of reactions, one concentrates on the pacemaker of a chain of reactions. An example is the control of a chain of metabolic enzymes, for example glycolysis, by a single enzyme, serving as pacemaker [134], or the control of many genes in an operon by a promoter. These stripped down models may not represent the much more involved dynamics of the actual system in the living cell, but by comparing such simplified models with the actual situation in the cell, one can at least see whether and when, and to what extent, such binary on/off switch models mimic the dynamics of the real complex network in the living cell. For example, such simplified models, could correctly predict the dynamics of the complex genetic network, controlling segment polarity in the development of *Drosophila melanogaster* [135], and a simplified binary model of the genetic network controlling the yeast cell cycle [136] realistically modelled the actual situation.

I have commented on some of the problems functional genetics and *in situ* proteomics will face in the future. The extent to which the new *in vivo* experimental biology of the 21st century will be able to handle these challenges is yet to be seen. The solution of these and other problems will depend on whether we shall be able to handle, both the stochasticity of gene expression and the complexity of the genetic readout. This requires large amounts of reliable *in vivo* data and algorithms, capable of transforming huge amounts of data into realistic models. Despite all uncertainties and all the problems that remain to be solved, there is no doubt that a new era has begun. On the one side, today's systems



biology is part of and profits from today's exponentially growing information technology, that has already proved its ability to handle large amounts of data. (Google and similar search engines are examples). But, systems biology and *in vivo* proteomics also represent a new experimental biology, for which the living cell and its functions have become research objects. To study processes in the living cell, old techniques, such as microscopy have been refined and new techniques developed. The exponential and continuing progress in the development of techniques in the last years, and the optimism in handling complexity [137] make me optimistic that some, but not all, of the ambitious goals of systems biology and proteomics may be attained. The new area today reminds me of the beginnings of cell biology in the first half of the 19th century, which the microscope in the hands of the pioneers of cell biology, of Rudolf

Virchow, [138] and of O. and R. Hertwig, [139,140], helped to make possible.

In vivo observation of the molecular organisation of living systems will do away with the persuasive argument that the insights obtained by biochemists from *in vitro* studies are uncertain and tentative, because living systems can only be observed under conditions in which their integrity is preserved. This view was shared by Niels Bohr [141] (1885-1962. Nobel prize for Physics, 1922), who is probably the best-known example of a famous physicist interested in biology. (An interest that may have had its origin in the fact, that his father was a renowned physiologist). The functions and the material properties of a living organism were for Bohr only complementary and could not be derived from one other, a view still shared today by some chemists and physicists. Would Niels Bohr still be alive, he

would certainly welcome the new *in vivo* biochemistry of the 21st century.

Acknowledgements

I dedicate this essay to Max Perutz, (1914-2002), to remember his pivotal contributions to the study of structure and function of regulatory proteins. I am greatly indebted to my colleague Utz Fischer, Professor and Chairman of Biochemistry I at the Biocenter of the University of Würzburg, for critically reading and correcting my manuscript and for sponsoring my publication in Cell Communication and Signaling. Thanks also to one of the referees, who also suggested inclusion of the role of intrinsically disordered proteins in the discussion of internal dynamics of proteins, advice which I have followed.

Received: 19 July 2010 Accepted: 17 September 2010

Published: 17 September 2010

References

1. Alberts B, et al: **Molecular Biology of the Cell**. New York. N.Y. USA, 4.
2. Boeger H, Bushnell DA, Avis R, Griesenbeck J, Lorch Y, Strattan JS, Westover KD, Kornberg RD: **Structural basis of eukaryotic gene transcriptions**. *FEBS Letters* 2005, **579**:899-903.
3. **What is a Gene?**. *Nature* 2006, **441**:399-401, News Feature.
4. Benzer S: **The fine structure of the gene**. *Scientific American* 1962, **70**:84.
5. Gilbert W: **Why genes in pieces?** *Nature* 1978, **271**:501.
6. Carr K: **Nobel goes to discoverers of "split genes"**. *Nobel Prize for Medicine or Physiology*. *Nature* 1993, **365**:597.
7. Dermitzakis ET, Reymond A, Antonarakis SE: **Conserved non-genic sequences - an unexpected feature of mammalian genomes**. *Nat Rev Genet* 2005, **6**:151-157.
8. Ambros V: **The function of animal micro RNA's**. *Nature* 2004, **431**:350-355.
9. Avery OT, MacLeod CM, McCarty M: **Studies on the chemical nature of the substance inducing transformation of pneumococcal types. Induction of transformation by a deoxyribonucleic acid fraction isolated from Pneumococcus Type III**. *J Exp Med* 1944, **79**:137-158.
10. Hershey AD, Chase M: **Independent functions of viral protein and nucleic acid in growth of bacteriophage**. *J Gen Physiol* 1952, **36**:39-56.
11. Watson JD, Crick FHC: **Molecular structure of nucleic acid. A structure of deoxyribose nucleic acid**. *Nature* 1953, **171**:737-738.
12. Watson JD, Crick FHC: **Genetical implications of the structure of Deoxyribonucleic Acid**. *Nature*. 1953, **171**:964-967.
13. Meselson M, Stahl FW: **The replication of DNA in Escherichia coli**. *Proc Natl Acad Sci USA* 1958, **44**:671-682.
14. Crick FHC: **On protein synthesis**. *Symposium of Soc Expt Biol* 1957, **12**:138-163.
15. Horace Freeland Judson: **The eight day of creation**. *Makers of the revolution in Biology. Expanded Edition* Cold Spring Harbor Laboratory Press. NY. USA 1996.
16. Weismann A: **Das Keimplasma. Eine Theorie der Vererbung**. *Gustav Fischer Verlag, Jena, Deutschland* 1892, (Engl. Ed. 1993).
17. Temin H: **The DNA provirus hypothesis. The establishment and implications of RNA directed DNA synthesis**. *Science* 1976, **192**:1075-1080.
18. Baltimore D: **Viruses, polymerases and cancer**. *Science* 1976, **192**:632-636.
19. Helmreich EJM: **Structural flexibility of small GTPases. Can it explain their functional**. *Biol Chem* 2004, **355**:1121-1136.
20. Gilbert W: **Why genes in pieces?** *Nature* 1978, **271**:501.
21. Carr K: **Nobel goes to discoverers of "split genes"**. *Nobel Prize for Medicine or Physiology*. *Nature* 1993, **365**:597.
22. Cech TR, Zaugg AJ, Grabowski PJ: **In vitro splicing of the ribosomal RNA precursor of Tetrahymena. Involvement of a guanosine nucleotide in the excision of the intervening sequence**. *Cell* 1981, **27**:487-496.
23. Thomas R: **Cech and Sidney Altman were awarded with the 1989 Nobel prize for Chemistry**. *Nature* 271:501, See also: Exploring the New RNA World. Nobel-lecture by Thomas R. Cech.
24. Spilianakis CG, Lalitoti MD, Town T, Lee GR, Flavell RA: **Interchromosomal associations between alternatively expressed loci**. *Nature* 2005, **435**:637-645.
25. Feinberg AP, Ohlsson R, Henikoff S: **The epigenetic progenitor origin of human cancer**. *Nature Rev Genetics* 2006, **7**:21-33.
26. Osborne CS, Chakalova L, Mitchell JA, Horton AM, Wood AL, Bolland DJ, Corcoran AE, Fraser P: **Myc dynamically and preferentially relocates to a transcription factory occupied by Igh**. *PLoS Biology* 2007, **192**:1763-1772.
27. Pinchas A, Toporik A, Edelleit S, Peretz Y, Diber A, Shemesh R, Novik R: **Transcription-mediated gene fusion in the human genome**. *Genome Res* 2006, **16**:30-36.
28. Kapranov P, Willingham AT, Gingeras TR: **Genome-wide transcription and the implications for genomic organization**. *Nature Reviews Genetics* 2007, **8**:413-423.
29. Tompa P: **Structure and Function of Intrinsically Disordered Proteins**. CRC Press. Taylor and Francis Group. Boca Raton, London, New York 2010, A Chapman and Hall Book.
30. Bird A: **Perceptions of epigenetics**. *Nature* 2007, **447**:396-398.
31. Buc H, et al: **Allosteric Proteins: Forty years with Monod, Wyman and Changeux**. In *Proceedings of the Accademia Nazionale dei Lincei. Rome*. Edited by: Brunori M. Accademia dei Lincei. Italy, and personal communication with Henri Buc; 2005:17:31-49.
32. Pauling L, Niemann C: **The structure of proteins**. *Science* 1939, **61**:1860-1897.
33. Anfinsen CB, Haber E, Sela M, White FH Jr: **The kinetics of formation of native ribonuclease during oxidation**. *Proc Natl Acad Sci USA* 1961, **47**:1309-1314.
34. Richards FM, Vithayathil PJ: **The preparation of subtilisin-modified ribonuclease and the separation of the peptide and protein components**. *J Biol Chem* 1959, **234**:1459-1465.
35. Lazarides T, Karplus M: **Thermodynamics of protein folding**. *Biophys Chem* 2003, **100**:367-395.
36. Goldberger RF, Epstein CJ, Anfinsen CB: **Acceleration of Reactivation of reduced Bovine Pancreatic Ribonuclease by a Microsomal System from Rat Liver**. *J Biol Chem* 1963, **238**:628.
37. Fanghänel J, Fischer G: **Insights into the Catalytic Mechanism of Peptidyl Prolyl cis/trans Isomerase**. *Frontiers in Bioscience* 2004, **9**:4553-4587.
38. Reimer U, Fischer G: **Local structural changes caused by peptidyl prolyl cis/trans isomerization in the native state of proteins**. *Biophysical Chemistry* 2002, **96**:203-212.
39. Walsh KA, Neurath H: **Trypsinogen and chymotrypsinogen as homologous proteins**. *Proc Natl Acad Sci USA* 1964, **52**:884-889.
40. Brennan SO, Peach RJ: **The processing of human Proinsulin and Chicken Proalbumin by hepatic vesicles suggests a convertase, specific for X-Y Arg-Arg or Arg-X-Y-Arg sequences**. *J Biol Chem* 1991, **266**:21504-21508.
41. Ron D, Walter P: **Signal integration in the endoplasmic reticulum. Unfolded Protein Response**. *Nature Rev Mol Cell Biol* 2007, **8**:519-529.
42. Collinge J, Clark AR: **A general model of Prion strains and their pathogenicity**. *Science* 2007, **318**:930-936.
43. Ming Y, Cheng F, Hartl FU, Jörg M, Pollock RA, Kalousek Fr, Neupert W, Hallberg EM, Hallberg RL, Horwich AL: **Mitochondrial heat-shock protein hsp60 is essential for assembly of proteins imported into yeast mitochondria**. *Nature* 1989, **337**:620-625.
44. Hartl FU, Hayer-Hartl M: **Molecular chaperones in the cytosol. From nascent chain to folded protein**. *Science* 2002, **295**:1852-1858.
45. Polier S, Dragovic Z, Bracher A, Hartl FU: **Structural basis for cooperative protein folding by HSP 70 and HSP 110 molecular chaperones**. *Cell* 2008, **133**:1068-1079.
46. Schiene-Fischer C, Habazettl J, Schmid FX, Fischer G: **The hsp70 chaperone DnaK is a secondary amide peptide bond cis/trans isomerase**. *Nature Structural Biology* 2002, **9**:419-423.
47. Meister G, Hannus S, Plöttner O, Lagerbauer B, Baars T, Hartmann E, Fakan S, Fischer U: **SMNrp is an essential pre-mRNA splicing factor required for the formation of the mature spliceosome**. *EMBO J* 2001, **20**:2304-2314.
48. Chari A, Golas MM, Klingenhäger M, Neuenkirchen N, Sander B, Englbrecht C, Sickmann A, Stark H, Fischer U: **An Assembly Chaperone collaborates with the SMN Complex to generate Spliceosomal SnRNPs**. *Cell* 2008, **135**:497-509.
49. Leroux MR, Hartl FU: **Cellular functions of molecular chaperones**. In *Mechanisms of Protein Folding*. Published in: *Frontiers in molecular Biology*. Edited by: Pain RH, Hames BD, Glover DM. Oxford University Press. Oxford. UK; , Second 2000:Chapter 14:364-392.
50. Perutz M: **Mechanisms of cooperativity and allosteric regulation in proteins**. Cambridge University Press. Cambridge, UK 1990.

51. Pauling L: **A theory of the structure and the process of formation of antibodies.** *J Am Chem Soc* 1940, **62**:2643-2657.
52. Lancet D, Pecht I: **Kinetic evidence for hapten-induced conformational transition in immunoglobulin MOPC 460.** *Proc Natl Acad Sci USA* 1976, **73**:3549-3553.
53. Metzger H: **The effect of antigen on antibodies: Recent studies.** *Contemporary Topics in Molecular Immunology* 1978, **7**:119-152.
54. Foote J, Milstein C: **Conformational isomerism and the diversity of antibodies.** *Proc Natl Acad Sci USA* 1994, **91**:10370-10374.
55. Kramer A, Keitel T, Winkler K, Stöcklein W, Höhne W, Schneider-Mergener J: **Molecular basis for the binding promiscuity of an anti-p24 (HIV-1) monoclonal antibody.** *Cell* 1997, **91**:799-809.
56. James LJ, Roversi P, Tawfi S: **Antibody multispecificity mediated by conformational diversity.** *Science* 2003, **299**:1362-1367.
57. Wilson IA, Stanfield RL: **Antibody-antigen interactions: New structures and new conformational changes.** *Curr Opin Struct Biol* 1994, **4**:857-867.
58. Pecht I: **Dynamic aspects of antibody function.** In *The Antigens*. Edited by: Sela M. Academic Press, New York, USA; 1982:6:1-49.
59. Michaelides MC, Sherman R, Helmreich E: **The Interaction of Muscle Phosphorylase with Soluble Antibody Fragments.** *J Biol Chem* 1964, **239**:4171-4181.
60. Perutz M: **Mechanisms of cooperativity and allosteric regulation in proteins.** Cambridge, UK 1990.
61. Helmreich EJM: **The Biochemistry of Cell signalling.** Oxford University Press, Oxford, UK 2001, 190-209.
62. Finlay BB, Falkow S: **Common themes in microbial pathogenicity revisited.** *Microbiol Mol Biol Rev* 1997, **61**(2):136-169.
63. Meylan E, Tschopp J, Karin M: **Intracellular pattern recognition receptors in the host response.** *Nature* 2006, **442**:39-44.
64. Stroud RM, Wells JA: **Mechanistic Diversity of Cytokine Receptor Signaling across Cell Membranes.** *Science* 2004, **7**:1-17.
65. Bazan JF: **Emerging families of cytokines and receptors.** *Curr Biol* 1993, **3**:603-606.
66. Wang Y, Shen BJ, Sebald W: **A mixed-charge pair in human interleukin 4 dominates high-affinity interaction with the receptor chain.** *Proc Natl Acad Sci USA* 1997, **94**:1657-1662.
67. Hage T, Sebald W, Reinemer P: **Crystal structure of the interleukin-4 receptor chain complex reveals a mosaic binding interface.** *Cell* 1999, **97**:271-281.
68. Zhang JL, Buehner M, Sebald W: **Functional epitope of common γ chain for interleukin-4 binding.** *Eur J Biochem* 2002, **269**:1490-1499.
69. Cunningham BC, Wells JA: **Comparison of a structural and a functional epitope.** *J Mol Biol* 1993, **234**:554-563.
70. Weljie AM, Hwang PM, Vogel HJ: **Solution structures of the cytoplasmic tail complex from platelet integrin α IIb- and β 3-subunits.** *Proc Natl Acad Sci USA* 2002, **99**:5878-5883.
71. Hogg N, Leitinger B: **Shape and shift changes related to the function of leukocyte integrins LFA-1 and Mac-1.** *J Leukoc Biol* 2001, **69**:893-898.
72. Schwartz MA, Ginsberg MH: **Networks and crosstalk: Integrin signalling spreads.** *Nat Cell Biol* 2002, **4**:65-68.
73. Milligan G: **Oligomerisation of G-protein-coupled receptors.** *J of Cell Science* 2001, **114**(7):1265-1271.
74. Palczewski K, Kumasaka T, Hori T, Behnke CA, Motoshima H, Fox BA, Le Trong I, Teller DC, Okada T, Stenkamp RE, Yamamoto M, Miyano M: **Crystal structure of rhodopsin: A G protein-coupled receptor.** *Science* 2000, **289**:739-745.
75. Subramaniam S, Lindahl M, Bullough P, Faruqi AR, Tittor J, Oesterheld D, Brown L, Lanyi J, Henderson R: **Protein conformational changes in the bacteriorhodopsin photocycle.** *J Mol Biol* 1999, **287**:145-161.
76. Scheerer P, Hildebrand PW, Kim YJ, Krauss N, Choe HW, Hofmann KP, Ernst O: **Crystal structure of opsin in its G-protein-interacting conformation.** *Nature* 2008, **455**:497-502.
77. Rasmussen SG, Choi HJ, Rosenbaum DM, Kobilka TS, Thian FS, Edwards PC, Burghammer M, Ratnala VR, Sanishvili R, Fischetti RF, Schertler GF, Weis WI, Kobilka BK: **Crystal structure of the human beta 2 adrenergic G protein-coupled receptor.** *Nature* 2007, **450**:383-387.
78. Helmreich EJM, Hofmann KP: **Structure and function of proteins in G-protein coupled signal transfer.** *Biochim Biophys Acta* 1996, **1286**:285-322.
79. Strange PG: **G-protein coupled receptors: Conformations and states.** *Biochem Pharmacol* 1999, **58**:1081-1088.
80. Pawson T, Nash P: **Assembly of cell regulatory systems through protein interaction domains.** *Science* 2003, **300**:445-452.
81. Böhme I, Beck-Sickinger AG: **Review: Illuminating the life of GPCR's, (G-protein-coupled receptors).** *Cell Communication and Signaling* 2009, **7**:16.
82. Austin RH, Beeson KW, Eisenstein L, Frauenfelder H, Gunsalus IC: **Dynamics of ligand binding to myoglobin.** *Biochemistry* 1975, **14**:5355-5373.
83. Frauenfelder H, Sligar SG, Wolynes PG: **The energy landscapes and motions of proteins.** *Science* 1991, **254**:1598-1603.
84. Böckmann RA, Grubmüller H: **Nanoseconds molecular dynamics simulation of primary mechanical energy transfer steps in F1-ATP synthase.** *Nat Struct Biol* 2002, **9**:198-202.
85. **Special issue on Single Molecules.** *Science* 1999, **283**:1593-1804.
86. Haas J, Vöhringer-Martinez E, Bögehold A, Matthes D, Pelah A, Abel B, Grubmüller H: **Primary steps of pH-dependent insulin aggregation kinetics are governed by conformational flexibility.** *Chem Biol Chem* 2009, **10**:1816-1822.
87. Wagner G, Wüthrich K: **Dynamic model of globular protein conformations based on NMR studies in solution.** *Nature* 1978, **275**:247-248.
88. Palmer III, Arthur G: **Characterization of the Dynamics of Biomacromolecules.** *NMR Chem Rev* 2004, **104**:3623-3640.
89. Mulder FA, Mittermaier A, Hon B, Dahlquist FW, Kay LE: **Studying excited states of proteins by NMR spectroscopy.** *Nature Structural Biology* 2001, **8**:932-935.
90. Huang YJ, Montelione GT: **News and views: Structural Biology: Proteins flex to function.** *Nature* 2005, **438**:36-37.
91. Eisenmesser EZ, Millet O, Labeikovsky W, Korzhnev DM, Wolf-Watz M, Bosco DA, Skalicky JJ, Kay LE, Kern D: **Intrinsic dynamics of an enzyme underlies catalysis.** *Nature* 2005, **438**:117-121.
92. Eisenmesser EZ, Bosco DA, Akke M, Kern D: **Enzyme Dynamics during Catalysis.** *Science* 2002, **295**:1520-1523.
93. Fischer G: **Chemical aspects of peptide bond isomerisation.** *Chem Soc Rev* 2000, **29**:119-127.
94. Tjandra N, Feller SE, Pastor RW, Bax A: **Rotational diffusion anisotropy of human ubiquitin from 13 N NMR relaxation.** *J Am Chem Soc* 2005, **127**:12562-12566.
95. Johnson EC, Lazar GA, Desjarlais JR, Handel TM: **Solution structure and dynamics of a designed hydrophobic core variant of ubiquitin.** *Struct Fold Des* 1999, **7**:967-976.
96. Schwieters CD: **How much backbone motion in ubiquitin is required to account for dipolar coupling data measured in multiple alignment media as assessed by independent cross validation?** *J Am Chem Soc* 2004, **126**:2923-2938.
97. Best RB, Clarke J, Karplus M: **The origin of protein side chain order parameter distributions.** *J Am Chem Soc* 2004, **126**:7734-7735.
98. Walsh CT: **Posttranslational Modifications of Proteins: Expanding Nature's Inventory.** Roberts and Co. Publ 2006.
99. Sickmeier M, Hamilton JA, Le Gall T, Vacie V, Cortese MS, Tantos A, Szabo B, Tompa P, Chen J, Uversky VN, Obradovic Z, Dunker AK: **DisProt: the Database of Disordered Proteins.** *Nucleic Acids Res. Special Issue on Synthetic Biology.* **35**:1-18 2007, , Database: D786-D793.
100. Helmreich EJM: **Environmental influences on signal transduction through membranes. A retrospective mini-review.** *Biophys, Chem* 2003, **100**:519-534.
101. Helmreich EJM, Elson EL: **Mobility of proteins and lipids in the membrane.** Edited by: Greengard P, Robinson GA. *Advances in Cyclic Nucleotide and Protein Phosphorylation Research.* New York; 1984:1-62.
102. Simons K, Ikonen E: **Functional rafts in cell membranes.** *Nature* 1997, **387**:569-572.
103. Simons K, Toomre D: **Lipid rafts and Signal transduction.** *Nature Reviews Molecular Cell Biology* 2000, **1**:31-39.
104. Nichols B: **Without a raft.** *Nature News and Views* 2005, **436**:637-639.
105. Cori CF: **Influence of hormones on enzymatic reactions.** *Proc 1st Internat Congr Biochem* London, UK. Cambridge University Press 1949, **9**.
106. Helmreich E, Cori CF: **The distribution of pentoses between plasma and muscle.** *J Biol Chem* 1957, **224**:663-679.
107. Helmreich E: **Control of Synthesis and Breakdown of Glycogen.** Edited by: Florkin M, Stotz EH. Elsevier, Amsterdam, Netherlands; 1969:17-92, Chapter II in: *Starch and Cellulose.* Comprehensive Biochemistry.
108. Kipnis DM, Cori CF: **Studies of Tissue Permeability. V. The penetration and phosphorylation of 2-deoxyglucose in the rat diaphragm.** *J Biol Chem* 1959, **234**:171-177, Clark SL Jr.: *The synthesis and storage of protein*

- by isolated lymphoid cells, examined by autoradiography with the electron microscope. *Am. J. Anat.* 1966, **119**:375-404.
109. Heithier HD, Hallmann F, Boege H, Reilander C, Deed KA, Jaeggi DJ, Arndt-Jovin TM, Jovin Helmreich EJM: **Synthesis and properties of fluorescent beta-adrenoceptor ligands.** *Biochemistry* 1994, **33**:9126-9134.
110. Henis YI, Hekman M, Elson EL, Helmreich EJM: **Lateral motion of G-protein-coupled receptors in membranes of cultured liver cells.** *Proc Natl Acad Sci USA* 1982, **79**:2907-2911.
111. Tsien RY: **Fluorescent probes of cell signalling.** *Annu Rev Neurosci* 1989, **12**:227-253.
112. Tsien RY: **The green fluorescent protein.** *Annu Rev Biochem* 1998, **67**:509-544.
113. Thompson RE, Larson DR, Webb WW: **Precise nanometer localization analysis for individual fluorescent probes.** *Biophys J* 2002, **82**:2775-2783.
114. Lohse MJ, Nikolaev VO, Hein P, Hoffmann C, Vilardaga JP, Bünemann M: **Optical techniques to analyze real-time activation and signaling of G-protein-coupled receptors.** *Trends in Pharmacological Sciences* 2008, **29**(3):159-165.
115. Vilardaga JP, Bünemann M, Krasel C, Castro M, Lohse MJ: **Measurement of the millisecond activation switch of G protein-coupled receptors in living cells.** *Nat Biotechnol* 2003, **21**:807-812.
116. Hein P, Rochais F, Hoffmann C, Dorsch S, Nikolaev VO, Engelhardt S, Berlot CH, Lohse MJ, Bünemann M: **Gs activation is time-limiting in initiating receptor-mediated signaling.** *J Biol Chem* 2006, **281**:33345-33351.
117. Tsien RY: **Imaging imaging's future.** *Nature Rev Mol Cell Biol* 2003, **4**:16-21.
118. Hell SW: **Toward fluorescence nanoscopy.** *Nat Biotechnol* 2003, **21**:1347-1355.
119. Hell SW: **Far-Field Optical nanoscopy.** *Science* 2007, **316**:1153-1158.
120. Huang B, Wang W, Bates M, Zhuang X: **Three-Dimensional Super-Resolution Imaging by Stochastic Optical Reconstruction Microscopy.** *Science* 2008, **319**:810-813.
121. Baumeister W: **Electron tomography: Towards visualizing the molecular organization of the cytoplasm.** *Curr Opin Struct Biol* 2002, **12**:679-684.
122. Medalia O, Weber I, Frangakis AS, Nicastro D, Gerisch G, Baumeister W: **Macromolecular architecture in eukaryotic cells visualised by cryoelectron tomography.** *Science* 2002, **298**:1209-1213.
123. Inomata K, Ohno A, Tochio H, Isogai S, Tenno T, Nakase I, Takeuchi T, Ikuhiko N, Futaki S, Ito Y, Hiroaki H, Shirakawa M: **High-resolution multi-dimensional NMR spectroscopy of proteins in human cells.** *Nature* 2009, **458**:106-109.
124. Sakakibara D, Sasaki A, Ikeya T, Hamatsu J, Hanashima T, Mishima M, Yoshimasu M, Hayashi N, Mikawa T, Wälchli M, Smith BO, Shirakawa M, Guntert P, Ito Y: **Protein structure determination in living cells by in-cell NMR spectroscopy.** *Nature* 2009, **458**:102-105.
125. Thibault P, Dierolf M, Menzel A, Bunk O, David C, Pfeiffer F: **High-Resolution Scanning X-ray Diffraction Microscopy.** *Science* 2008, **321**:379-382.
126. Chapman HN: **Focus on X-ray Diffraction.** *Science* 2008, **321**:352-353.
127. White SH: **Perspectives Biochemistry.** *Science* 2007, **317**:1045-1046.
128. The ENCODE Project Consortium: *Science* 2004, **306**:636-640.
129. Bornholdt S: **Less Is More in Modeling Large Genetic Networks. Perspectives: Systems biology.** *Science* 2005, **310**:449-451.
130. Gierer A: **Relation between neurophysiological and mental states. Possible limits of decodability.** *Naturwissenschaften* 1983, **70**:282-287.
131. Brenner S: **The genetics of *Caenorhabditis elegans*.** *Genetics* 1974, **77**:71-94.
132. Fire A, Xu SQ, Montgomery MK, Kostas SA, Driver SE, Mello CC: **Potent and specific genetic interference by double stranded RNA in *Caenorhabditis elegans*.** *Nature* 1998, **391**:806-811.
133. Brandman O, Ferrell JE Jr, Li R, Meyer T: **Interlinked Fast and Slow Positive Feedback Loops Drive Reliable Cell Decisions.** *Science* 2005, **310**:496-498.
134. Helmreich EJM, Cori CF: **Regulation of glycolysis in muscle.** *Advances in Enzyme Regulation* 1965, **3**:91-107.
135. Dassow G, Meir E, Munro EM, Odell GM: **The segment polarity network is a robust developmental module.** *Nature* 2000, **406**:188-192.
136. Li F, Long T, Lu Y, Ouyang Q, Tang C: **The yeast cell-cycle network is robustly designed.** *Proc Natl Acad Sci USA* 2004, **101**:4781-4786.
137. Gilbert SF, Sarkar S: **Embracing complexity: Organicism for the 21st century.** *Developmental Biology* 2000, **219**:1-9.
138. Virchow R: **Rudolf Virchow.** *Ludwig Aschoff Hoffmann & Campe, Verlag, Hamburg* 1940.
139. Hertwig O: **Das Problem der Befruchtung und der Isotropie des Eies. Eine Theorie der Vererbung.** *Jenaische Zeitschrift für Naturwissenschaft* 1884, **18**:21-23.
140. Hertwig R: **Lehrbuch der Zoologie.** 9. Aufl., Jena, Deutschland. 1910.
141. Bohr N: **Light and life.** *Naturwissenschaften* 1933, **21**:245-250.

doi:10.1186/1478-811X-8-26

Cite this article as: Helmreich: Ways and means of coping with uncertainties of the relationship of the genetic blue print to protein structure and function in the cell. *Cell Communication and Signaling* 2010 **8**:26.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

