Available online at www.sciencedirect.com
ScienceDirect
BioSystems 87 (2007) 289–298
BioSystems
www.elsevier.com/locate/biosystems

Protein folding and the robustness of cells

G.J. Shawn Tan, Ferran D. Revilla, Klaus-Peter Zauner*

School of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, United Kingdom

Received 28 February 2005; received in revised form 8 July 2006; accepted 15 July 2006

Abstract

The intricate intracellular infrastructure of all known life forms is based on proteins. The folded shape of a protein determines both the protein's function and the set of molecules it will bind to. This tight coupling between a protein's function and its interconnections in the molecular interaction network has consequences for the molecular course of evolution. It is also counter to human engineering approaches. Here we report on a simulation study investigating the impact of random errors in an abstract metabolic network of 500 enzymes. Tight coupling between function and interconnectivity of nodes is compared to the case where these two properties are independent. Our results show that the model system under consideration is more robust if function and interconnection are intertwined. These findings are discussed in the context of nanosystems engineering.

Keywords: Self-isolation; Organic devices; Complex networks; Nanoengineering; Evolvability; Modularity

1. Nature's nanoengineering

Nature's astounding integration density has its roots in an enormous number of macromolecular structures that implement a wide variety of highly specific functions. Most prominent among them are proteins. These ubiquitous functional nanocomponents are macromolecules consisting of several thousand atoms. Biomimicry in the field of materials sciences is increasingly exploring the molecular and nanolevel detail that yields the remarkable properties of animate matter. A technology that would enable the deliberate design and fabrication of similar materials and systems would be highly desirable for a broad variety of applications (Lehn, 2002). Established engineering processes, however, are not well suited to the development of organic devices (Ball, 2002; Luo, 2003). In the context of such a technology biological means of achieving robust systems are of interest. In the following we will explore a potential role of protein folding for the robustness of networks of interacting proteins.

1.1. Protein folding

Proteins are linear chains of typically a few hundred building blocks taken since two thousand million years from the same set of about 20 amino acids. The linear chain is assembled according to an edited molecular copy of the coding region of a gene on the DNA. In principle the amino acids can be combined in arbitrary order. This opens up a vast space of possible macromolecules that can be assembled from the amino acid building blocks. During and following the assembly, the protein will spontaneously curl up under the electrostatic interaction of its atoms into a defined but agile three-dimensional form.

In proteins function follows form. Folding of the amino acid chain into a spacial structure is essential to the capabilites of a protein. The resulting spatial

* Corresponding author. Tel.: +44 23 8059 3089; fax: +44 23 8059 9179.
E-mail addresses: nwahs81@hotmail.com (G.J. Shawn Tan), fdr04r@ecs.soton.ac.uk (F.D. Revilla), kpz@ecs.soton.ac.uk (K.-P. Zauner).

0303-2647/$ – see front matter © 2006 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.biosystems.2006.09.025

structure typically fits in a sphere of 4–8 nm diameter and is largely determined by the amino acid sequence. Amino acids differ in geometry, electrostatic characteristics, and chemical properties. Not all amino acid chains readily fold (Sosnick et al., 2002), but many randomly created sequences do (Davidson and Sauer, 1994). More often than not, however, random sequences aggregate and are insoluble in water (Prijambada et al., 1996). Proteins that occur in cells have sequences that allow them to fold under physiological conditions. Due to a considerable energy gap between the native and any competing fold, natural proteins fold in a cooperative all-or-none transition. This energy gap is the key characteristic that distinguishes a natural protein from a random amino acid sequence (Finkelstein and Galzitskaya, 2004). Electrostatic interactions among amino acid residues that approach closely in space facilitate correct folding (Dobson, 2003; Lindorff-Larsen et al., 2005). An important feature of the folded three-dimensional protein structure is the coupling of amino acids distant on the linear sequence but in close proximity in space. Through this coupling local changes in the protein are propagated through the protein's structure, a process common in the transduction of signals (Luque et al., 2002).

The final folded protein has a characteristic compact shape. A precise spatial placement and orientation of specific amino acid residues gives rise to the broad spectrum of functions exhibited by proteins. The highly selective and efficient catalytic properties of enzymes, for instance, rely on the precise positioning of atom groups in the catalytic centre. Furthermore, the shape of the protein is also crucial for its capability of interacting with other molecules (Fischer, 1894; Friedrich, 1984). The spatial conformation of the protein's amino acid chain thus determines both the function and the interactions of the protein.

1.2. Networks of proteins

In the cell proteins are embedded in molecular networks. If two proteins have mutual compatible docking surfaces, they can interact directly. Proteins can also interact through exchange of molecular signals. Enzymes, i.e. catalytically active proteins, can form metabolic networks where the reaction product of one enzyme serves as substrate for a following enzyme. In this paper we focus on metabolic networks. Metabolic networks, in contrast to signalling networks, are connected through precise steric fit of enzymes and substrates and therefore considered more difficult to evolve (Kirschner and Gerhart, 1998). Our interest here is the robustness of metabolic networks to faults in its components, i.e. in the enzymes. The robustness of networks is frequently discussed from a solely topological perspective, for example, the connectedness of the network after random deletion of nodes is considered (Barabási and Oltvai, 2004). On the contrary, the approach taken here is concerned with dynamic properties of the metabolic network. To simulate the network dynamics we use an abstract artificial chemistry model (Dittrich et al., 2001) described in detail in Section 2.

1.3. Faulty proteins

Proteins are complex molecular machines the function of which typically hinges on positioning of atomic groups with Å-precision. This positioning is achieved through the self-organisation of folding. A modification of the amino acid sequence of a protein will affect the cooperative folding process, the spatial arrangement of atoms in the folded protein, and the conformational dynamics of the protein structure. Over the course of evolution the complex protein networks of cells are formed through diversification and specialisation of protein structures as a consequence of alterations in their amino acid sequences. Changes to the DNA will lead to inheritable modifications of protein sequences that form the basis of molecular evolution as well as genetic disease (Wang and Moult, 2001). Errors can also be introduced at later stages in the production of proteins. An erroneous transcription of DNA into mRNA, for instance, will give rise to a "production run" of faulty proteins over the lifetime of the incorrect mRNA molecule. If a correct mRNA is incorrectly translated then individual proteins with faults are produced. The cell employs an elaborate molecular machinery for the quality control of proteins (Netzer and Hartl, 1998) and eliminates proteins that cannot fold correctly (Dobson, 2003).

Altering an amino acid in a protein can affect the protein in three basic ways:

(1) It can derail the folding process, e.g. through steric obstruction or electrostatic repulsion or attraction.
(2) It can destabilise the folded protein, e.g. by eliminating hydrogen bonds or salt bridges or by disrupting the non-polar core area.
(3) It can eliminate a signal sequence pattern, e.g. a recognition signal for protein targeting or post-translational modification.

The first two mechanisms can impact ligand binding, catalytic activity, and regulation of the protein. Depending on the protein, the sequence position, and the amino acid substituted, the effect can range in severity