Physica A 388 (2009) 851–862

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Protein folding as an evolutionary process

Vivek Sharma [a], Ville R.I. Kaila [a], Arto Annila [a,b,c,*]

[a] *Institute of Biotechnology, University of Helsinki, Finland*
[b] *Department of Biosciences, University of Helsinki, Finland*
[c] *Department of Physics, University of Helsinki, Finland*

ARTICLE INFO

Article history:
Received 1 October 2008
Available online 7 December 2008

Keywords:
Dissipation
Evolution
Energy landscape
Entropy
Kinetics
Natural process
Natural selection
Self-assembly

ABSTRACT

Protein folding is often depicted as a motion along descending paths on a free energy landscape that results in a concurrent decrease in the conformational entropy of the polypeptide chain. However, to provide a description that is consistent with other natural processes, protein folding is formulated from the principle of increasing entropy. It then becomes evident that protein folding is an evolutionary process among many others. During the course of folding protein structural hierarchy builds up in succession by diminishing energy density gradients in the quest for a stationary state determined by surrounding density-in-energy. Evolution toward more probable states, eventually attaining the stationary state, naturally selects steeply ascending paths on the entropy landscape that correspond to steeply descending paths on the free energy landscape. The dissipative motion of the non-Euclidian manifold is non-deterministic by its nature which clarifies why it is so difficult to predict protein folding.

1. Introduction

Protein folding has attracted intellectual interest ever since 1968 when Cyrus Levinthal pointed out the conceptual difficulty [1]: Why do proteins fold much faster than it would appear to take them to find native states among all conceivable conformations? Today Levinthal's paradox is no longer regarded that much of a mystery [2]. In general, protein folding is pictured to direct down along paths on a free energy funnel landscape, rather than sampling conformational space randomly [3–7]. Conformational entropy of a polypeptide is viewed to decrease along with increasing entropy of the surrounding solvent [8,9] in agreement with the understanding that the *total* entropy must be increasing.

In this study protein folding is described as a *natural process* [10]. Natural processes are evolutionary courses that direct toward more probable states by dispersal of energy [11,12]. They follow the 2nd law of thermodynamics that was recently given as an equation of motion [13] and associated subsequently with the principle of least action [14]. The general formalism establishes correspondence between increasing entropy and decreasing free energy. It has recently given insight to many puzzling natural phenomena and resulting distributions [15–20]. The results are in agreement with earlier findings based on the maximum entropy principle [21–30]]. Thus, the description of protein folding by the ubiquitous imperative provided by this study is not new as such but brings protein folding within the general theory of evolution by natural selection [31].

Protein folding, from random conformational disorder to hierarchical structural order, when formulated properly, follows also the general principle of increasing entropy without an *ad hoc* exemption that entropy of the polypeptide chain would

* Corresponding address: Institute of Biotechnology, University of Helsinki, POB 65, FIN-00014 Espoo, Finland. Tel.: +358 9 191 50629; fax: +358 9 191 50639.
E-mail address: arto.annila@helsinki.fi (A. Annila).

0378-4371/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.physa.2008.12.004

be decreasing at the expense of increasing entropy in the surroundings. The new view is in contrast to the common but obscure interpretation that an increase in entropy would invariably coincide with increasing disorder. There is no compelling reason to associate high entropy only with high disorder because entropy $S = k_B \ln P$ is a mere logarithmic measure for the probability P of a state [32,33]. In the light of the 2nd law orderly structures are functional mechanisms of energy dispersal that have emerged during evolution to decrease free energy.

Considering protein folding by the 2nd law of thermodynamics is not only of intellectual interest, but sheds light on the practical folding problem and protein structure prediction. However, it is not claimed that structures could be predicted from amino acid sequences using thermodynamics. On the contrary, it will be clarified that the protein folding is a very hard problem because the evolution of a non-Euclidian energy landscape is non-deterministic and non-integrable [13,14].

2. Protein as a thermodynamic system

It is insightful to picture protein folding as a sequence of chemical reactions using statistical physics. Although covalent bonds, apart from disulphide linkages, do not form during folding, other interactions like hydrogen bonds, ion pairs and also weaker interactions are being established within the polypeptide as well as with and within the surrounding solvent contributing to entropy as hydrophobic effect [34,35]. According to the basic chemical thermodynamics maxim, chemical reactions proceed toward the maximum entropy state, which is equivalent to the minimum free energy state where the chemical potentials of substrates ($\sum \mu_k$) and the product (μ_j) are equal. Protein folding can be formulated in the same way. The simple description will clarify *why* proteins fold, but not *how* specific mechanistic steps lead to the native fold [36].

When a natural process evolves via chemical reactions, it is obvious that the reactants must actually interact with each other for the system to change its state. Energy flows in interactions. Thus, the principle of increasing entropy is a meaningful imperative for a system of interacting entities. This is in contrast with the common misconception that entropy would be a valid concept only for an isolated system accommodating non-interacting or weakly interacting constituents in equilibrium. For the system to change its state, at least a quantum must be expelled to or acquired from the surroundings [13,14]. Of course, one may imagine of isolating an evolving system in a closed space that would accommodate also photons that result from the evolutionary process. However, there is no such a thing as an empty space [37] without any energy, e.g. photons forming the space.

To apply the universal principle of increasing entropy it is immaterial how one wishes to label some energy densities as constituting the system and others as forming the surroundings. Irrespective of the choice the energy differences are diminishing by energy flows that direct down along gradients. Thus, regardless of the viewpoint, entropy of the system, just as entropy of its surroundings, is increasing when free energy is decreasing. The flows connect the system to its surroundings. Therefore the common postulate for the emergence of orderly structures, that the entropy of a system would decrease at the expense of increasing entropy of surroundings, does in fact violate conservation of energy.

For the statistical description of folding, the concept of identity, i.e., ability to distinguish is central [13,14]. Identities are distinguished in interactions. In other words, a hypothetical external observer cannot distinguish polypeptide conformations from each other without interacting with them. First a measurement that couples the observer to the system via a dissipative process [38,39] is able to distinguish one conformation from another. The identity of a functional moiety, e.g., a carboxylic group of an aspartate or a hydroxyl group of a serine, is distinguished in its interactions with other residues and other constituents of the system such as solvent molecules. The change in energy is the measure of distinguishing. In case there is no difference in interaction energy among two moieties, the two are statistically indistinguishable. Of course one may imagine two peptide configurations with identical energy to differ from each other, however, in practice it will turn out to be impossible to separate the inter-converting conformations from each other.

To describe folding as an evolutionary course in the simplest way using the previously described formalism [13,14], the constituents of the initial random coil in numbers N_1 are regarded essentially as indistinguishable substrates with a chemical potential [12] $\mu_1 = RT \ln[N_1 \exp(G_1/RT)]$. The Gibbs free energy G_1 is compared to the average energy RT per mole ($R = k_B N_A$) that includes all other interacting ingredients of the system, e.g., solvent and external fields. Under native conditions the unfolded protein is high in energy compared with its surroundings. Obviously, we realize that the random coil state does not quite make a degenerate pool of conformations but in fact spans a distribution of interaction energies. As long as the band of closely spaced sublevels is narrow in energy there is no significant bias for any particular conformation. Then the diverse random coil conformations distinguish poorly from each other in mutual interactions because their energies are nearly degenerate. This justifies the crude but illustrative approximation of nearly indistinguishable constituents. In this sense, the random coil polypeptide resembles a homopolymer.

Once some mutual interactions, stronger than RT, happen to form, the moieties will begin to distinguish from each other. A folding pathway is opening up in analogy to a reaction pathway opening up (Fig. 1). By pair-forming interactions the substrate pool of N_1 is used in syntheses of products N_2, which can be considered as nascent motifs. Their formation is a way, i.e., a mechanism to disperse energy from the initial random coil state to the surroundings that are lower in energy. Customarily the key residues of nascent motifs that initiate folding are revealed by the Φ-value analysis [40,41]. The chemical potential of products N_2 is $\mu_2 = RT \ln[N_2 \exp(G_2/RT)]$. The pool of N_2 may, in turn, act as the substrates to drive further assembly by energy dispersal to subsequent structural motifs. Differences in the chemical potentials drive the syntheses of larger and larger assemblies (Fig. 1). The number of interactions and their strengths are increasing during the evolutionary

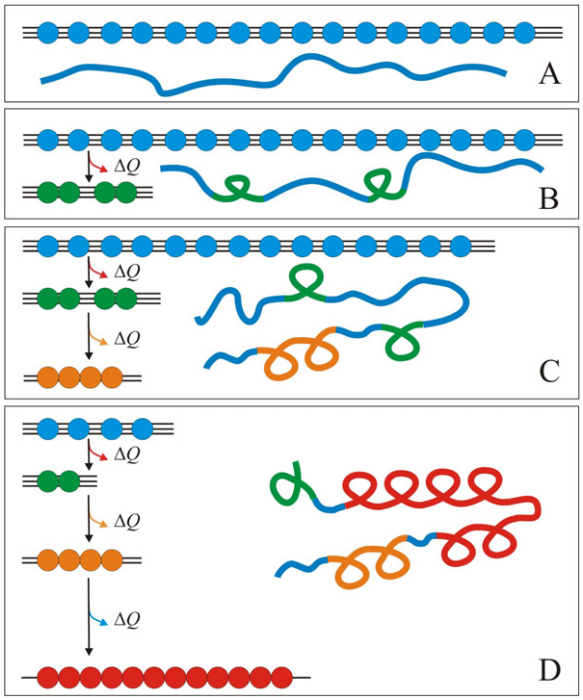

Fig. 1. Cartoon presentation of folding by changes in partitions on the energy level diagram. A: The initial random coil state comprises only weakly interacting moieties (blue) and is high in energy. B: When moieties interact with each other, quanta (ΔQ) are dissipated to the surroundings that are lower in energy density. Formation of nascent structures (green) is a mechanism of dissipation. C: Native-like structures (orange) grow in interactions to speed up the dissipation process. D: The quest for the equilibrium with the surroundings by dissipation results in a mature fold (red), a maximum entropy partition that has exhausted all free energy of folding. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

course which is consistent with the present understanding of the folding process [42]. Eventually, when no further means of energy dispersal by dissipation are found, the folding process comes to its end to the stationary state.

3. Entropy of a chemical system

A state along the folding pathway is associated with the general form of entropy which is the logarithm of the total probability $P = \prod P_j$ [13]

$$S = R\sum_{j=1} \ln P_j = \frac{1}{T}\sum_{j=1} N_j \left(\sum_k \mu_k + \Delta Q_{jk} - \mu_j + RT\right)$$
$$= R\sum_{j=1} N_j \left(\frac{A_j}{RT} + 1\right). \tag{1}$$

This equation sums up all interacting moieties, each is considered as a product or a substrate. The products are indexed by j and substrates by k together with co-products which enter in the summation (Σ_k) with an opposite sign. The free energy, in this case the difference in chemical potentials $\Sigma\mu_k + \Delta Q_{jk} - \mu_j$, including the surrounding energy density ΔQ_{jk} that couples to the reaction, experienced by j, is also known as the affinity A_j, sometimes also referred to as exergy. The free energy terms are thermodynamic driving forces of folding and evolution in general. In other words, entropy as a convenient statistical measure that sums up from the free energy and energy contained in the repository of N_j.

The probability P_j of an energy level G_j associates with its occupancy N_j and potential energy differences A_j/RT in respect to other interacting densities-in-energy [11] $\Phi_k = N_k \exp(G_k/RT)$. This is in contrast to the common understanding that entropy as $S = k_B \ln W$ stands for conformational entropy [43] that is counted from the number of conformational degrees of freedom W that do not necessarily associate with distinct energies. According to Eq. (1) the conformations N_k that have identical energy may disperse on an isoenergy surface by relative phases but altogether add up to the probability P_k of the state. As pointed above, the random coil conformations do not clearly distinguish from each other due their nearly degenerate energy levels. Thus the conformational space measured by energetically distinct states given by Eq. (1) is not particularly large. Of course the isergonic conformational disorder can still be large but it does not contribute to entropy.

4. Dispersal of energy

During the probable course of folding, entropy increases at the rate [13]

$$\frac{\mathrm{d}S}{\mathrm{d}t} = \frac{1}{T}\sum_{j=1}\frac{\mathrm{d}N_j}{\mathrm{d}t}\left(\sum_k \mu_k + \Delta Q_{jk} - \mu_j\right) = \frac{1}{T}\sum_{j=1} v_j A_j \geq 0 \tag{2}$$

when the differences in chemical potentials, i.e., A_j are diminished by interactions that form and assemble diverse moieties k to structural elements j at the rates $v_j = \mathrm{d}N_j/\mathrm{d}t$. During folding energy flows from a high-energy repository μ_k to another lower-energy repository μ_j. The chemical reactions are non-conserved transformations of identities because photons are emitted. Thus the energy content of the polypeptide is decreasing but for a sufficiently statistic system the average energy RT is a meaningful concept at each moment of time during the evolutionary course [44].

When the polypeptide system evolves toward stationary state by exergonic reactions, potential energy differences diminish, i.e., free energy is consumed when quanta are dissipated (e.g. as heat) from the open system to its surroundings. If the dissipation is prevented, there is no progress. When the system moves toward a non-equilibrium stationary state by endergonic jk-reactions, the process is coupled to an influx of external energy, ΔQ_{jk} which is included in the A_j term to describe, e.g., actively driven refolding by chaperons [45]. In all cases, the evolution eventually reaches the free energy minimum corresponding to the maximum entropy state. In this stationary state of dynamic equilibrium no more interactions will mount or dismount on the average. Then the maximum entropy state $S = R\Sigma N_j$ contains a sum of the steady-state populations of N_j, corresponding to the native state partition of conformations.

The Eq. (2) has been ascribed earlier to the entropy expelled *from* the producing system to its surroundings to maintain a presumed low-entropy state of an orderly (living) system [10]. However, when $\mathrm{d}S/\mathrm{d}t$ is derived from the first principle probability calculation, it clearly means that entropy of any system, irrespective how one chooses its entities, increases during the evolutionary course toward free energy minimum [13,14]. There is no need, for example, to account for the existence of orderly structures, by dividing the total entropy to the entropy exchanged between the system and its environment, and to the entropy produced by irreversible processes within the system. When $A_j > 0$, energy flows ($v_j > 0$) into the system; and when $A_j < 0$, energy flows ($v_j < 0$) out of the system. In both situations, the entropy of the system, just as its entropy of its surroundings, increases ($\mathrm{d}S > 0$) by diminishing the potential energy differences $A_j \neq 0$ or in general thermodynamic gradients $\nabla\mu_j/RT \neq 0$ until a stationary state, where $\mathrm{d}S = 0$, has been reached. The entropy increase by dispersal of energy is the essence of the 2nd law.

Mathematically speaking Eq. (2) describes a time-dependent tangential vector field [46], i.e., an evolving manifold of potential energy differences A_j that diminish by flows v_j down along the gradients. When a protein folds, the evolution naturally selects mechanisms, associated with diverse constituents j, that allow the system to advance along steepest gradients $(\mathrm{d}S/\mathrm{d}N_j)(\mathrm{d}N_j/\mathrm{d}t) > 0$ to abolish $A_j \neq 0$ most rapidly. The free energy provides the "evolutionary pressure" in its biological meaning. Flows of energy naturally select [13,14] among diverse mechanisms of energy dispersal those that allow to reach the stationary state most rapidly. In other words, abiotic molecular species, just as biotic cellular species are considered as mechanisms of energy transduction that 'compete' for common sources of free energy. The orderly structures are not improbable, i.e., low in entropy when they function to consume available free energy.

Energy is dispersed via interactions. When structures j form from moieties k, the rate is proportional to the thermodynamic driving force A_j/RT [13])

$$v_j = \frac{\mathrm{d}N_j}{\mathrm{d}t} = r_j\frac{A_j}{RT} \tag{3}$$

to satisfy continuity. The coefficient r_j depends on the particular mechanisms, e.g., autocatalysis, co-operative and chaperon assisted folding. Different mechanisms direct to alternative folding trajectories all consuming free energy that they may access. During folding various paths are explored and those, where the energy landscape descends steeply down, providing high rates of entropy increase, will be taken with high probabilities [14].

Customarily, a reaction is modeled by the law of mass action [47] where concentrations are multiplied by distinct forward and backward rates that may vary during the course. However, it is not the concentrations but the potential energy differences of reactants, according to Eq. (3), that drive the reaction toward the stationary state. By the evolutionary description new flows are opening up and old ones are closing down when the energy landscape is changing. A particular flow v_j via a particular mechanism, denoted to by r_j, is changing when A_j is changing.

In addition to the driving free energy, the flow is, of course, affected by changes in the mechanistic capacity. A particular reaction mechanism itself is a result from an earlier evolutionary process. It may continue to evolve, i.e., change during an autocatalytic reaction, to affect v_j. The kinetics due to varying mechanistic capacities is intricate but it still follows from thermodynamics. In other words, the energy flows are downhill, not crossing barriers, but restricted by various capacities.

Although Eq. (3) appears linear, it describes folding with an overall sigmoidal behavior which is the typical indication of a cooperative process. During the evolutionary succession the energy landscape evolves so that nascent motifs facilitate the rise of secondary structures that, in turn, promote assembly to a tertiary structure. The overall course is sigmoidal. When a flow v_j begins to exhaust a particular source of free energy A_j, via a particular mechanism r_j, the overall dissipation turns to use other sources of free energy via other mechanisms until all driving forces have been consumed. These mechanistic changes contribute as well to the nonlinearity of the overall dissipation process.

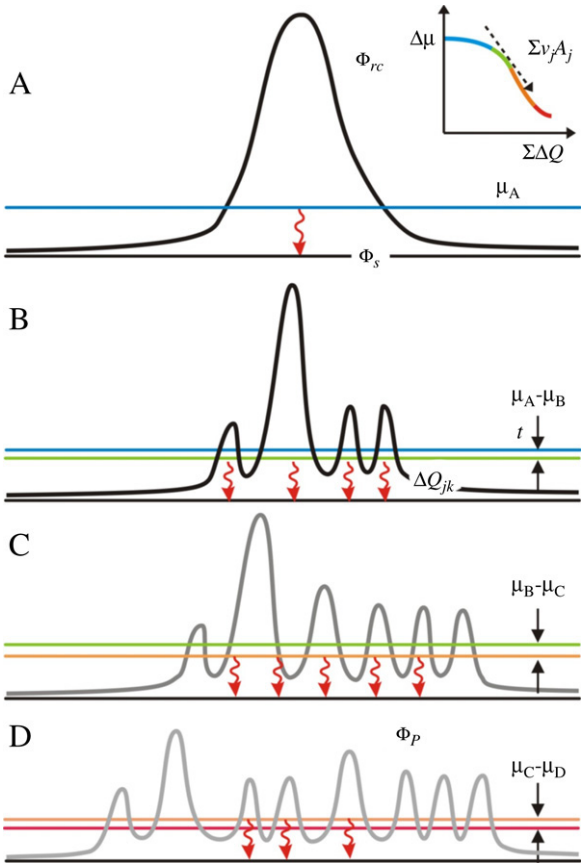

Fig. 2. Diverse densities-in-energy along the folding pathway, the random coil conformation Φ_{rc}, folded protein Φ_p and solvent Φ_s, compose an energy landscape along the (reaction) coordinate. Protein folding is a series of dissipative (wave arrows) transformations from A $\rightarrow$ D that diminish diverse potential differences $\Delta\mu = \mu_k - \mu_j$, i.e., the free energy until the stationary state is attained. As a result the energy landscape flattens and disperses. The flows that transform Φ_k to Φ_j, take the deepest descents, given by tangent vectors $\Sigma v_j A_j$ corresponding the shortest paths on the free energy landscape (inset). Since densities-in-energy are in relation to each other via interactions, the transforming flows keep redirecting on the changing landscape. This leaves no invariants of motion to predict precisely folding trajectories.

5. The equation of evolution

The general equation of motion for dissipative processes in terms of probability is obtained from Eq. (2) using the definition $S = R \ln P$ and differentiating it, $\mathrm{d}S/\mathrm{d}t = [R/P](\mathrm{d}P/\mathrm{d}t)$. This gives [13]

$$\frac{\mathrm{d}P}{\mathrm{d}t} = LP \geq 0, \qquad L \equiv \sum_{j=1} v_j \frac{A_j}{RT} \tag{4}$$

where L denotes the tangential vector field, i.e., the energy landscape [46,14]. The manifold is non-Euclidian. Since the reaction coordinate along the free energy landscape is asymmetric, i.e., A_j carries a sign, the distance in free energy is not a proper distance but should be accurately referred to as divergence as is in the context of information theory [44]. Furthermore, entropy does not have all the topological properties of a distance. The triangle inequality is not satisfied. For example, the potential energy difference between two moieties is affected when a third one comes within interaction range. Then a distance between two repositories of energy k and j may even become shorter via the third density-in-energy k' and energy flows from k to j via k', the catalyst.

The seemingly simple equation of motion (Eq. (4)) is surprisingly insightful. For the first, it lacks an analytical solution because the driving forces keep changing during the evolution leaving no invariants of motion to find a solution by transformation. Therefore, per definition, the course of folding is non-deterministic and trajectories are unpredictable in detail. The landscape is not preset (Fig. 2) but the manifold 'unfolds' during the evolutionary process. For example, the free energy barriers that are customarily viewed as kinetic bottlenecks of folding, are not intact obstacles but channels for energy flows open up during the folding while others narrow down due to dissipation to prevent backtracking toward unfolding. We emphasize that there is nothing moving on the landscape or crossing the barriers but it is the landscape itself which is the system and its surroundings that keeps molding during folding.

The natural process is also sensitive to the initial state, i.e., evolution is per definition chaotic [48]. Thus, there is no guarantee that all trajectories will wind up to a narrow native distribution, but the maximum entropy partition (population) may house also partially folded, completely unfolded and even misfolded proteins. The resulting ensemble is also referred to as conformational sub-states [49] that may differ, e.g., by their catalytic properties r_j to result in dynamic heterogeneity [50].

When protein folding is understood as a dissipative process, the dispersal of energy from the open system to its surroundings (or vice versa) is the reason for the directional, irreversible course. Typically, a folded protein at the equilibrium is stabilized compared to its denaturated states only by a free energy 5–15 kcal/mol [40]. Thus the folding-associated dissipation is minuscule, e.g., in comparison to gigantic flows of energy that native proteins generate as parts of present-day organisms' energy transduction machineries. Nevertheless, the principle is the same and provides understanding to the evolution of proteins. Primordial polypeptides folded, just as contemporary proteins do, to attain thermodynamic equilibrium with their surroundings. The ubiquitous imperative to level differences in energy drives matter in functional structures that disperse energy. Machinery of life has high energy content because its surrounding is high in energy (e.g. solar radiation). Thus dissipative structures maintain not low but high entropy by dissipation of free energy.

Thermodynamic systems attempt to attain stationary states with respect to their surroundings by minimizing free energy. This is in accordance with Le Chatelier's principle [12,51] that the equilibrium depends on conditions. First at the stationary state $\mathrm{d}P/\mathrm{d}t = 0$, the system is conserved. The steady-state motions are fluctuations, e.g. protein dynamics, due to sporadic influx and efflux or cyclic motions in a phase space but there is no net flow of energy to the system from the surroundings or vice versa hence no evolution either.

The evolving energy landscape is described as probabilities that keep changing during folding. Boltzmann's factors are not preset, but they will change during the course of evolution toward the stationary state values that emerge when the flows have equalized differences among diverse potentials. It makes sense to talk about the potential energy differences rather than absolute values themselves because the driving forces are relative to one and other. Obviously fluxes from a small open system to an overwhelmingly dominating surrounding energy density, do not practically change the surroundings at all. Then also the average energy RT remains practically constant during folding.

The flattening non-Euclidian manifold composed of energy densities is the general way to describe the non-integrable evolution without invariants of motion [13]. For any non-trivial system, it is difficult to foresee all possible potentials and their conversions to all imaginable mechanisms that may, in turn, tap into new conceivable potentials. Owing to this non-deterministic and irreversible nature of evolution, also protein folding as a non-unitary dynamics remains a difficult problem.

6. Folding kinetics

Customarily it is assumed that there is a kinetic course along the folding funnel which progressively leads to energetically more stable states [40]. When the course is viewed as crossing high-energy barriers, kinetics and thermodynamics are not formally linked together. However, the $\mathrm{d}S/\mathrm{d}t$ equation (Eq. (2)) and its associated flow equation (Eq. (3)) describe kinetic courses as motion of the energy manifold down along gradients. Narrow passages are bottlenecks, not barriers. Importantly, due to the dissipation the thermodynamic gradients change during the course of folding. New paths open up while others close down. Although the equation of evolution (Eq. (4)) cannot be solved analytically, a kinetic course can be studied numerically or an ensemble of trajectories can be simulated.

Under native conditions the initial random coil state is high in energy, but it has no particularly efficient mechanisms to conduct energy to its surroundings by its transient and weak interactions. Once some stronger interactions, e.g., a nascent helix or a hydrophobic cluster forms, due to random conformational variations, the rate of further dissipation coupled to the growth of structures increases. The formation of a new structure that provides a new mechanism for dissipation, a path for energy to flow, appears initially as a bottleneck (Figs. 1 and 2). Its appearance is customarily pictured as a formation of a high-energy transition state. Such a description of motion across barriers regards kinetics as independent from thermodynamics and makes it difficult to understand the process. It is emphasized that flows of energy are down along the gradients but the paths must first open up and they may remain as narrow bottlenecks.

The folding curve has an overall sigmoidal shape (Fig. 3) closely resembling a logistic curve [52] familiar from many growth processes. Consistently with our results, interactions have previously been found to accumulate in a sigmoidal manner [53]. The sigmoidal form is also familiar from simple phenomenological statistical models for folding that are parameterized for initiation and propagation [54]. The form is also recognized from many ecological models [55]. Oscillatory behavior that is familiar from the Lotka–Volterra equations, may also appear during protein folding [56]. This "strange kinetics" [57], analogous to the Belousov–Zhabotinsky reaction [10], is accommodated in the flow Eq. (3), when rapid mechanisms of energy transduction emerge and momentarily over-deplete sources and subsequently must decay.

The sigmoidal kinetics can be understood from structural considerations. For example, when the first few hydrogen bonds form to make a nascent helix, further dissipation by bonding becomes more probable. This punctuation is quantified by an increase of S_j, or equivalently by a decrease of G_j, which further increases the rate of folding. The dissipative structures function to increase entropy. They are favored and appear in a co-operative or autocatalytic manner. The emergence of a new mechanism of dissipation goes hand in hand with the rise of structural hierarchy. However, the growth begins to slow down and level to a stasis as soon as the helix grows longer and its chemical potential becomes comparable to those of remaining random coil segments. The dissipation rises again when a new mechanism, e.g., when two helices zip together, and again

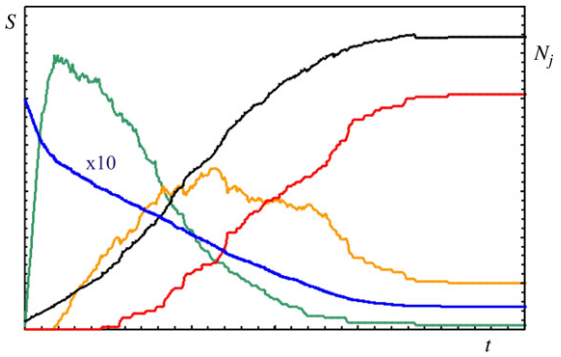

Fig. 3. Time t course of simulated folding. The initial random coil state comprising N_1 residues (blue, right scale) turned to a folding pathway when dissipative interactions formed nascent motifs $N_{j>1}$ (green) (see Fig. 1). Concurrently entropy S (black) begins to increase. As soon as the nascent motifs evolved to more efficient mechanisms, i.e., 'secondary' structures (orange), dissipation increased. The course leveled off to the 'tertiary' structure $N_{j\gg1}$ (red) when interactions resulted in no more dissipation. The folded-state equilibrium partition houses structural diversity that still includes the simplest moieties (blue) which can be considered to resemble 'flexible tails', nascent motifs (green) such as 'loops' and 'secondary' structures (orange) or eventually domains. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

when several helices pack on each other to dissipate energy from side chain interactions and hydrophobic effects. Formation of new structures provides new means to access free energy and devour it. Thus the structure-activity relationship is also valid in protein folding. The general condition of integration to more effective energy transducers is $\mathrm{d}S > 0$. Eventually when no new means of dispersal are found the system settles to the stasis of a folded partition. Punctuations and stasis identify protein folding to the general characteristics of evolutionary courses [58].

Folding was simulated as an evolutionary process according to the flow equation (Eq. (3)). The simulation served to demonstrate the concept, but it was not intended to model any particular protein. Interactions were described on a formal level, i.e., as transitions in an energy level diagram, and not by computing Gibbs free energies G_j explicitly, e.g., from molecular dynamics simulations parameterized for particular potentials. In the simulations moieties k interacted randomly which led to folding or unfolding of structures j depending on the gradient $A_j = \Sigma\mu_k - \mu_j$ according to Eq. (3). The decreasing free energy directed the evolutionary course. Thus the 'memory' of the past trajectory was not modeled in, e.g., as a Markovian process [59,41]. The probabilities P_j kept changing, in a Bayesian manner, during the process and converge to a stationary state that is described by the Boltzmann factors.

Increasingly larger rate constants r_j were assigned for increasingly larger structures j, to model that larger structures have more interaction sites. In addition the rates r_j were randomly varied (up to 10%) to mimic fluctuations in energy densities. The particular values assigned for the variables N_j, G_j and r_j were important for the outcome, however in all cases those particular entities that provided the highest rates $(\mathrm{d}S/\mathrm{d}N_j)(\mathrm{d}N_j/\mathrm{d}t) > 0$ along the steepest gradients were naturally selected. The emerging functional order, i.e. dissipative structures served to abolish the gradients, the thermodynamic bias, that underlies irreversibility.

The folding simulations were launched from the random coil state housing moieties N_1 associated with chemical potential μ_1. A step of random syntheses was executed by allowing any two moieties to form N_2. Chemical potentials were updated and next step of random syntheses and degradations was performed. The surroundings was assumed to absorb all dissipative quanta without a marked change in its energy density. At each step diverse pairwise interactions led to integration or disintegration depending on A_j. As long as $A_j > 0$, the process continued and advanced toward larger assemblies.

The simulation was cursory but it captured the characteristic features of folding. Entropy and the number of interactions were found to increase in a sigmoidal manner. Initially process was slow because the first mechanisms of folding r_j were not efficient in dissipation. Later, more powerful mechanisms integrated from the primitive mechanisms and were able to access more efficiently larger potentials. The evolution leveled off to a dynamic stationary state, when no new mechanism appeared to exhaust new potentials. Then random integration and disintegration resulted in no net dissipation. The maximum entropy state was a partition dominated by the large folded assembly, but it still housed some random coil segments and nascent motifs. This is in accordance with the statistical picture of an ensemble of conformations.

7. The folded-state partition

Once the critical developmental phases of folding have terminated to the maturity [60], the native state is stable for fluctuations in interactions as long as the surrounding densities-in-energy remain steady. When using mathematical methods of nonlinear dynamics [48], entropy as the Lyapunov functional reveals, that the maximum entropy state is robust against internal fluctuations δN_j in the number of various structural motifs. The same concept accounts also for ecosystem stability. A partial unfolding is a transient improbable event because it would require a substantial input of energy. The particular moieties that gave rise to most dissipation are mostly conserved residues in regular structures which form high-energy interaction network. Non-conserved residues in loops and tails contributed less to dissipation because they formed only few interactions. Also, some nascent motifs, the early species of succession that paved the folding pathway,

disintegrated during later phases of folding when secondary and tertiary species in succession took over (Fig. 3). This general characteristic of evolution is in agreement with experimental Φ-value analyses [40].

The stationary state partition is obtained from Eq. (2) by the condition

$$\mathrm{d}S = 0 \Leftrightarrow \sum_k \mu_k + \Delta Q_{jk} - \mu_j = 0 \Leftrightarrow N_j = \prod_k N_k \mathrm{e}^{-(\Delta G_{jk} - \Delta Q_{jk})/RT}. \tag{5}$$

The equilibrium partition, devoid of free energy, is obtained directly from Eq. (2) without introducing Lagrange's multipliers *ad hoc*. Under native conditions the entropy landscape is convex about the most stable native state against perturbations in the structural elements (δN_j) according to the Lyapunov stability criterion [10,48], i.e., $\mathrm{d}S(\delta N_j) < 0$ and $\mathrm{d}S(\delta N_j)/\mathrm{d}t > 0$. A folded protein will resist unfolding. This is customarily described as the heat capacity, defined as $C = T(\mathrm{d}S/\mathrm{d}T)$. Upon differentiation S (Eq. (1)) with respect to T and multiplication by T, one obtains $C = \sum(\mathrm{d}N_j/\mathrm{d}T)(A_j/T) + \sum N_j \Delta G_{jk}/T$. The first term corresponds to the changes $\mathrm{d}N_j$ in populations N_j to accommodate the change $\mathrm{d}T$ in the average energy imposed by the influx of energy. The second term expresses the reservoir of interaction energies $\Delta G_{jk} = \sum G_k - G_j$ within the populations [61,62]. When this term is large, a substantial influx of energy will first be absorbed to raise the energy density of the polypeptide toward that of the denaturating surroundings.

The kinetic course may also lead to a partition that contains partially or misfolded structures. Folding may also end up with a state that turns out to be metastable later, i.e., susceptible to further changes when new dissipation mechanisms and potential differences appear due to a change in the surroundings. For example, an appearance of a misfolded partner may impose an additional thermodynamic force and provide mechanisms of dissipation that drive and induce further conformational changes. This may lead to a dissipative dimerization and subsequent aggregation to filaments that are familiar from prion diseases [63].

8. Unfolding kinetics

Protein unfolding, just as the folding, is described by Eq. (2), but the surrounding conditions (A_j) that drive denaturation are different from native conditions that drive folding. Reversible folding [64] implies that the course can be reversed by changing the conditions. This means, in terms of the formalism considered here, simply changing A_j for folding and unfolding. However, for many proteins it may not be possible to reverse the course simply by changing the surrounding solvent densities-in-energy because other densities-in-energy present *in vivo* might be missing or not introduced in a proper sequence to drive refolding.

Let us picture that the aforementioned helical bundle is subject to an instantaneous change in conditions for unfolding, e.g., due to a denaturant pulse (Fig. 4). Consequently the helical bundle system that was in the equilibrium ($G_{\min}$, $S_{\max}$), suddenly experiences a large potential gradient. According to Eq. (1) entropy that used to be high, is all of a sudden low. The change in conditions causes the Gibbs free energy to increase drastically. In the denaturing conditions the helical bundle is simply high in energy relative to individual helices and solvent. Thus it is improbable.

As a response to the reversed gradients, the helical bundle system evolves toward more probable states by structural disintegration. Helices break loose from the bundle one after another in the quest to lower free energy. When all interactions among helices have vanished, the helical bundle system has disassembled, and per definition entropy associated with it has lost meaning. This leads to a system lower down in hierarchy. If the helices in this partially unfolded system are still high in energy relative to the random coil, they continue to fragment and deteriorate. Cold denaturation [40,65] can also be understood to result from changes in the relative densities-of-energy of protein and solution.

The overall course of the unfolding is also a sigmoid. The initial break down is limited by the mechanistic capacity. For example, in the helical bundle there are not that many interfaces accessible to denaturant molecules. When the bundle opens up, the rate of unfolding will increase. Finally, near the random coil state the thermodynamic driving forces are nearly exhausted and the process is again slow.

9. Entropy enigma

The entropy law, by the up-to-date words of Georgescu-Roegen, *is still surrounded by many conceptual difficulties and equally numerous controversies* [66]. It seems that much of the confusion arise when increasing entropy is erroneously equated with increasing disorder [67]. Here it is emphasized that the entropy increase results from the free energy decrease, whereas an increasing disorder results from isergonic phase dispersal, i.e., loss of coherence in motions.

The 2nd law, following Carnot, is simple: an energy difference is a motive force [68]. The overall entropy of the system is given by Eq. (1), its rate of change depends on the energy differences according to Eq. (2) and the rates by which interactions form is given by Eq. (3). The self-similar thermodynamic description Eq. (2) clarifies the emergence of nested hierarchical organization 'systems within systems' [69], in this case protein structural organization. The Gibbs free energy of the system composed of structural moieties indexed with j can be written as,

$$G = \sum_{j=1} G_j = \sum_{j=1} \left(H_j - TS_j\right) = \sum_{j=1} \left(U_j - N_j A_j\right) \tag{6}$$

where the definition for entropy (Eq. (1)) has been employed. The enthalpy $H_j = U_j + N_j RT$ sums up internal energy U_j and mutual interactions among N_j.

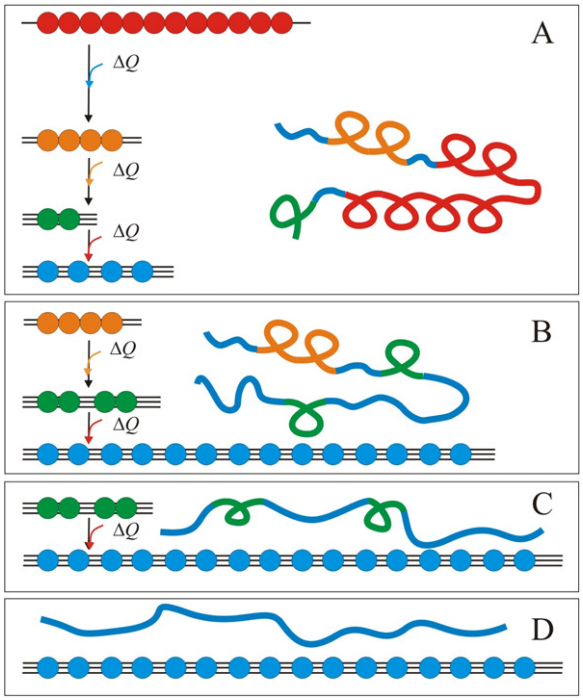

Fig. 4. Cartoon presentation of unfolding by changes in partitions on the Gibbs free energy level diagrams. A: The native fold (red) due to its high energy content in the denaturing condition experiences large gradients. It is forced to absorb by disintegrating to lower energy partitions. B: Also substructures (orange) face gradients albeit not as high and thus they continue to degrade. C: Rudimentary structural elements (green) gradually melt away by accepting last quanta of energy to complete unfolding. D: Finally, the random coil state comprises only weakly interacting residues (blue). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

The equation (6) explicitly states that an increase in entropy means a decrease in free energy by dispersal of energy. If $A_j > 0$, N_j is below its equilibrium partition, hence G_j is low. In contrast, if $A_j < 0$, N_j is above its equilibrium partition, hence G_j is high. The difference in the chemical potential A_j sums up the interaction fields which N_j experiences with other constituents N_k in the system, e.g., those that give rise to the hydrophobic effect. The interaction field among N_j themselves, given by the N_jRT-term in TS_j, is included also in enthalpy H_j. However, among indistinguishable N_j, the interactions are equal and there is no potential difference to drain and drive motion. Therefore, the N_jRT-term does not affect G_j which is the relative 'pricing' among the constituents of thermodynamic system. At the equilibrium $S_{\max}$ and the minimum free energy G_j equals internal energy U_j. Then the thermodynamic 'prices' G_j are 'right' for all entities. It is emphasized that during evolution the system is changing in its energy content. Therefore the average energy RT is changing and so are also other thermodynamic quantities but their mutual relations are valid at any given time for a sufficiently statistic system.

The dissipated heat of folding and the absorbed energy of unfolding can be obtained from the $\mathrm{d}S/\mathrm{d}t$ equation (Eq. (2)) via integration, to give the familiar result $\Delta Q = \int C\mathrm{d}T = T\Delta S$ due to Carnot. These relations emphasize that entropy (Eq. (1)) that was derived from the first principle probability calculation [13], yields the basic thermodynamics properties. According to thermodynamics, entropy neither denotes configurational degrees of freedom nor does it identify with the width of the free energy funnel. Instead, increase of entropy results from the probable motion due to dispersal of energy (Fig. 5).

10. Discussion

The view of protein folding as a dissipative process may at first appear as a naïve and superficial description of the complex and intricate phenomenon. In a sense it is. The statistical description by its nature, sums up the complexity of numerous microscopic interaction fields, however, not by approximations but by abstraction. The equation that links increasing entropy with decreasing free energy contains implicitly every quantum of energy, also in units of matter, and all interactions. The flows of energy channel along the steepest descents equivalent to the shortest paths in energy to reach the stationary state most rapidly. The dispersal of energy down along gradients is the bias that governs irreversible process in general and specifically guides folding. The universal criterion of natural selection directs energy flows along the steepest descents. It funnels folding of contemporary proteins and it has also selected polypeptides for fast folding during the eons of global evolution.

The holistic view of nature provided by the 2nd law of thermodynamics reminds us from the obvious: all we seen in proteins cannot be understood from proteins alone but from the activities they participate in. Such an ambitious overall

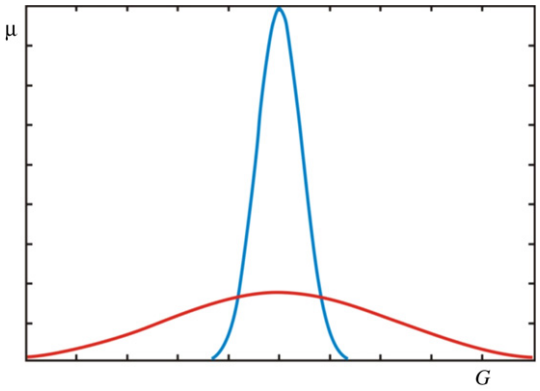

Fig. 5. The dispersal of energy densities, i.e., potential energy μ, among energetically distinct conformations G are schematically illustrated for random coil (blue) and folded (red) states under native conditions. The area under the random coil distribution equals that of the folded distribution and the energy dissipated to the surroundings as heat ΔQ consistently with the conservation of energy, i.e. the 1st law of thermodynamics. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

modeling of an organisms or even an entire ecosystem would be possible within the outlined thermodynamic framework but obviously it would be an extremely tedious mission when elaborated into details.

Admittedly, our examples for folding and unfolding of a helical bundle are cursory. However, numerous terms in the entropy sum over diverse interactions can be included in the statistical formalism. Emergence of interaction fields along folding pathways can also be seen as increasing dispersion of NMR lines along the frequency axis, i.e., on the energy scale. The random coil conformations give rise to a narrow band of overlapping broad signals corresponding to nearly indistinguishable conformations whereas the native protein displays a well-dispersed spectrum with distinct lines that identify to numerous distinguishable coordinate loci in a native protein. Also our abilities to distinguish those lines from each other have improved during the past few decades due to increased energy resolution in dissipative detection processes [70,71]. For example, conformational heterogeneity of folded proteins is nowadays routinely detected.

The 2nd law of thermodynamics given as a differential [13] and integral [14] equation of motion has been applied to describe other puzzling biological phenomena, e.g., emergence of life [20], chirality consensus and other standards of life [16], evolution of eukaryote genomes to diversity [17], skewed distribution of genes and population of species [15] and their cumulative sigmoid curves [19] as well as global evolution and homeostasis [18]. Many biological processes, unlike protein folding *in vitro*, are driven by *high* surrounding densities-in-energy such as sunlight and food. In these cases the Gibbs free energy includes external fields, e.g., radiation, flows of matter or electric potential gradients that couple to the dissipative processes. These surrounding fields drive the system toward non-equilibrium stationary states where the net dissipation ends. Accordingly, when the vital external fields disappear, the metastable states will collapse by disintegration toward the equilibrium. Therefore structural order is not improbable, when it associates with a functional mechanism that disperses energy from high-energy sources. The 2nd Law states the universal tendency to equate differences between system's energy densities and between the system and its surroundings irrespective of the system being high or low in energy density relative to its surroundings. Specifically, the ubiquitous imperative does not speak about order or disorder as being a factor that would be directing natural processes. Disorder, e.g., as phase decoherence may increase within a stationary system due to sporadic exchange of quanta with incoherent surroundings.

The old harmony about entropy being the mere probability measure for a system sliding down along the potential energy gradients may appear bizarre to the contemporary consent about entropy containing the dichotomy between structure and randomness. According to the statistical physics, probabilities relate to energy, not merely to the number of permutations without acknowledging energetic 'costs'. In other words, an efflux of energy turns an improbable non-equilibrium state to a probable equilibrium, whereas an influx of energy turns an improbable equilibrium to a probable non-equilibrium state. The 2nd law of thermodynamics works both ways. Entropy, as a statistical measure, increases without exemptions.

The non-deterministic equation of motion for the probability and the associated flow equations are not intended to replace Newtonian molecular mechanics and its many powerful variants in folding simulations [72,73], but to draw attention to the dissipative nature of the folding process. The thermodynamic argumentation of folding by dissipation is in agreement with calorimetric data; a protein 'cools down' when it folds. The established connection between entropy and free energy reveals that the dissipative motion of the entropy landscape is not different from the one pictured to take place on the free energy landscape. However, it is important to realize that the non-Euclidian landscape is not predetermined but it is forming and deforming during the folding. Hence the prediction of folding, in addition to technical problems, contains principal difficulties associated with the lack invariants of motion. The non-integrable nature of protein folding and its definite solution as the free energy minimum are in agreement with the folding problem being non-deterministic polynomial time (NP) complete [74]. Driving forces change with flows that in turn affect the forces hence *Ceteris paribus* does not hold [75].

This study was carried out to emphasize that the protein folding follows from the same principle as all evolutionary processes in nature. The consilience brought about by the general principle is inspiring. Studies of protein folding contribute

to the understanding of evolution in general, just as studies on other evolutionary processes contribute to the understanding of folding specifically. Protein folding is a sequence of succession, just as other natural processes that lead to integrated structural and functional hierarchies. Protein folding as a molecular process displays already an intricate interplay between internal processes and surrounding conditions that is characteristic of evolutionary processes in general. Also the concepts of nonlinear mathematics such as bifurcations, catastrophes and motions on manifolds as well as fixed points and separatrices, which are widely used to describe evolution, may be insightful also when giving an account on protein folding and unfolding.

Although it is in general acknowledged that evolution is blind, we tend to inadvertently deem many functions as 'meaningful' and to contemplate their 'incentives'. The unorthodox intentional viewpoint is often surprisingly insightful [76] – for a good reason – when we realize that all activities have the same motive, to increase entropy by abolishing free energy [77]. This may be hard to judge for each and every activity but consequences, in the form of distributions, are easy to examine. For example, protein length distributions from diverse organisms are skewed [78], just as population distributions of species [79], resembling closely log-normal distributions that accumulate in a sigmoidal manner. Thus, evolutionary processes not only select for native proteins on the basis of phenotype of an individual organism, but the protein folding phenomenon itself is an evolutionary process. Natural selection does not only work on genetic material but on all matter.

Acknowledgments

We thank Ilja Honkonen, Päivi Ruokola and Peter Würtz for enlightening discussions and valuable comments. V.R.I.K is supported by the Finnish Cultural Foundation and the Graduate School of Biotechnology and Molecular Biology.

References

[1] C. Levinthal, Are there pathways for protein folding? J. Chim. Phys. 65 (1968) 44–45.
[2] M. Karplus, The Levinthal paradox: Yesterday and today, Fold. Des. 2 (1997) S69–75.
[3] J.N. Onuchic, P.G. Wolynes, Z. Luthey-Schulten, N.D. Socci, Toward an outline of the topography of a realistic protein-folding funnel, Proc. Natl. Acad. Sci. USA 92 (1995) 3626–3630.
[4] K.A. Dill, H.S. Chan, From Levinthal to pathways to funnels, Nature Struct., Mol. Biol. 4 (1997) 10–19.
[5] A.R. Dinner, A. Sali, L.J. Smith, C.M. Dobson, M. Karplus, Understanding protein folding via free-energy surfaces from theory and experiment, Trends Biochem. Sci. 25 (2000) 331–339.
[6] V. Daggett, A. Fersht, The present view of the mechanism of protein folding, Nature Rev. Mol. Cell Biol. 4 (2003) 497–502.
[7] J.D. Bryngelson, J.N. Onuchic, N.D. Socci, P.G. Wolynes, Funnels, pathways, and the energy landscape of protein folding: A synthesis, Proteins: Structure, Function, and Genetics 21 (2004) 167–195.
[8] K.A. Dill, Dominant forces in protein folding, Biochemistry 29 (1990) 7133–7155.
[9] K.A. Dill, S.B. Ozkan, T.R. Weikl, J.D. Chodera, V.A. Voelz, The protein folding problem: When will it be solved? Curr. Opin. Struct. Biol. 17 (2007) 342–346.
[10] D. Kondepudi, I. Prigogine, Modern Thermodynamics, Wiley, New York, 1998.
[11] J.W. Gibbs, The Scientific Papers of J. Willard Gibbs, Ox Bow Press, Connecticut, 1993–1994.
[12] P.W. Atkins, J. de Paula, Physical Chemistry, 8th ed., Oxford University Press, New York, 2006.
[13] V. Sharma, A. Annila, Natural process – Natural selection, Biophys. Chem. 127 (2007) 123–128.
[14] V.R.I. Kaila, A. Annila, Natural selection for least action, Proc. R. Soc. A. (2008) doi:10.1098/rspa.2008.0178.
[15] T. Grönholm, A. Annila, Natural distribution, Math. Biosci. 210 (2007) 659–667.
[16] S. Jaakkola, V. Sharma, A. Annila, Cause of chirality consensus, Curr. Chem. Biol. 2 (2008) 53–58.
[17] S. Jaakkola, S. El-Showk, A. Annila, The driving force behind genomic diversity, Biophys. Chem. 134 (2008) 232–238. arXiv:0807.0892.
[18] M. Karnani, A. Annila, Gaia again, Biosystems (2008).
[19] P. Würtz, A. Annila, Roots of diversity relations, J. Biophys. (2008) doi:10.1155/2008/654672.
[20] A. Annila, E. Annila, Int. J. Astrobiol. 7 (2008) 293–300. doi:10.1017/S1473550408004308.
[21] E.T. Jaynes, Information theory and statistical mechanics, Phys. Rev. 106 (1957) 620–630.
[22] S.N. Salthe, Evolving Hierarchical System, Columbia University Press, New York, 1985.
[23] R.E. Ulanowicz, B.M. Hannon, Life and the production of entropy, Proc. R. Soc. B 232 (1987) 181–192.
[24] D.R. Brooks, E.O. Wiley, Evolution as Entropy: Toward A Unified Theory of Biology, 2nd ed., The University of Chicago Press, Chicago, 1988.
[25] S.N. Salthe, Development and Evolution: Complexity and Change in Biology, MIT Press, Cambridge, MA, 1993.
[26] E.D. Schneider, J.J. Kay, Life as a manifestation of the second law of thermodynamics, Math. Comp. Model. 19 (1994) 25–48.
[27] K. Matsuno, R. Swenson, Thermodynamics in the present progressive mode and its role in the context of the origin of life, BioSystems 51 (1999) 53–61.
[28] D.R. Brooks, The nature of the organism: Life takes on a life of its own, Proc. N.Y. Acad. Sci. 901 (2000) 257–265.
[29] R.D. Lorenz,Planets, life and the production of entropy, Int. J. Astrobiol. 1 (2002) 3–13.
[30] R. Dewar, Information theory explanation of the fluctuation theorem, maximum entropy production and self-organized criticality in non-equilibrium stationary states, J. Phys. A 36 (2003) 631–641.
[31] C. Darwin, On the Origin of Species, John Murray, London, 1859.
[32] L. Boltzmann, The second law of thermodynamics, in: B. McGinnes (Ed.), Ludwig Boltzmann, 1974 Theoretical Physics and Philosophical Problems, Reidel, Dordrecht, 1886.
[33] R. Feynman, R. Leighton, M. Sands, The Feynman Lectures on Physics, vol. I, Addison-Wesley, Reading, MA, 1963 (chapters 44–46).
[34] A. Ben-Naim, Hydrophobic Interactions, Plenum, New York, 1980.
[35] G. Hummer, S. Garde, A.E. Garcia, A. Pohorille, L.R. Prat, An information theory model of hydrophobic interactions, Proc. Natl. Acad. Sci. USA 93 (1996) 8951–8955.
[36] O. Bieri, T. Kiefhaber, Elementary steps in protein folding, Biol. Chem. 380 (1999) 923–929.
[37] J. Foster, J.D. Nightingale, A Short Course in General Relativity, Springer, New York, 1994.
[38] L. Brillouin, Science and Information Theory, Academic Press, New York, 1963.
[39] A.S. Eddington, The Nature of Physical World, MacMillan, New York, 1928.
[40] A.R. Fersht, Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding, Freeman, New York, 1999.
[41] A.R. Fersht, V. Daggett, Protein folding and unfolding, Cell 108 (2002) 1–20.
[42] M. Andrec, A.K. Felts, E. Gallicchio, R.M. Levy, Protein folding pathways from replica exchange simulations and a kinetic network model, Proc. Natl. Acad. Sci. USA 102 (2005) 6801–6806.
[43] P.J. Flory, Principles of Polymer Chemistry, Cornell University Press, Ithaca, 1953.

[44] S. Kullback, Information Theory and Statistics, Wiley, New York, 1959.
[45] Fenton Horwich, GroEL-mediated protein folding, Protein Sci. 6 (1997) 743–760.
[46] J.M. Lee, Introduction to Smooth Manifolds, Springer-Verlag, New York, 2003.
[47] P. Waage, C.M. Guldberg, Forhandlinger 35, Videnskabs-Selskabet i Christiana, Christiana, 1864.
[48] S.H. Strogatz, Nonlinear Dynamics and Chaos with Applications to Physics, Biology, Chemistry and Engineering, Westview, Cambridge, MA, 2000.
[49] H. Frauenfelder, Function and dynamics of myoglobin, Ann. N.Y. Acad. Sci. 504 (1987) 152–167.
[50] H.P. Lu, L. Xun, S. Xie, Single-molecule enzymatic dynamics, Science 282 (1998) 1877–1882.
[51] E. Fermi, Thermodynamics, Dover Publications, New York, 1956.
[52] P.F. Verhulst, Recherches mathématiques sur la loi d'accroissement de la population, Nouv. Mém. Acad. Roy. Sci. Belleslett. Bruxelles 18 (1845) 1–38.
[53] P.P. Chapagain, J.L. Parra, B.S. Gerstman, Y. Liu, Sampling of states for estimating the folding funnel entropy and energy landscape of a model alpha-helical hairpin peptide, J. Chem. Phys. 127 (2007) 075103.
[54] C.R. Cantor, P.R. Schimmel, Biophysical Chemistry Part III The Behavior of Biological Macromolecules, Freeman, San Francisco, 1980.
[55] M. Kot, Elements in Mathematical Ecology, Cambridge University Press, Cambridge, 2001.
[56] J. Sabelko, J. Ervin, M. Gruebele, Observation of strange kinetics in protein folding, Proc. Natl. Acad. Sci. USA 96 (1999) 6031–6036.
[57] M.F. Shlesinger, G.M. Zaslavsky, J. Klafter, Strange kinetics, Nature 363 (1993) 31–37.
[58] N. Eldredge, S.J. Gould, in: T. J. M Schopf (Ed.), Models in Paleobiology, Freeman, Cooper, San Francisco, 1972, pp. 82–115.
[59] D.T. Gillespie, Markov Processes: An Introduction for Physical Scientists, Academic Press, Boston, 1992.
[60] I. Prigogine, J.M. Wiame, Biologie et thermodynamique des phenomenes irreversibles, Experientia 2 (1946) 451–453.
[61] B.N. Provotorov, Zh. Exsp. Teor. Fiz. 42 (1962) 882; Sov. Phys. JETP 15 (1962) 611.
[62] A.S. Oja, A.J. Annila, Y. Takano, Cross relaxation by single spin flips in silver at nanokelvin temperatures, Phys. Rev. Lett. 65 (1990) 1921–1924.
[63] C.M. Dobson, Protein misfolding, evolution and disease, Trends Biochem. Sci. 9 (1999) 329–332.
[64] C.B. Anfinsen, Principles that govern the folding of protein chains, Science 181 (1973) 223–230.
[65] A. Pastore, S.R. Martin, A. Politou, K.C. Kondapalli, T. Stemmler, P.A. Temussi, Unbiased cold denaturation: Low- and high-temperature unfolding of yeast frataxin under physiological conditions, J. Am. Chem. Soc. 129 (2007) 5374–5375.
[66] N. Georgescu-Roegen, Analytical Economics, Harvard University Press, Cambridge, Mass, 1967.
[67] D. Sagan, From producing entropy to reducing gradients to spreading energy: Understanding the essence of the second law, in: G.P. Beretta, A. Ghoniem, G. Hatsopoulos (Eds.), Meeting the entropy challenge: An International Thermodynamics Symposium in Honor and Memory of Professor Joseph H. Keenan, in: AIP Conference Proceedings, vol. 1033, Cambridge, MA, 2007, pp. 194–197.
[68] S. Carnot, Reflections on the Motive Force of Fire, in: E. Mendoza (Ed.), Reflections on the Motive Force of Fire by Sadi Carnot and other Papers on the Second Law of Thermodynamics by E. Clapeyron and R. Clausius, Peter Smith, Gloucester, MA, 1977.
[69] S.N. Salthe, Evolving Hierarchical Systems, Columbia University Press, New York, 1985.
[70] M. Saunders, A. Wishnia, J.G. Kirkwood, The nuclear magnetic resonance spectrum of ribonuclease, J. Am. Chem. Soc. 79 (1957) 3289–3290.
[71] L.E. Kay, K.H. Gardner, Solution NMR spectroscopy beyond 25 kDa, Curr. Opin. Struct. Biol. 7 (1997) 722–731.
[72] Y. Sugita, Y. Okamoto, Replica-exchange molecular dynamics method for protein folding, Chem. Phys. Lett. 314 (1999) 141–151.
[73] Y.M. Rhee, V.S. Pande, Multiplexed-replica exchange molecular dynamics method for protein folding simulation, Biophys. J. 84 (2003) 775–786.
[74] A.S. Fraenkel, Complexity of protein folding, Bull. Math Biol. 55 (1993) 1199–1210.
[75] S.J. Gould, The Structure of Evolutionary Theory, Harvard University Press, Cambridge, Mass, 2002.
[76] M. Gell-Mann, The Quark and the Jaguar: Adventures in the Simple and the Complex, Henry Holt, New York, 1995.
[77] E.D. Schneider, D. Sagan, Into the Cool: Energy Flow, Thermodynamics, and Life, The University of Chicago Press, Chicago, 2005.
[78] J. Zhang, Protein-length distributions for the three domains of life, Trends Genet. 16 (2000) 107–109.
[79] M.L. Rosenzweig, Species Diversity in Space and Time, Cambridge University Press, Cambridge, 1995.