



# Equilibrium thermodynamics and the genesis of protein–protein complexes in cells

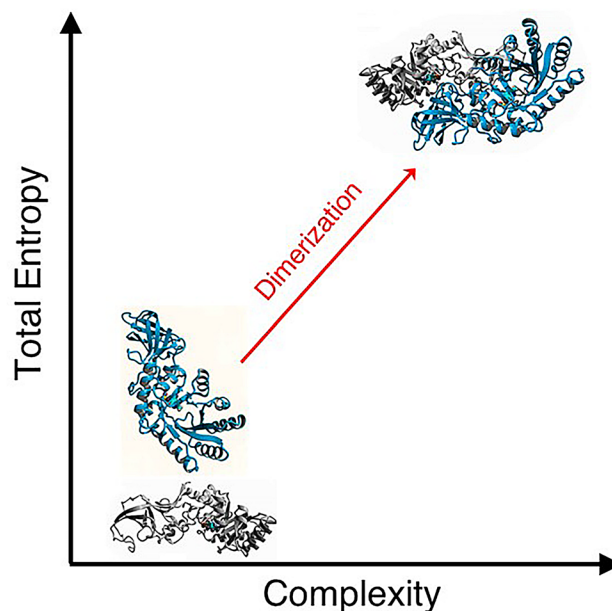
Robert C. Jennings<sup>1,2</sup> · Erica Belgio<sup>3</sup> · Giuseppe Zucchelli<sup>1,2</sup>

Received: 22 March 2021 / Accepted: 17 May 2021 / Published online: 19 June 2021  
© The Author(s) 2021

## Abstract

It is often thought that the structural complexity of living organisms places Life outside the laws of Physics. According to the Second Law of Thermodynamics, inanimate matter tends towards ever-increasing randomness. Most thermodynamic studies on the living system are course-grained in the sense that it is the whole organism which is considered and they lack microscopic details. In these studies, as the living system is an open system, non-linear thermodynamics have been used. This requires that a number of assumptions be made concerning the living system itself, which may not be correct in organisms living under natural environmental conditions. In the present study, we depart from this approach and use a fine-grained analysis of the genesis of subcellular protein complex structures. The analysis is performed in terms of classical equilibrium thermodynamics using the acquired knowledge of protein/protein interactions. In this way, it is demonstrated that the spontaneous creation of ordered subcellular structures occurs in accordance with the Second Law of Thermodynamics. We specifically consider the simple example of protein dimer and trimer formation from its monomer components, both in vitro and with chaperone assistance in vivo. The entropy decrease associated with protein complex assembly, on which the continuing debate is founded, is shown to be a relatively small component in the overall and positive entropy increase.

## Graphic abstract



**Keywords** Thermodynamics in biology · Biological complexity · Subcellular complexes · Multi-protein complexes · Thermodynamics second law

Extended author information available on the last page of the article

## 1 Introduction

The living system is characterized by the organization of matter in the most elaborate and complex forms known. At a macroscopic, morphological level systematic attempts to describe this date from Aldrovandi in the sixteenth century and culminated in the monumental classification works of Linnaeus and the Theory of Evolution based on Natural Selection of Darwin. At a more microscopic level, Hooke, in the seventeenth century, first observed the biological cell. In the twentieth century, with the development of the electron microscope and X-ray crystallographic techniques, the microscopic description of cellular structures made extraordinary progress. This laid the grounds to increasing awareness by physicists, mathematicians, and philosophers of the unique nature of biological structural complexity and attempts were initiated to understand the basic physical laws which govern this complexity. This led to an active debate over the last 80 years, which continues to this day. What is this “vital force” which propels the living system to move towards ever-increasing levels of complexity? In this Introduction, we present a brief overview of the extensive variety of approaches and thoughts concerning this matter of biological complexity. As will be seen, no general consensus has emerged. It is well known that in the physical universe, matter and energy are spontaneously degraded into more simple and more random states, as is predicted by the Second Law of Thermodynamics. However, at first sight, this appears not to be the case for the living system *in so much as order is apparently produced from less ordered states*, where “order” may be intuitively understood in terms of the complexity of biological structure (e.g., Azua-Bustos and Vega-Martinez 2013) which decreases the degrees of freedom of the molecular and multimolecular constituents of the components and the entropy of the system. We may also express this concept by the apparent *concentration* of energy in the structures of the living system, rather than it being dispersed, as in physical systems. This has led many, over the years, to view the living system as being in some way “outside” the accepted Laws of Physics and in particular the Second Law of Thermodynamics.

From the late nineteenth century, occasional speculations on this “apparent contradiction” were made. Over the years, other generalizations continue to appear in the scientific literature, but it seems that the first serious attempt to come to grips with the problem of the ever-increasing complexity of the living system in terms of physical principles was initiated by the mathematician Fantappiè (1942) and subsequently by physicists around the middle of the twentieth century. The best known example of an early discussion on this point is the book written by Schrödinger (1944), “What is Life?”, in which he reached the highly

unconventional and highly controversial conclusion that “we must be prepared to find it working in a manner that cannot be reduced to the ordinary laws of physics. And that not on the ground that there is any ‘new force’ or what not, directing the behavior of the single atoms within a living organism, but because the construction is different from anything we have yet tested in the physical laboratory”. He does however suggest, in very general terms, that Life is a process whereby energy exchanges lead, in the long run, to an increase in entropy, and thus, Life does not represent an exception to the Second Law. This latter point of view is generally accepted in Physics, though the statement “cannot be reduced to the ordinary laws of physics” seems to set living organisms apart from “ordinary matter” (Martyushev 2013). In this context, Fantappiè (1942) introduced the concept of “syntropy” as an alternative to the Second Law and as a principle governing the genesis of structural order in the animate. Syntropy is supposed to govern all those phenomena which are *attracted towards causes* (attractors), where the “causes” are pre-existing. In approximate terms, syntropy leads to the spontaneous creation of order, much as the negentropy of Schrödinger (1944), i.e., syntropy is considered to be a principle which is symmetrical to entropy and some suggest that it characterizes the living system (e.g., Fantappiè 1942; Levins and Lewontin 1985; Vannini 2005; Di Corpo and Vannini 2011). On the other hand, psychologists, while not going down the “syntropic path”, have expressed surprise and perplexity in so much as evolution proceeds along a path of increasingly ordered structures (e.g., Levins and Lewontin 1985; Tooby et al. 2003; Beichler 2016). The latter authors stated

Given the belief that the physical universe is moving toward a static death rather than a thermodynamic equilibrium in which molecular motion continues, it is no surprise that evolutionists believe organic evolution to be the negation of physical evolution.

A clear example of this line of thought was also expressed by the well-known Physiologist Szent-Gyorgyi (1977) who summarized his thoughts on the matter with the rather extremist statement “A major difference between amoebas and humans is the increase of complexity that requires the existence of a mechanism that is able to counteract the law of entropy. In other words, there must be a force that is able to counter the universal tendency of matter towards chaos and energy towards dissipation. Life always shows a decrease in entropy and an increase in complexity, in direct conflict with the law of entropy”. Tooby et al. (2003) write “Thus, to study organisms scientifically is to be confronted with the following questions: Why is it that living things exhibit a miraculously high level of order not found among the nonliving? Where does

this high level of order come from?”, a question posed but not answered.

Azua-Bustos and Vega-Martinez (2013) quantified the high level of surface complexity/order of lichens growing on rocks by fractal analysis and referred to this in terms of system entropy.

Another line of thought, generally considered to have a more solid scientific and propounded by physicists, is based on the principles of non-equilibrium thermodynamics to explain the development of biological complexity. In his review, the physicist Martyushev (2013) states “However, the above question still remains: what forces life to continuously develop and become more complex, and can this question be solved within the scope of thermodynamics? Can it be true that the supporters of Schrödinger who believe that the problems of origin and evolution of biological structures are beyond the scope of physics because its laws are insufficient for understanding thereof, are right?” In fact, the above citation of Martyushev (2013) concludes with the following words “In our opinion, it is not the physics of “animate” that is required but a deeper investigation of the properties of entropy and, first of all, the rate of its change”. This statement refers to the concept that the generation of system complexity may be based on the same system spontaneously “evolving down the path” of maximum entropy production (MEP) which is greater than the entropy decrease associated with the creation of highly “ordered structures”, with no violation of the Second Law of Thermodynamics (e.g., Martyushev 2013; Swenson 1977; Kondepudi and Prigogine 1998; Ziegler 1963; Toussaint and Schneider 1998) to name just a few proponents. In these terms, the basic concept points in the direction that the “vital force” which “forces life to continuously develop and become more complex” is *entropy generation or production*. We shall refer to these concepts as MEP (maximum entropy production).

The basic concept of MEP views the living system as an open system in which energy and/or matter is continually “imported” and “exported”. This excludes the use of conventional equilibrium thermodynamics and MEP is, in fact, based on non-equilibrium thermodynamics. Instead of reactions proceeding towards equilibrium, as occurs in a closed system, in MEP, they are conceived to proceed towards the steady state (SS). MEP views the SS as representing a stable, or metastable state, for open thermodynamic systems, in much the same way as equilibrium does for closed systems and equilibrium thermodynamics. In MEP, the dynamics of steady-state systems are thought to adjust themselves to achieve a state in which the entropy production rate is maximized, given the constraints. Non-equilibrium thermodynamics and MEP are based on equilibrium thermodynamic functions and even the rigorous definition of entropy and temperature are lacking (e.g., Lebon et al. 2008). The

application of MEP theory to the living system has recently been strongly criticized (Jennings et al. 2020).

Intimately related to MEP is the concept of “dissipative structures”, structures which increase the entropy production rate over that which it would be in their absence. These structures are based on the non-equilibrium formation of steady states and are considered to have the capacity to increase structure and order. They characterize complex systems and are well established for such inanimate “structures” as hurricanes and Bénard Cells, where the convection currents generate structure. They have also been applied to the generation of structure and complexity in the living system (e.g., Zotin 2014; Dewar 2010).

As mentioned above, the living state, considered in its entirety, is an open system, as is thermodynamically defined, in so much as energy and matter enter and exit. Fluxes are formed and steady states may, in principle, be attained. This is the basic assumption of the non-equilibrium thermodynamics in general and to its application to the living system. The present case of protein–protein complex formation, basic to the formation of subcellular complexes, is however different and a strict use of the “open system” definition hides these differences. If we consider the formation of subcellular complexes, we are obliged to recognize that this process may not be represented in terms of an open system as the genesis of a cellular structure, if stable on a physiological time scale is, in fact, a “*cul de sac*” in which matter enters, but does not pass through (see Discussion below). This is a central point to the present study as the generation of structural complexity may then be considered in terms of classical equilibrium thermodynamics as a useful investigative tool.

Protein chemists who study protein/protein interactions usually start out by determining the equilibrium constant, which allows calculation of the standard free energy change ( $\Delta G^\circ$ ). This is mostly achieved in vitro by measurement of the kinetic constants of association and dissociation. Equilibrium thermodynamics is employed. For the in vivo situation, very few studies exist due to the extreme difficulty of measurement in the dense cytoplasmic environment of the cell (Rivas and Minton 2018). The equilibrium approach used for in vitro measurements is, however, the same as the cellular membranes are impermeable to most proteins and the cellular volume is constant. This is a central point to the present study as the generation of structural complexity may then be considered in terms of classical equilibrium thermodynamics both in vitro and in vivo.

Another line of thought, which overlaps with the two above-mentioned hypotheses, is that of the “shaping” of the evolutionary development of biological complexity by the presence of pre-existing environmental factors (King 1996; England 2013), both taking the photosynthetic process as an example. King suggested that, given the existence

of photons, photosynthetic organisms will inevitably arise through mutation and natural selection, though the development of complexity is not directly addressed. England, on the other hand, proposed in detailed terms that the development of biological complexity is associated with entropy production (England 2013) and is reported to have made the rather extreme statement (Wolchover 2014) “You start with a random clump of atoms, and if you shine light on it for long enough, it should not be so surprising that you get a plant”.

Carrà (2020) has discussed the problem in terms of information theory.

The supporters of these various viewpoints seem to have side stepped the possible role of the so-called “high energy” compounds, in which the free energy of phosphate diester hydrolysis may be coupled to and “drive” biological systems against the gradient of thermodynamic potential and so produce complexity (order) from simpler molecular states. Massive use of the free energy associated with nucleotide triphosphates, and also pyrophosphate hydrolysis, in the synthesis of such fundamental and “low entropy” macromolecules as the nucleic acids and proteins is very well known. It is probably with the synthesis of these complex molecular states that biological complexity begins.

From this brief overview, it is evident that the evolution of complexity which characterizes the living system is subject to extremely different, and sometimes divergent, interpretations which, to some measure, depend of the reference background of the researcher involved. These reference backgrounds are, in fact, very varied and includes biologists, protein chemists, mathematicians, physicists, and also philosophers (e.g., Heylighen et al. 2007; Santos 2013).

In the light of these divergent opinions, and taking into account the hierarchical scale of decreasing complexity that characterize the living system, coarsely going from ecosystems, complex organisms, cells, subcellular complexes to molecular structures, it is the purpose of this article to attempt to clarify only one limited aspect of the structural complexity problem of the living system. Our approach specifically considers the formation or assembly of simple cellular substructures which are due to protein/protein interactions, known as protein complexes, in terms of their entropic footprint. Cellular protein complexes may consist of just two non-covalently bound proteins or, as in the case of, for example, proteasomes, many protein subunits. Protein complexes are considered by biochemists to be involved in most cellular biochemistry, which underlines the importance of the analysis. Thus, our approach differs from almost all other attempts to understand biological complexity in terms of Physical Chemistry, in so much as it considers the specific example of the development of structural complexity at the molecular/cellular level, i.e., a well-defined system. The importance of this approach was recently recognized by Andrieux and Gaspard (2008) who wrote “... biological

systems have structures and functions at every scale down to the molecular level, and the understanding of their origin is a challenge”. Most thermodynamics studies on biological complexity (biothermodynamics) attempt to consider and explain “Life itself” where the living system is considered as being some kind of non-defined, structured “black box”.

## 2 Discussion

Biological complexity at a cellular level is characterized by the multiproteic structures which regulate most of the metabolic activity of the living cell (e.g., Hartwell et al. 1999; Kastiris and Bonvin 2013). We consider an example of the thermodynamics of protein/protein interactions, fundamental to the formation of the complexes which make up the ordered subcellular structures. In this hypothetical example of the increase in complexity (order) of a cellular structure, the complex may consist of just two polypeptide chains, or some tens of proteins.

The concept of entropy as thermodynamic order is derived from statistical mechanics and is illustrated by the well-known Boltzmann equation  $S = k_B \ln W$ , written for an isolated system, i.e., non-interacting with the environment.  $W$  represents the number of accessible microstates and  $k_B$  is the Boltzmann’s constant. The function  $W$  can be factorized into different contributions due, e.g., to particle spatial distribution, momenta or others defining the system state, but also between the contributions due to different objects comprising a heterogeneous system, as can be shown when an ideal gas in volume  $V$  is considered as an example (Appendix A). As we are interested in structural order, just the position coordinates are considered and then (Appendix A)

$$S_x = k_B \ln W_x. \quad (1)$$

The entropic decrease due to order creation is simply illustrated by the following example of protein/protein dimer formation,  $D$ , from two non-identical polypeptides  $P_1$  and  $P_2$  in an isolated volume  $V$ , i.e.,  $P_1 + P_2 \leftrightarrow D$ . The difference in statistical entropy, when the accessible configurational microstates for the two sets of macromolecules are taken into account, is given by

$$S_{x,D} - S_{x,\{P_1,P_2\}} = k_B \left( \ln W_{x,D} - \ln W_{x,\{P_1,P_2\}} \right). \quad (2)$$

The number of accessible microstates is greater for the two polypeptide system when compared to the protein dimer system (Appendix A). This is also intuitive, due to the decreased number of particles when protein dimers are established, and leads to a negative configurational entropy of dimerisation. Thus, as expected, the creation of

a more “ordered” state from a less “ordered” state leads to a decrease in entropy. It is this negative entropy component which has attracted the attention of many, leading often to the suggestion that Life violates the Second Law. However, other entropy contributions exist, which in the simple example above, have not been considered. These “other” entropy contributions, often ignored in the relevant literature, are briefly mentioned below.

Protein–protein binding decreases rotational degrees of freedom and this also yields a second negative entropy contribution. On the other hand, biological complexes, as is well known, are held together by a number of distinct interactions (e.g., Sowmya et al. 2015). Important contributions to binding strength are made by non-covalent van der Waals forces, electrostatic interactions, and hydrogen bonds in which the van der Waals forces are suggested to be the dominant force (Nilofer et al. 2017), at least in some cases. These interactions are spontaneous and exothermic and are important in stabilizing the complex (Eq. 7). Bond formation releases heat and thus produces thermodynamic entropy in the surroundings. In the exposed hydrophobic domains of both protein complexes and monomeric proteins in aqueous solution, water molecules are thought to form a “cage” of structured water molecules (Kastritis and Bonvin 2013; Chen et al. 2013) in which the translational and rotational degrees of freedom are reduced with respect to bulk water. This idea, introduced by Tanford (1973) and commonly invoked, is not however supported by experimental evidence (Kastritis and Bonvin 2013). Neutron scattering experiments (e.g., Turner et al. 1990; Buchanan et al. 2005a, b) found no evidence for the “structured water cage”. Thus, the often invoked release of “structured water” molecules into the bulk phase leading to an entropy increase remains unclear. It is the balance between these entropy contributions which determines cellular complex formation. Recent studies, using crystallographic structures, have been directed at understanding the relative contributions of these factors to the binding free energy (e.g., Sowmya et al. 2015; Nilofer et al. 2017) and molecular dynamics calculations for specific heterodimers are moving towards an increasingly accurate description of experimental protein/protein-binding data (Liu et al. 2019), though the differences between calculated and experimental binding free energies are often considerable.

We wish to emphasize that the genesis of protein /protein complexes is not analogous to the chemical polymerisation as covalent bonds are not involved, as briefly discussed above.

In the following, several general examples of complex formation are considered. First, cytosolic heterodimer formation, where the single protein “building blocks”,  $P_1$  and  $P_2$ , bind non-covalently to form the nascent dimer polypeptide complex,  $D$ , i.e.,  $P_1 + P_2 \leftrightarrow D$ . While we realize that the single proteins are synthesized by coupling to reactions which provide free energy, e.g., nucleotide triphosphates,

the above chemical equation is that which describes *complex formation itself*. This model system may be analyzed in terms of classical equilibrium thermodynamics (Eq. 3).

Though the main thrust of this article treats entropy, in the section which goes from Eqs. 3 to 6, the discussion is in terms of Gibbs free energy,  $G$ , as this is the parameter used by protein chemists in studies on protein/protein interactions, where

$$\Delta G = -T\Delta S_{\text{tot}} = -T\left(\Delta S_{\text{sys}} - \frac{\Delta H}{T}\right), \quad (3)$$

$T$  is the temperature (Kelvin) and  $H$  is the enthalpy. The subscript “tot” indicates the total entropy, i.e., that of the system *plus* that of the environment in which the proteins are embedded,  $\frac{\Delta H}{T}$ , which, for an exothermic spontaneous reaction, is associated with the heat released into the environment at temperature  $T$ . Equation 4 gives the free energy change,  $\Delta G$ , as a function of the reaction quotient for the binding of  $P_1$ ,  $P_2$  to form the heterodimer  $D$

$$\Delta G = \Delta G^\circ + RT \ln \frac{[D]}{[P_1][P_2]}. \quad (4)$$

The Standard Gibbs Free Energy,  $\Delta G^\circ$ , for protein/protein binding is experimentally determined, in vitro, from the equilibrium constant  $K_a$  (Eq. 5) (Kastritis and Bonvin 2013; Chen et al. 2013)

$$\Delta G^\circ \equiv \Delta H^\circ - T\Delta S^\circ = -RT \ln K_a. \quad (5)$$

Equation 5 describes the “intrinsic” binding tendency, which is modulated by the substrate/reaction concentrations (Eq. 4). Substituting Eq. 5 in Eq. 4, the effective free energy change is

$$\Delta G = RT \ln \left( \frac{[D]}{[P_1][P_2]K_a} \right). \quad (6)$$

Under physiological conditions, stable binding occurs spontaneously (i.e.,  $\Delta G < 0$ ) when  $[D]/[P_1][P_2] < K_a$ .

In an analysis of over 100 protein heterodimers from the Protein Data Bank, Chen et al. (2013) estimated the standard Gibbs free energy for protein/protein binding in an aqueous solvent using the experimentally determined values for  $K_a$  (Eq. 5). They observed a direct relationship between the buried interfacial surface area and binding affinity: as the buried surface area increases, binding affinity increases, a conclusion confirmed using computational techniques (Sowmya et al. 2015; Nilofer et al. 2017).

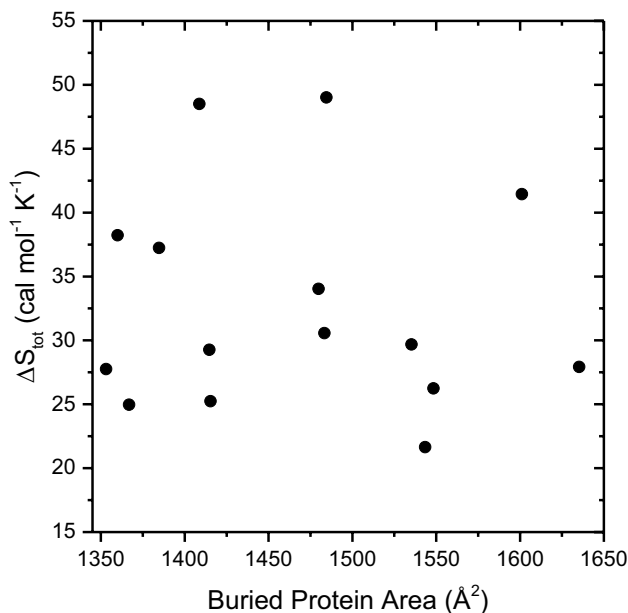
The  $\Delta G^\circ$  values per unit protein buried area,  $\Delta G_{\text{uA}}^\circ$ , over the entire binding surface area considered for protein–protein complexes, ranging from approximately 880 to above 3400  $\text{\AA}^2$ , were estimated to be in the range  $-10 \leq \Delta G_{\text{uA}}^\circ \leq -4 \text{ cal mol}^{-1} \text{\AA}^{-2}$ .

If the “moderate” buried binding surface of around  $1500 \text{ \AA}^2$  is taken as an example, the protein–protein complex  $\Delta G^\circ$  data spread lies approximately in the interval  $-15 \leq \Delta G^\circ \leq -6 \text{ kcal mol}^{-1}$ . As  $\Delta S_{\text{tot}}^\circ = -\Delta G^\circ/T$ , this interval may also be written in terms of the total entropy change, i.e.,  $20 \leq \Delta S_{\text{tot}}^\circ \leq 50 \text{ cal mol}^{-1} \text{ K}^{-1}$ , with  $T=300 \text{ K}$  (Fig. 1).

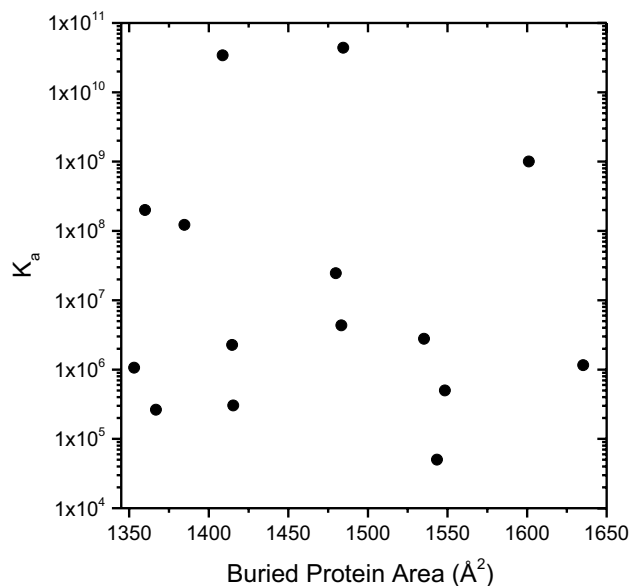
This clearly shows that the total entropy change on dimerisation is positive, in all cases, which in turn indicates that the configurational (ordering) term does not dominate.

For the lower limit of  $\Delta S_{\text{tot}}^{\text{uA}} = 13 \times 10^{-3} \text{ cal mol}^{-1} \text{ K}^{-1}$  ( $\Delta G_{\text{uA}}^\circ = -4 \text{ cal mol}^{-1} \text{ \AA}^2$ ,  $T=300 \text{ K}$ ), the entropy change per unit protein buried area, in the present case of a binding surface area of  $1500 \text{ \AA}^2$ , the corresponding value of  $K_a$  is about 10,000 for peptide/peptide binding (Fig. 2). In this case, spontaneous binding is expected when the cellular  $[D]/[P_1][P_2] < 10,000$  which, in words, means that even at low substrate concentrations with respect to the dimer product, spontaneous binding may occur. In those cases where the  $\Delta S_{\text{tot}}^\circ$  values are more positive, then protein binding would be even more favored, even at extremely low concentrations of the protein monomers. It is therefore clear that the increase in molecular complexity of the many protein dimeric cellular structures taken into account by Chen et al. (2013) is spontaneous, modulated by the product/substrate ratio, and occurs with an increase in the total entropy ( $\Delta S_{\text{tot}}^\circ$ ).

The total entropy change due to increased molecular complexity is illustrated in Eq. 7 in terms of the various entropy



**Fig. 1** Total entropy changes in the assembly of protein/protein dimers as a function of the buried binding surface area in the  $150 \text{ \AA}^2$  range around  $1500 \text{ \AA}^2$ . These values are obtained using the data published by Chen et al. (2013)



**Fig. 2** Equilibrium association constant values for protein dimerization as a function of the buried binding surface area in  $150 \text{ \AA}^2$  range around  $1500 \text{ \AA}^2$ . These values are obtained using the data published by Chen et al. (2013)

production contributions:  $\Delta S_{\text{er}}^\circ$ , due to environmental rearrangement (e.g., the hydrophobic effect);  $\Delta S_{\text{b}}^\circ = \Delta H^\circ/T$ , the heat (entropy) released into the environmental bath due to bond formation;  $\Delta S_{\text{SCC}}^\circ$ , the side-chain configurational entropy term, which Zhang and Liu (2006) suggested may increase in the extra-interface domains of protein complexes;  $\Delta S_{\text{vr}}^\circ$ , the entropy decrease due to vibrational and rotation restrictions, upon dimerization, in the interface area;  $\Delta S_{\text{c}}^\circ$ , the configurational entropy decrease (order formation) upon complex formation discussed above. *It is specifically this negative  $\Delta S_{\text{c}}^\circ$  term (Appendix A, for the simple case of a gas) which is associated with biological complexity and much academic perplexity, perplexity which is due to the failure to consider the other entropy terms*

$$\Delta S_{\text{tot}}^\circ = (\Delta S_{\text{er}}^\circ + \Delta S_{\text{b}}^\circ + \Delta S_{\text{SCC}}^\circ) - (\Delta S_{\text{vr}}^\circ + \Delta S_{\text{c}}^\circ). \quad (7)$$

For the moderate binding surface of around  $1500 \text{ \AA}^2$  ( $T=300 \text{ K}$ ) considered above  $20 \leq \Delta S_{\text{tot}}^\circ \leq 50 \text{ cal mol}^{-1} \text{ K}^{-1}$ . It is these positive entropy changes (exothermic processes) which drive dimer formation. The “order formation” term ( $\Delta S_{\text{c}}^\circ$ ) does not dominate in the present case of dimer formation.

As  $\Delta S_{\text{c}}^\circ$  is a configurational entropy term, we discuss the matter comparing the formation of simple protein complexes which are commonly present in cells, i.e., dimers and trimers.

Following the ideal gas modeling in the Appendix A, it is readily shown that the configurational (order) entropy

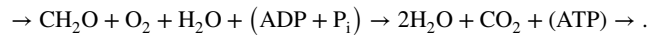
term is greater for trimer formation than for dimer formation, ( $\Delta S_T - \Delta S_D$ ) is negative, as expected. In the case of protein dimers, there is just one binding interface, whereas in the case of most trimers, there are three. For the lower  $\Delta G^\circ$  limit and considering a binding interface of  $1500 \text{ \AA}^2$ , as assumed for the dimer, the total entropy increase is expected to be approximately 3 times that for the dimer, i.e.,  $\Delta S_{\text{tot}} \approx 60 \text{ cal mol}^{-1} \text{ K}^{-1}$ . This simple dimer/trimer assembly example illustrates that even though trimer formation leads to a further decrease in the negative entropy “order” term  $\Delta S_c^\circ$ , trimerization is nonetheless thermodynamically more favourable than dimer assembly. This serves to underline the previous conclusion that the “order formation” term does not have a major impact on cellular complex formation. This simple conclusion is important in the context of the century long debate on the “antientropic” nature of complex formation and the concept that Life may lie outside the Laws of Thermodynamics. As far as we are aware, it is the first time this has been demonstrated.

In the above discussion, no mention is made of the possible role of molecular chaperones and chaperonins in the *in vivo* assembly of complex biological structures (e.g., Ellis 2007). Chaperones are themselves multiproteic structures which, in the dense cellular environment, seem to “assist” complex assembly, in many cases. The word “assist” means that the chaperone role is that of screening reactive protein surfaces from non-specific interactions in the dense cellular environment, allowing them to be transferred from their site of synthesis to the binding area of the developing complex. In particular, the Hsp family of chaperones are considered to play a role in the insertion of some proteasome proteins (e.g., Schmidtke et al. 1997; Mayer et al 2002; Makhnevych and Houry 2012). Though the “assist” activity of many chaperones is ATP-dependent, this is not expected to modify the thermodynamics of the protein/protein interactions involved in complex assembly. This is because the ATPase activity plays a fundamental role in the binding of the polypeptide substrate(s) to the chaperone by promoting unfolding of the chaperone, to be subsequently released upon by refolding. This “assists” the formation of the nascent complex, without being involved in the protein/protein interaction as such (e.g., Makhnevych and Houry 2012; Clare and Saibil 2013; Saibil 2013). In other words, the free energies of the substrates and products of the complex assembly reaction are not expected to be affected by chaperone activity.

That the chaperone role in protein/protein interactions in complex assembly is largely “passive”, in thermodynamic sense, for protein/protein assembly this is not surprising as, from the above discussion, it is evident that protein/protein interactions are themselves thermodynamically spontaneous.

It is universally accepted that the living system is a non-equilibrium, open system. This concept is clearly illustrated by the chemical equation which summarizes the central

oxidative phosphorylation process of respiration in which high energy substrates ( $\text{CH}_2\text{O}$ ) and oxygen enter, ATP is synthesized, and both water and  $\text{CO}_2$  exit



The above overall representation is due to Nishiyama et al. (2009) and Yang et al. (2021) who revealed that  $\text{H}_2\text{O}$  is indispensably involved in the reaction. This non-equilibrium, open system process, in those cases where the chaperone is ATP-dependent, is coupled to cellular complex formation, as discussed above. The coupling via ATP is, however, not thermodynamic, as ATP does not modify the protein/protein interaction, as discussed above.

On the basis of these considerations, we conclude that the living system is overall, an open system, as is common knowledge. However, the formation of cellular complexity itself is not. This point is interesting when the ATP-dependent chaperone involvement is considered. In this case, ATP is formed in an open system, which couples to complex formation via the “assist” mechanism and is non-thermodynamic. This point will be further developed in a subsequent study.

Finally, it should be noted that the recent suggestion that primary processes in plant photosystems may consume entropy under certain conditions (Jennings et al. 2017) concerns function and not the genesis of increasing structural complexity, and is therefore not in contradiction with the present study.

### 3 Conclusions

We address the question of biological complexity in thermodynamic terms. Over that past century a considerable debate as to whether the highly ordered structure of living systems is in contradiction to the Second Law of Thermodynamics has developed, and continues to the present day. Most physicists consider that no contradiction exists and suggest that non-equilibrium thermodynamics may be used to demonstrate this (see Introduction). Maximum Entropy Production is assumed to constitute the driving force which produces the complexity of living systems. However, this has yet to be proven. In non-equilibrium thermodynamic theory, both entropy and temperature lack a rigorous definition and are based on equilibrium concepts (e.g., Lebon et al. 2008). Furthermore, its application to living systems has been contested (see Introduction). Thus, the question and the nature of the “vital force” leading to ever-increasing complexity in the living system remain unresolved.

Most studies on biological complexity attempt to consider and explain “Life itself”, where the living system is considered as being some kind of ill-defined, structured “black box”. Usually, a precise biological model is lacking. In the present study,

an alternative approach is adopted in which the specific case of structural complexity at a cellular level is examined in terms of the formation of protein–protein dimer and trimer complexes. This is achieved by employing standard equilibrium thermodynamics, commonly used in protein chemistry. The aim was to examine the impact of the negative configurational entropy contribution (order formation), which describes the genesis of complexity, to the other positive entropy changes associated with protein–protein interactions, both in vitro and in vivo (Eq. 7). The positive entropy changes are shown to dominate over the negative entropy contributions. This is an unambiguous demonstration that living system complexity is not in violation of the Second Law in this particular case.

## Appendix A

Consider, in the classical limit, an isolated ideal gas of  $N$  identical molecules with mass  $m$ , enclosed in a volume  $V$ . The mutual interaction between molecules is assumed negligibly small as well as is, for simplicity, the energy contribution due to rotational and vibrational degrees of freedom. The total energy,  $E$ , of this system is then the total kinetic energy due to the translational motion of the molecules. The system is described, using the center-of-mass of each molecule, by the couples  $\{\mathbf{x}_i, \mathbf{p}_i\}$  of position,  $\mathbf{x}_i$  and momentum,  $\mathbf{p}_i$ , each defined by their three coordinates  $\{x, y, z\}$  and  $\{p_x, p_y, p_z\}$  for a total of  $3N=f$  degrees of freedom. The states of the system are described in the  $2f=6N$  dimensional phase space defined, by  $\{\mathbf{x}, \mathbf{p}\}$  and discretized in  $M$  cells, with  $M \gg N$ , borrowing from quantum mechanics the lower limit to the size  $\delta x \delta p = \hbar$ , the reduced Planck's constant, so that the volume of the elementary cell of the entire phase space is  $\hbar^f = \hbar^{3N}$ . The set  $\{E, V, N\}$  of the macroscopic parameters, describing the physical macro-state of this model system, constrains the possible number,  $W$ , of microscopic configurations that enters the Boltzmann definition of entropy

$$S = k_B \ln W, \quad (\text{A.1})$$

where  $k_B$  is the Boltzmann constant.

The number of microstates,  $W$ , can be factorized as the product of two independent contributions (e.g., Reif 1965), one due to the spatial,  $\mathbf{x}$ , coordinates, defined here as configurational, and the other to the momenta,  $\mathbf{p}$ , of the system particles

$$W = W_{\mathbf{x}} \cdot W_{\mathbf{p}}, \quad (\text{A.2})$$

so that

$$S = k_B \ln W_{\mathbf{x}} + k_B \ln W_{\mathbf{p}} = S_{\mathbf{x}} + S_{\mathbf{p}}. \quad (\text{A.3})$$

We start by considering the configurational contribution,  $W_{\mathbf{x}}$ , as, in this paper, the interested is in structural

changes only. To this end, we calculate the number of microstates for a system of  $N$  identical molecules distributed into  $M$  identical and distinguishable cells of the phase space. That is equivalent to finding the number of modes in which  $N$  identical objects are distributed in  $M$  boxes without superposition

$$W_{\mathbf{x}} = \frac{M!}{N!(M-N)!}. \quad (\text{A.4})$$

To simplify, and taking into account that  $M \gg N$

$$\begin{aligned} \frac{M!}{N!(M-N)!} &= \frac{M(M-1)(M-2) \cdots (M-N+1)(M-N)!}{N!(M-N)!} \\ &= \frac{M(M-1)(M-2) \cdots (M-N+1)}{N!} \cong \frac{M^N}{N!}. \end{aligned} \quad (\text{A.5})$$

The number  $M$  of cells is

$$M = \frac{V}{\hbar^3}, \quad (\text{A.6})$$

and the configurational contribution,  $S_{\mathbf{x}}$ , to the total entropy is

$$S_{\mathbf{x}} = k_B \ln W_{\mathbf{x}} = k_B \ln \left( \frac{1}{N!} \frac{V^N}{\hbar^{3N}} \right). \quad (\text{A.7})$$

As a point of interest, though it is not of direct relevance in the present case, adding the system particles momenta contribution,  $S_{\mathbf{p}}$ , to the configurational contribution,  $S_{\mathbf{x}}$ , (Eq. 10), calculated according to the constraint given by the total kinetic energy

$$E = \frac{1}{2m} \sum_{i=1}^N \sum_{j=1}^3 p_{ij}^2, \quad (\text{A.8})$$

which defines a 3  $N$ -sphere with radius  $R = \sqrt{2mE}$ , leads to the total entropy expression given by the Sackur–Tetrode equation (e.g. Sommerfeld 1955).

We now return to the configurational entropy,  $S_{\mathbf{x}}$ , and consider a mixture of two different ideal gases, one of  $N_A$  identical molecules A and the other of  $N_B$  identical molecules B, in the same volume  $V$ . The total number of molecules is  $2N$ . The configurational contribution,  $W_{\mathbf{x},\{A,B\}}$ , can be written in terms of the independent contribution,  $W_{\mathbf{x},A}$  and  $W_{\mathbf{x},B}$ , for each category of particle

$$W_{\mathbf{x},\{A,B\}} = W_{\mathbf{x},A} \cdot W_{\mathbf{x},B} = \left( \frac{V}{\hbar^3} \right)^{N_A+N_B} \frac{1}{N_A! N_B!}. \quad (\text{A.9})$$

When an ideal gas of  $N$  dimers,  $D$ , of molecules A and B in the volume  $V$ , neglecting rotational and vibrational degrees of freedom, is considered, the configurational contribution,  $W_{\mathbf{x},D}$ , for this molecular ensemble is



$$W_{x,D} = \left(\frac{V}{\hbar^3}\right)^N \frac{1}{N!}. \quad (\text{A.10})$$

We are now in the position to determine the configurational entropy difference,  $\Delta S_x$ , between the gas of  $N$  dimer,  $D$ , and that of the two molecules  $A$  and  $B$ , with  $N_A = N_B = N$

$$\begin{aligned} \Delta S_x &= k_B (\ln W_{x,D} - \ln W_{x,\{A,B\}}) = k_B \ln \left( \frac{W_{x,D}}{W_{x,\{A,B\}}} \right) \\ &= k_B \ln \left( \frac{N_A! N_B!}{N!} \frac{V^N}{V^{N_A+N_B}} \frac{\hbar^{3(N_A+N_B)}}{\hbar^{3N}} \right) \\ &= k_B \ln \left( \frac{N! \hbar^{3N}}{V^N} \right) = -k_B \ln \frac{1}{N!} \left( \frac{V}{\hbar^3} \right)^N, \end{aligned} \quad (\text{A.11})$$

which is negative, as expected.

**Funding** Open access funding provided by Università degli Studi di Milano within the CRUI-CARE Agreement. This research was partially funded by the Czech Science Foundation (GACR—19-11494S + ALGAMIC CZ.1.05/2.1.00/19.0392) to EB. RCJ and GZ did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

## Declarations

**Conflicts of interest/competing interests** The authors declare that they have no conflict of interest or competing interest.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Andrieux D, Gaspard P (2008) Nonequilibrium generation of information in copolymerization processes. *Proc Natl Acad Sci USA* 115:9516–9521
- Azua-Bustos A, Vega-Martinez C (2013) The potential for detecting ‘life as we don’t know it’ by fractal complexity analysis. *Int J Astrobiol* 12:314–320
- Beichler JE (2016) Consciousness of unification: the mind-matter phal-lacy bites the dust. In: Amoroso RL, Kauffman LH, Rowlands P (eds) *Unified field mechanics: natural science beyond the veil of spacetime*. Proceedings of 9th Vigier Symposium on Unified Field Mechanics—Natural Science Beyond the Veil of Spacetime, Morgan State Univ, Baltimore, MD, Nov 16–19, 2014. World Scientific Publ Singapore. p 564.

- Buchanan P, Aldiwan N, Soper AK, Creek JL, Koh CA (2005a) Decreased structure on dissolving methane in water. *Chem Phys Lett* 415:89–93
- Buchanan P, Soper AK, Thompson H, Westacott RE, Creek JL, Hobson G, Koh CA (2005b) Search for memory effects in methane hydrate: Structure of water before hydrate formation and after hydrate decomposition. *J Chem Phys* 123:164507
- Carrà S (2020) Reaction kinetics: scientific passion or applicative tool? *Rend Fis Acc Lincei* 31:269–284
- Chen J, Sawyer N, Regan L (2013) Protein-protein interactions: general trends in the relationship between binding affinity and interfacial buried surface area. *Protein Sci* 22:510–515
- Clare DK, Saibil HR (2013) ATP-driven molecular chaperone machines. *Biopolymers* 99:846–859
- Dewar RC (2010) Maximum entropy production and plant optimization theories. *Phil Trans R Soc B* 365:1429–1435
- Di Corpo U, Vannini A (2011) The EVOLUTION OF LIFE ACCORDING TO THE LAW OF Syntropy. *Syntropy* 1:39–49
- Ellis RJ (2007) Protein misassembly: macromolecular crowding and molecular chaperones. *Adv Exp Med Biol* 594:1–13
- England JL (2013) Statistical physics of self-replication. *Chem Phys* 139:121923–121931
- Fantappiè L (1942) Sull’interpretazione dei potenziali anticipati della meccanica ondulatoria e su un principio di finalità che ne discende. *Rend D Accad D’Italia* 7:81–86
- Hartwell LH, Hopfield JJ, Leibler S, Murray AW (1999) From molecular to modular cell biology. *Nature* 402:C47–52
- Heylighen F, Cilliers P, Gershenson C (2007) Complexity and Philosophy. In: Brogg J, Geyer R (eds) *Complexity, science and society*. Radcliffe Publishing, Oxford
- Jennings RC, Belgio E, Zucchelli G (2017) Photosystem I, when excited in the chlorophyll  $Q_y$  absorption band, feeds on negative entropy. *Biophys Chem* 233:36–46
- Jennings RC, Belgio E, Zucchelli G (2020) Does maximal entropy production play a role in the evolution of biological complexity? A biological point of view. *Rend Fis Acc Lincei* 31:259–268
- Kastritis PL, Bonvin AMJJ (2013) On the binding affinity of macromolecular interactions: daring to ask why proteins interact. *J R Soc Interface* 10:20120835
- King CC (1996) Quantum mechanics, chaos and the conscious brain. *Journal of Mind and Behavior* 18:155–170
- Kondepudi D, Prigogine I (1998) *Modern thermodynamics: from heat engines to dissipative structures*. Wiley, New York
- Lebon G, Jou D, Casas-Vázquez J (2008) *Understanding non-equilibrium thermodynamics: foundations, applications, frontiers*. Springer-Verlag, Berlin
- Levins R, Lewontin R (1985) *The Dialectical Biologist*. Harvard University Press, USA
- Liu X, Peng L, Zhang J (2019) Accurate and Efficient Calculation of Protein–Protein Binding Free Energy-Interaction Entropy with Residue Type-Specific Dielectric Constants. *J Chem Inf Model* 59:272–289
- Makhnevych T, Houry WA (2012) The role of Hsp90 in protein complex assembly. *Biochim Biophys Acta* 1823:674–682
- Martyushev LM (2013) Entropy and entropy production: old misconceptions and new breakthroughs. *Entropy* 15:1152–1170
- Mayer MP, Nikolay R, Bukau B (2002) Aha, another regulator for Hsp90 chaperones. *Mol Cell* 10:1255–1256
- Nilofer C, Sukhwil A, Mohanapriva A, Kanguene P (2017) Protein-protein interfaces are VdW dominant with selective H-bonds and (or) electrostatics towards broad functional specificity. *Bioinformatics* 13:64–173
- Nishiyama M, Kleijn S, Aquilanti V, Kasai T (2009) Mass spectrometric study of the kinetics of  $O_2$  consumption and  $CO_2$  production by breathing leaves. *Chem Phys Lett* 470:332–336

- Reif F (1965) Fundamentals of statistical and thermal physics. McGraw-Hill International Editions, Physics Series, Singapore
- Rivas G, Minton AP (2018) Toward an understanding of biochemical equilibria within living cells. *Biophysics Rev* 10:241–253
- Saibil H (2013) Chaperone machines for protein folding, unfolding and disaggregation. *Nat Rev Mol Cell Biol* 14:630–642
- Santos GC (2013) Philosophy and Complexity *Found Sci* 18:681–686
- Schmidtke G, Schmidt M, Kloetzel PM (1997) Maturation of mammalian 20 S proteasome: Purification and characterization of 13S and 16S proteasome precursor complexes. *J Mol Biol* 268:95–106
- Schrödinger E (1944) *What is Life*. Cambridge University Press, London
- Sommerfeld A (1955) *Thermodynamics and Statistical Mechanics, Lectures on Theoretical Physics, vol V*. Academic Press, New York
- Sowmya G, Breen EJ, Ranganathan. (2015) Linking structural features of protein complexes and biological function. *Protein Sci* 14:1486–1494
- Swenson S (1977) Autocatakinetics, evolution, and the law of maximum entropy production: a principled foundation towards the study of human ecology. *Adv Hum Ecol* 6:1–47
- Szent-Gyorgyi A (1977) Drive in Living Matter to Perfect Itself. *Synthesis* 1:14–26
- Tanford C (1973) *The hydrophobic effect: formation of micelles and biological membranes*. Wiley, New York
- Tooby J, Cosmides L, Barrett HC (2003) The second law of thermodynamics is the first law of psychology: Evolutionary developmental psychology and the theory of tandem, coordinated inheritances: Comment on Lickliter and Honeycutt. *Psychol Bull* 129:858–865
- Toussaint O, Schneider ED (1998) The thermodynamics and evolution of complexity in biological systems. *Comp Biochem Physiol a: Mol Integr Physiol* 120:3–9
- Turner J, Soper AK, Finney JL (1990) A neutron diffraction study of tetramethylammonium chloride in aqueous-solution. *Mol Phys* 70:679–700
- Vannini A (2005) Entropy and Syntropy, from mechanical to life sciences. *Neuroquantology* 3:88–110
- Wolchover NA (2014) New Physics Theory of Life. *Quanta*. <https://www.quantamagazine.org/a-new-thermodynamics-theory-of-the-origin-of-life-20140122/>
- Yang H-C, Ge Y-C, Su K-H, Chang C-C, Lin K-C, Aquilanti V, Kasai T (2021) Temperature effect on water dynamics in tetramer phosphofructokinase matrix and the super-Arrhenius respiration rate. *Sci Reports* 11:1–14
- Zhang J, Liu JS (2006) On side-chain conformational entropy of proteins. *PLoS Comput Biol* 2:e168
- Ziegler Z (1963) Some extremum principles in irreversible thermodynamics with application to continuum mechanics. In: Sneddon J, Hill R (eds) *Progress in solid mechanics*, vol 4. North Holland, Amsterdam, p 91
- Zotin AA (2014) Why Linear Thermodynamics Does Describe Change of Entropy Production in Living Systems? *Nat Sci* 6:495–502

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Authors and Affiliations

Robert C. Jennings<sup>1,2</sup> · Erica Belgio<sup>3</sup> · Giuseppe Zucchelli<sup>1,2</sup>

✉ Robert C. Jennings  
robert.jennings@unimi.it

Erica Belgio  
erica.belgio@gmail.com

Giuseppe Zucchelli  
giuseppe.zucchelli@unimi.it

<sup>1</sup> Consiglio Nazionale delle Ricerche, Istituto di Biofisica, via Giovanni Celoria 26, 20133 Milan, Italy

<sup>2</sup> Dipartimento di Bioscienze, Università degli Studi di Milano, via Giovanni Celoria 26, 20133 Milan, Italy

<sup>3</sup> Institute of Microbiology, Academy of Sciences of Czech Republic, Centre Algatech, Opatovický mlýn, 37981 Trebon, Czech Republic