Biophysics and Physicobiology
https://www.jstage.jst.go.jp/browse/biophysico/

Review Article

Statistical mechanics of protein structural transitions: Insights from the island model

Yukio Kobayashi

1Department of Information Systems Science, Faculty of Science and Engineering, Soka University, Hachioji, Tokyo 192-8577, Japan

Received August 2, 2016; accepted September 5, 2016

The so-called island model of protein structural transition holds that hydrophobic interactions are the key to both the folding and function of proteins. Herein, the genesis and statistical mechanical basis of the island model of transitions are reviewed, by presenting the results of simulations of such transitions. Elucidating the physicochemical mechanism of protein structural formation is the foundation for understanding the hierarchical structure of life at the microscopic level. Based on the results obtained to date using the island model, remaining problems and future work in the field of protein structures are discussed, referencing Professor Saitô’s views on the hierarchical structure of science.

Key words: protein structural transition, island model, protein folding, protein structure prediction, hydrophobic interaction

Professor Nobuhiko Saitô was a Japanese theoretical physicist almost the same age as the well-known American physicist Richard P. Feynman and a Japanese theoretical physicist Ryogo Kubo, and contributed significantly to the development of interdisciplinary work in statistical mechanics and biophysics. The research method of Professor Saitô was based on his wide range of insight, and thus transcended conventional boundaries between and within physics and its related areas. Until just before his passing, Professor Saitô was focused on establishing a theoretical method for the prediction of protein structures predicated on physicochemical mechanisms, one of the most important goals in both biophysics and physicobiology. Biophysics is an interdisciplinary field that has developed as a combination of traditional sciences such as biology, physics, chemistry, applied mathematics, computer science, and genetics. The ultimate purpose of biophysics is to reveal the mechanisms of biological phenomena, which can be roughly classified into ecology and evolutionary biology, as well as problems explained in terms of the physical and chemical properties of biomolecules. The latter is closely related to molecular biology, which aims to decipher the genetic information of genomes and to develop methods for regulating the expression of genetic information. The current biological classification systems show that hierarchical levels of life can be recognized. The smallest units of biological systems are molecules such as amino acids and nucleotides, but macromolecules also play an essential role as the functional and structural building blocks of cells and their organelles. Ecosystems are also organized in a hierarchical fashion, albeit at a higher level.

This review summarizes the genesis and statistical mechanical basis of the island model of protein structural transitions. The results obtained to date using the island model, remaining problems, and future work in the field of protein structures are discussed, referencing the late Professor Saitô’s views on the hierarchical structure of science. The island model holds that hydrophobic interactions are the key to both the folding and function of proteins.
The physicochemical properties of biopolymers such as proteins and DNA are vital to elucidating the origin of life, since these molecules govern the processes of life. For this reason, both experimental and theoretical approaches have been used in attempts to explain the relationship between the structure and function of biomolecules. Professor Saitô and his collaborators have elucidated the mechanisms by which protein molecules are formed at the microscopic level in a hierarchical fashion, based on a theoretical approach.

The principal methods of theoretical biophysics are continuum, statistical, and quantum mechanics, and cooperation among these approaches as well as the involvement of computer science and other technologies are also important. In the 1960s and 1970s, Professor Masao Kotani’s laboratory in Japan actively studied the physical properties of biological molecules by directly comparing numerical values obtained by quantum mechanical methods with experimental data, and related these physical properties to biological activities [1]. Their achievements showed that molecular biology has some common ground with physics at the microscopic level. The functional properties of biomolecules are known to depend on their three-dimensional structures [2]. In this regard, the specific spatial conformation of a protein results from the folding of particular sequences of amino acids in polypeptide chains to generate three-dimensional structures with compact globular units. The three-dimensional structures of most proteins, however, have not yet been elucidated, and thus it remains difficult to understand how a protein performs its individual function. Since Kendrew et al. produced the first three-dimensional image of myoglobin in 1958, the X-ray analysis of protein molecules has been increasingly improved [3] and, as of 2016, the Protein Data Bank (http://www.rcsb.org/pdb/home.do) maintains a database of approximately 120,000 experimentally determined protein structures. In spite of these rapid advances in experimental techniques for solving protein structures, however, the structures of many proteins have not been determined, and so the assistance of programs that predict protein structures is still necessary, particularly in the fields of medicine and biotechnology. Methods for protein structure prediction can be roughly classified into three approaches: comparative modelling, fold recognition, and ab initio modelling. Although there have been great advances in computer science and related technologies, the accuracy of structural predictions is still not satisfactorily accurate to allow prediction of detailed structures. Since the 1980s, Professor Saitô and his collaborators have developed a physicochemical ab initio method for predicting the three-dimensional structure of a protein from its amino acid sequence. Professor Saitô, who studied at Kotani’s laboratory at Tokyo Imperial University, focused on assessing protein structures via a statistical mechanical approach, while Professor Kotani studied the physical properties of biological molecules by a quantum mechanical approach. Both these approaches had a common basis in that they took advantage of recent developments in computer technology, although their research subjects in the field of theoretical biophysics were different, since Professor Saitô concentrated on the thermodynamic mechanisms of the structural transitions of proteins instead of quantum mechanical phenomena in biological systems. The details of this work are presented herein, in the section concerning the statistical mechanical foundation of protein structural formations. Professor Saitô also studied the physical properties of polymers in his early works, and contributed Sections 31 and 73 (concerning the thermodynamics of dielectric substances) to a textbook on the theory of dielectric substances by Syoten Oka [4], although he is not formally listed as a coauthor. Professor Saitô’s insights into protein functions will be described in the concluding remarks section herein.

This review article focuses primarily on the initial development and achievements of the statistical mechanical theory termed the island model, with regard to predictions of protein structures. Professor Saitô contended that the tertiary structures of globular proteins for which the positions of secondary structures on the polypeptide chains are known can be predicted by the island model, although the development of this prediction method is still a work in progress. He believed that future work would consist of the development of a statistical mechanical method for secondary structure predictions based on the physicochemical mechanisms of secondary structure formation, even though this opinion has been criticized as premature [5]. The purpose of the present manuscript is to describe the core theories of Professor Saitô through presenting the island model, as well as to provide guidance for the future development of protein structure prediction methods.

Genesis of the island model: from polymers to proteins

Professor Saitô’s early work was focused primarily on the study of polymer physics using a statistical mechanics approach. In this respect, he was inspired by Oka’s suggestion that polymer physics would be the basis of future developments in biophysics [6]. Professor Saitô studied irreversible macromolecular phenomena involving viscosity, viscoelasticity, sedimentation, and electrical conductivity, until the early 1960s [7–10]. Subsequent research subjects primarily consisted of the mechanisms by which polypeptide conformations were generated through helix-coil transitions [11–14]. This prior work is summarized in Professor Saitô’s exceptional textbook on polymer physics [15], which provides some information concerning his research into protein conformations. In the 1970s, his research interests expanded to include the statistics of random coil chains, the stability of α-helical polymer structures, and the widths of polymer chains [16–22]. These studies were followed by investigations of intramolecular α-helix-β-structure-random coil transitions in polypeptides [23,24].
Professor Saitô had foresight and eventually began to consider protein structures, likely because he was able to utilize the ever-increasing information regarding protein structures provided by X-ray analysis. He also took an interest in Anfinsen’s thermodynamic hypothesis, proposed in the 1950s [25], and in the helix-coil transition of polypeptides discovered by Wada et al. [26], although Oka had also previously studied proteins such as silk fibroin [6]. Professor Saitô et al. developed a map representing the native conformations of proteins [27] as described in the next section. Advances in the representation and computation of protein conformations are applicable to fundamental studies of the tertiary structures of proteins. The basis of Professor Saitô’s work on protein structures was the statistical mechanical treatment of \( \alpha \)-helices and extended structures with short-range and medium-range interactions [28]. The statistical mechanics of a one-dimensional system undergoing first-order diffuse phase transitions [29] was developed to obtain a statistical mechanical theory of the folding pathways of proteins [30,31], thus generating the island model [28].

In the 1980s and 1990s, Professor Saitô and his collaborators investigated the unfolding pathways of protein structures [32], and then applied the resulting prediction method based on the island model to typical proteins to explain the folding pathways of these proteins [33–39]. His various coworkers wrote programs to assist in these studies, using languages such as Fortran, BASIC, and Pascal. The coworkers simulated the folding process of an extended polypeptide chain with secondary structures at the same positions as in its native structure. The development of statistical mechanical methods for secondary structure prediction continues even now. In the next section, we examine this work in detail from the viewpoint of the statistical mechanics of transition phenomena, since Professor Saitô claimed that we must return to the foundation of statistical mechanics to elucidate the mechanisms of protein folding [40,41].

Statistical mechanical foundation of protein structure formation

Professor Saitô described the genesis of the island model on July 24, 1992, in a seminar at the Hitachi Advanced Research Laboratory. He had the idea of applying statistical mechanics to protein architecture from work performed on the helix-coil transition that he encountered during his stay at the University of Oregon in the 1960s [42]. At that earlier time, it was believed that quantum mechanics was essential to the understanding of living systems, while statistical mechanics was not considered very effective because living beings each had their own unique characteristics. Eventually, however, it was anticipated that various phenomena associated with living organisms, such as the adjustment of body temperature and blood electrolyte concentrations, could be compared to the helix-coil transition. This idea was also related to Anfinsen’s thermodynamic hypothesis [25,43]. In the 1950s, Anfinsen proposed that the chemical properties of the amino acid sequence of a protein determine its tertiary structure, based on experimental investigations of the reversible denaturation of several proteins. Professor Saitô considered that the elucidation of the mechanism of helix-coil transitions would also provide new insights into the above phenomena.

Phase transitions indicate a particular change in the thermodynamic variables at a transition point, such as in temperature or concentration, which reflect the characteristics of the substance. Essentially, the transition point is an effective means of maintaining a thermodynamic system at a specific temperature, and so phase transitions tend to induce stability and have a regulatory effect on biological functions. Phase transitions result from cooperative phenomena; that is, changes in the molecular states of a system take place rapidly in response to relatively small changes in conditions such as temperature and concentration that induce only slight variations in molecular states. Phase transitions are also associated with the thermodynamic limits, that is, the limit of an infinitely large number of particles and the volume with the density kept constant. No phase transitions can be expected to occur in a one-dimensional system with finite range interactions [44]. A biological molecule is a polymer chain with a finite degree of polymerization, and thus the helix-coil transition in polypeptides and the formation of tertiary structures in proteins are not, strictly speaking, phase transitions but rather transition-like phenomena known as diffuse transitions [40,41,45]. These phenomena are connected with sharp changes in state and result in biological functions such as catalysis and oxygen dissociation. Thus, a protein conformation can be considered to be a specific ordered structure favorable to a specific biological function. Transitions yield new, stable structures with different physical properties and functions. A decrease in the free energy is always guaranteed by the law of thermodynamics and the chain entropy decreases due to the formation of an ordered structure. Thus the energy also decreases inevitably. Professor Saitô considered proteins to be masterpieces of the natural world [46], since they are exceptional molecular machines responsible for various aspects of life, such as biosynthesis, heredity, and evolution. In addition, even a single protein molecule or a congregation of organized protein molecules can exhibit biological functions.

Based on Anfinsen’s thermodynamic hypothesis [25,42], Professor Saitô proceeded to develop the use of statistical mechanics to model the formation of protein structures. His main goal was to establish a method for finding the minimum free energy state of a protein, considered as a thermodynamic system. The essence of this problem is closely related to the ergodic hypothesis in statistical mechanics, which states that the average of a physical quantity over time is equal to the average of the quantity over the statistical ensemble. This hypothesis suggests that all accessible microstates with the same energy in the phase space are equally
probable when observed over a reasonably long period of time. Levinthal noted that conventional statistical thermodynamics is not applicable to the conformation space of a protein [47], even though the conformation of a protein is governed by equilibrium thermodynamics. Professor Saitô explained Levinthal’s paradox by estimating the time required to survey all states of a protein composed of 100 amino acid residues to be $10^{75}$ seconds, working only in the configuration space rather than the momentum space [48]. This simple estimation indicates that the configuration space of a protein is restricted from the dynamical point of view, and thus the minimum free energy state of a protein cannot be searched for over the entire configuration space but rather must be approached through a specific pathway over a restricted space [49]. Professor Saitô explained that the statistical mechanics of an ensemble of identical particles such as an ideal gas or a homopolymer also holds true for a small phase space over a reasonable amount of observation time [41].

In the early stages of this work, there were two significant problems. The first challenge was to find the folding pathway so as to restrict the configuration space, while the second was to elucidate the principles that determine the folding pathway. Professor Saitô considered that the key was to determine which interactions among the amino acid residues are essential to protein structure formation. The essence of his approach can be summarized by the phrase “Look deep into the native structure of a protein” [46], which is similar to Einstein’s maxim “Look deep into nature, and then you will understand everything better.” He took note of the mutual range interactions between amino acid residues, because interactions among residues can be roughly expressed as a function of distance [50]. Nishikawa et al. collaborated with Professor Saitô and developed a method of representing the native conformations of proteins [27]. They introduced a two-dimensional map representing the patterns of the mutual distances between the $\text{C}^\alpha$-atoms of $i$th and $j$th residues, which is not affected by viewing direction. The original map was of the form shown in Figure 1 (a), although this version is little known. In a triangle distance map, the mutual distances are listed against the residue numbers in rows and columns, as in Figure 1 (b). The triangle distance map is a convenient means of comparing a native conformation with a simulation and comparing conformations among native homologous proteins. However, a more important issue than the comparison of conformations is elucidating the driving force behind folding, and the triangle distance map provides essential information that is useful in this respect.

Professor Saitô examined the growth types of folding processes to explain rapid folding as a means of solving the first problem regarding understanding the folding pathway to restrict the configuration space. He noted amino acid residues between the diagonal and parallel line associated with small separation distances in the triangle distance map, where the separation distance represents the number of amino acid residues intervening along a polypeptide chain. The folding pathway is determined by intramolecular interactions, and so he also investigated the types of interaction that are relevant to the helix-coil transition. As such, it was determined that protein structural transitions can be explained by developing a statistical mechanical theory for the helix-coil transitions of homopolypeptides, regarding a protein as a heteropolypeptide composed of amino acid sequences [40, 41, 50, 51].

Onuki and Professor Saitô explained the helix-coil transition using a model for a one-dimensional lattice gas with repulsive potentials between nearest neighbors and attractive potentials between next-nearest neighbors [29]. This model exhibits a diffuse first-order phase transition and thus can be regarded as a lattice gas version of the helix-coil transition in a polypeptide. The $\alpha$-helix undergoes hydrogen bonding that limits the dihedral angles to definite values. The overall energy is decreased by this hydrogen bonding and, simultaneously, the entropy of the polypeptide chain also decreases. The decrease of the chain entropy can be modeled as a lattice gas with repulsive potential between the nearest neighbors.

Kosuge et al. calculated the interaction energies for the $\alpha$-helix of polyalanine and determined that medium range interactions play an essential role in the helix-coil transition [17]. Nonbonded interactions do not change appreciably with variations in the dihedral angles, and so are not associated with the stability of the $\alpha$-helix. This result provides an important clue as to the factors that stabilize the tertiary structure of a protein. The lattice gas model incorporating long-range interactions is appropriate for predicting protein folding, and later was generalized to include interactions of remote particles to allow its application to globular proteins.

Wako and Professor Saitô extended the statistical mechanical theory of the helix–coil transition to generate a theory of protein structural transitions by considering only interac-
tions among amino acid residues in the native definite structures [30,31]. They termed these structures the “island,” meaning the portions of the polypeptide chains with definite native structures formed via local interactions among amino acid residues [28]. It should be noted that this model is different from the island model used in population genetics, although the two models share the same name. The folding of proteins is considered to proceed first by the formation of small islands of amino acid residues that lie close along a polypeptide chain, after which these small islands grow and merge into larger islands by incorporating neighboring islands and amino acid residues in a random state [48]. The results of computer calculations using this model indicated that the folding order is related to a hierarchy of secondary and tertiary structures. The initial stage of protein folding is the formation of secondary structures, followed by which secondary structures merge into a specific tertiary structure. All-or-none type transitions between two states (a random coil state and a native state) and the presence of intermediate structures similar to the molten globule state in equilibrium were also identified. In this manner, Wako and Professor Saitô successfully elucidated the folding pathway based on the island model.

The second challenge related to determining the folding pathway, however, remained. The mechanism of secondary structure formation is still unsolved, although secondary structures are also known to form definite island structures. If protein folding proceeded only by hydrogen bonding, we would expect the entire polypeptide chain to be folded into a helical configuration. Thus, Professor Saitô turned his attention to the packing of secondary structures already formed and focused on finding the driving force for packing. In this regard, he obtained insights into a particular type of interaction required for rapid folding [48]. Professor Saitô determined that long-range interactions between specific residues was essential to determine a folding pathway, and Professor Saitô et al. noted that hydrophobic residue pairs at short distances were situated near the diagonal in the triangle distance map [34]. Short-range interactions, such as the Lennard-Jones potential, will occur between any given pair of atoms or groups of atoms, and thus this type of interaction is not effective at determining the folding pathway. In contrast, electrostatic forces represent short-range interactions resulting from shielding effects by ions in solution. Hydrophobic residue pairs close to one another on the polypeptide chain are essential to packing of the secondary structures, and small islands are formed at an early stage by interactions among those residues close to the diagonal in the triangle distance map. These hydrophobic residue pairs act to bind two neighboring secondary structures. The hydrophobic residues far from the diagonal become close during the folding process, such that larger islands are formed by their interactions. Professor Saitô recognized the long-range nature of hydrophobic interactions, in contrast to other researchers who had pointed out the importance of hydrophobic interactions but had not noted their long-range nature. The legitimacy of these long-range nature of hydrophobic interactions was supported by a timely publication from Israelashvili and Pashley [52].

Simulations of protein folding based on the island model

In this work, Professor Saitô relied heavily on computer calculations, likely because he had previously succeeded in computer simulation of nonlinear lattice oscillations with Hirooka et al. [53–56]. This work was closely related to the ergodic hypothesis in statistical mechanics, and he considered that a random dynamic process was at work and so probabilistic solutions were possible even when working with mechanics, because statistical mechanics is based on mechanics. This early work identified non-periodic random solutions to the high energy states of nonlinear lattice oscillations. Those solutions complemented work on the Fermi-Pasta-Ulam problem [57], in which complicated dynamical systems exhibit almost exactly periodic behavior instead of ergodic behavior. The simulation developed by Professor Saitô et al. indicated a chaotic state, long before the technical term “chaos” became routine in the literature. At this time, he developed the concept that dynamical systems with reversibility and thermodynamic systems with irreversibility can coexist through the existence of chaos.

He had anticipated that the development of the computer would contribute to the study of protein folding problems, and simulated the folding of typical globular proteins based on the island model [58,59]. The problem was to explain the packing of secondary structures which were known from native structures or other information. This approach assumed that secondary structures have a tendency to form earlier than the specific conformation resulting from hydrophobic interactions among close hydrophobic residues between neighboring secondary structures. There have been some discussions about the validity of this assumption [60], as described in the section concerning the remaining problems to be solved. However, the mechanism of secondary structure formation was ignored, and the amino acid residues were approximated by spheres of appropriate radii located at the necessary positions. The simulations searched for the minimum energy state of the proteins rather than the overall minimum free energy state, because the change in the chain entropy is due to small vibrations around the equilibrium position in the tertiary protein structure [35,50], also described later. They applied hydrophobic interaction between the side chains of the hydrophobic residues. They also introduced the Lennard-Jones potential between non-bonded atoms, with the exception of hydrogen atoms, to avoid steric hindrance. Other interactions were also not considered, although disulfide bonding between the side chains of relevant cysteines was introduced. In this manner, the effect of the chain entropy was automatically included as a result of the choice.
of the nearest specific hydrophobic residues, as described later, although the folding simulation is kinetic in essence [61]. The entropic effect of the water structure is also taken into consideration in hydrophobic interactions [35]. Thus the entropic effects of the thermodynamic origin could be neglected in the folding simulation.

The triangle distance map acted as a guide to the folding pathway during simulations of folding. During the initial stage of refolding, hydrophobic residues close to one another on a polypeptide chain were assessed, because the hydrophobic interactions among these residues are responsible for packing of neighboring secondary structures. Interactions among residues distant from each other were not considered at this stage, even if some of these residues approached one another in threedimensional space due to fluctuations of the chain. The packing order of the secondary structure was determined using this approach. The minimum energy state was searched for at every stage of the conformation by considering the hydrophobic residues responsible for packing. This folding process is consistent with a scenario in which a random coil region incorporated into an existing island is an isolated structure obtained by simulation [61]. The entropic effect of the water structure is also taken into consideration in hydrophobic interactions [35]. Thus the hydrophobic interactions responsible for the packing of secondary structures are essential to folding, independent of the sequential homology. Lysozyme and phospholipase [35,36] were also assessed as a guide to selecting the correct cysteine pairs when working with proteins having disulfide bonds. Bovine pancreatic trypsin inhibitor [37] was simulated and the folding pathways in intermediate structures with various disulfide bonds were explained by the island model. Ferredoxin [39] was simulated by using the secondary structures that were predicted by the three-state model described in the next section. The packing of the predicted secondary structures indicated the possibility of predicting the tertiary structure from an amino acid sequence.

On November 17, 1990, Professor Saitō gave a lecture entitled “Protein: reading and decoding the language of amino acid” as part of the Nishina Memorial Lectures at Waseda University. During these lectures, he noted that researchers working in thermodynamics [41] had been “heart-shocked” by the discovery of heat shock proteins, a jest that was indicative of his warm and humorous personality. Some heat shock proteins assist in vivo in establishing the proper conformation by preventing unwanted protein aggregation during folding. However, Anfinsen’s thermodynamic hypothesis based on in vitro experiments states that protein folding proceeds spontaneously without any enzymatic assistance. Professor Saitō believed that the restriction of the free motion of a polypeptide chain does not cause any difficulty in achieving the proper folding because hydrophobic residues close to one another on the chain are bound at an early stage of the folding process.

Encouraged by this insight, Professor Saitō et al. applied the procedure shown in Figure 2 to the prediction of the unknown structures of a parathyroid hormone-related protein (residue 1–34) [62] and the mutant protein of cytochrome $b_{562}$, with a disulfide bond introduced [63]. The latter protein had been researched during work on bioelectronic materials as part of the Frontier Research Program of RIKEN, at which Kobayashi worked during that time period.

In concluding this section, we note the physical meaning of the island model from both equilibrium and kinetic perspectives. Professor Saitō used the term “island model” from either point of view [28,40,41,61,62]. As described in the previous section, an island is an ordered local structure with no interactions between residues existing in different islands. This structure is an isolated structure obtained by simulation or a part of the native structure such as known secondary structure. In the statistical mechanical treatment of the island model [30,31,61], the specific intraisland-interactions in the native protein and the entropic effect in binding remote hydrophobic residues were taken into consideration. The structural transitions of proteins were explained as a result of using the method to follow the folding pathway based on the island model. The hydrophobic interaction between the specific hydrophobic residues is essential to the formation of tertiary structure. The stability or instability of intermediate
structures was understood by the effect of the chain entropy. The results indicate the validity of the island model for describing the process of protein folding, and thus the island model was implemented to the computer simulation of protein folding. In this kinetic simulation, the effect of the chain entropy was taken into account as the folding order of stepwise mechanism determined by the hydrophobic interaction between the specific hydrophobic residues. The initial conformation was taken as an extended chain with the secondary structures of the native protein, because the method for predicting secondary structures of proteins has not yet been established. In this treatment, the native interactions were automatically taken into consideration in the secondary structures. The secondary structures incorporate nearer islands to merge into larger islands. In these processes, hydrophobic interactions and nonbonded interactions are applied between short-distance residues at an early stage and then between long-distance residues. In other words, nearer residues in each larger island interact with each other, and thus these islands become stable by primarily hydrophobic interactions as well as nonbonded interactions.

Remaining problems to be solved

The remaining problems associated with the island model can be classified into two main categories: the search for the minimum energy state and the establishment of a method for secondary structure prediction. Both of these remain matters of controversy among researchers.

The search for the energy minimum state is a problem of multivariable optimization. Professor Saitô claimed that Anfinsen’s thermodynamic hypothesis [25,43] holds true in a restricted space dominated by hydrophobic interactions among hydrophobic residues at key positions on a polypeptide chain. The restricted space is determined by proceeding from the formation of local structures among amino acid residues located at short distances along the chain. This approach requires significant computational time to obtain a final structure when using the island model, because energy minimization is performed in associated with steps of local structure formation. At the time, Professor Saitô commented on the physical meaning of the energy minimum states that other researchers were attempting to identify. Firstly, he stated that the role of hydrophobic interactions is not considered in the method of simulated annealing [41]. Secondly, he felt that rapid folding cannot be explained by the funnel hypothesis based on the energy landscape theory [6,61,64,65]. However, the statistical mechanical theory developed by Wako and Professor Saitô [30,31] played an important role in the research of protein folding funnels [66], and it appears that their theory was a useful means of identifying interacting amino acid residues. Thirdly, Professor Saitô contended that the physicochemical mechanism of folding is not properly considered in most research studies, even if the prediction accuracy of protein tertiary structures is high. This opinion did not necessarily agree with the ab initio prediction method based on the cooperation of protein 3D-1D compatibility analysis [67].

Professor Saitô did not necessarily object to a funnel structure in the energy landscape of a protein for rapid folding. He interpreted the funnel as the collection of the folding pathways over all the initial random conformations [61]. The folding pathways through the step-by-step growth process shown in Figure 2 are determined by the specific hydrophobic interaction at every stage of the local structure formation. Thus the folding pathway and the folding order depend on the initial state at every stage of a conformation. Professor Saitô’s view means that the funnel for rapid folding does not yield the energy minimum state but the folding pathway constructs the funnel [61]. The pair of hydrophobic residues is not assessed at random, but the minimum energy state at every stage of folding is searched for among the possible deformations determined by the specific hydrophobic residues. The conformation space of a protein is restricted to a valley-and-hill region with a small part by the dihedral angles concerned in the deformation of the conformation at every stage of folding, although the whole space is spanned by all the dihedral angles in kinematic perspective. In equilibrium perspective, the probability of the binding of residues distant from each other is low due to the effect of the chain entropy. The chain entropy decreases due to contact between two hydrophobic residues by the long-range nature of hydrophobic interaction. This view indicates that the folding order in the kinetic process and the statistical thermodynamic treatment are taken into account simultaneously. The stepwise mechanism in folding agrees with the stability or instability of the intermediate structures, because two-state transition is not observed experimentally during folding of some proteins.

As pointed out by Mitaku, the mechanism of secondary structure formation remains unsolved [68], and an essential question remains as to why some amino acid sequences exhibit a propensity for the formation of $\alpha$-helices while other sequences tend to generate $\beta$-strands. The relationship between secondary structures and amino acid sequences are closely related to the mechanism of tertiary structure formation, and Mitaku claimed that these relationships found by a statistical approach should be explained on a physicochemical basis. Professor Saitô therefore devised a statistical mechanical theory of the helix-coil transition [30,31] as a method for secondary structure prediction. He presented formalisms based on recurring relationships [42] rather than the formulation of the island model performed using a matrix method [28,30,31]. At this time, Yura wrote a program allowing computer calculations of the partition function of a protein and the probabilities of each amino acid residue in an $\alpha$-helix or $\beta$-strand. Based on the central concept of the island model, the $i$th residue was assumed to interact with the $(i+k)$th residues in the same $\alpha$-helix or $\beta$-strand, where the distance of the interactions was taken to be $0 \leq k \leq 4$. In
this formulation, a set of 3240 parameters representing statistical weights was necessary, and the values of these parameters were determined so as to predict the native secondary structures of the proteins without duplication of amino acid sequences. Yura et al. used a two-state model to predict α and non-α (β or coil) states or β and non-β (α or coil) states, and developed a three-state model to predict α, β, and coils simultaneously [48]. Together, they estimated the prediction accuracy of the secondary structures of proteins other than the reference proteins used for determination of the values of the parameters, and it was found that the parameters of the amino acid residue pairs not observed in the secondary structures of the reference proteins could not be optimized properly. Kobayashi et al. took over the optimization by adding the reference proteins and consequently the prediction accuracy of the reference proteins was improved to 78.282% [69–71]. To further improve the prediction accuracy, the interaction distance was taken to be ≤4 for α-helices and 0≤k≤2 for β-strands, where the parameters of the i th and (i±k)th residue pairs (k=3, 4) may not be different from 1. The determination of the values of the parameters, however, was not yet complete, and so the prediction accuracy of the reference proteins was 70.486% at this stage of the optimization process (unpublished data).

Professor Saitô next turned his attention to the growth of secondary structure elements during the folding process. He considered that the prediction of the secondary structures initially formed in the unfolded state was necessary for the simulation of folding. Thus the essence of secondary structure prediction is not to predict secondary structures in a completed tertiary structure. His view was that secondary structures cannot be predicted based solely on amino acid sequences but rather a prediction of a tertiary structure is required. This insight is based on a statistical mechanical theory of protein structural transition [61], which was developed in accordance with the island model by considering hydrophobic interactions and entropic factors while connecting two hydrophobic residues. In previous works [30,31], hydrophobic interactions and chain entropy were not considered, because the role of hydrophobic interactions was not yet completely understood [41]. In contrast to the Lifson-Roig model [72] and the Zimm-Bragg model [73], the role of hydrophobic interactions in protein structural transitions was taken into account in the island model. Chain entropy is a small quantity dependent on the volumes of interacting hydrophobic residues, and thus brings about the possible occurrence of unstable hydrophobic interactions and unstable intermediate structures [61]. The intermediate structures of proteins with many hydrophobic residues are, however, stable because the neighboring hydrophobic residues at shorter distances from one another on the polypeptide chain result in a relatively high entropy and many hydrophobic interactions from a standpoint of thermodynamic treatment. This process is related to the formation of secondary structures and the order of the packing of secondary structures [60]. The island model does not assume that secondary structures assemble after the formation of all secondary structures [74]. During the process of folding, the i th and (i+k)th hydrophobic residues are bound, after which the secondary structure is formed between the (i−l)th and (i+k+l′)th residues, with the ith and (i+k)th hydrophobic residues unbound. This theory is supported by the fact that unbound hydrophobic residue pairs are observed in the secondary structures of the native structures of some proteins. The ith and (i+k+l′+k′)th hydrophobic residues are bound easily after the formation of this secondary structure in the kinetic folding process, because the possibility of bonding depends on the number of residues in between. After secondary structure is formed between the (i−l)th and (i+k+l′)th residues, the ith and (i+k+l′+k′)th residues can be bound in the process of deformation of the part consisting of (k−1) residues instead of (k+l′+k′−1) residues. The dihedral angles of (k−1) residues are distorted in this process. Thus the island model assumes that secondary structures have a tendency to form earlier than the specific conformation resulting from the hydrophobic interaction between these hydrophobic residues.

Based on the above theory of protein structural transitions, Professor Saitô considered that hydrophobic interactions also play a significant role in secondary structure formation. Hydrophobic interactions, in addition to hydrogen bonding, are present between the ith and (i+4)th hydrophobic residues in an α-helix, and the hydrophobic interaction between the i th and (i+3)th hydrophobic residues is also important to the stability of the helix. Interactions between amino acid residues in a β-strand must also be considered, because hydrophobic residue pairs are not observed in some strands. In contrast to the determination of a huge number of parameters in previous works [70,71], the optimization now required only the determination of the parameters of hydrogen bonding between the ith and (i+4)th residues and the ith and (i+k)th hydrophobic residue pairs (k=3, 4) in regard to the α-helix. The key information required is the characteristics of the hydrophobic residues on the α-helices of proteins with known crystal structures selected from the Protein Data Bank (http://www.rcsb.org/pdb/home/home.do). In this manner, Professor Saitô developed a predictive method based on rules concerning the hydrophobic residue pairs in secondary structures. Unfortunately this method was never published. The rules regarding the formation of antiparallel β-sheets from neighboring β-strands were also never solved. In February 2012 Seimei no Butsuri was published by Iwanami in a new format, and Professor Saitô added two references concerning recent progress in protein structure prediction.

Professor Saitô as an educator

All students begin their studies looking up to their professor, and Professor Saitô was an excellent educator, although he himself may not have said so. During lectures and semi-
ners, he would often become silent as he was deep in thought. Many of us enjoyed these opportunities to observe his contemplation, although it is true that not all students necessarily took a favorable view of such pauses. The purpose of education is not necessarily to impart knowledge but to provide intellectual training to students, and in this regard we received a proper education from the style of his lectures. Through his lectures, Professor Saitō definitely communicated the concepts of physics and indeed science in general. Unfortunately, modern university education in Japan runs counter to this style.

In his essay [75], he stated that the world of knowledge always excites and impresses people beyond description, and expressed his doubts concerning the lyrics of the Japanese singer Kei Ogura, which indicates that all things are vacuous after youth. In addition, this essay looked back on his high school days, and expresses his idea that studying means encountering professors who take the time to prepare for lectures.

He also described two views on elementary mathematics education in this essay [75]. Firstly, that teaching finite sets using the technical terms of set theory does not have any educational benefits, because the concept of an infinite set is concerned with essential ideas such as countability and one-to-one correspondence. Secondly, the essence of geometry is not the intuitive generation of auxiliary lines but rather the logical construction of definitions, axioms, theorems, and corollaries. He considered that Euclidean geometry is effective at teaching the construction of concepts and that physics is difficult for junior high school students. In contrast to these views, however, a geometric analogue based on pictures or diagrams can allow students to see at a glance why the theorems of algebra, calculus, and other fields of mathematics are true [76–78]. Einstein and Poincaré held the view that we should use visual intuition [76], and Professor Saitō in fact devised the concept of the island model by visual intuition based on the triangle distance map. Logical configuration and visual intuition therefore have complementary roles, although training in logical thinking is essential to elementary mathematics education.

Concluding remarks

Professor Saitō took a broad view of science on the basis of statistical mechanics, and contributed much to establishing the foundations for a wide range of theoretical biophysics studies. Among other researchers, John Gamble Kirkwood also studied the fundamental principles of statistical mechanics and applied these to polymer physics and, in addition, Professor Saitō maintained close relations with Terrell L. Hill, the American theoretical physicist, physical chemist, and molecular biologist. Hill also worked in a wide variety of research areas, mainly in statistical mechanics, theoretical molecular biology, and biochemistry [42]. There were two main research groups in Professor Saitō’s laboratory: that working with protein in the field of biophysics and that studying chaos in dynamical systems. When I asked him about the relationship between these two research subjects in 1982, his answer was noncommittal, although he published an essay on his views in October 1983 [79]. He felt that both subjects are necessary for an understanding of the natural world, because both are at the boundary of statistical mechanics: biology and dynamics are located at the upper and lower levels of thermodynamics, respectively. He had an interest in phenomena associated with living entities, such as the movement of microorganisms, as well as the exploration of animals and the human thought process. These biological activities indicate that living creatures actively make use of fluctuations in physical variables at the macroscopic level, in contrast to the behavior of systems that passively obey the fluctuation-dissipation theorem. Professor Saitō considered that these fluctuations are actively generated by chaos due to taxes taking part in chemical reactions and neural integration [80]. From the standpoint of the chaos of dynamical systems, he also explained the stability and instability of the coexistence of two well-characterized and phylogenetically different species, such as Escherichia coli and Dictyostelium discoideum [81]. Thus biology and chaos were connected in his mind.

For this reason, his insights into the biological world were not restricted to the microscopic level of biomolecules in the hierarchic structure of science, as described in Introduction, although he primarily studied the mechanism of protein folding in the field of theoretical biophysics. He believed that living things do not preserve order solely through the equilibrium structures of biomolecules [45]. In fact, the thermodynamic equilibrium of a living body equates with death, which indicates that its order is preserved by a supply of energy and matter from the external surroundings. Stationary systems with characteristics independent of time yield ordered structure different from equilibrium systems, and morphosis in a living body is associated with dissipative structure formation. This, in turn, results from chemical reactions in an open system operating out of thermodynamic equilibrium with the external surroundings via the exchange of energy and matter. He expected that Prigogine’s works on dissipative structures, complex systems, and irreversibility could be applicable to the mechanism of morphosis. He also believed that irreversible morphosis with a selection direction is to dynamical molecular evolution what irreversible thermodynamics with a signpost of change is to reversible dynamical systems showing chaotic behavior on an infinitesimal time scale [82]. Morphosis, however, can be observed only on large scales in time and space, whereas molecular evolution frequently occurs over infinitesimal time intervals [49]. His view was that the formation and evolution of life were propagated by the so-called selfish gene through the chaotic process of morphosis due to the molecular evolution of many proteins. He also explained the movement of bacteriophage χ based on fluid dynamics with the help of the
numerical calculations by Takasugi [83–85]. Fluid dynamics operates on macroscopic time and space scales, in the same manner as the irreversible changes involved in diffusion and heat transfer.

Studies of protein structures at the microscopic level are a starting point for the elucidation of the science of hierarchy: dynamical systems, thermodynamic systems, and biological systems accompanied with morphosis and molecular evolution. In this regard, chaos provides a connection between dynamics and thermodynamics and, in fact, Atkins explains that the ordered structures of proteins are also formed by the constructive power of chaos [86]. The emergence of the specific structure of a protein indicates a local abatement of chaos due to a behavior similar to that of oil molecules in water and, in compensation, an equivalent amount of chaos is generated in the surrounding water so as to increase the chaos of the universe. Thus Atkins also concentrates on the important role of hydrophobic interactions in the formation and stability of the tertiary structure of a protein although, in contrast to Professor Saitô, he does not consider the long-range nature of hydrophobic interactions. Only hydrogen bonds were considered in the early theories of the helix-coil transition of a polypeptide chain [87], because the role of hydrophobic interactions in the structural stability of polymers was not fully understood. Based on the statistical mechanical theory of protein conformations and transitions [61], however, Professor Saitô determined that hydrophobic interactions between hydrophobic side chains are more important than hydrogen bonds.

Professor Saitô expected that the elucidation of protein functions would constitute the future work on proteins at the microscopic level, in a hierarchical fashion. Both the dynamics of a protein structure and the quantum mechanical behavior of electrons and protons take part in protein functioning. According to his view, the physicochemical mechanism of protein structure formation has not been properly considered in previous works on protein dynamics. Hydrophobic interactions are the driving force behind protein folding, and thus he anticipated that the deformation required to sever strong hydrophobic bonding was essential to the dynamical behavior of protein molecules. He also suggested that the roles of hydrophobic interactions are different in structure formation and functional expression. The island model therefore holds that hydrophobic interactions are vital to both protein folding and function, indicating that the constructive power of chaos governs the processes of life.

Acknowledgements

I would like to dedicate this review article to the late Professor Saitô, whose motivation, enthusiasm for research, and immense knowledge were truly inspiring. Even now, his work on a wide variety of subjects continues to inform our research methods. I also wish to express my sincere thanks to Professor Key Yura of Ochanomizu University and Professor Motonori Ota of Nagoya University for providing information about the publication of the special issue. Finally, I would like to thank Professor Yasushi Matsumaga of Waseda University for providing the list of publications of Professor Saitô.

Conflict of Interest

The author declares that he has no conflict of interests.

Author Contributions

Y. K. contributed to review the work of the late Professor Saitô and wrote the manuscript.
[67] Nishikawa, K. PRC No. 99.21 http://www.prcnews.org/ (1999) [in Japanese].
[68] Mitaku, S. PRC No. 99.34 http://www.prcnews.org/ (1999) [in Japanese].
[69] Kobayashi, Y(ukio). & Saitô, N. Prediction of structures of globular proteins I. secondary structure. *Genome Informatics Series* No. 4 (T. Takagi, H. Imai, S. Miyano, S. Mitaku and M. Kanehisa eds.) (Universal Academy Press, Tokyo, 1993) 293–299.
[70] Kobayashi, Y(ukio). & Saitô, N. Statistical mechanical method for predicting secondary structures. *Protein Eng.* 6, 1027 (1993).
[71] Kobayashi, Y(ukio). & Saitô, N. Protein secondary structure prediction using a statistical mechanical method. *Protein Eng.* 7, 1164 (1994).
[72] Lifson, S. & Roig, A. On the theory of helix-coil transition in polypeptides. *J. Chem. Phys.* 34, 1963–1974 (1961).
[73] Zimm, B. H. & Bragg, J. K. Theory of the phase transition between helix and random coil in polypeptide chains. *J. Chem. Phys.* 31, 526–535 (1959).
[74] Saitô, N. PRC No. 99.34 http://www.prcnews.org/ (1999) [in Japanese].
[75] Saitô, N. Shodana no naka no issatsu no hon. *Mathematical Sciences* 5, 101–104 (1977) [in Japanese].
[76] Nelsen, R. B. *Proofs without Words* (MAA, Washington, 1993).
[77] Kobayashi, Y. Visual thinking of mathematical concepts. *Far East J. Math. Educ.* 10, 157–174 (2013).
[78] Kobayashi, Y. Simple and intuitive mathematics for learning elementary physics. *JPS Conference Proceedings* 1, 017036 (2014).
[79] Saitô, N. Preface. *Seibutsu Butsuri* 23 (1983) [in Japanese].
[80] Nagai, Y., Hara, R., Tsuchiya, T. & Saitô, N. Existence of low-dimensional chaos in paramecium membrane-potential suggested by the correlation integral method. *J. Phys. Soc. Jpn.* 57, 3305–3308 (1988).
[81] Saitô, N. The society of E. coli and D. discoideum. *Seibutsu Butsuri* 55, 41–42 (2014) [in Japanese].
[82] Saitô, N. Forum: Bunshi shinka to keitai shinka. *Seibutsu Butsuri* 35, 216–217 (1995) [in Japanese].
[83] Saitô, N. & Takasugi, M. Behavior and movement of bacterio-phage χ. *Seibutsu Butsuri* 40, 281–283 (2000) [in Japanese].
[84] Saitô, N. *Ryutai Rikigaku Handbook* (The Japan Society of Fluid Mechanics ed.) (Maruzen, 1998) [in Japanese].
[85] Saitô, N. Chisana ryuushi no ryutai rikigaku. *Butsuri* 23, 822–828 (1968) [in Japanese].
[86] Atkins, P. W. *The Second Law* (Freeman, New York, 1994).
[87] Bixon, M., Scheraga, H. & Lifson, S. Effect of hydrophobic bonding on the stability of poly-L-alanine helices in water. *Biopolymers* 1, 419–429 (1963).