Syst Synth Biol (2013) 7:221–227
DOI 10.1007/s11693-013-9122-6

RESEARCH ARTICLE

An attempt to construct a (general) mathematical framework to model biological "context-dependence"

Anirban Banerji

Received: 24 June 2013 / Accepted: 4 August 2013 / Published online: 28 August 2013

Abstract Context-dependent nature of biological phenomena is well documented in every branch of biology. While there have been few previous attempts to (implicitly) model various (particular) facets of biological context-dependence, a formal and general mathematical construct to model the wide spectrum of context-dependence, eludes the students of biology. Such an objective model, from both 'bottom-up' as well as 'top-down' perspective, is proposed here to serve as the template to describe the various kinds of context-dependence that we encounter in different branches of biology. Interactions between biological contexts was found to be transitive but non-commutative. It is found that a hierarchical nature of dependence among the biological contexts models the emergent biological properties efficiently. Reasons for these findings are provided in a general model to describe biological reality. Scheme to algorithmically implement the hierarchic structure of organization of biological contexts was proposed with a construct named 'Context tree'. A 'Context tree' based analysis of context interactions among biophysical factors influencing protein structure was performed.

Keywords Biological contexts · Mathematical model · Hierarchical organization · Emergence · Thread-mesh model · Context tree · Protein structure · Structural dependencies

A. Banerji (✉)
Bioinformatics Centre, University of Pune, Pune 411007, Maharashtra, India
e-mail: anirbanab@gmail.com

Introduction

'Context-dependence' is omnipresent in Biology. From the realm of substitution of nucleotides (Siepel et al. 2004; Zhang et al. 2007) to the paradigm of protein structure-function (Main et al. 1998; Nobeli et al. 2009), from the sphere of cellular dynamics (Hagan and Sharrocks 2002) to that in virulence studies in host-parasite systems (Brown et al. 2003) and evolutionary dynamics (Jablonski et al. 2006)—one encounters events and processes that are "context-dependent". While various attempts have been made from differing perspectives to somehow quantify context-sensitiveness of particular biological events (Andrianantoandro et al. 2006; Torney et al. 2009; Banerji and Ghosh 2011), a general mathematical framework that attempts to capture and describe the ubiquitous 'context-dependence', eludes the students of Biology. One notes that the need to engineer a scheme to model biological context-dependence was felt by many in recent past (Doboli et al. 2000; Hoare et al. 2004; Loewe 2009; Haseltine and Arnold 2007; Marguet et al. 2007; Platzer and Meinzer 2002; Dhar and Giuliani 2010); the present work attempts to take these concerns to a tangible outcome by proposing the template of a general theoretical framework to model biological organisation from top-down perspective with algorithmically implementable construct. (A recent work that attempted constructing a mathematical model to unambiguously describe the concept 'evolvability' (Valiant 2009), underlines the necessity of present genre of works.) Unlike some previous attempts, the framework proposed here do not tangentially touch upon context-dependence modelling (Standish 2001; Edmonds 1999; Yartseva et al. 2007), but concentrates solely on it. On the other hand, it does not attempt to construct a computational structure that helps in retrieval of biological data from some repository in a

context-dependent manner (Yu et al. 2009; Boeckmann et al. 2005), nor does it propose some (effective) visualization tool to observe context-dependent interactions between biological properties (Gopalacharyulu et al. 2008).

Model

Present work suggests the triad of the form ⟨Structure of biological goals (F), Biological contexts (C), Physical structure of the system (P)⟩, ($\langle F, C, P\rangle$) to describe the structure of any biological process. Components of this triad are (of course) not independent; F determines the suitable choice of C; while C, in its turn, operates upon some particular subset of P with definite features. To ensure F is achieved, C engages only certain elements of P. Such triad-based structure is necessary because same physical structures can be subjected to different contexts to achieve different biological goals; for example, same proteins under different set of contexts may be involved in different biological processes, so that the goals of these processes (each different) are achieved (Gopalacharyulu et al. 2008). Other application of this triad can be found in (Singh and Banerji 2011). The current work attempts to formulate general and formal principles of contextual interactions, by modeling the nature of dependencies between biological contexts that operate upon physical (structural) parameters to ensure that biological goals are achieved.

Since the motivation of any biological process is solely to accomplish a set of necessary biological goals, and the structure of biological goals is hierarchic (Troyanskaya et al. 2003; Camon et al. 2004); we propose a hierarchic organizational structure for biological contexts too. Thus, structure of C will assume that of a tree (namely, the 'Context-Tree'(CT)), where the root-vertex will denote the biological context necessary to achieve the global goal of the system under consideration. The lowest level of CT will be occupied by basic, elementary contexts (-for example, the genes. A systematic combination of the molecular function of their products (proteins) is studied with respect to achieving a specific biological goal (Camon et al. 2004)), and their set $A = \{a_1, a_2, \ldots, a_n\}$ will constitute the leaves of CT. This form of hierarchic structure for CT, helps in analyzing the measure of performance of any component of the physical structure of the system under specific context to achieve any particular F, in the form $\sum_i \frac{\partial F}{\partial a_i}\left(= \sum_i \frac{\partial F}{\partial C_i} \cdot \frac{\partial C_i}{\partial a_i}\right)$. In this manner, every context, including the global context, may be described as a function of composition of basic contextual elements of CT. However, to accomplish a definite set of biological goal, the contexts describing various facets of a biological system need to interact between themselves. We define a set of rules $U(U = \{\alpha_1, \alpha_2, \ldots, \alpha_n\})$ that governs these interactions. Taken in entirety, they form the framework for composition between the base elements of CT, denoted by α. We note here that set U might be infinite, and the compositions $a_i\ \alpha\ a_j$ with $\alpha \in U$ need not necessarily be defined for all $a_i, a_j \in A$ (that is, some of these context interactions might well be mere theoretical possibilities, unrealized in biological paradigm).

Importance of the set U is paramount; it is this set of rules that governs how the various elements of CT will interact to ensure the necessary dependencies between contexts, which in turn will (ultimately) ensure that the system achieves the desired goal F. Let us assign to each element α of the set U, an $n \times n$ incidence matrix (Skiena 1990). A^α with entries a_{ij}^α is unity if the composition $a_i \alpha\ a_j$ is defined, or zero otherwise. We can then introduce a matrix A^U whose elements are given by:

$$a_{ij} = \bigvee_{\alpha \in U} a_{ij}^\alpha \tag{1}$$

stating that $a_{ij} = 1$ if there exists a valid composition rule between a_i and a_j, taken in order.

Matrix A^U conforms to constraints of biological reality because it suggests that for certain magnitudes of i, the ith row and ith column of the matrix A^U can be all zeros. These cases describe algorithmically the fact that interactions between certain basic elements of context-set $A = \{a_1, a_2, \ldots, a_n\}$ are not allowable biologically. We demonstrate such non-allowable biological contexts with two examples:

Example 1 When λ phage (a virus that infects the bacteria *Escherichia coli*) encounters a bacterium, it attaches itself only to certain particular receptors with specific structural features, on the bacterial membrane. This process implies that, relevant biological contexts make sure that binding of λ phage to various other candidate receptor sites with slightly varying structural aspects, is not allowed. Subsequently, when the virus genome enters the bacterium, only two pathways (out of theoretically infinite number of pathways) of alternative nature, namely the 'lytic pathway' or the 'lysogenic pathway' are allowed biologically (Yartseva et al. 2007); although a theoretical thermodynamic study of the situation can suggest many possible pathways with (almost) similar efficiencies. This entire process, in the context-space description can be modeled with non-zero entries for the aforementioned two pathways, while the rest of the entries in A^U will be assigned zero to represent the fact that U specifies the contexts that ultimately ensures certain biological goals.

Example 2 Out of the entire spectrum of possible mRNAs that can be generated from a single gene, only one or a few are created at a time. This means that the nature of A^U makes sure that the other possibilities do not come to

being; although these theoretically possible elements of context-set A might have operated upon the same structural parameters that constitute P. This act of ensuring the interplay among set of allowable contexts to achieve any particular goal, form the so-called 'regulatory mechanism', which describes a delicate balance between concentration magnitudes of pertinent entities, the destination, the sequence variety, structural diversity and the functional options of the resulting protein, where the last one is in turn dependent upon the type of tissue, the stage of development, etc.. (Boeckmann et al. 2005).

Discussion of above suggest that we may as well remove corresponding rules to describe the biological impossibilities in U, making it sure thereby that these compositions do not participate at all in the construction of the tree CT. Removing these aforementioned entries, from the set A, we obtain a minimal set of basic rules of context composition. From here onwards, we will denote this minimal set as set A.

A recent work (Smaldon et al. 2010) has proposed a synthetic biology-driven bottom-up framework to describe cellular processes that take place within liposome. However, we note that the (bottom-up) paradigm of description of interplay of biological contexts can be generalized by describing the entire biological universe with the Thread-Mesh (TM) model (Banerji 2009). The TM model segments entire biological space-time into a series of different biological organizations, viz. biological organization at the level of the nucleotides; that, at the level of amino acids, macromolecules, biochemical pathways, network of pathways, biological cells, tissues, organs, organisms, and ecosystem; - where each one of these organizational schemes is called a threshold level. To define the concept (somewhat) formally, emergence of a single biological property (compositional and/or structural and/or functional) creates a new biological threshold level in the TM model. Thus, if any arbitrarily chosen ith biological threshold level is denoted as TH_i, the succeeding one, viz. TH_{i+1} will be containing at least one biological property that TH_i didn't possess. Schemes with similar philosophy to identify biological threshold levels were proposed previously (Testa and Kier 2000; Dhar 2007) too, but representation of emergence of any biological property and subsequent classification of biological organization with respect to the emergent behavior was not done in either of these models. Basic principles for subsequent discourse are general and can be applied to any threshold level. Every possible property that a threshold level is endowed with, is represented by a 'thread' in the TM model. Thus an environmental property capable of influencing biological action will be called as an 'environmental thread' in the present parlance. Threads can be compositional, structural or functional in their nature. For example, for the biological threshold level representing enzymes (viz., an object belonging to the threshold level representing the macromolecules ($TH_{macromolecules}$)), one of the compositional threads may be its primary structure (viz., the amino acid sequence); whereas enzyme's radius of gyration, its resultant backbone dipole moment and each of its bond lengths, bond angles, torsion angles may be counted as some examples of 'structural threads' of $TH_{macromolecules}$. Finally, the values for K_m, V_{max}, K_{cat} may be counted as some examples of an enzyme's functional threads. We observe that, although a particular amino acid is described by the threshold preceding to that of $TH_{macromolecules}$ (viz.—$TH_{amino\text{-}acids}$); the property set describing an enzyme differs significantly from that describing an amino acid. Indeed, since a threshold level is defined by the presence of (at least) one emergent property that the previous threshold did not possess, the example involving enzyme and an amino acid implies that there may exist several threshold levels between the one representing an individual amino acid and an individual enzyme. Examples of possible (intermediary) threshold levels can be $TH_{molten\text{-}globules}$ or $TH_{secondary\text{-}structures}$, or $TH_{super\text{-}secondary\text{-}structures}$, etc. (Reader is referred to (Banerji 2009), for further details on TM model.) Suffice to say that it seems advantageous to work with the TM model because it can attempt describing context-dependence and emergence from the framework of an invariant template.

However, while the (template of the) framework discussed above (alongside example-1 and example-2) describe's the nature of multilevel organization of CT, such description is 'bottom-up' in nature. Hence, while it is helpful to describe the context-mapping between any two particular adjacent biological threshold levels 'l' and '$l + 1$', (say between threshold levels representing nucleotides and amino acids, amino acids and proteins, or between proteins and biochemical pathways, etc..) the general mode of dependency among systemic and environmental threads may (invariably) tend to become intractable. Ths will imply that, general mode of dependency within CT with a birds-eye ('top-down') view of the organization of it, can hardly be guessed from such bottom-up approach. Moreover, we note that, any attempt to construct a framework to describe biological context dependence will be extremely difficult (if not impossible) from a bottom-up perspective. On the other hand, the top-down perspective attempts to describe the system in steady-states. Thus, it has the potential to not consider the multifarious dependencies among interacting threads and yet can attempt to describe the patterns of emergent properties at various threshold levels.

The top-down framework

We start construction of the top-down scheme of description of dependencies between biological

contexts, by enlisting the assumptions involved therein. Hence:

Assumption 1 In absence of random external disturbances and without a failure of any component belonging to physical structure of the system (P), all the rules of multilevel interactions between the contexts representing any biological threshold level can be constructed in suitably deterministic manner. (Success of recent attempts with deterministic modeling of various biological phenomena from diverse backgrounds (Janda and Gegina 2008; Kim and Maly 2009; Ferreira and Azevedo 2007) suggest that such assumption is not ill-founded, and that too in absence of possible perturbations.)

Assumption 2 The necessary and sufficient condition for these deterministic rules of inter-level context interactions to hold true, is that they should account for the accomplishment of certain biological goals (F). (Previous studies (Yartseva et al. 2007; Troyanskaya et al. 2003; Camon et al. 2004) vindicate such assumption.)

Assumption 3 Although biological systems will be exposed to randomly varying magnitudes of external parameters, the essence of the deterministic criteria of context interactions in order to accomplish any set of required biological function, will not be perturbed by significant margin. This assumption implies that deterministic manner of context interactions will not be undergoing significant change when the magnitudes of components of underlying physical structures $\{p_i\}(p_i \in P)$, comprised of relevant biological parameters, are altered within some allowable range. We describe this allowable range of assumed magnitude of some arbitrarily chosen parameter π by an interval $[\pi_0, \pi_1]$.

Relevance of the last assumption can be easily understood when one analyzes the nature of some previous results in depth. Since every biological property operates within a specified bound of magnitude, something that has been referred to as 'fluctuation' in an earlier study (Testa and Kier 2000), the mathematical functions that represent them will also be bounded within their respective ranges. Examples of aforementioned fluctuation are many. Say, in the threshold level representing living cells, for the mitogen-activated protein kinase cascade studies, the total concentrations of MKKK, MKK and MAPK have been found to be in the range 10–1,000 nm and the estimates for the K_{cat} values of the protein kinases and phosphatases have been found to range from 0.01 to 1 s^{-1} (Kholodenko 2000). Similarly, for the proteins, the mass fractal dimension and hydrophobicity fractal dimension representing compactness of mass and hydrophobicity distribution, have been found to be in the range between 2.18 to 2.37 and 2.22 to 2.43 respectively (Banerji and Ghosh 2009).

Based on these assumptions, we propose that the functional that defines the probability of attaining the biological goal (F) under consideration, will assume the form:

$$F_0 = \int_{x_i \in S} \phi(x_1, x_2, \ldots, x_n) dx_i \qquad (1 \le i \le n) \tag{2}$$

where ϕ is the probability density of attaining the objective (biological goal) and X is the feasibility domain of the contexts x_i.

Since, to achieve every biological goal, many (say, m) successive stages of context interactions are required, we can express the last equation at a higher resolution, as:

$$\phi(x_1, x_2, \ldots, x_n) = \prod_{j=1}^{m} \phi_j | \phi_{j-1}(x_1, x_2, \ldots, x_n) \tag{3}$$

where $\phi_j|\phi_{j-1}$ represent the conditional probability associated with context $\{x_i\}$ interactions, while attempting to achieve a particular biological goal.

However, we note that individual physical parameters $p_i(p_i \in P)$, upon which the contexts are working, may not always be strongly correlated and although they are related to each other, can be considered independent when viewed individually with respect to their functional contribution to the system. For example, the time-dependent and context-dependent fluctuations in individual bond lengths, bond angles and torsion angles in the protein interior, although might be related in some intricate way to the resultant dipole moment for the protein; can be considered, for all practical purposes, in terms of their individual (and not linked) contributions in ensuring proteins stability and functionality. Hence we attempt to partition the relevant contexts into a sum of disjoint domains; such that: $(x_i \in X_i)$ and $\sum_i X_i = S$.

Considering this partition we can re-write Eq. 2 as:

$$F_0 = \int_{x_1 \in X_1} \int_{x_2 \in X_2} \cdots \int_{x_n \in X_n} \phi_1(x_1, x_2, \ldots, x_n) \times \phi_{2|1}(x_1, x_2, \ldots, x_n) \ldots \phi_{m|m-1}(x_1, x_2, \ldots, x_n) dx_1 dx_2 \ldots dx_m \tag{4}$$

In other words, purely in terms of achievement of biological goals:

$$F_0 = F_1 F_{2|1} \ldots F_{m|m-1} \tag{5}$$

where

$$F_{j|j-1} = \int_{x_1 \in X_1} \int_{x_2 \in X_2} \ldots \int_{x_n \in X_n} \phi_{j|j-1}(x_1, x_2, \ldots, x_n) dx_1 dx_2 \ldots dx_n \tag{6}$$

are the conditional probabilities of context-interactions of the system realizing the successive stages of the task. It is

necessary to mention here that to achieve any biological function, the domain of integration for every x_i in the last equation must be within their respective permissible range, say $[x_{\pi_0},\ x_{\pi_1}]$.

While it is difficult to assume that every context-interaction necessary to realize certain biological goal will always be operating in deterministic manner with perfect efficiency, observation suggests that biological goals are seldom compromised with. Hence we assume that, the reliability of any arbitrarily chosen context interaction at any arbitrarily chosen *j*th state in the realization of certain biological function, is statistically independent of the probability of the realization of that particular biological function. In that case, the integrand of the last equation can be expressed as a product $\phi_{j|j-1}(x_1, x_2, \ldots, x_n)$ $r_j(x_1, x_2, \ldots, x_n)$, where $r_j, \left(r_j \in R\right)$ describes the probability of reliability of any arbitrarily chosen context-interaction at *j*th state in the realization of certain biological function.

Hence the last equation can be expressed more realistically as:

$$\phi(x_1, x_2, \ldots, x_n) = \prod_{j=1}^{m} \phi_{j|j-1}(x_1, x_2, \ldots, x_n) r_j(x_1, x_2, \ldots, x_n) \tag{7}$$

Thus, when the reliability of context-interactions are taken into account, Eq. 5 can be re-written as:

$$F_0 = \prod_{j=1}^{m} F_j | F_{j-1} R_j \tag{8}$$

where

$$R_j = \int\limits_{x_1 \in X_1} \int\limits_{x_2 \in X_2} \cdots \int\limits_{x_n \in X_n} r_j(x_1, x_2, \ldots, x_n) dx_1 dx_2 \ldots dx_n \tag{9}$$

In other words, Eq. 5, can be re-written (in the final form) as:

$$F_0 = F_1 F_{2|1} \ldots F_{n|n-1} R_1 R_2 \ldots R_m \tag{10}$$

Result

Modeling hierarchical organization with 'context tree'

Case-study with protein structure

The effectiveness of hierarchical organizational structure to model interactions among biological contexts was touched upon in the last section. To describe the 'Context-Tree' (*CT*) in such hierarchic paradigm under a generalized scheme we introduce a construct *C*, which is a family of embedded partitions of contexts $C = \langle C^1, C^2, \ldots, C^r \rangle$ that in turn operate upon any relevant subset of structural threads (*J*) representing the physical structure (*P*) of any arbitrarily chosen threshold level *S*. $J \subset P$ and $J = \{1, 2, \ldots, m\}$. For example, it has been found (Main et al. 1998) that at the threshold level of proteins $\mathrm{TH}_{\mathrm{Proteins}}$(which is just a subset of $\mathbf{TH}_{\mathrm{Macromolecules}}$), in an urea-induced media (the 'environmental thread' influencing *J*), the extent of stability of mutant proteins are highly dependent on the contexts (*C*) which operate upon the various structural parameters (*J*), that form a subset of (*P*) describing (*S*). Thus, for $(S : \mathrm{TH}_{\mathrm{Proteins}})$, we can describe the situation as:

$$\begin{aligned} C^S &= \langle C_1^S, C_2^S, \ldots, C_l^S \rangle \quad \cup_{j=1}^{l} C_j^S = J, \\ C_i^S &\cap C_j^S = \emptyset (i \neq j), \quad S = \overline{1, r} \end{aligned} \tag{11}$$

The embedding refers to any element of the partition of the *S*th biological threshold level; i.e., the set C_j^s represents the union of several sets $C_{i_1}^{S-1}, C_{i_2}^{S-1}, \ldots, C_{i_z}^{S-1}$ of the $(S-1)$th biological threshold level. Such description of (*CT*) conforms to a previous study on similar topic (Andrianantoandro et al. 2006). Findings from a recent study (Haseltine and Arnold 2007) vindicates $C_i^S \cap C_j^S = \varnothing$. To elaborate the hierarchic structure, we can write $C \longleftrightarrow \langle C_1^{S-1}, C_2^{S-1}, \ldots, C_l^{S-1} \rangle$, if $C^S = \cup_{i=1}^{l} C_i^{S-1}$. Since the entire set of interactions between various contexts is ultimately geared to satisfy biological goals and since the nature of organization of biological goals is hierarchic, we attempt to describe it by defining $l(C^S) = l$ and $C^{root} = \langle \{1, 2, \ldots, m\} \rangle$; i.e., the partition at the highest (root) level consists of one set, namely *J*.

We can associate each element $C_j^S \left(s = \overline{2, r}\right)$ of the partition to the context-interaction function, namely $f_j^S\left(\alpha_1, \alpha_2, \ldots, \alpha_{l\left(C_j^S\right)}\right)$, where $\alpha \in \{0, -1, +1\}$, conforming to the previously defined $U(U = \{\alpha_1, \alpha_2, \ldots, \alpha_n\})$. We associate each element C_j^1 of the first level to a binary relation R_j(the previously defined elementary contexts are related by this, say *aRb*, where $A = \{a, b, \ldots, z\}$) on the biological sub-space EC_j^1.

These concepts can formally be described as: Let $C_j^2 \longleftrightarrow \langle C_1^1, \ldots, C_l^1 \rangle$ and define a relation R_j^2 on $E_{C\_j}^2$ using the formula (for $l > 1$):

$$a_{C_j^2} R_{bC_j^2}^2 \Longleftrightarrow f_j^2(\alpha_1, \alpha_2, \ldots, \alpha_l) = 1 \tag{12}$$

$$\begin{aligned} \alpha_i &= +1 \quad \text{if} \quad a_{C_i^1} R_i b_{C_i^1} \\ \alpha_i &= -1 \quad \text{if} \quad a_{C_i^1} \bar{R} b_{C_i^1} \\ \alpha_i &= 0 \quad\ \ \text{if} \quad a_{C_i^1} = b_{C_i^1} \end{aligned}$$

In case of $l = 1$, $R_j^2 = R_j$. If all the relations R^{S-1} of the $(S-1)$th biological threshold level are defined, then the

relations R_j^S of the Sth threshold level with $(l > 1)$ can be defined by the following construct:

If $C_j^S \longleftrightarrow \langle C_1^{S-1}, C_2^{S-1}, \ldots, C_l^{S-1}, \rangle$, then

$$a_{C_j^S} R^S_{bC_j^S} \Longleftrightarrow f_j^S(\alpha_1, \alpha_2, \ldots, \alpha_l) = 1 \tag{13}$$

$$\begin{aligned} \alpha_i &= +1 \quad \text{if} \quad a_{C_i^{S-1}} R_i^{S-1} b_{C_i^{S-1}} \\ \alpha_i &= -1 \quad \text{if} \quad a_{C_i^{S-1}} \overline{R^{S-1}} b_{C_i^{S-1}} \\ \alpha_i &= 0 \quad \text{if} \quad a_{C_i^{S-1}} = b_{C_i^{S-1}} \end{aligned}$$

If $C^S = C^{S-1}$, then $R^S = R^{S-1}$; in other words the construction requires the relation R to coincide with R^r, which is a single relation at the rth upper level.

To describe the entire (bottom-up) paradigm of description of interaction scheme between biological contexts, we consider an example where we describe the contextual constraints on the active site of an enzyme in simplistic terms. For this case, without any loss of generality, we consider the threshold level representing proteins to be the root level in this case. The goal of the system (F) is to make the enzyme functional. We assume the elementary contexts that can influence functionality of the enzyme active site to be represented with three basic partitions; namely, first, the contextual differences originating out of protein 'internal coordinates'; second, contextual differences arising out of interaction profile of the active site atoms with water; and third, contextual differences arising out of the capability of the active site to undergo a shape change. Hence, we may describe the family of partitions C, as $C^1 = \langle \{1,2,3\}, \{4,5\}, \{6\} \rangle$ and $C^2 = \langle J \rangle$; where element-1 denotes (possible) contextual difference arising out of the fluctuation of bond lengths, element-2 denotes (possible) contextual difference arising out of the fluctuation of bond angles, element-3 denotes (possible) contextual difference arising out of the fluctuation of torsion angles. Similarly, element-4 stands for (possible) contextual difference arising out of the hydrophobicity of active site patch, element-5 denotes the (possible) contextual difference arising out of the local electrostatic profile of the active site patch. Element-6 denotes the extent of (possible) contextual difference arising out of the change in the local shape of the active site patch. Denoting the set $\{1,2,3\}$ as $D_1^2, \{4,5\}$ as D_2^2 and $\{6\}$ as D_3^2, the hierarchic nature of these contextual dependencies can easily be described as:

$$D_2 R D_1 \Longleftrightarrow [D_2 \geq D_1], \text{recalling the relation } R;$$

similarly,

$$D_3 R D_2 \Longleftrightarrow [D_3 \geq D_2] \quad \text{and} \quad D_3 R D_1 \Longleftrightarrow [D_3 \geq D_1]$$

Implying that a (possible) contextual difference arising out of the hydrophobicity of active site patch, or a (possible) contextual difference due to the local electrostatic profile of the active site patch will surely account for some change in the distribution profile of bond length, bond angle and torsion angle distribution. But inverse of this case, viz., a (secondary) change in the local electrostatics profile and local hydrophobic profile due to a (primary) change in bond-length, bond-angle or torsion angle might or might not be observed in reality. Similarly, in case of a possible change in local shape the local electrostatic profile, local hydrophobicity profile, local distribution of bond length, bond angle, torsion angle will surely be taking place; but the other way round might or might not be observed. This vindicates and generalizes our previous finding that interactions between biological contexts, are transitive but are not commutative (hence, if D_2RD_1 and D_3RD_2 are defined, $D_3RD_2RD_1$ can be defined; but merely the existence of D_3RD_2 doesn't imply that D_2RD_3 exists too).

Conclusion

Taken together, the top-down and bottom-up set of equations present a possible framework to quantitatively model the omnipresent "context-dependence" in biology. Since contemporary biology, as never before, is attempting to be objective in its philosophy, the necessity of a mathematical framework to describe the "context-dependent" nature of it can hardly be ignored. The model proposed here may not be complete; however, it probably is the (necessary) first step.

References

Andrianantoandro E, Basu S, Karig D, Weiss R (2006) Synthetic biology: new engineering rules for an emerging discipline. Mol Syst Biol 2:0028

Banerji A (2009) Existence of biological uncertainty principle implies that we can never find 'THE' measure for biological complexity. arXiv:0902.0490.[q-bio.OT]

Banerji A, Ghosh I (2009) Revisiting the myths of protein interior: studying proteins with mass-fractal hydrophobicity-fractal and polarizability-fractal dimensions. PLoS One 4(10):e7361

Banerji A, Ghosh I (2011) Mathematical criteria to observe mesoscopic emergence of protein biochemical properties. J Math Chem 49(3):643–665

Boeckmann B, Blatter M, Famiglietti L, Hinz U, Lane L, Roechert B, Bairoch A (2005) Protein variety and functional diversity: Swiss-Prot annotation in its biological context. Comptes Rendus Biologies 328:882–899

Brown M, Schmid-Hempel R, Schmid-Hempel P (2003) Strong context-dependent virulence in a host-parasite system: reconciling genetic evidence with theory. J Anim Ecol 72(6):994–1002

Camon E, Magrane M, Barrell D, Lee V, Dimmer E, Maslen J, Binns D, Harte N, Lopez R, Apweiler R (2004) The gene ontology annotation (goa) database: sharing knowledge in uniprot with gene ontology. Nucleic Acids Res 32:D262–D266

Dhar P (2007) The next step in biology: a periodic table? J Biosci 32:1005–1008

Dhar PK, Giuliani A (2010) Laws of biology: why so few? Syst Synth Biol 4(1):7–13

Doboli S, Minai A, Best P (2000) Latent attractors: a model for context-dependence place representations in the hippocampus. Neural Comput 12(5): 1009–1043

Edmonds B (1999) Syntactic measures of complexity. Ph.D. thesis, University of Manchester. http://www.cpm.mmu.ac.uk/~bruce/thesis

Ferreira P, Azevedo P (2007) Evaluating deterministic motif significance measures in protein databases. Algorithms Mol Biol 2:16

Gopalacharyulu PV, Lindfors E, Miettinen J, Bounsaythip CK, Oresic M (2008) An integrative approach for biological data mining and visualization. Int J Data Min Bioinforma 2:54–77

Hagan I, Sharrocks A (2002) Understanding cancer: from the gene to the organism. Conference on genes and cancer. EMBO Rep 3(5):415–419

Haseltine E, Arnold F (2007) Synthetic gene circuits: design with directed evolution. Ann Rev Biophys Biomol Struct 36:1–19

Hoare D, Couzin I, Godin J, Krause J (2004) Context-dependent group size choice in fish. Anim Behav 67:155–164

Jablonski P, Lee S, Jerzak L (2006) Innate plasticity of a predatory behavior: nonlearned context dependence of avian flush-displays. Behav Ecol 17(6): 925–932

Janda J, Gegina G (2008) A deterministic model for the processing and presentation of bacteria-derived antigenic peptides. J Theor Biol 250(3): 532–546

Kholodenko B (2000) Negative feedback and ultrasensitivity can bring about oscillations in the mitogen-activated protein kinase cascades. Eur J Biochem 267(6):1583–1588

Kim M, Maly I (2009) Deterministic mechanical model of T-killer cell polarization reproduces the wandering of aim between simultaneously engaged targets. PLoS Comp Biol 5(1):1–12

Loewe L (2009) A framework for evolutionary systems biology. BMC Syst Biol 3:27

Main E, Fulton K, Jackson S (1998) Context-dependent nature of destabilizing mutations on the stability of FKBP12. Biochemistry 37(17):6145–6153

Marguet P, Balagadde F, Tan C, You L (2007) Biology by design: reduction and synthesis of cellular components and behaviour. J Royal Soc Interface 4:607–623

Nobeli I, Favia A, Thornton J (2009) Protein promiscuity and its implications for biotechnology. Nat Biotechnol 27(2):157–167

Platzer U, Meinzer H (2002) Simulation of genetic networks in multicellular context. In: Polani D, Kim J, Martinez T (eds) 5th German workshop on artificial life: abstracting and synthesizing the principles of living systems, Berlin: Akad. Verl.-Ges, pp 43–51

Siepel A, Haussler D (2004) Phylogenetic estimation of context-dependent substitution rates by maximum likelihood. Mol Biol Evol 21(3):468–488

Skiena S (1990) Implementing discrete mathematics: combinatorics and graph theory with mathematica. Reading: Addison-Wesley, pp 135–136

Singh PP, Banerji A (2011) Case for an RNA-prion world: a hypothesis based on conformational diversity. J Biol Phys 37(2):185–188

Smaldon J, Romero-Campero FJ, Fernández Trillo F, Gheorghe M, Alexander C, Krasnogor N (2010) A computational study of liposome logic: towards cellular computing from the bottom up. Syst Synth Biol 4(3):157–179

Standish R (2001) On complexity and emergence, arXiv:nlin/0101006v1 [nlin.AO]

Testa B, Kier L (2000) Emergence and dissolvence in the self-organisation of complex systems. Entropy 2:1–25

Torney C, Neufeld Z, Couzin I (2009) Context-dependent interaction leads to emergent search behavior in social aggregates. Proc Nat Acad Sci USA 106(52):22055–22060

Troyanskaya O, Dolinski K, Owen A, Altman R, Botstein D (2003) A Bayesian framework for combining heterogeneous data sources for gene function prediction (in Saccharomyces cerevisiae). Proc Nat Acad Sci 100(14):8348–8353

Valiant L (2009) Evolvability. J Assoc Comp Mach 56,1, 56:1, 3:1–3:21

Yartseva A, Klaudel H, Devillers R, Kepes F (2007) Incremental and unifying modelling formalism for biological interaction networks. BMC Bioinformatics 8:433

Yu C, Zavaljevski N, Desai V, Reifman J (2009) Genome-wide enzyme annotation with precision control: catalytic families (CatFam) databases. Proteins. 74:449–460

Zhang W, Bouffard G, Wallace S, Bond J (2007) Estimation of DNA sequence context-dependent mutation rates using primate genomic sequences. J Mol Evol 65:207–214