MATCH
Communications in Mathematical and in Computer Chemistry

MATCH Commun. Math. Comput. Chem. **81** (2019) 177-192

ISSN 0340 - 6253

The Significance of Noise in the Evolution of Negative and Positive Cooperativity in Protein Complexes

Mostafa Taheri[1], Kazem Nouri[1*], Yazdan Asgari[2], Zahra Zabihinpour[3], Mehdi Sadeghi*[4]

[1]*Department of Mathematics, Faculty of Mathematics, Statistics and Computer Sciences, Semnan University, Semnan, P.O.Box 35195-363, Iran*

[2]*Department of Medical Biotechnology, School of Advanced Technologies in Medicine, Tehran University of Medical Sciences, Tehran, Iran*

[3]*School of Biological Sciences, Institute for Research in Fundamental Sciences (IPM), Tehran, Iran*

[4]*National Institute of Genetic Engineering and Biotechnology, Tehran, Iran.*

knouri@semnan.ac.ir , sadeghi@nigeb.ac.ir

(Received April 17, 2018)

Abstract

Evolution, tends to build complex structures out of protein molecules which has advantages for protein function. Few reasons have been proposed to explain this event. Some protein multimers show cooperativity in which binding of a ligand to a chain affect on the binding of other ligands. But it has remained unclear why cooperativity arises in positive or negative forms and what is the advantage of each of these forms. Our analysis, using deterministic and stochastic approaches, shows that the average and the standard deviation of ligand binding capacity indeed differ between positive and negative cooperativity. Considering standard deviation as a measure of noise in the system, our results demonstrate that the transition from negative to positive cooperativity is accompanied by an increase in the level of noise in the system. This source of noise enables the system to adapt in a fluctuating environment and could be subject to selection during evolution.

1 Introduction

It would be easier to contemplate the evolutionary process at different levels of life by taking into account the molecular diversity of vital biological elements. Since cellular scaffold is mostly composed of proteins, dealing with the development of protein structures during evolution would be a worthy starting point [1]. Protein molecules have been evolved to be more [2]. Evolutionary modifications in ribosome [3], cytoskeleton [4], spliceosome [5], ion channels [6], and nuclear pore complex [7] are some examples in which higher structural stability and significantly longer half-lives have evolved [8-10]. Proteins transition rates from monomers to multimers have also increased, and more than 50% of proteins in eukaryotes and prokaryotes are multimer complexes [11]. Why does a protein molecule should become more complex? During past decades, a myriad of studies have been performed to answer this question; being easier for multiple number of small proteins to be folded compared to folding a single long protein, increasing opportunities for allosteric regulation of protein activity, reducing protein's defenselessness to denaturation because of smaller surface to volume ratio, increasing the frequency of productive encounters due to removal of unnecessary protein surfaces, and protection of unstable proteins from aggregation are some of the reasons espoused for the observed increase in protein complexity in nature [12-15]. In addition, gaining new features would be also possible after oligomerization, which affects both stability and function [16]. For example, enhancing diversity could create some opportunities for proteins to represent allosteric regulation functions, enabling them to perform cooperativity behaviors [11, 14, 17, 18]. Cooperativity is a biochemical phenomenon in which identical or near-identical elements change their performance. This was first discovered in hemoglobin by C. Bohr et al., where they found that the affinity of hemoglobin binding sites increases when an oxygen atom binds to one of four binding sites (called positive cooperativity) [19]. Later, it was discovered that cooperativity could also be negative in some proteins. However, transition into multimers (in order to gain cooperativity function) inevitably enforces an energetic cost and increases the probability of certain diseases related to inappropriate protein aggregation [14]. Consequently, one cannot simply conclude that positive cooperativity should always be favored by selection [20, 21]. Some studies suggested a thermodynamic-based selective preference for cooperativity [22, 23]. Another interesting method is dealing with the kinetic models of cooperativity using calculated experimental parameters for kinetic constant rates. However, finding analytical solution (deterministic approach) for the differential equations derived in a kinetic model is always challenging, especially when there is a paucity of experimental data [24]. It becomes

more problematic when dealing with low number of proteins in different cells where stochastic effects play an important role [25]. Therefore, various features seem to be relevant to proteins to adjust their cooperativity: some proteins prefer positive cooperativity which increases affinity at a lower cost, some choose negative cooperativity which decreases affinity but enhances the transition speed, and some proteins utilize a more advanced strategy that involves choosing cooperative behaviors (negative and positive) based on the environmental conditions [20].

So why does cooperativity arise in positive or negative forms? What is the advantage for each forms? Are they dependent on the role of proteins in the cell? Is a negative or positive cooperative behavior of a protein related to the effect of natural selection in a fluctuating environment?

In this study, we have explored some possible reasons for the development of multimer proteins with cooperative functions, using different approaches such as exploring frequency distribution of cooperative proteins, structural stability investigation in the context of thermodynamics, and performing kinetic modeling to measure variation in positive/negative cooperativity using deterministic and stochastic approaches.

2 Material and Methods

2.1 Deterministic Approach

For a deterministic insight into the cooperative binding of a protein, we have applied a mass action kinetic model. Here we use the simplest model, which consists of a protein complex (*M*) with two binding sites where 2 ligand molecules (*L*) could attach [26]:

$$\begin{aligned} M + L &\underset{k_{-1}}{\overset{k_1}{\leftrightarrows}} ML \\ ML + L &\underset{k_{-2}}{\overset{k_2}{\leftrightarrows}} ML_2 \end{aligned} \qquad (1)$$

where *ML* and ML_2 are proteins contain one and two ligands, respectively.

If $K_1=K_2$ (where $K_i = \frac{k_i}{k_{-i}}$, $i=1,2$, and k_i and k_{-i} are forward and reverse rate constants for reactions in the eq. 1, respectively), a protein is non-cooperative. In this case, the value of Hill coefficient (n_H) would be 1. One could obtain n_H value by plotting $\log\left(\frac{Y}{1-Y}\right)$ versus log (*L*) where

$$Y = \frac{1}{2} \times \frac{ML + 2ML_2}{M + ML + ML_2}. \quad (2)$$

The Hill coefficient would be the slope of the graph [27]. If $K_1 < K_2$, a protein contains positive cooperativity and the Hill coefficient (n_H) would be greater than 1. Finally, if $K_1 > K_2$, a protein contains negative cooperativity and the Hill coefficient (n_H) would be less than 1.

One could also write Ordinary Differential Equations (ODEs) for eq. 1 as follow:

$$\frac{d[M]}{dt} = -k_1[M][L] + k_{-1}[ML]$$

$$\frac{d[L]}{dt} = -k_1[M][L] + k_{-1}[ML] - k_2[ML][L] + k_{-2}[ML_2] - k_3[ML_2][L]$$

$$\frac{d[ML]}{dt} = k_1[M][L] - k_{-1}[ML] - k_2[ML][L] + k_{-2}[ML_2]$$

$$\frac{d[ML_2]}{dt} = k_2[ML][L] - k_{-2}[ML_2], \quad (3)$$

where [*M*], [*L*], [*ML*], and [ML_2] are concentrations of different species, and k_1, k_{-1}, k_2, and k_{-2} are equilibrium constants of forward and reverse reactions.

Similar strategy could be used for proteins with more binding sites. All equations related to protein complexes with 3 and 4 binding sites are available in the supplementary file S1-A.

It should also be noted that a statistical correction on the rate constants is necessary since the binding sites are indistinguishable – i.e., it is only possible to count total number of the occupied binding sites regardless of their positions [28]. The correction equation would be:

$$K_i^{'} = \frac{n - i + 1}{i} K_i, \quad (4)$$

where $K_i^{'}$ is statistically corrected rate constant, n is total number of protein binding sites, $K_i = \frac{k_i}{k_{-i}}$ (for $i=1,\ldots,n$), and k_i and k_{-i} are forward and reverse rate constants for every reactions in the model, respectively.

For solving ODEs, Copasi software (ver. 4.16) [29] was used on a computer with the Intel Core-i7-3770k CPU and 8 GB of RAM.

2.2 Stochastic Approach

To explore the stochastic behavior of a biochemical reaction, we utilized Chemical Master Equations (CMEs). CMEs belong to a class of statistical Markov formalism in which the

number of species in each state is considered as random variables where their transition matrices are defined by probability methods. Therefore, number of species in each step only depends on the number of species in the previous step [30, 31]. Here a stochastic modeling approach for the simplest case (a protein complex with two binding sites) is show. The reactions are identical to eq. (1). Transition probabilities for each kind of molecule is shown in Table 1 [30].

Table 1. Transition probability of different molecules in eq. (1) where *m*, *l*, *ml*, and ml_2 are the number of free proteins, free ligands, proteins with one ligand, and proteins with two ligands, respectively.

$t + dt$	t	*reaction*	*probability*
M (t + dt) = m	$m - 1$	k_{-1}	$k_{-1}(ml+1)\,dt$
	$m + 1$	k_1	$k_1(m+1)(l+1)\,dt$
L (t + dt) = l	$l - 1$	k_{-1} *or* k_{-2}	$k_{-1}(ml+1)\,dt + k_{-2}(ml_2+1)dt$
	$l + 1$	k_1 *or* k_2	$k_1(l+1)(m+1)\,dt + k_2(l+1)(ml+1)\,dt$
ML (t + dt) = ml	ml - 1	k_1 *or* k_{-2}	$k_1(m+1)(l+1)\,dt + k_{-2}(ml_2+1)dt$
	ml + 1	k_{-1} *or* k_2	$k_{-1}(ml+1)\,dt + k_2(l+1)(ml+1)\,dt$
ML_2 (t + dt) = ml_2	ml_2 - 1	k_2	$k_2(l+1)(ml+1)\,dt$
	ml_2 + 1	k_{-2}	$k_{-2}(ml_2+1)\,dt$

Based on Table 1, the probability that none of the reactions occur would be:

$$1-\left(k_{-1}ml + k_1 ml + k_{-2}ml_2 + k_2 ml.l\right)dt \tag{5}$$

Now we could write:

$$\begin{aligned} P(m,l,ml,ml_2;t+dt) = P(m,l,ml,ml_2;t)\left[1-\left(k_{-1}ml + k_1 ml + k_{-2}ml_2 + k_2 ml.l\right)dt\right] \\ + k_{-1}(n+1)P(m-1,l-1,ml+1,ml_2;t)\,dt \\ + k_1(m+1)(l+1)P(m+1,l+1,ml-1,ml_2;t)\,dt \\ + k_{-2}(ml_2+1)P(m,l-1,ml-1,ml_2+1;t)\,dt \\ + k_2(l+1)(ml+1)P(m,l+1,ml+1,ml_2-1;t)\,dt \end{aligned} \tag{6}$$

where $P(m,l,ml,ml_2;t + dt)$ is the probability of being in a state in which there are *m*, *l*, *ml*, and ml_2 number of molecules at time $t + dt$. So, this could lead to:

$$\begin{aligned}\frac{dP_{m,l,ml,ml_2}}{dt} &= -k_{-1}mlP_{m,l,ml,ml_2} + k_{-1}(n+1)P_{m-1,l-1,ml+1,ml_2} \\ &\quad -k_1mlP_{m,l,ml,ml_2} + k_1(m+1)(l+1)P_{m+1,l+1,ml-1,ml_2} \\ &\quad -k_{-2}ml_2P_{m,l,ml,ml_2} + k_{-2}(ml_2+1)P_{m,l-1,ml-1,ml_2+1} \\ &\quad -k_2n_1lP_{m,l,ml,ml_2} + k_2(l+1)(ml+1)P_{m,l+1,ml+1,ml_2-1}\end{aligned} \tag{7}$$

where *m*, *l*, *ml*, and ml_2 are the number of free proteins, free ligands, proteins with one ligand, and proteins with two ligands, respectively.

Equation (7) is treated as Stochastic Differential Equation (SDE) that is continues in time and discontinues in regards to the number of species.

Similar strategy could be used for protein complexes with 3 and 4 binding sites. (For all the transition probabilities and SDEs see the supplementary file S1-B.)

We used Numerical Stochastic Simulation Algorithm (SSA) to solve the eq. (7), since finding an analytical solution is daunting and has not been feasible in more complicated cases [32, 33]. In addition, using the standard SSA method (Gillespie Algorithm) is time-consuming. So, we have used the most recent version called the Tau-leaping algorithm which is much faster and more suitable for our case [34]. We have used MATLAB software (ver. R2015b) on a computer with the Intel Core-i7-3770k CPU and 8 GB of RAM for running the Tau-leaping algorithm.

3 Results and Discussion

3.1 Cooperativity Distributions and Ligand Concentrations

First, we browsed the literature (data not shown) to figure out whether any particular type of cooperativity (positive/negative) is more prevalent in nature. Then, we explored ligand concentrations where cooperativity occurs. Our results show that neither types of cooperativity (positive and negative) is more common in nature. In addition, both positive and negative cooperativity exist in a wide range of ligand concentrations. However, it seems that negative cooperativity occurs more between 10^{-5}-10^{-1} (mol/litre) of ligand concentrations, while positive cooperativity occurs more between 10^{-10}-10^{-6} (mol/litre) of ligand concentrations (Figure 1).

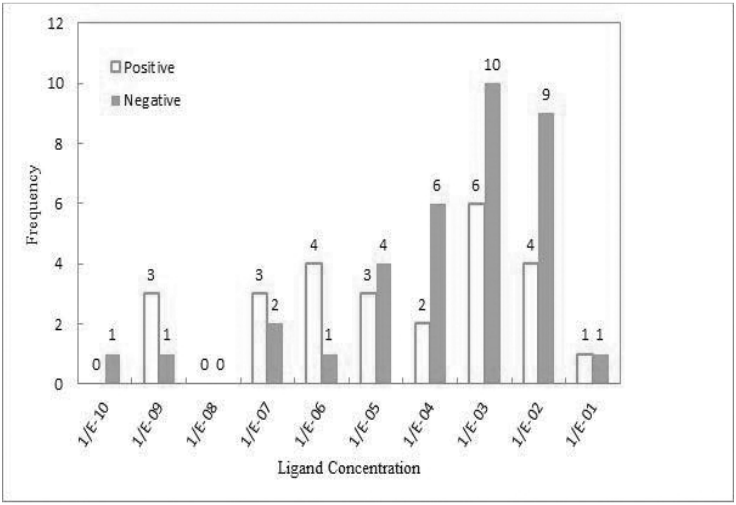

Figure 1. Frequency of the experimentally-verified positive and negative cooperativity behaviors in different ligand concentrations.

3.2 Deterministic Approach

In the deterministic method, proteins with four binding sites (detailed equations are available in supplementary file S1-A) were considered. Simulations were done for 3 different cases; negative cooperativity ($K_1 > K_2 > K_3 > K_4$), non-cooperativity ($K_1 = K_2 = K_3 = K_4$), and positive cooperativity ($K_1 < K_2 < K_3 < K_4$), where $K_i = \frac{k_i}{k_{-i}}$ (for $i=1,2,3,4$) are statistically corrected rate constants, and k_i and k_{-i} are forward and reverse rate constants for every reaction in the model, respectively. Results are shown in Table 2 and Figure 2.

All models can be found in the supplementary file S2, and the results are found in the supplementary file S3_Table2.

As illustrated in Table 2, the average ligand binding number (M_{ave}) and standard deviation (*std. dev.*) increase when moving from negative to positive cooperativity.

In another simulation, we explored the effect of rate constant values on M_{ave} and *std. dev.* values while a protein retains its cooperativity type.

For the first case, a with non-cooperative protein with different rate constants values was considered. Results are shown in Table 3. The results show an incremental behavior in M_{ave} and *std. dev.* while increasing rate constant values.

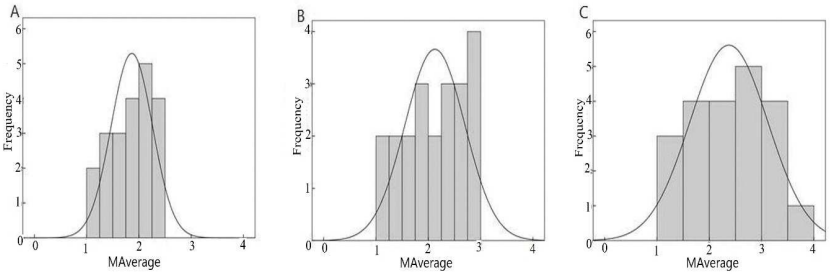

Figure 2. Distribution histogram of values obtained for average ligand binding number at the end of every simulation, i.e. for 21 different initial ligands number: A) Negative cooperativity, B) Non-cooperativity, C) Positive cooperativity. The fitted line indicates normal distribution.

Table 2. Results of different types of cooperativity using deterministic approach. N_H is Hill coefficient, M_{ave} is average ligand binding number of proteins, and *std. dev.* is standard deviation value. Parameters for the simulations are: initial protein concentration (M_0) = 8×10^{-5} mol/litre, ligand concentrations (L_0) = 0.0001 to 0.0003 mol/litre, number of points=21.

Type of Cooperativity	k_1 *(ml/mmol)*	k_2 *(ml/mmol)*	k_3 *(ml/mmol)*	k_4 *(ml/mmol)*	n_H	M_{ave}	*std. dev.*
Negative	1.4×10^5	7.5×10^4	4.5×10^3	3.6×10^3	0.57	1.86	0.395
Non-cooperativity	4.5×10^4	4.5×10^4	4.5×10^4	4.5×10^4	1	2.13	0.572
Positive	3.6×10^3	4.5×10^4	7.5×10^5	1.4×10^6	3.78	2.37	0.746

Table 3. Results of effect of rate constant values on a protein with non-cooperativity, using deterministic approach. N_H is Hill coefficient, M_{ave} is average ligand binding number of proteins, and *std. dev.* is standard deviation value. Other parameters for the simulations are the same as Table 2.

Type of Cooperativity	k_1 *(ml/mmol)*	k_2 *(ml/mmol)*	k_3 *(ml/mmol)*	k_4 *(ml/mmol)*	n_H	M_{ave}	*std. dev.*
Non-cooperative	4.5×10^2	4.5×10^2	4.5×10^2	4.5×10^2	1	0.29	0.955
Non-cooperative	4.5×10^3	4.5×10^3	4.5×10^3	4.5×10^3	1	1.22	1.84
Non-cooperative	4.5×10^4	4.5×10^4	4.5×10^4	4.5×10^4	1	2.13	1.95

The effect of adding a constant value (ε) to the rate constants of a protein on positive cooperativitywas investigated. Results are shown in Table 4. The results show an incremental behavior in M_{ave} and *std. dev.* while increasing rate constant values.

Table 4. Results of effect of rate constant values on a protein with positive cooperativity using deterministic approach. The ε value was 10,000 which means k_i values in lower rows are 10,000 more that the above ones. N_H is Hill coefficient, M_{ave} is average ligand binding number of proteins, and *std. dev.* is standard deviation value. Other parameters for the simulations are the same as Table 2.

Type of Cooperativity	k_1 *(ml/mmol)*	k_2 *(ml/mmol)*	k_3 *(ml/mmol)*	k_4 *(ml/mmol)*	n_H	M_{ave}	*std. dev.*
Positive	3.6×10^3	4.5×10^4	7.5×10^5	1.4×10^6	3.78	2.37	0.746
Positive	1.36×10^4	5.5×10^4	7.6×10^5	1.41×10^6	3.625	2.41	0.754
Positive	2.36×10^4	6.5×10^4	7.7×10^5	1.42×10^6	3.526	2.43	0.756

Similarly, we investigated the effect of the rate constant values on a protein with negative cooperativity. We subtracted a constant value (ε) from the rate constants of a protein in a negative cooperativity behavior while the protein retained its negative copperativity. Results are shown in Table 5. In this case, the results show that M_{ave} and *std. dev.* would decrease while the rate constant values decreased.

Table 5 Results of effect of rate constant values on a protein with negative cooperativity using deterministic approach. The ε value was 1,000 which means k_i values in lower rows are 1,000 less than the upper ones. N_H is Hill coefficient, M_{ave} is average ligand binding number of proteins, and *std. dev.* is standard deviation value. Other parameters for the simulations are the same as Table 2.

Type of Cooperativity	k_1 *(ml/mmol)*	k_2 *(ml/mmol)*	k_3 *(ml/mmol)*	k_4 *(ml/mmol)*	n_H	M_{ave}	*std. dev.*
Negative	1.4×10^5	7.5×10^4	4.5×10^3	3.6×10^3	0.57	1.86	0.395
Negative	1.39×10^5	7.4×10^4	3.5×10^3	2.6×10^3	0.52	1.83	0.368
Negative	1.38×10^5	7.3×10^4	2.5×10^3	1.6×10^3	0.466	1.79	0.336

All models and results are found in the supplementary files S2 and S3_Table3-5.

3.3 Stochastic Approach

In the stochastic simulations, a protein complex with four binding sites was considered (detailed equations are available in supplementary file S1-B). The Tau-leaping algorithm was used for the simulation process. Similar to the deterministic approach, we performed simulations for 3 different cases; negative cooperativity ($K_1 > K_2 > K_3 > K_4$), non-cooperativity ($K_1 = K_2 = K_3 = K_4$), and positive cooperativity ($K_1 < K_2 < K_3 < K_4$), where K_i is statistically corrected rate

constant. First, we adjusted the rate constants values, similar to the deterministic simulation, and explored the behavior of the system on just one trajectory (Figure 3).

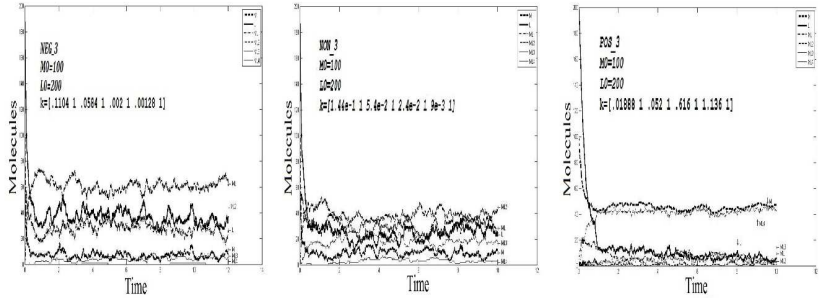

Figure 3. Changes in molecules of a stochastic simulation in different types of cooperativity in one trajectory. (L_0=200, M_0 = 100, M is total number of free proteins, L is total number of free ligands, ML, ML_2, ML_3, and ML_4 are total number of proteins with 1, 2, 3, and 4 ligands attached to, respectively.)

As it is shown in Table 6, the results do seem to be inaccurate at first, since, for instance, the value of Hill coefficient should be equal to 1 in a non-cooperativity situation. Although the results are not accurate enough, the system still shows a similar transition from negative to positive cooperativity (M_{ave} and *std. dev.* values demonstrate increasing behavior from negative to positive cooperativity).

Table 6. Results of different types of cooperativity using stochastic approach (tau-leaping algorithm) for just one trajectory. N_H is Hill coefficient, M_{ave} is average ligand binding number of proteins, and *std. dev.* is standard deviation value. Parameters for the simulations are: initial protein numbers (M_0) = 100, initial ligand numbers (L_0) = 125 to 375, number of points=21.

Type of Cooperativity	k_1 *(ml/mmol)*	k_2 *(ml/mmol)*	k_3 *(ml/mmol)*	k_4 *(ml/mmol)*	n_H	M_{ave}	*std. dev.*
Negative	0.1104	0.0584	0.002	0.00128	1.239	1.8538	0.40064
Non-cooperativity	0.036	0.036	0.036	0.036	1.78	2.1314	0.56639
Positive	0.01888	0.052	0.616	1.136	2.567	2.421	0.74746

The reason why there is a difference between stochastic and deterministic simulations (where all rate constants values have been chosen the same based on converting concentration into number unit), is probably due to stochastic behavior of the system, since the results were calculated for a single trajectory. Unlike a deterministic simulation of a system, one would expect some differences between the results for every run of a stochastic simulation. To see

similar results, number of replicates (here number of trajectories) should increase. Therefore, the stochastic simulations were repeated for 10 trajectories. The results are shown in Table 7 and one could see closer similarity to the deterministic results, because increasing the number of trajectories would reduce stochastic behavior of the system.

Table 7. Results of different types of cooperativity using stochastic approach (tau-leaping algorithm) for 10 trajectories. N_H is Hill coefficient, M_{ave} is average ligand binding number of proteins, and *std. dev.* is standard deviation value. Parameters for the simulations are: initial protein number (M_0) = 100, ligand numbers (L_0) = 125 to 375, number of points=21.

Type of Cooperativity	k_1 *(ml/mmol)*	k_2 *(ml/mmol)*	k_3 *(ml/mmol)*	k_4 *(ml/mmol)*	n_H	M_{ave}	*std. dev.*
Negative	0.1104	0.0584	0.002	0.00128	0.5678	1.8678	0.3987
Non-cooperativity	0.036	0.036	0.036	0.036	0.9657	2.1261	0.57
Positive	0.01888	0.052	0.616	1.136	3.36	2.3736	0.7499

Similar to the deterministic approach, we explored the effect of the rate constant values on M_{ave} and *std. dev.* values while a protein retains its cooperativity type. Here, results for 10 trajectories, to compare them the deterministic results, are shown. Results of just one trajectory can be found in the supplementary file S1-C.

For a protein with non-cooperativity behavior, results are shown in Table 8, which show an incremental behavior in M_{ave} and *std. dev.* while increasing the rate constant values.

Table 8. Results of effect of rate constant values on a protein with non-cooperativity using stochastic approach for 10 trajectories. N_H is Hill coefficient, M_{ave} is average ligand binding number of proteins, and *std. dev.* is standard deviation value. Other parameters for the simulations are the same as Table 7.

Type of Cooperativity	k_1 *(ml/mmol)*	k_2 *(ml/mmol)*	k_3 *(ml/mmol)*	k_4 *(ml/mmol)*	n_H	M_{ave}	*std. dev.*
Non-cooperative	3.6×10^{-4}	3.6×10^{-4}	3.6×10^{-4}	3.6×10^{-4}	1.1	0.29	0.0926
Non-cooperative	3.6×10^{-3}	3.6×10^{-3}	3.6×10^{-3}	3.6×10^{-3}	0.9949	1.2221	0.3161
Non-cooperative	3.6×10^{-2}	3.6×10^{-2}	3.6×10^{-2}	3.6×10^{-2}	0.9657	2.1261	0.57

For a protein in a positive cooperativity behavior, we added a constant value (ε) to the rate constants. Results are shown in Table 9, which show an incremental behavior in M_{ave} and *std. dev.* while increasing the rate constant values.

Table 9. Results of effect of rate constant values on a protein with positive cooperativity using stochastic approach for 10 trajectories. The ε value was 0.008 which means k_i values in lower rows are 0.008 more that the above ones. N_H is Hill coefficient, M_{ave} is average ligand binding number of proteins, and *std. dev.* is standard deviation value. Other parameters for the simulations are the same as Table 7.

Type of Cooperativity	k_1 *(ml/mmol)*	k_2 *(ml/mmol)*	k_3 *(ml/mmol)*	k_4 *(ml/mmol)*	n_H	M_{ave}	*std. dev.*
Positive	2.88×10^{-3}	3.6×10^{-2}	6×10^{-1}	1.12	3.36	2.3736	0.7499
Positive	1.088×10^{-2}	4.4×10^{-2}	6.08×10^{-1}	1.128	2.842	2.4118	0.7539
Positive	1.888×10^{-2}	5.2×10^{-2}	6.16×10^{-1}	1.136	3.064	2.4193	0.755

Finally, for a protein with negative cooperativity, we subtracted a constant value (ε) from the rate constants. Results are shown in Table 10 which showed M_{ave} and *std. dev.* would be decreased while there is a reduction in the rate constant values.

It is possible to compare the results of Tables 8, 9, and 10 to Tables 3, 4, and 5, respectively, to see similar behaviors between stochastic and deterministic approaches.

The complete results of all the stochastic simulations are available in the supplementary file S1-C.

Table 10. Results of effect of rate constant values on a protein with negative cooperativity using stochastic approach for 10 trajectories. The ε value was 0.0008 which means k_i values in lower rows are 0.0008 less than the upper ones. N_H is Hill coefficient, M_{ave} is average ligand binding number of proteins, and *std. dev.* is standard deviation value. Other parameters for the simulations are the same as Table 7.

Type of Cooperativity	k_1 *(ml/mmol)*	k_2 *(ml/mmol)*	k_3 *(ml/mmol)*	k_4 *(ml/mmol)*	n_H	M_{ave}	*std. dev.*
Negative	1.12×10^{-1}	6×10^{-2}	3.6×10^{-3}	2.88×10^{-3}	0.5678	1.8678	0.3978
Negative	1.112×10^{-1}	5.92×10^{-2}	2.8×10^{-3}	2.08×10^{-3}	0.5336	1.8228	0.3714
Negative	1.104×10^{-1}	5.84×10^{-2}	2×10^{-3}	1.28×10^{-3}	0.4626	1.7808	0.3364

4 Discussion

One of the properties of proteins is their ability to regulate their activities, in order to function appropriately in the face of varied physiological environment that might arise during the life of a cell [35]. Among different kinds of proteins, multimer proteins have the interesting property of cooperativity. More than 50% of proteins are in complex forms and lots of them have more

than just one binding site for ligands [27,28, 36]. It has been demonstrated that various binding behaviors exist in multimer proteins; a) non-cooperativity in which binding of ligands is completely independent, that is, the binding of the first ligand does not affect the affinity of the next ligand, b) positive cooperativity in which binding of each ligand increases the affinity for the next binding, and c) negative cooperativity in which binding of each ligand decreases the affinity for the next binding. In this study, we attempted to answer why multimer proteins are needed and why a specific cooperativity type is seen?

First, we explored the different kinds of cooperativity. Our study indicated that there is not any significant preference for a specific type (Figure 1). In the next step, we considered stability advantage as a key criterion for cooperativity selection. Thus, we investigated cooperativity from a thermodynamic point of view [22].

A few studies illustrated that ΔG would be decrease in a positive cooperativity situation, which makes the system more stable [37]. However, it not clearly obvious that such decrement is due to the decrease in enthalpy term or the increase in entropy term or both [38,39]. If a protein choose its type of cooperativity solely based on stability, it would always be suitable to select positive cooperativity. But based the literature on this subject, positive cooperativity is not dominant in nature (Figure 1). Also, few studies demonstrated that low ligand concentrations would result in negative cooperativity[37]. Again, based the literature, there is not any preference based on ligand concentrations (Figure 1).

We believe that cooperativity, introduced by changing in rate constant, should result in some advantages beyond stability, ligand concentrations, and likes. Consequently, we performed deterministic simulations to explore the effect of rate constant fora specific kind of cooperativity. When the system reached steady-state, we calculated the average number and the standard deviation (*std. dev.*) of bound protein molecules (M_{ave}), Here, we considered *std. dev*. values as a measure of noise in the cooperative system. One could see three reasonable conclusions based on the simulation results; 1) M_{ave} and *std. dev.* values increase from negative to positive cooperativity, 2) in a positive and non-cooperativity behaviors, M_{ave} and *std. dev.* would increase while increasing the rate constant values, and 3) in negative cooperativity, M_{ave} and *std. dev.* values decrease when there is a reduction in the rate constant values.

However, when molecules are few (which is the norm in cells), deterministic approaches can not capture all the characteristics of the system and stochasticity plays an important role [37]. Therefore, simulations were repeated using a stochastic approach. Interestingly, we have seen the same behavior; increasing in M_{ave} and *std. dev.* values from negative to positive cooperativity, increasing in M_{ave} and *std. dev*. values when increasing the rate constant values

for positive and non-cooperativity conditions, and decreasing in M_{ave} and *std. dev.* values when reducing the rate constant values for negative cooperativity condition.

Based on the simulation results, it could be inferred that for cooperative proteins, the rate constants might be the regulators. Rate constant values can be changed to obtain a desired cooperativity behavior, not only because of stability or ligand concentrations, but also influencing the *std. dev.* values. It seems rate constant values are a way to tune the noise (represented as *std. dev.* in our study) of a system. In a more stable environment, a system needs to be less flexibility, so it prefers negative cooperativity, which leads to reduction in noise (i.e. *std. dev.* value). In a fluctuating environment, positive cooperativity and increasing the noise would be preferred. In biology, noise is not necessarily a destructive force; instead, it could be advantageous in the hands of natural selection. A wider range of *std. dev.* value would guarantee robustness and stability of a system in different situations [41]. Our study reveals that for model of cooperativity, the biochemical rate constant values act as a source of intrinsic noise. So, changing intrinsic noises (*std. dev.* values) would affect the cell function and provide the necessary variation in a population of cells.

Finally, we came back to our previous questions: why multimer proteins are needed and why a specific cooperativity type is chosen? Our answer to this question is that, although multimer proteins require more energy to assemble and increase the possibility of flaws (such as aggregations), but they enable the system to tune the noise in fluctuating environments, which can be a valuable trait during evolution. So, it would be beneficial for a protein to be capable of changing cooperativity behavior under different conditions, e.g., positive cooperativity might lead to an increase in stability and a protein could bind more ligands simultaneously, whereas negative cooperativity would be useful where ligand concentration is low. This study suggests that the rate constants can be viewed as tools for adjusting the noise in a cooperative protein. An adjustment that affect the flexibility of the protein in each type of cooperativity. However, our results raises more questions, which would benefit from further studies.

References

[1] T. M. Greco, I. M. Cristea, The biochemical evolution of protein complexes, *Trends Biochem. Sci.* **41** (2016) 4-6.

[2] J. A. Marsh, S. A. Teichmann, Structure, dynamics, assembly, and evolution of protein complexes, *Annu. Rev. Biochem.* **l84** (2015) 551-575.

[3] P. Smits, J. A. Smeitink, L. P. van den Heuvel, M. A. Huynen, T. J. Ettema, Reconstructing the evolution of the mitochondrial ribosomal proteome, *Nucleic Acids Res*. **35** (2007) 4686-4703.

[4] J. Lowe, L. A. Amos, Evolution of cytomotive filaments: The cytoskeleton from prokaryotes to eukaryotes, *Int. J. Biochem. Cell Biol*. **41** (2009) 323-329.

[5] D. G. Scofield, M. Lynch, Evolutionary diversification of the Sm family of RNA-associated proteins, *Mol. Biol. Evol*. **25** (2008) 2255-2267.

[6] J. A. Dent, The evolution of pentameric ligand-gated ion channels, *Adv Exp Med. Biol*. **683** (2010) 11-23.

[7] F. Alber, S. Dokudovskaya, L. M. Veenhoff, W. Zhang, J. Kipper, D. Devos, A. Suprapto, O. Karni-Schmidt, R. Williams, B. T. Chait, A. Sali, M. P. Rout, The molecular architecture of the nuclear pore complex, *Nature* **450** (2007) 695-701.

[8] C. H. Schein, Controlling oligomerization of pharmaceutical proteins, *Pharm. Acta. Helv*. **69** (1994) 119-126.

[9] M. H. Ali, B. Imperiali, Protein oligomerization: How and why, *Bioorganic Med. Chem*. **13** (2005) 5013-5020.

[10] P. Wong, S. Althammer, A. Hildebrand, A. Kirschner, P. Pagel, B. Geissler, P. Smialowski, F. Blöchl, M. Oesterheld, T. Schmidt, N. Strack, F. J. Theis, A. Ruepp, D. Frishman, An evolutionary and structural characterization of mammalian protein complex organization, *BMC Genomics* **9** (2008) #629.

[11] M. Lynch, The evolution of multimeric protein assemblages, *Mol. Biol. Evol*. **29** (2012) 1353-1366.

[12] M. V. Smoluchowski, Attempt for a mathematical theory of kinetic coagulation of colloid solutions, *Z. Phys. Chem*. **92** (1917) 129–168.

[13] S. Bershtein, W. M. Mu, and E. I. Shakhnovich, Soluble oligomerization provides a beneficial fitness effect on destabilizing mutations, *Proc. Natl. Acad. Sci. U.S.A.* **109** (2012) 4857-4862.

[14] M. Lynch, Evolutionary diversification of the multimeric states of proteins, *Proc. Natl. Acad. Sci. U.S.A.* **110** (2013) E2821-E2828.

[15] K. G. Phillips, P. Kuhn, O. J. T. McCarty, Physical biology in cancer. 2. The physical biology of circulating tumor cells, *Am J Physiol. Cell Physiol*. **306** (2014) C80-C88.

[16] S. Pechmann, J. Frydman, Interplay between chaperones and protein disorder promotes the evolution of protein networks, *PLOS Comput. Biol*. **10** (2014) #e1003674.

[17] K. Hashimoto, H. Nishi, S. Bryant, A. P. Panchenko, Caught in self-interaction: evolutionary and functional mechanisms of protein homooligomerization, *Phys. Biol.* **8** (2011) #035007.

[18] N. J. Marianayagam, M. Sunde, J. M. Matthews, The power of two: protein dimerization in biology, *Trends Biochem. Sci*. **29** (2004) 618-625.

[19] C. Bohr, K. Hasselbalch, A. Krogh, Concerning a biologically important relationship - The influence of the carbon dioxide content of blood on its oxygen binding, *Skand. Arch. Physiol.* **16** (1904) 402–412.

[20] D. E. Koshland, K. Hamadani, Proteomics and models for enzyme cooperativity, *J. Biol. Chem*.. **277** (2002) 46841-46844.

[21] H. Abeliovich, An empirical extremum principle for the Hill coefficient in ligand-protein interactions showing negative cooperativity, *Biophys. J*. **89** (2005) 76-79.

[22] T. K. Dam, R. Roy, D. Pagé, C. F. Brewer, Thermodynamic binding parameters of individual epitopes of multivalent carbohydrates to concanavalin a as determined by "reverse" isothermal titration microcalorimetry, *Biochemistry* **41** (2002)1359-1363.

[23] C. J. Tsai, A. Del Sol, R. Nussinov, Protein allostery, signal transmission and dynamics: a classification scheme of allosteric mechanisms, *Mol. BioSyst*. **5** (2009) 207-216.

[24] M. Grinfeld, J. P. Bennett, J. Hubble, Application of computer algebra to affinity-binding equations, *IMA J. Manag. Math.* **8** (1997) 157-166.

[25] R. Milo, What is the total number of protein molecules per cell volume? A call to rethink some published values, *Bioessays* **35** (2013) 1050-1055.

[26] P. Lecca, I. Laurenzi, F. Jordan, *Deterministic Versus Stochastic Modelling in Biochemistry and Systems Biology*, Woodhead Publishing, 2013.

[27] J. E. Ferrell, Jr., Q&A: cooperativity, *J. Biol.* **8** (2009) #53.

[28] M. I. Stefan, N. LeNovere, Cooperative binding, *PLoS Comput. Biol*. **9** (2013) #1003106.

[29] F.T. Bergmann, S. Hoops, B. Klahn, U. Kummer, P. Mendes, J. Pahle S. Sahle, COPASI and its applications in biotechnology, *J. Biotechnol.* **261** (2017) 215-220.

[30] H. Qian, L. M. Bishop, The chemical master equation approach to nonequilibrium steady-state of open biochemical systems: linear single-molecule enzyme kinetics and nonlinear biochemical reaction networks, *Int. J. Mol. Sci.* **11** (2010) 3472-3500.

[31] S. S. Andrews, T. D. Adam, P. Arkin, Stochastic models of biological processes, in: R. A. Meyers (Ed.), *Encyclopedia of Complexity and System Science*, Springer-Verlag, New York, 2009, pp. 8730-8749.

[32] T. C. Meng, S. Somani, P. Dhar, Modeling and simulation of biological systems with stochasticity, *In Silico Biol.* **4** (2004) 293-309.

[33] D. J. Higham, Modeling and simulating chemical reactions, *SIAM Review* **50** (2008) 347-368.

[34] K. B. Tamás, J. Székely, Stochastic simulation in systems biology, *Comput. Struct. Biotechnol. J.* **14** (2014) 14-25.

[35] G. M. Cooper, *The Cell: A Molecular Approach*, Sinauer Associates, Sunderland, 2000.

[36] C. A. Hunter, H. L. Anderson, What is cooperativity? *Angew Chem. Int. Ed. Engl.* **48** (2009) 7488-7499.

[37] E. M. Y. Suzuki, D. Tsuchiya, H. Jingami, Negative Cooperativity of glutamate binding in the dimeric metabotropic glutamate receptor subtype, *J. Biol. Chem*. **279** (2004) 35526-35534.

[38] P. N. Romasanta, L. M. Curto, N. Urtasun, M. B. Sarratea, S. Chiappini, M. V. Miranda, J. M. Delfino, R. A. Mariuzza, M. M. Fernandez, E. L. Malchiodi, A positive cooperativity binding model between Ly49 natural killer cell receptors and the viral immunoevasin m157, *J. Biol. Chem.* **289** (2014) 5083–5096.

[39] A. Brown, Analysis of cooperativity by isothermal titration calorimetry, *Int. J. Mol. Sci.* **10** (2009) 3457-3477.

[40] R. Schrodinger, E. Schrödinger, *What Is Life?: With Mind and Matter and Autobiographical Sketches*, Cambridge Univ. Press, Cambridge, 1992.

[41] L. S. Tsimring, Noise in biology, *Rep. Prog. Phys.* **77** (2014) #026601.