ISSN: 0975-8585

Research Journal of Pharmaceutical, Biological and Chemical Sciences

Mathematical Modeling and Protein Structure Prediction for Amino Acid Sequences through Protein Plots and Ecological changes.

R Senthamarai and S Balamuralitharan*.

Faculty of Engineering and Technology, Department of Mathematics, SRM University, Kattankulathur-603 203, Tamil Nadu, India.

ABSTRACT

This paper deals with amino acid sequences a-c-d-e-f-g-h-i-k-l-m-n-p-q-r-s-t-v-w-y in testing the protein linear plots. We observe that the relation between ecological fluctuations and protein changeover is studied. The model is tested further with the amino acid value and found that the relation changes to properties as well as different methods. We used protein plot tool in MATLAB for the purpose to find out the fluctuation of protein turnover and ecological change. Finally, we conclude that protein changeover rate is relatively higher than ecological change.

Keywords: *Ecological change, protein plot, amino acid sequences.*

Corresponding author

ISSN: 0975-8585

INTRODUCTION

The genomics revolution is providing gene sequences in exponentially increasing numbers. Converting this sequence information into functional information for the gene products coded by these sequences is the challenge for post-genomic biology. The first step in this process will often be the interpretation of a protein sequence in terms of the three dimensional structure into which it folds [1].

We may not know how long the atmosphere has been changed from one state to another which does not alter the expected waiting time until it shifts to the other state. This means that the cell has nothing to gain by storing information about the amount of time exhausted in the current atmosphere [2]. A more complex rigid control structure would be optimal in the simple atmosphere and our analysis focuses simple guideline of transcription rates [3-9]. Therefore more rational ecological scenario may impress selection for more complex approach.

A metabolic system has not effectively give any atmospheric changes at a slow rate, which has no response to the stimulus changing at rapid rates in the same study report of researchers [10-25]. This type of variation in response to atmospheric study which was observed from a different point of view in stimulation of a simple model, switching between like and dislike protein production [26]. Similarly the stimulation should protein synthesis varied over a wide range when the switch rate was very slow but was essentially steady at a rapid rate of switching. Major examples of the biochemical functions of proteins include binding; catalysis; operating as molecular switches; and serving as structural components of cells and organisms [27-29].

Proteins may bind to other macromolecules, such as DNA in the case of DNA polymerases or gene regulatory proteins, or to proteins in the case of a transporter or a receptor that binds a signaling molecule [30]. This function exploits the ability of proteins to present structurally and chemically diverse surfaces that can interact with other molecules with high specificity. Catalysis requires not only specific binding, to substrates and in some cases to regulatory molecules, but also specific chemical reactivity. Regulated enzymes and switches, such as the signaling G proteins, require large-scale conformational changes that depend on a delicate balance between structural stability and flexibility.

Structural proteins may be as strong as silk or as tough and durable as keratin, the protein component of hair, horn and feathers; or they may have complex dynamic properties that depend on nucleotide hydrolysis, as in the case of actin and tubulin [30 - 42]. This extraordinary functional diversity and versatility of proteins derives from the chemical diversity of the side chains of their constituent amino acids, the flexibility of the polypeptide chain, and the very large number of ways in which polypeptide chains with different amino acid sequences can fold.

The works of researchers provide a clearer understanding and the study of protein synthesis evolves a bio-modal distribution. In protein synthesis or mRNA synthesis we show the time dependent distributions converge to the moment's stationary distributions with exponential rates depending on the damage and switching rates [43-45]. We also consider the two time step process of transcription and translation from gene to protein and give some arguments as how the relative rates can affect the cells ability to hope with the atmosphere.

Both steps are also consider and compared from a different modeling approach, with different goals. Again the relative rates play on important role in the analysis of protein plots and mRNA. We the relative rates of ecological fluctuations, mRNA and protein turnover determine that the ability of a cell to track its atmosphere. Our stimulation indicates protein concentration will track changes in atmosphere provided both RNA and protein changeover rates are higher than rates of ecological change [11].

MATERIALS AND METHODS

Proteins are polymers of 20 different amino acids joined by peptide bonds. At physiological temperatures in aqueous solution, the polypeptide chains of proteins fold into a form that in most cases is globular [6]. The sequence of the different amino acids in a protein, which is directly determined by the

sequence of nucleotides in the gene encoding it, is its primary structure. This in turn determines how the protein folds into higher level structures [13].

The secondary structure of the polypeptide chain can take the form either of alpha helices or of beta strands, formed through regular hydrogen-bonding interactions between N–H and C=O groups in the invariant parts of the amino acids in the polypeptide backbone or main chain [17]. In the globular form of the protein, elements of either alpha helix, or beta sheet, or both, as well as loops and links that have no secondary structure, are folded into a tertiary structure. Many proteins are formed by association of the folded chains of more than one polypeptide; this constitutes the quaternary structure of a protein [21].

For a polypeptide to function as a protein, it must usually be able to form a stable tertiary structure under physiological conditions. On the other hand, the demands of protein function require that the folded protein should not be too rigid. Presumably because of these constraints, the number of folds adopted by proteins, though large, is limited [33]. Whether the limited number of folds reflects physical constraints on the number of stable folds, or simply the expedience of divergent evolution from an existing stable fold, is not known, but it is a matter of some practical importance: if there are many possible stable folds not represented in nature, it should be possible to produce completely novel proteins for industrial and medical applications [35].

The amino-acid side chains have different tendencies to participate in interactions with each other and with water [36]. These differences profoundly influence their contributions to protein stability and to protein function. Alanine and leucine are strong helix-favoring residues, while proline is rarely found in helices because its backbone nitrogen is not available for the hydrogen bonding required for helix formation [39]. The aromatic side chain of phenylalanine can sometimes participate in weakly polar interactions.

Hydrophilic amino-acid residues are able to make hydrogen bonds to one another, to the peptide backbone, to polar organic molecules, and to water. This tendency dominates the interactions in which they participate [40]. Some of them can change their charge state depending on their pH or the microenvironment.

MATHEMATICAL MODELING

We often use a distribution associated with the gamma function defined for p > 0 by the formula

$$\Gamma_p = \int_0^\infty x^{p-1} e^{-x} dx \tag{1}$$

Substituting x = ay, a > 0 in (1)

$$\frac{\Gamma_p}{a^p} = \int_0^\infty y^{p-1} e^{-ay} dy \tag{2}$$

$$\varphi(t) = \frac{b^p \Gamma_p}{\Gamma\left(p^{(b-it)p}\right)} = \frac{1}{\left(1-\frac{it}{b}\right)^p} \tag{3}$$

The random variable with density f(x), defined by the formula

$$f(x) = \begin{cases} 0, & x < 0 \\ \lambda e^{-\lambda x}, & x \geq 0 \end{cases} \tag{4}$$

$$f(x) = \begin{cases} \frac{1}{B(p,q)} x^{p-1}(1-x)^{q-1}, & x < 0 \\ 0, & x \geq 0 \end{cases} \tag{5}$$

If is convenient to obtain the moments of the Beta distribution directly from the formula

$$\pi = \frac{(r_1 - r_0)^{\alpha+\beta}}{\mu} B(\alpha, 1+\beta) \tag{6}$$

where *B(α, β)* is the Beta function. Similarly

Let ϕ_0, ϕ_1 be the solutions of the distribution, and π_0, π_1 be the unique stationary distribution. We consider time varying moments of the differences $T = \int_a^b x^k(\phi_i - \pi_i)dx, \quad i = 0,1.$

(7)

$$S(t) = e^{-\mu t}\int_0^t e^{\mu s}e^{-\lambda s}ds = \frac{e^{-\lambda t} - e^{-\mu t}}{\mu - \lambda} = te^{-\theta} \le te^{-\lambda t}, t \ge 0. \tag{8}$$

$$\begin{aligned} M(t) &= F(t) + \sum_{n=1}^{\infty}\int_0^{\tau} F^{(n)}(t-x)\,f(x)dx, \\ &= F(t) + \int_0^{\tau} M(t-x)\,f(x)dx. \end{aligned} \tag{9}$$

We use stochastic simulations to compute realizations of $\{Z_k(t)\}_{k=1}^{\infty}$, for large *N* and then an approximation $\rho_N(x, y)$ of the density $\rho(x, y)$ as follows. Let *[a, b]* be subdivided into M_x subintervals of equal width Δx, and *[c, d]* be subdivided into M_y subintervals of equal width Δy. This determines M_xM_y sub rectangles $R_{i,j}$ of *R*. We define $\rho_N(x, y)$ as a piecewise constant function, with $\rho_N(x, y) = \rho_{ij}$, for all *(x, y)* ϵ $R_{i,j}$, where $\rho_{ij}\Delta x\Delta y = \frac{1}{N}\sum_{k=1}^{N}\delta_{zk}(R_{ij})$ [4]. Piecewise constant marginal densities are then determined by $\sigma_N^y(x) = \sigma_i^y \in [x_{i-1}, x_i],\ \sigma_N^x(x) = \sigma_j^x \in [y_{i-1}, y_i]$ with

$$\begin{aligned} \sigma_i^y &= \sum_{j=1}^{M_y}\rho_{ij}\Delta y, i = 1,2,3,\ldots,M_x, \\ \sigma_j^x &= \sum_{j=1}^{M_x}\rho_{ij}\Delta x,\ j = 1,2,3,\ldots,M_y. \end{aligned} \tag{10}$$

One of the marginal densities of $\rho(x, y)$ can be computed analytically [6]. Integrating the system with respect to *y* over the interval *[c, d],* and using the fact ρ_0, ρ_1 should be zero on the edges *y = c*, *y = d* (a consequence of the integral identities ρ_0, ρ_1 must satisfy), leads to the conclusion that, $\pi_0(x) = \int_c^d \rho_0(x, y)dy,\ \pi_1(x) = \int_c^d \rho_1(x, y)dy$ [7].

Relationship between the DNA and the Amino-Acid Sequence of the Protein

The genetic code is the formula that converts hereditary information from genes into proteins. Every amino acid in a protein is represented by a codon consisting of three consecutive nucleotides in the gene. DNA contains four different nucleotides, with the bases adenine (A), guanine (G), thymidine (T) and cytosine (C), whose sequence in a gene spells out the sequence of the amino acids in the protein that it specifies: this is the primary structure of the protein [41].

The nucleotide sequence of the DNA is transcribed into messenger RNA (mRNA), with uridine (U) replacing thymine (T). It shows the correspondence between the 64 possible three-base codons in mRNA and the 20 naturally occurring amino acids. Some amino acids are specified by only one codon, whereas others can be specified by as many as six different codons: the genetic code is degenerate [42]. There are three codons that do not code for amino acids, but signal the termination of the polypeptide chain.

In bacteria and other lower organisms, the relationship between the base sequence of the gene and the amino acid sequence of the corresponding protein is strictly linear: the protein sequence can be read directly from the gene sequence. In higher organisms, however, genes are typically segmented into coding regions (exons) that are interrupted by noncoding stretches (introns). These non-coding introns are transcribed into RNA, but are enzymatically excised from the resulting transcript, and the exons are then spliced together to make the mature mRNA [43].

The process of intron removal and exon ligation has been exploited in the course of evolution through alternative splicing, in which exon segments as well as intron segments may be differentially excised from the primary transcript to give more than one mRNA and thus more than one protein [44-45]. Depending on the arrangement of the introns, alternative splicing can lead to truncated proteins, proteins with different stretches of amino acids in the middle, or frame shifts in which the sequence of a large part of the protein is completely different from that specified by an in-frame reading of the gene sequence.

Coding sequences can also be modified by RNA editing. In this process, some nucleotides are changed to others, and stretches of additional nucleotides can be inserted into the mRNA sequence before translation occurs. Modification of the coding sequences by RNA processing in these ways complicates the interpretation of genomic sequences in terms of protein structure, though this complication does not apply to cDNA sequences, which are artificially copied by reverse transcription from mRNA [21].

The genetic code is the formula that converts hereditary information from genes into proteins.
Every amino acid in a protein is represented by a codon consisting of three consecutive nucleotides in the gene. Amino acids are crucial components of living cells because they are easy to polymerize [12]. Alpba Amino acids are preferable to beta-amino acids because the latter are too flexible to form spontaneously folding polymers. The amino acids of a protein chain are covalently joined by amide bonds, often called peptide bonds: for this reason, proteins are also known as polypeptides [15].

The amide bonds in the backbone are the only covalent bonds that hold the residues together in most proteins. In proteins that are secreted, or in the extracellular portions of cell-surface proteins, which are not exposed to the reducing environment in the interior of the cell, there may be additional covalent linkages present in the form of disulfide bridges between the side chains of cysteine residues [26]. Except for cross-links like these, however, the remainder of the stabilization energy of a folded protein comes not from covalent bonds but from noncovalent weakly polar interactions.

Weakly polar interactions depend on the electrostatic attraction between opposite charges. The charges may be permanent and full, or fluctuating and partial. But in principle, all polar interactions are electrostatic and the effect is the same: positively polarized species will associate with negatively polarized ones. Such interactions rarely contribute even one-tenth of the enthalpy contributed by a single covalent bond, but in any folded protein structure there may be hundreds to thousands of them, adding up to a very large contribution [13-17].

RESULTS

In contrast to the alpha helix, the beta pleated sheet, whose name derives from the corrugated appearance of the extended polypeptide chain, involves hydrogen bonds between backbone groups from residues distant from each other in the linear sequence [2]. In beta sheets, two or more strands that may be widely separated in the protein sequence are arranged side by side, with hydrogen bonds between the strands [7]. The strands can run in the same direction (parallel beta sheet) or antiparallel to one another; mixed sheets with both parallel and antiparallel strands are also possible (see Fig.1, & Appendix Table 1).

Parallel sheets are always buried and small parallel sheets almost never occur. Antiparallel sheets by contrast are frequently exposed to the aqueous environment on one face (see Fig.1). These observations suggest that antiparallel sheets are more stable, which is consistent with their hydrogen bonds being more linear. Silk, which is notoriously strong, is made up of stacks of antiparallel beta sheets [19]. Antiparallel sheets most commonly have beta turns connecting the strands, although sometimes the strands may come from

ISSN: 0975-8585

discontiguous regions of the linear sequence, in which case the connections are more complex and may include segments of alpha helix. Parallel sheet strands are of necessity always discontiguous, and the most common connection between them is an alpha helix that packs against a face of the beta sheet.

The polypeptide chain in a beta sheet is almost fully extended. The distance between consecutive residues is 3.3 Å and the phi and psi angles for peptides in beta sheets are approximately –130° and +125° respectively [20]. Beta strands usually have a pronounced right-handed twist, due to steric effects arising from the L-amino acid configuration (see Fig.1). Parallel strands are less twisted than antiparallel ones. The effect of the strand twist is that sheets consisting of several long strands are themselves twisted. Because the polypeptide chain in a beta sheet is extended, amino-acid side chains such as those of valine and isoleucine, which branch at the beta carbon, can be accommodated more easily in a beta structure than in a tightly coiled alpha helix where side chains are crowded more closely together. Although unbranched side chains can fit in beta structures as well, branched side chains appear to provide closer packing so they are found more frequently in sheets than other residues.

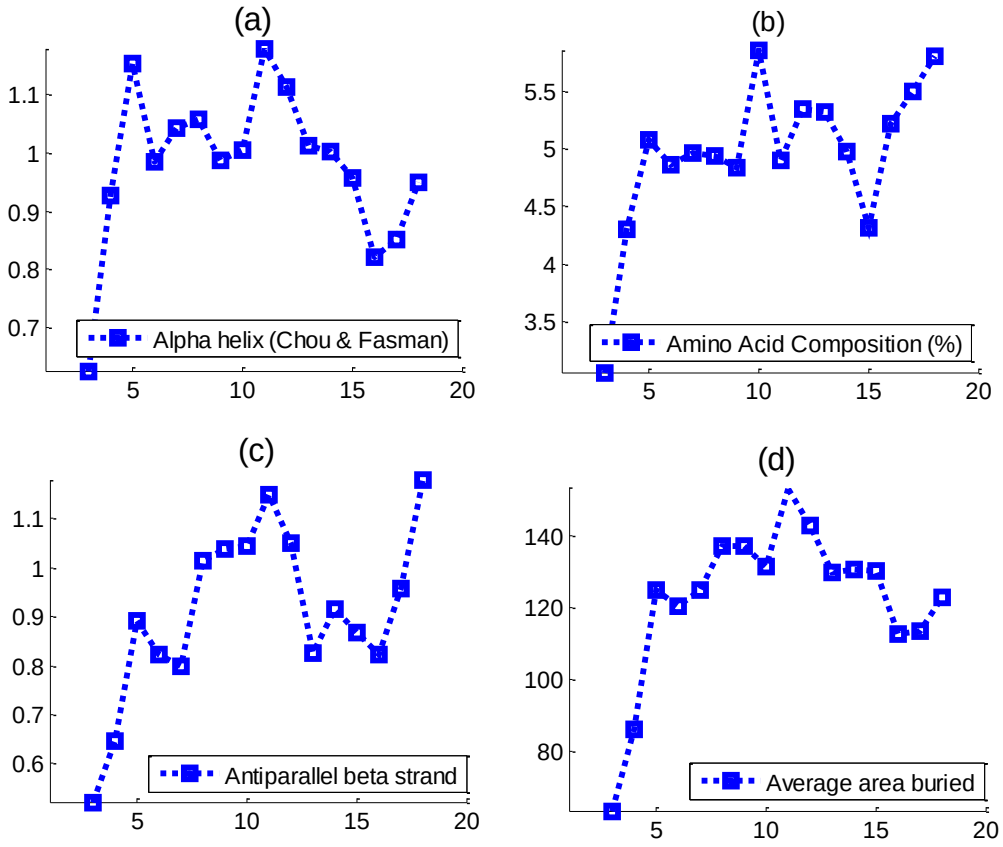

Fig. 1. Amino acid full sequence (a-c-d-e-f-g-h-i-k-l-m-n-p-q-r-s-t-v-w-y) linear fit properties of (a) alpha helix, (b) amino acid composition, (c) anti parallel beta strand and (d)average area buried.

However, it is generally easier to identify helical stretches in sequences than to identify sections of beta structure, and locating the ends of beta sections from sequence alone is particularly difficult (see Fig.2, & Appendix Table 1, 2). Analysis of the frequency with which different amino acids are found in different types of structure shows some general preferences. For example, long side chains such as those of leucine, methionine, glutamine and glutamic acid are often found in helices, presumably because these extended side chains can project out away from the crowded central region of the helical cylinder.

In contrast, residues whose side chains are branched at the beta carbon, such as valine, isoleucine and phenylalanine, are more often found in beta sheets, because every other side chain in a sheet is pointing in the

ISSN: 0975-8585

opposite direction, leaving room for beta-branched side chains to pack. Such tendencies underlie various empirical rules for the prediction of secondary structure from sequence, such as those of Chou and Fasman [11].

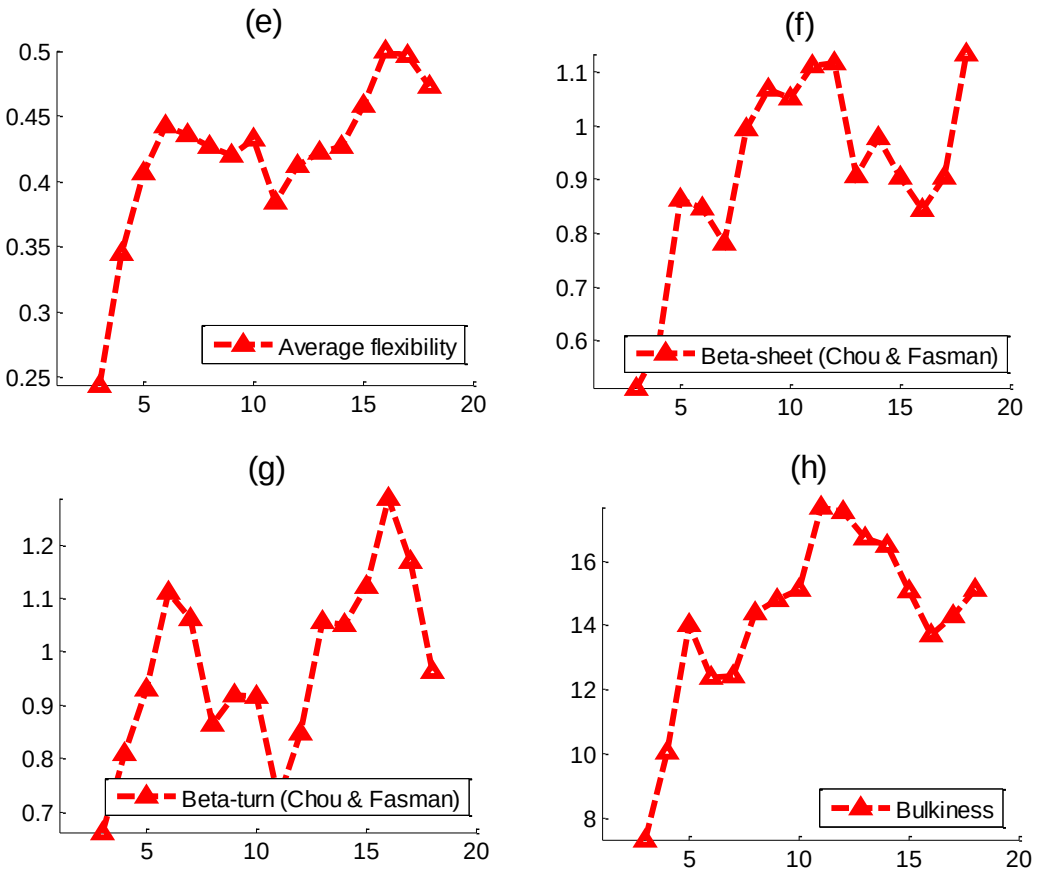

Fig. 2. Amino acid full sequence (a-c-d-e-f-g-h-i-k-l-m-n-p-q-r-s-t-v-w-y) linear fit properties of (e) average flexibility, (f) beta sheet, (g) beta turn and (h) bulkiness.

Domains that contain only beta sheet, tight turns and irregular loop structures are called beta domains. Proteins made up of beta domains include immunoglobulins, several enzymes such as superoxide dismutase, and proteins that bind to sugars on the surfaces of cells [5]. Because there are no helices to make long connections between adjacent strands of the beta sheet, all-beta domains contain essentially nothing but antiparallel beta structure, the strands of which are connected with beta turns and larger loops (see Fig.1).

The patterns of connections between strands give rise to beta sheets with two distinct topologies. The directionality of the polypeptide chain dictates that a strand in an antiparallel beta sheet can only be linked to a strand an odd number of strands away (see Fig.1). The most common connections are to an immediately adjacent strand or to one three strands away. If all the connections link adjacent strands, the beta sheet has an up-and-down structural motif. A particularly striking example is found in the enzyme neuraminidase from the influenza virus, which consists of a repeating structural motif of four antiparallel strands (see Fig.1). Each up-and-down motif forms the blade of a so-called beta-propeller domain.

Average flexibility sheets in beta domains tend to be oriented with one face on the surface of the protein, exposed to the aqueous surroundings, and the other face oriented toward the hydrophobic core (see Fig.2). This internal face is packed against another section of beta sheet with the inward-facing side chains of

both packing together to form a hydrophobic core. Thus, in beta domains, the sheet tends to be amphipathic, with one face predominantly hydrophilic while the other is almost entirely composed of hydrophobic amino acids. This characteristic may make it possible to recognize such domains from the distribution of polar and nonpolar residues in the amino-acid sequence if secondary structure prediction methods become more accurate.

In Bulkiness domains the beta sheet is composed of parallel or mixed strands; the parallel strands must be joined by long connections because the linking segment has to traverse the length of the sheet, and these connections are usually made by alpha helices connecting parallel adjacent strands, giving rise to beta-alpha-beta-alpha units (see Fig.2). As illustrated in Fig. 2, the crossover connection between the two parallel beta strands can be either right-handed or left-handed. The right-handed twist of the beta strand, however, produces an enormous bias toward the right-handed crossover topology: it is observed in more than 95% of alpha/beta structures [12]. This crossover rule is obeyed even when the connected strands are not adjacent or when the connecting segment is a loop, not a helix.

Just as two motifs predominate in antiparallel barrels and sandwiches, two motifs also account for nearly all alpha/beta domains. One of these is a closed structure called an alpha/beta barrel (see Fig.2). The other is an open twisted beta structure that looks somewhat like a saddle; we will call it an alpha/beta twist.

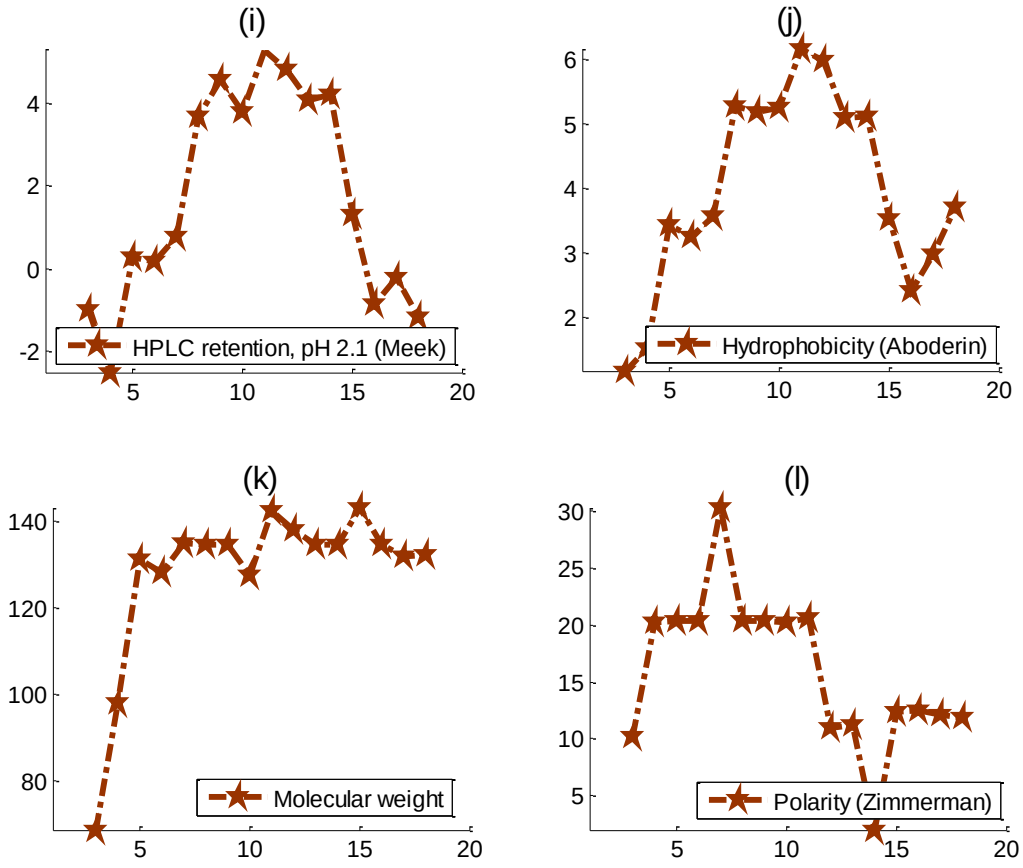

Fig. 3. Amino acid full sequence (a-c-d-e-f-g-h-i-k-l-m-n-p-q-r-s-t-v-w-y) linear fit properties of (i) HPLC retertion, (j) Hydrophobicity, (k) molecular weight and (l) polarity.

Alpha+beta domains contain both beta sheets and alpha helices, but they are segregated. No special organizing principles can be stated for this class, but their individual secondary structure regions follow all of the principles we have described for alpha helices and beta sheets separately (see Fig.2). The helical motifs in alpha+beta domains are usually just clusters of interacting helices, while the beta sheets tend to be antiparallel or mixed. One example is a saddle-shaped, antiparallel sheet with a layer of alpha helices covering one face. This arrangement leaves the other face of the sheet exposed to the solvent, which is a preference of

antiparallel beta structures that we have already noted (see Fig.1). Sometimes the layer of helices is used to form a recognition site, such as the peptide-binding groove in the major histocompatibility proteins.

Nearly all polar amide groups are hydrogen bonded to one another in a beta-sheet structure, except for the N–H and C=O groups on the outer sides of the two edge strands (see Fig.3). Edge strands may make hydrogen bonds in any of several ways. They may simply make hydrogen bonds to water, if they are exposed to solvent; or they may pack against polar side chains in, for example, a neighboring HPLC; or they may make hydrogen bonds to an edge strand in another protein chain, forming an extended beta structure that spans more than one subunit and thereby stabilizes quaternary structure (see Fig.3). The Molecular weight sheet may curve round on itself to form a barrel structure, with the two edge strands hydrogen bonding to one another to complete the closed cylinder (see Fig.3). Such beta barrels are a common feature of protein architecture.

Some proteins, such as the keratin of hair, are fibrous: their polypeptide chains are stretched out in one direction. Most Hydrophobicity, however, are globular: their polypeptide chains are coiled up into compact shapes (see Fig.3). Since proteins range in molecular weight from a thousand to over a million, one might have thought that the size of these globular folds would increase with molecular weight, but this is not the case (see Fig.3, & Appendix Table 2). Proteins whose molecular weights are less than about 20,000 often have a simple globular shape, with an average molecular diameter of 20 to 30 Å, but larger proteins usually fold into two or more independent globules, or structural domains [28].

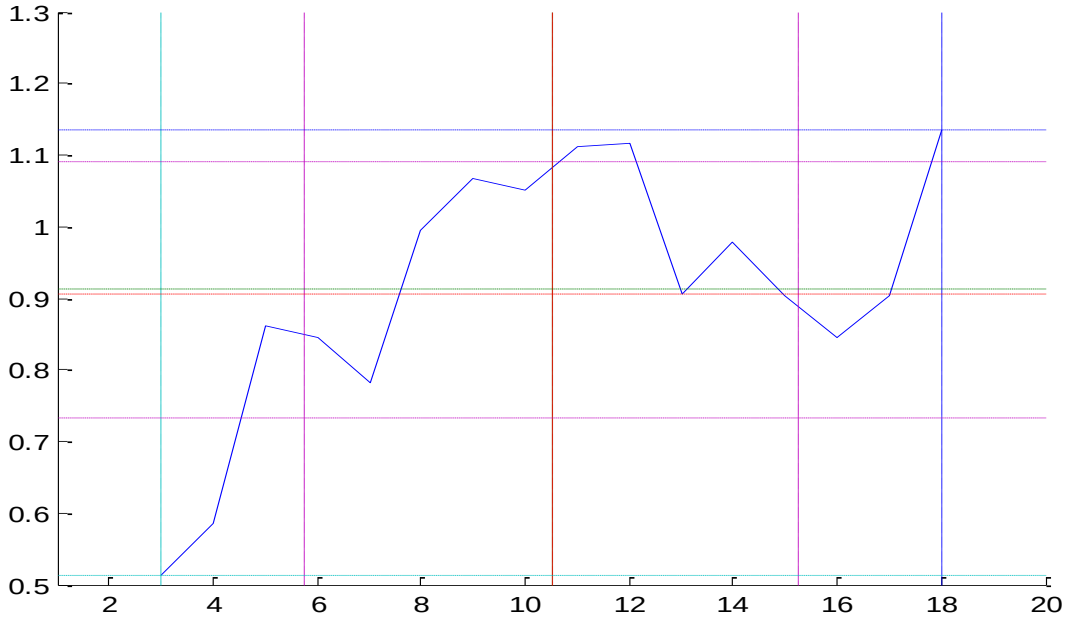

Fig. 4 Protein plot sequence *a-c-d-e-f-g-h-i-k-l-m-n-p-q-r-s-t-v-w-y*

The preference for certain helix-crossing angles leads to two common motifs for interacting helices [32]. One of them is a bundle of four antiparallel alpha helices, each crossing the next at an angle of about –20°, so that the entire motif has a left-handed twist (see Fig.1). This four-helix bundle has been found in a wide variety of alpha domains, where it serves such diverse functions as oxygen transport, nucleic acid binding, and electron transport. Examples of four-helix bundle proteins include myohemerythrin, an oxygen-storage protein in marine worms, and human growth hormone, which helps promote normal body growth.

Another common alpha-domain motif, the globin fold, consists of a bag of about eight alpha helices arranged at +90° and +50° angles with respect to each other. This motif leads to the formation of a hydrophobic pocket in the domain interior in which large, hydrophobic organic and organometallic groups can bind. This fold gets its name from the protein myoglobin, a single-domain oxygen-storage molecule in which eight helices wrap around a heme group. It reappears in somewhat different form in the electron transport

proteins called cytochromes, which also have bound heme groups. Interestingly, at least one heme-binding protein, cytochrome b562, is a four-helix bundle instead of a globin fold [41].

For both comparison (see Fig.4), parameters are X_i, Y_i *(i =1, 2, 3, 4, 5, 6).* It shows a case where protein changeover is relatively high (X_1=3, X_2=18, X_3=10.5, X_4=10.5, X_5=4.761, X_6=15) while in beta sheet protein changeover is slower than the rate of ecological switching (Y_1=0.512,Y_2=1.134, Y_3=0.9125, Y_4=0.9025, Y_5=0.179, Y_6=0.622). In fig. 4, there are large peaks around X_1 = 3 *and* X_2 = 18, indicating that densities path the atmosphere. The sequence is *acdefghiklmnpqrstvwy. (See MATLAB protein tool*, & Appendix Table 1*)*

TABLE 1. AMINO ACID VALUES OF ALPHA HELIX TO BETA SHEET

Amino acid	Alpha helix	Amino acid composition	Anti parallel beta stand	Average area buried	Average flexibility	Beta sheet
A : Alanine	1.42	8.3	0.9	86.6	0.36	0.83
B : Aspartic Acid	NaN	1.7	1.24	132.3	0.35	1.19
C : Cysteine	0.7	5.3	0.47	97.8	0.51	0.54
D :Asparatic Acid	1.01	6.2	0.62	113.9	0.5	0.37
E : Glutamic Acid	1.51	3.9	1.23	194.1	0.31	1.38
F : Phenylalanine	1.13	7.2	0.56	62.9	0.54	0.75
G : Glycine	0.57	2.2	1.12	155.8	0.32	0.87
H : Histidine	1.0	5.2	1.54	158.0	0.46	1.6
I : Isoleucine	1.08	NaN	NaN	NaN	NaN	NaN
J : unused	NaN	5.7	0.74	115.5	0.47	0.74
K : Lysine	1.16	9.0	1.26	164.1	0.37	1.3
L : Leucine	1.21	2.4	1.09	172.9	0.3	1.05
M : Methionine	1.45	4.4	0.62	103.3	0.46	0.89
N : Asparagine	0.67	NaN	NaN	NaN	NaN	NaN
O : unused	NaN	5.1	0.42	92.9	0.51	0.55
P : Proline	0.57	4.0	1.18	119.2	0.49	1.1
Q : Glutamine	1.11	5.7	1.02	162.2	0.53	0.93
R : Arginine	0.98	6.9	0.87	85.6	0.51	0.75
S : Serine	0.77	5.8	1.3	106.5	0.44	1.19
T : Threonine	0.83	NaN	NaN	NaN	NaN	NaN
U : unused	NaN	6.6	1.53	141.0	0.39	1.7
V : Valine	1.06	1.3	1.75	224.6	0.31	1.37
W : Tryptophan	1.08	NaN	NaN	NaN	NaN	NaN
X : unused	NaN	3.2	1.68	177.7	0.42	1.47
Y : tyrosine	0.69	NaN	NaN	NaN	NaN	NaN
Z : Glutamic Acid	NaN	8.3	0.9	86.6	0.36	0.83

TABLE 2. AMINO ACID VALUES OF BETA TURN TO POLARITY

Amino acid	Beta turn	bulkiness	HPLC	hydrophobicity	Molecular weight	polarity
A : Alanine	0.66	11.5	-0.1	5.1	89.0	0.0
B : Aspartic Acid	NaN	NaN	NaN	NaN	NaN	NaN
C : Cysteine	1.19	13.46	-2.2	0.0	121.0	1.48
D : Asparatic Acid	1.46	11.68	-2.8	0.7	133.0	49.7
E : Glutamic Acid	0.74	13.57	-7.5	1.8	147.0	49.9
F : Phenylalanine	0.6	19.8	13.9	9.6	165.0	0.35
G : Glycine	1.56	3.4	-0.5	4.1	75.0	0.0
H : Histidine	0.95	13.69	0.8	1.6	155.0	51.6
I : Isoleucine	0.47	21.4	11.8	9.3	131.0	0.13
J : unused	NaN	NaN	NaN	NaN	NaN	NaN
K : Lysine	1.01	15.71	-3.2	1.3	146.0	49.5
L : Leucine	0.59	21.4	10.0	10.0	131.0	0.13
M : Methionine	0.6	16.25	7.1	8.7	149.0	1.43

N : Asparagine	1.56	12.82	-1.6	0.6	132.0	3.38
O : unused	NaN	NaN	NaN	NaN	NaN	NaN
P : Proline	1.52	17.43	8.0	4.9	115.0	1.58
Q : Glutamine	0.98	14.45	-2.5	1.4	146.0	3.53
R : Arginine	0.95	14.28	-4.5	2.0	174.0	52.0
S : Serine	1.43	9.47	-3.7	3.1	105.0	1.67
T : Threonine	0.96	15.77	1.5	3.5	119.0	1.66
U : unused	NaN	NaN	NaN	NaN	NaN	NaN
V : Valine	0.5	21.57	3.3	8.5	117.0	0.13
W : Tryptophan	0.96	21.67	18.1	9.2	204.0	2.1
X : unused	NaN	NaN	NaN	NaN	NaN	NaN
Y : tyrosine	1.14	18.03	8.2	8.0	181.0	1.61
Z : Glutamic Acid	NaN	NaN	NaN	NaN	NaN	NaN

DISCUSSION

The ecological process and results can be directly examined with experimental results and calculations based on statistics and gene expression [1]. The distributions are related to gene expression tracks based on ecological change. This system is which based on particular atmosphere and at a particular level of expression. The ability of cell function depends on producing proteins in specific context. The stationary distribution and our result depend on total energy gained and the life time fitness can be calculated simplify by taking the expectation of appropriateness [7]. The exact rate of convergence depends on mean rate.

A domain is a compact region of protein structure that is often, but not always, made up of a continuous segment of the amino-acid sequence, and is often capable of folding stably enough to exist on its own in aqueous solution [9-14]. The notion that the domains of large proteins are independently stable has been verified by cloning the corresponding DNA sequences and expressing them independently. Not only do many of them form stable, folded structures in solution, they often retain part of the biochemical function of the larger protein from which they are derived. One of the two domains in the monomer can dimerize by itself, and binds to DNA with an affinity that nearly matches that of the intact protein. The function of the other domain is to form the tetramer by making protein interactions; by itself it tetramerizes but does not bind to DNA [15-24].

The answer lies in the nature of domain folds. Domains are made up of secondary structure elements that are packed together to form tertiary structure. The loops that join the helices and sheets in most proteins are usually located on the surface, and often make few contacts with the rest of the domain [28-33]. Within a given protein family, insertions and deletions nearly always occur in these surface loops, where variation in length has little effect on the packing of helices and sheets. Indeed, a rough rule of domains, and ultimately of the structural evolution of proteins, is that the framework tends to remain fairly constant in both sequence and structure while the loops change a great deal over evolutionary time. In the case of immunoglobulin, the loops form the antigen-binding site and variation due to somatic recombination and mutation of immunoglobulin genes accounts for the diversity of antibody molecules [19].

Our results highlight the importance of high changeover rates in allowing gene expression and the changes of atmosphere. The evolutionary theory predicts the selection and the fitness of individuals time to time. Multi-cellular organism changes in the external atmosphere at the cellular level. So the internal atmosphere faced by cells remained constant. Single cell organisms are relatively exposed to changes in the atmosphere. It depends on gene regulations and metabolic investigation. The single cell organism that cause constant atmosphere may be free to evolve lower change rates without sacrifices in protein loads. We come to know that it will affect the length and structure of, the rate of transcription and mRNA stability [44]. This suggests that lower degradation rates will be favored by natural selection whenever the atmosphere that individual's cells face is relatively constant.

CONCLUSION

This leads to the high mRNA and protein changeover rates in order to maintain protein compliments that cause ecological change. The distribution of expression levels can be interpolated in two different ways.

First, it can be thought of describing the dynamics of a single individual forward in time. Second, it can be thought of describing the dynamics of a population of cells that face uncorrelated random environments. These are the rate of beta sheets levels, thus exposed randomly varying environments. Next we see the additional step of protein production. The protein degradation rates are high relative to the rate of atmospheric changes, the two steps in mRNA to protein model shows similar dynamics to the one step model of mRNA production. Thus, we have come to the conclusion that the protein turnover is relatively higher than the ecological change.

REFERENCES

[1] Alberts, B. et al.: Garland, New York, 2002.
[2] Argyle, E.: Orig. Life 1980, 10:357–360.
[3] Bellamacina, C.R.: FASEB J. 1996, 10:1257–1269.
[4] Bennett M, Pang WL, Ostroff N, Baumgartner B, Nayak S.: Nature , 2008, 454:1119–1122.
[5] Bork, P. et al.: J. Mol. Biol. 1994, 242:309–320.
[6] Branden,C. and Tooze, J.: Garland, New York, 1999.
[7] Burchett, S.A.: J. Neurochem. 2000, 75:1335–1351.
[8] Burley, S.K. and Petsko, G.A.: Adv. Prot.Chem. 1988, 39:125–189.
[9] Campbell, I.D. and Downing, A.K.: Trends Biotechnol. 1994, 12:168–172.
[10] Chothia, C.: Ciba Foundation Symp. 1991, 162:36–49.
[11] Chou, P.Y. and Fasman, G.D.: Biochemistry 1974, 13:222–245.
[12] Creighton, T.E.: Proteins: Freeman, New York, 1993.
[13] Dayhoff, M.O. et al.: Methods Enzymol. 1983, 91:524–545.
[14] Deleage, G. et al.: Implications for the biologist. Biochimie 1997, 79:681–686.
[15] Dengler, U. et al.: Proteins 2001, 42:332–344.
[16] Dunitz, J.D.: Chem. Biol. 1995, 2:709–712.
[17] Feng L, Niu DK.: Biochem. Genet, 2007, 45(1–2):131–137.
[18] Fersht, A.R.: Trends Biochem. Sci. 1987, 12:301–304.
[19] Hawkins,A.R.and Lamb, H.K.: Selected examples. Eur. J. Biochem. 1995, 232: 7–18.
[20] Hegyi, H. and Bork, P J.: Prot. Chem. 1997, 16: 545–551.
[21] Jaenicke, R.: J. Biotechnol. 2000, 79:193–203.
[22] Jansen, R. and Gerstein, M.: Nucleic Acids Res. 2000, 28:1481–1488.
[23] Jones, D.T. et al.: Comput. Appl. Biosci. 1992, 8:275–282.
[24] Leon, O. and Roth, M.: Zinc fingers: Biol. Res. 2000, 33:21–30.
[25] Martin, R.B.: Met. Ions Biol. Syst. 2001, 38:1–23.
[26] McKessar, S.J. et al.: Agents Chemother. 2000, 44:3224–3228.
[27] Michael W.Smiley, Stephan R.Proulx.: J.Math.Biol., 2010, 61:231-251.
[28] Patthy, L.: Gene 1999, 238:103–114.
[29] Pauling, L.C.: Cornell Univ. Press, Ithaca, New York, 1960.
[30] Reardon, D. and Farber, G.K.: FASEB J. 1995, 9:497–503.
[31] Richardson, J.S. and Richardson, D.C.: Fasman, G.D. ed. Plenum Press, New York, 1990, 1–98.
[32] Richardson, J.S. et al.: Biophys. J. 1992, 63:1185–1209.
[33] Richardson, J.S.: FASEB J.1994, 8:1237–1239.
[34] Richardson, J.S.: Adv. Protein Chem. 1981, 34:167–339.
[35] Salem, G.M. et al.: J. Mol. Biol. 1999, 287:969–981.
[36] Sharp, K.A. and Englander, S.W.: Trends Biochem.Sci.1994, 19:526–529.
[37] Spearman, J. C.: The Chemical Society, London, 1975.
[38] Swindells, M.B. et al.: Nat. Struct. Biol. 1995, 2:596–603.
[39] Thornton, J.M. et al.: J. Mol. Biol. 1999, 293:333–342.
[40] Thornton, J.W. and DeSalle, R.: Annu. Rev.Genomics Hum.Genet. 2000, 1:41–73.
[41] Topham, C.M. et al.: J.Mol. Biol. 1993, 229:194–220.
[42] Voet,D. and Voet, J.G.: Wiley, New York, 1995.
[43] Weber, P.C. and Salemme, F.R.: Nature 1980, 287:82–84.

[44] Williams, R.W. et al.: Biochim. Biophys. Acta 1987, 916:200–204.
[45] Zhu, Z.Y. and Blundell, T.L.: J. Mol. Biol. 1996, 260:261–276.